SIMULATION OF PROTEIN MOLECULE FLUCTUATIONS BY IRREGULAR IMPEDANCE NETWORK METHOD

V.I. AKIMOV, A.V. SHESTOPALOVA[1], V.M. TULUPENKO, L. REGGIANI[2]

UDC 577

Donbass State Engineering Academy
(72, Shkadinov Str., Kramatorsk 84313, Ukraine; e-mail: avi@laser.donetsk.ua),

[1]**A. Usikov Institute for Radiophysics and Electronics, Nat. Acad. Sci. of Ukraine**
(12, Academician Proskura Str., Kharkiv 61085, Ukraine),

[2]**Dipartimento di Ingegneria dell'Innovazione, Università di Lecce**
(Via Arnesano, Lecce 73100, Italy)

The method of irregular impedance network developed by the authors has been applied to studying the electrical fluctuations of a protein molecule located between two ohmic contacts. Specific calculations have been carried out for a bovine rhodopsin molecule. Random fluctuations of distances between amino acids (the link oscillation model) or random fluctuations of amino acids themselves about relevant fixed positions (the node oscillation model) have been analyzed. The mean network impedance and its dispersion have been calculated as functions of the fluctuation amplitude. The similarity and the difference between the results obtained in the framework of these two models have been discussed. A universal dependence of the impedance dispersion on the oscillation amplitude has been found.

1. Introduction

The forecasting and the interpretation of structural modifications in protein molecules ("protein folding") belong to the fundamental problems of modern biophysics, but they are far from being ultimately solved [1, 2]. For instance, the fluctuation dynamics of protein molecules was described in work [3] in the framework of the model of a network of elementary elastic linkers. On the other hand, complex networks made up of resistances and/or impedances comprise a useful and rather widespread approach to the simulation of a lot of phenomena [4], in particular, in the framework of the so-called percolation theory [5]. Earlier, we have proposed an analogous method for simulating the electric properties of a protein molecule [6] in order to solve the problem of creation of a nano-sized biosensor on the basis of an olfactory receptor molecule [7]. In this method, the molecule is considered as a set of nodes, which correspond to amino acids, and linkers, which represent the electric interaction between amino acids. Linkers are characterized by elementary impedances which contribute to the total network impedance. In this work, in the framework of the method described above, two complementary models of internal fluctuations of the total network impedance have been proposed and analyzed (the fluctuations are associated with oscillations of elementary components), as well as the topology of a structure. As a source that causes the values of elementary impedances to oscillate, there were selected the oscillations of either the lengths of linkers between nodes – the link oscillation model (LOM) – or the nodes themselves about relevant fixed positions – the node oscillation model (NOM). The models have been applied to a network which corresponds to a rhodopsin molecule located between two ohmic contacts. The dependences of the key input parameters (the link number, the impedance, and so on) on the model ones have been studied. Our main goals were the attempt to single out individual properties of the molecule of a specific protein (or its configuration) on the basis of the models proposed and searching for the ways to calibrate the model parameters.

2. Model of the Irregular Impedance Network and Its Application to a Protein Molecule

We consider the spatial structure of a protein molecule connecting two metal electrodes [6], which an external voltage is applied to, as is shown in Fig. 1,*a*. An elementary RC impedance is attributed to every pair of amino acids that are located from each other at a distance no more than a twice interaction radius $2R_a$ (Fig. 1.*b*). In such a way, a network of impedances [7,8], which simulates the electric properties of the rhodopsin molecule [9], is constructed. The total impedance of the network depends on the interaction radius R_a which is a model parameter and can be calculated making use of the Kirchhoff laws, if the network topology and the values of elementary impedances are known.

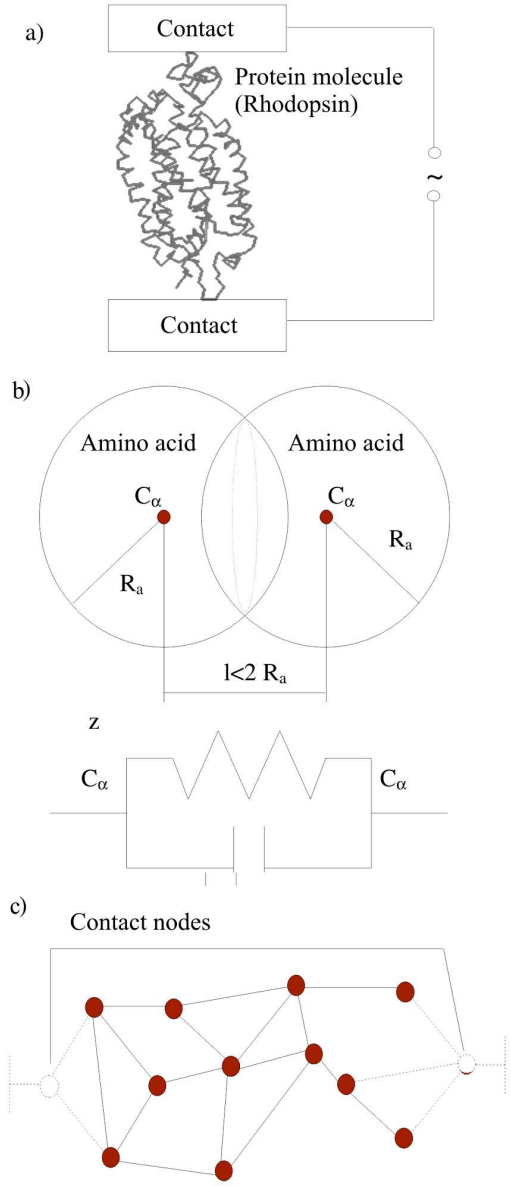

Fig. 1. *a*) A protein molecule between contacts. *b*) Overlapping between two amino acids and its equivalent circuit. C_a denotes the center of the sphere and corresponds to the alpha-carbon atom of the amino acid, the parameter R_a is the radius of interaction between amino acids, z is the impedance of the link between two nodes. *c*) An example of a graph for a network of elementary impedances

3. Fluctuation Models

We propose to consider two models of electric fluctuations in such a network. Both models use the distribution of probabilities over the coordinate of a classical harmonic oscillator:

– In the LOM, the network changes its characteristics – such as the link number and the impedance – through an independent variation of the effective distance between nodes in every pair.

– In the NOM, the characteristics are varied by changing isotropically the position of every node in the three-dimensional space.

3.1. Link oscillation model

In the LOM, the current state of the molecule can be entirely described by a matrix $\mathbf{D}_0$ of distances between nodes of each pair. Hence, $\mathbf{D}_0$ is a symmetric square matrix with the dimension equal to the number of amino acid residues in the molecule concerned; for rhodopsin, this is a 348×348-matrix. In order to calculate the total impedance of the structure, each distance is compared with a selected interaction length, and, if this distance does not exceed $2R_a$, a link between corresponding nodes is established. In order to describe fluctuations, a computer generates a series of $\mathbf{D}$-matrices on the basis of the ground state matrix $\mathbf{D}_0$; in so doing, the shift of each distance is selected randomly, on the basis of the probability distribution for a classical harmonic oscillator.

The state of a classical harmonic oscillator is described by the formula

$$x = x_{\max} \sin \omega t, \tag{1}$$

where $x_{\max}$ is the amplitude, and ω the angular frequency of oscillations. The probability for the coordinate x to have a value x_a is reciprocal to the rate of its variation at the corresponding point:

$$P(x_a) \sim \left(\frac{\partial x}{\partial t} \right)^{-1} \Bigg|_{x=x_a}. \tag{2}$$

For numerical simulations, we should obtain a dependence between a random number r, which is regularly distributed within the interval from 0 to 1, and the quantity x_a:

$$r = \int_{-x_{\max}}^{x_a} P(x)dx \quad / \int_{-x_{\max}}^{x_{\max}} P(x)dx. \tag{3}$$

By expressing t in terms of x in Eq. (1), substituting Eq. (1) into Eq. (2) and the result obtained into Eq. (3), and expressing x_a in terms of r, we obtain $x_a(r) = -x_{\max} \cos \pi r$ or, equivalently,

$$x_a(r) = x_{\max} \cos \pi r. \tag{4}$$

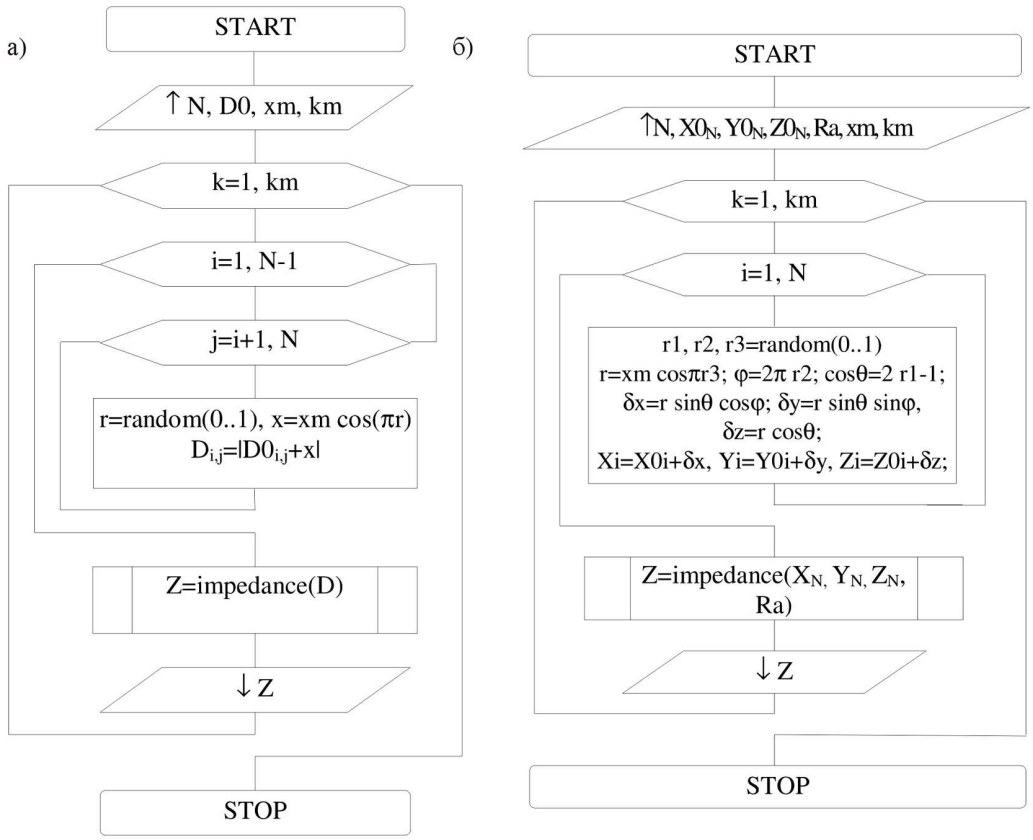

Fig 2. *a*) Block diagram of the link oscillation model: N is the node number, km the number of iterations, $D0$ the matrix of initial distances, D the matrix of current distances, and xm the amplitude of oscillations. *b*) Block diagram of the node oscillation model: $X0_N$, $Y0_N$, and $Z0_N$ the arrays of initial node coordinates; X_N, Y_N, and Z_N the arrays of current node coordinates; and R_a the maximal radius of interaction

Equation (4) can be used for the random choice of a state of the harmonic oscillator. Hence, the effective link length is determined as

$$l = |l_0 + x_a(r)| = |l_0 + x_{\max} \cos \pi r|. \tag{5}$$

Here, l_0 is the distance between amino acids (different for each amino acid pair) in the initial state of the molecule which is described by the matrix $\mathbf{D}_0$.

The next step of calculations includes the filling of the matrix $\mathbf{D}$ with the values of the effective lengths and the evaluation of the key characteristics of the structure – the number of links and the network impedance. The latter is calculated on the basis of the Kirchhoff laws and the known values of elementary impedances. Two models of elementary impedances were used: a steep – $z = \text{const}$ – and a smooth one – $\left(z = \frac{1}{\pi} \frac{l}{R_a^2 - (l^2/4)} \frac{1}{\rho^{-1} + i\varepsilon_0 \varepsilon \omega}\right)$ (see Fig. 1). Here, ρ is the specific resistance, ε is the dielectric permittivity (both are selected as in work [6]), and i is the imaginary unity. To calculate the required average values, such iterations are repeated until the necessary statistics is gathered (5 thousand iterations per every point). The details of the described Monte-Carlo procedure are illustrated by the block diagram of calculations (Fig. 2,*a*). Afterwards, the obtained set of values is treated statistically, and, as a result, the averaged values of key characteristics and the root-mean-square deviations are calculated.

In the case of thermal fluctuations, the amplitude of an oscillator, according to the classical statistics, is proportional to the square root of the temperature:

$$x_{\max} = A\sqrt{T}, \tag{6}$$

where A is a parameter. In principle, the value of A depends on the elastic properties of the whole structure; hence, it depends on the distance between amino acids. Nevertheless, as the first approximation, the value of A can be considered constant. Since the

absolute value of the impedance in our model can only decrease with increase in the interaction radius (see the formulas for the elementary impedance), the structures corresponding to the minimum or the maximum of the impedance can be created. However, the probability of the natural generation of such structures is very low and quickly vanishes, if the number of nodes grows.

It is worth noting that the approach proposed allows one to find – in both models (Fig. 2) – only statistically averaged values for the quantities concerned. However, the temporal characteristics of the noise generated by structure fluctuations cannot be determined, because every following state is calculated irrespective of the previous one; so that it is the statistics of entirely independent states, rather than the temporal dynamics, that is simulated. For the same reason, the application of a similar procedure does not allow one to simulate the dynamics of structural variations of a protein molecule. In the future, the method can be improved, if necessary, by introducing the parameter of elasticity for the interaction between amino acids of every pair (in the LOM case) and, correspondingly, by introducing the individual values of angular frequency for every oscillator, which would allow the process to be simulated in dynamics.

3.2. Node oscillation model

In this model, at every step, each node is described by a random shift in a random isotropically (i.e. uniformly distributed over the solid angle) selected direction around a relevant static position. Then, every new state of the molecule is described by the arrays including the absolute coordinates (X, Y, Z) of the nodes. For every i-th node, $X_i = X0_i + \delta x$, $Y_i = Y0_i + \delta y$, and $Z_i = Z0_i + \delta z$. Here, the shifts δx, δy, and δz are randomly selected in the spherical coordinates: $\delta x = r \sin\theta \cos\varphi$; $\delta y = r \sin\theta \sin\varphi$, and $\delta z = r \cos\theta$; the angles satisfy the relations $\varphi = 2\pi r_1$ and $\cos\theta = 2r_2 - 1$; and the absolute shift, according to formula (4), equals $r = r_m \cos \pi r_3$. The arrays $X0$, $Y0$, and $Z0$ describe the initial state; r_1, r_2, and r_3 are random numbers from the interval $[0, 1]$. Then, on the basis of the selected R_a and a new set of nodes, the matrix of states is constructed, and the network impedance is determined. The further steps of the algorithm are the same as those in the LOM case. For details, see the block diagram in Fig. 2,*b*.

4. Results of Averaging

In this section, we consider the quantities characterizing the network – namely, the link number and the absolute value of network impedance – averaged over all randomly selected states, as well as their dependences on the amplitude and the radius of interaction in two models. In the LOM, the average link number can be derived analytically. Proceeding from formula (5) and taking into account that a link exists if $l \leq 2 \times R_a$, it can be shown that the probability for a link to exist is determined as

$$P_l = \frac{1}{\pi} \left(\arcsin c_2 - \arcsin c_1 \right), \tag{7}$$

where $c_1 = \min\left[1, \max\left(\frac{-2R_a - l_0}{x_{\max}}, -1\right)\right]$ and $c_2 = \max\left[-1, \min\left(\frac{2R_a - l_0}{x_{\max}}, 1\right)\right]$, and each of the functions min and max depends on two arguments. Accordingly, the expected average link number is equal to the sum of P_l over all pairs of nodes. Figure 3 makes it evident that the link numbers obtained by these formulas entirely coincide with the corresponding values averaged by the Monte-Carlo method described in the previous section. This testifies to both the validity of formula (7) and the sufficiency of the statistical sample for the fulfilled numerical simulation. Note that the link number per one amplitude decreases with increase in the interaction radius.

In the NOM, the link number was calculated making use of the Monte-Carlo method only. From Fig. 3, one can see that now, in contrast to the LOM, the average link number decreases with the growth of interaction amplitude and increases with increase in the interaction radius.

The calculated dependences of the average impedance on the interaction amplitude in both fluctuation models and for two models of elementary impedance (the smooth and the steep one, see Fig. 1) are depicted in Fig. 4. One can see that the average impedance of a structure, which undergoes fluctuations, can differ from the impedance of the initial stable structure. The averaged impedance considerably decreases in the LOM case and substantially increases in the NOM one. Since the amplitude $x_{\max}$ depends on the temperature, the plots in Fig. 4 can be interpreted as a forecast of the temperature dependence of the molecular impedance. One can see that, in the NOM case, the dependences were obtained only in the range $x_{\max} < R_a$. This stems from the fact that, in this model, there appears the appreciable probability for an open random

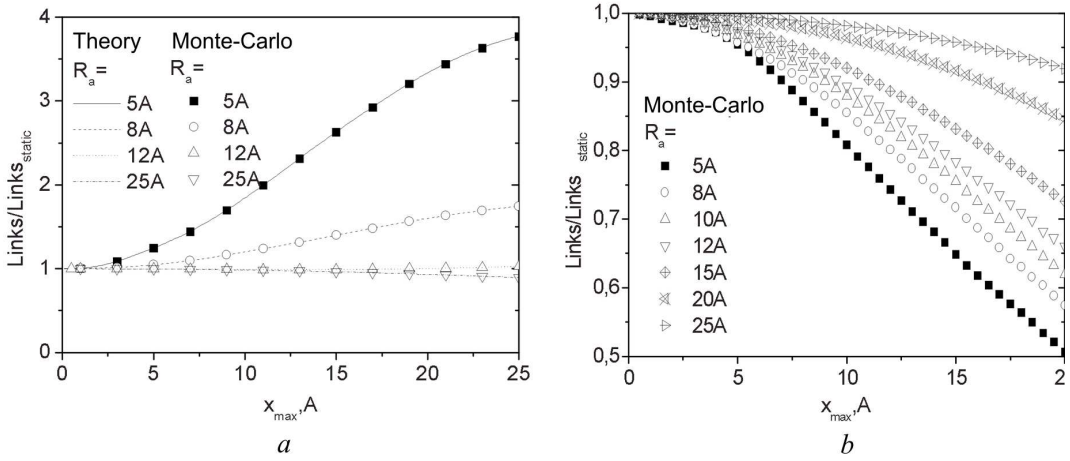

Fig. 3. Dependences of the normalized average link number on the amplitude for various values of interaction radius R_a

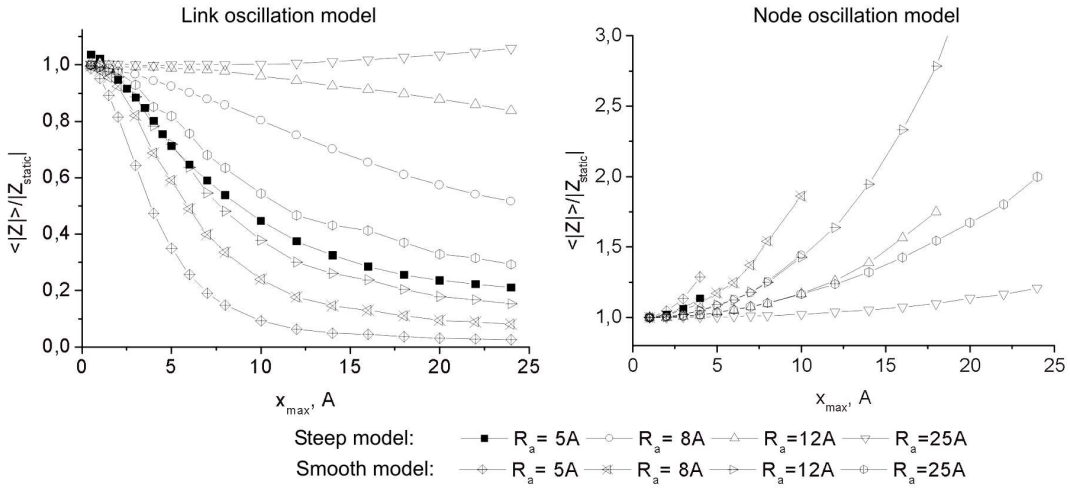

Fig. 4. Dependences of the average relative impedance on the amplitude for various values of interaction radius and different elementary impedance models

structure to be generated, so that we failed in gathering a large enough statistics correctly.

5. Dispersion of Impedance Fluctuations

In this section, we consider the dependences of the fluctuation dispersion (the root-mean-square deviation) of the network impedance (the total impedance of the structure) on the interaction amplitude and the radius for two models of fluctuations and two models of elementary impedance. The obtained dependences of the relative dispersion on the ratio $x_{\max}/R_a$ are shown in Fig. 5. One can see that, in the LOM and the smooth model of elementary impedance, the dependences of the relative dispersion of the network impedance Z are in good agreement with the formula $\frac{\langle \delta^2 |Z| \rangle}{\langle |Z| \rangle^2} = \left(\frac{x_{\max}}{5R_a} \right)^2$, provided that the $x_{\max}$-values are small. At the same time, in the steep model, the same dependences behave in different and unpredictable ways.

One of the reasons for such a difference is the presence of a random telegraph signal (as one of the terms) stimulated by the availability of "bottle necks"in the topology of the structure under consideration. As corresponding factors, such links can be those which give a significant contribution to the network impedance at given R_a-values and, at the same time, can be either switched on (if $l \leq 2R_a$) or switched off (if $l > 2R_a$) at given $x_{\max}$-values in the course of fluctuations. They can enhance the dispersion, especially at small $x_{\max}$, depending on the structure features.

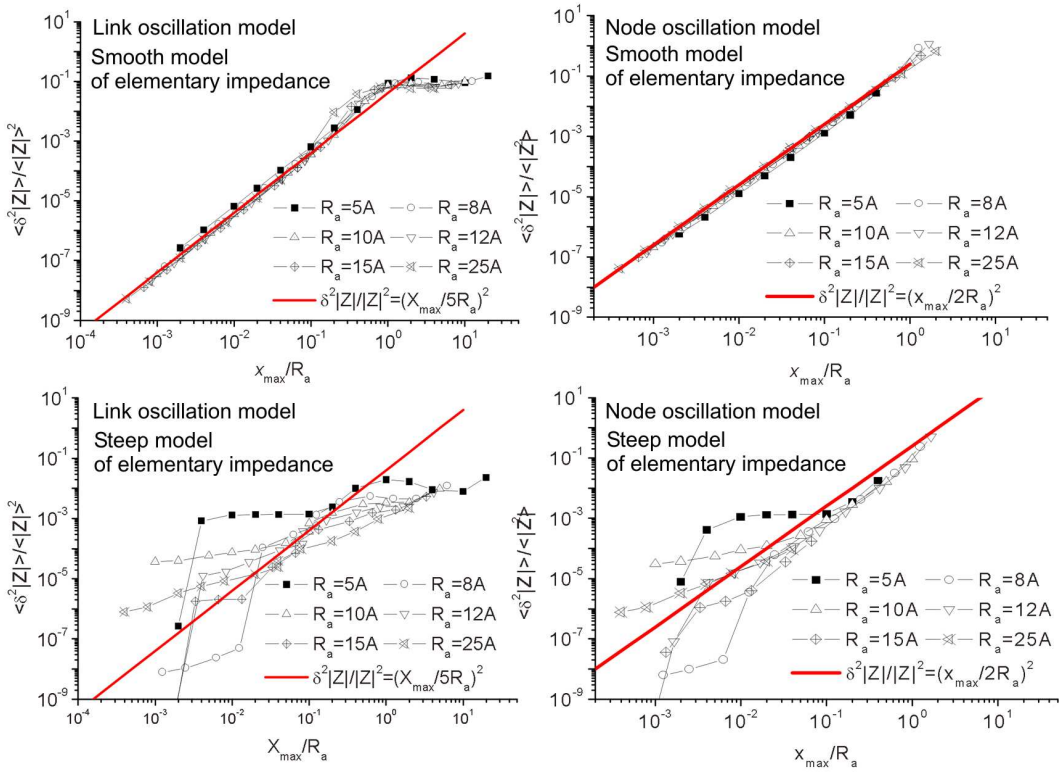

Fig 5. Dependences of the relative impedance dispersion on the ratio $x_{\max}/R_a$ for different models

Another reason is a weaker sensitivity of the smooth model of elementary impedance to fluctuations; it dominates, if $x_{\max}$ is large. Hence, the steep model is more sensitive to features of the structure topology and can be of interest from the viewpoint of recognition of various structures belonging to a definite type, e.g., different structural variants of a protein molecule. The smooth model of elementary impedance ensures a more general and universal behavior which is mainly associated with a whole class of similar structures.

In the NOM, a universal law for the smooth model of elementary impedances was also revealed: the relative dispersion behavior satisfies well the formula $\frac{\langle \delta^2 |Z| \rangle}{\langle |Z| \rangle^2} = \left(\frac{x_{\max}}{2R_a} \right)^2$. It is the same law as that in the LOM, but with a different multiplier. Such a difference can be explained, first, by the nonequivalence of the parameter $x_{\max}$ in two models: a symmetric shift of nodes in opposite directions by x generates a change $2x$ for the link length. Second, the number of links in the considered range of the interaction radii is several times larger than the number of nodes; therefore, the LOM provides a better averaging of the extreme values of elementary shifts in the total impedance than the NOM does and, hence, a lower impedance dispersion. This means that a larger shift of a node brings about larger shifts of several links connected with this node. Therefore, the NOM – in comparison with the LOM – produces a sort of the cumulative effect.

The steep model, similarly to the NOM, is characterized by an unpredictable individual behavior caused by features of the topology of a specific structure. It should be noted that, as the plots in Fig. 5 demonstrate, the relevant features of the curves obtained, in the framework of the steep model, for the LOM and the NOM qualitatively correlate with one another (making relevant corrections to the universal law). This means that the determination of specific structural features is robust with respect to the selection of the elementary impedance model.

6. Conclusions

The fluctuation models have been developed in the framework of the elementary impedance network, and their application to the protein molecule of rhodopsin located between two ohmic contacts has been considered.

The results of the corresponding simulation are as follows.

The specific features of a definite structure, which allow different molecules or different states of the same molecule to be distinguished, are most pronounced in the dependences of the total impedance dispersion on the amplitude of oscillations (the temperature dependences) which are calculated in the framework of the steep model for an elementary impedance. It has also been revealed that the dependences of the impedance dispersion demonstrate a similar behavior in different fluctuation models, which evidences for a stability of the dispersion concerned to the type of a fluctuation model, and can be described by a universal law for the smooth model of elementary impedance.

The averaged value of the link number reveals the opposite dependences on the amplitude in the link and node oscillation models and, at the same time, a similar behavior for the smooth and steep elementary impedance models.

The method can be made more accurate and allows a further development.

1. C.B. Anfinsen, Science **181**, 223 (1973).
2. M.J. Gething and J. Sambrook, Nature **355**, 33 (1992).
3. A.R. Atilgan, S.R. Durell, R.L. Jernigan, M.C. Demirel, O. Keskin, and I. Bahar, Biophys. J. **80**, 505 (2001).
4. R. Albert and A.-L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).
5. M. Sahimi, *Applications of Percolation Theory* (Taylor and Francis, London, 1994); G. Grimmett, *Percolation*, 2-nd edition (Springer, Berlin, 1999).
6. C. Pennetta, V. Akimov, E. Alfinito, L. Reggiani, T. Gorojankina, J. Minic, E. Pajot-Augy, M.A. Persuy, R. Salesse, I. Casuso, A. Errachid, G. Gomila, O. Ruiz, J. Samitier, Y. Hou, N. Jaffrezic, G. Ferrari, L. Fumagalli, and M. Sampietro, in *Nanotechnologies for the Life Sciences, Vol. 4: Nanodevices for the Life Sciences,* edited by S. Challa and S.R. Kumar (Wiley, New York, 2006), p. 221.
7. *SPOT-NOSED. Single Protein Nanobiosensor Grid Array. The 5-th Framework Programme of European Community,* http://www.nanobiolab.pcb.ub.es/projectes/Spotnosed/2003-2005.
8. C. Pennetta, V. Akimov, E. Alfinito, L. Reggiani, and G. Gomila, Proc. SPIE **5472**, 172 (2004).
9. K. Palczewski, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I. Le Trong, D.C. Teller, T. Okada, R.E. Stenkamp, M. Yamamoto, and M. Miyano, Science **289**, 739 (2000).

Received 19.09.07
Translated from Ukrainian by O.I. Voitenko

СПОСОБИ МОДЕЛЮВАННЯ ФЛУКТУАЦІЙ МОЛЕКУЛИ БІЛКА В РАМКАХ МЕТОДУ НЕРЕГУЛЯРНОЇ СІТКИ ІМПЕДАНСІВ

В.І. Акімов, А.В. Шестопалова, В.М. Тулупенко, Л. Реджиані

Р е з ю м е

Розроблений нами метод нерегулярної сітки імпедансів застосовано для дослідження електричних флуктуацій білкової молекули, розміщеної між омічними контактами. Розрахунки проведено для молекули коров'ячого родопсину, а як джерело флуктуацій прийнято випадкові осциляції відстаней між амінокислотами (модель зв'язків, що осцилюють) чи, як альтернатива, випадкові осциляції положення амінокислот в околі фіксованих точок (модель вузлів, що осцилюють). Розраховано середній імпеданс сітки і його дисперсію як функції амплітуди флуктуацій. Результат розрахунків для двох моделей порівняно між собою. Виявлено універсальний закон залежності дисперсії від амплітуди коливань.