Kinetic analysis of the search for damaged DNA bases by repair enzymes: theoretical investigation of diffusion-controlled steps*

N. N. Lukzen,[a] K. L. Ivanov,[a] and V. V. Koval[b]*

[a]*International Tomography Center, Siberian Branch of the Russian Academy of Sciences, 3a ul. Institutskaya, 630090 Novosibirsk, Russian Federation. Fax: +7 (383) 333 1399. E-mail: ivanov@tomo.nsc.ru*
[b]*Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, 8, prosp. Akad. Lavrent´eva, 630090 Novosibirsk, Russian Federation*

The kinetics of the search for damaged DNA bases by repair enzymes was theoretically studied. In particular, the characteristic features of the diffusive motion of the enzyme were investigated. Compact analytical expressions for the average search time and its dispersion were obtained for two models of the one-dimensional motion of the enzyme along the DNA fragment: continuous diffusion and stochastic jumps. It was shown that the kinetics of the search is essentially non-exponential due to characteristic features of the diffusive motion.

Key words: DNA repair, DNA damage, diffusion, kinetics.

Almost all biological processes in the cell begin with the binding of a protein (enzyme) to a specific nucleic acid site. It is this pathway that is repsonsible for the transcription of the genetic information encoded in the DNA and the retention of its integrity and stability. The question of how the protein finds a particular binding site has long been debated. It was suggested[1] that the efficacy of interactions in biological systems can be substantially increased with a decrease in the dimensionality of diffusion-controlled rate constants. This model was comprehensively developed during the last four decades,[2,3] which has resulted in the elaboration of several models for the description of the mechanism of the search for specific nucleic acid sites by a protein.

Nowadays, the BWH (Berg—Winter—von Hippel) model is the most well-developed and widely accepted one.[4—6] This model is based on the superposition of the one- and three-dimensional diffusion. It is commonly assumed that the search for a specific site by a protein consists of the diffusion of a protein molecule in solution and its one-dimensional motion along a nucleic acid molecule. The search process involves several sequential events: the protein binds to the DNA at a random site, moves along the DNA strand to search for the specific site, and, if the specific site is not found, dissociates from the DNA. After going to solution, the protein molecule can again bind at any site on the DNA with equal probability. The cyclic repetition of these events results in the finding of the specific site.

Besides, several search models were proposed: the electrostatic mechanism,[7,8] the mechanism of the simultaneous localization,[9,10] and the correlation model.[11,12]

The authors of the electrostatic mechanism quite reasonably assume that the experimentally observed high rate constants for the complex formation are achievable due to electrostatic interactions between charged molecules. According to the estimate,[11] the electrostatic interaction (in low-salt buffers) can occur at distances of ~2 nm. Hence, it can be concluded that this interaction is the final step resulting in the formation of the complex once its components approach each other.

An elegant approach is based on the colocalization model[9,10] for bacterial genes. An idea to decrease the number of cycles of the one- and three-dimensional diffusion in the BWH model led to an interesting hypothesis: bacterial genes encoding particular proteins are located in the genome adjacent to the binding sites of these proteins, which substantially reduces the search time. Unfortunately, this model does not work in eukaryotic cells, where the translation and transcription are separated in time and space. Besides, it does not explain the case of multiple protein-binding sites.

The correlation model[11] is the development of the BWH model. Like in the BWH model, the search process consists of successive steps of one- and three-dimensional diffusion. It is particularly important that the first step of three-dimensional diffusion at the instant of decomposition of the DNA—protein complex is considered as the one-dimensional diffusion. Hence, there is the probability that the protein will return on the target (at the same or

* Dedicated to Academician of the Russian Academy of Sciences R. Z. Sagdeev on the occasion of his 70th birthday.

Published in Russian in *Izvestiya Akademii Nauk. Seriya Khimicheskaya,* No. 12, pp. 2569—2572, December, 2011.
1066-5285/11/6012-2621

adjacent site) and the new search begins. In this approach, the fraction of the "slow" three-dimensional diffusion decreases.

All the above-mentioned approaches are well developed *in silico* and were often experimentally confirmed. The description of real biological systems presents serious problems.

The aim of this study was to develop a model for the description of the kinetics of the search for a damaged DNA base by repair enzymes. We will not restrict ourselves to the determination of the average search time for the damaged site and will try to find its dispersion and to directly study the kinetics of the repair. This statement of the problem is important because, in the case of the stochastic motion of the enzyme (one-dimensional jumps or continuous diffusion), the kinetics of the search should have a complex character, in particular, be essentially non-exponential. Since the complete solution of the problem (taking into account both the three- and one-dimensional diffusion, electrostatic interactions, *etc*.) presents difficulties, we restrict ourselves to the consideration of the step of one-dimensional diffusion of the enzyme along the damaged DNA fragment. In addition, let us assume that the mobility of the enzyme is independent of the direction of motion and its current site on the DNA fragment.

The scheme of the DNA repair through the catalytic repair of the damaged nucleotide by an enzyme is shown in Fig. 1. This process involves several steps, each resulting in conformational changes of the enzyme molecule:

$$\mathrm{E + S \leftrightarrow ES1 \leftrightarrow ES2 \leftrightarrow \ldots \leftrightarrow ES}k \leftrightarrow \ldots \leftrightarrow \mathrm{ES}n \rightarrow \mathrm{EP}, \quad (1)$$

where E is the enzyme, S is the DNA fragment, Sk is the complex of the enzyme with the kth nucleotide, and P is the product. Therefore, the enzyme jumps between successive nucleotides: once the enzyme recognizes the damaged base at the nth site, the ESn complex changes its conformational state and the enzyme repairs the DNA damage. The conformational changes and, consequently, successive repair steps can be studied, for example, by measuring intrinsic tryptophan fluorescence of the enzyme molecule.[13,14] The repair is preceded by the delivery of the enzyme from the bulk solvent to the damaged DNA fragment: E + S $\leftrightarrow$ ES1. The delivery of the enzyme occurs as the diffusion-controlled binding to the DNA or an oligonucleotide.

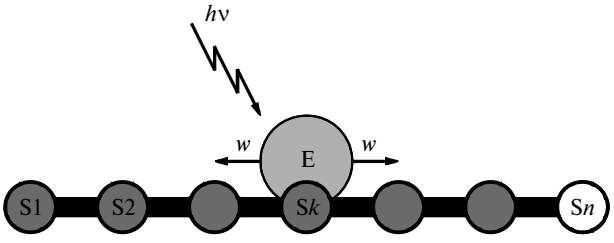

Fig. 1. Scheme of the enzymatic DNA repair; E is the enzyme; S1, S2, ..., Sk, ... are nucleotides of the DNA fragment, Sn is the damaged nucleotide; w is the jump rate to the adjacent base; intrinsic tryptophan fluorescence of the protein E is a standard method used to study conformational transitions.

Let us assume that the enzyme binds to the kth nucleotide and slides along the DNA *via* random jumps to adjacent nucleotides. The probability of the jump per unit time is denoted by w. Let us suppose that the recognition by the enzyme of the damaged nucleotide (the nth nucleotide) is accompanied by the fast conformational transition resulting in the irreversible binding of the enzyme to Sn and the repair of the damaged fragment. This model can be described by the following system of equations (C_i is the probability of the presence of the enzyme on the ith nucleotide):

$$\begin{cases} \mathrm{d}C_1/\mathrm{d}t = -wC_1 + wC_2, \\ \mathrm{d}C_i/\mathrm{d}t = -2wC_i + wC_{i-1} + wC_{i+1}, \\ \mathrm{d}C_{n-1}/\mathrm{d}t = -wC_{n-1} + wC_{n-2}, \\ C_n = 0. \end{cases} \quad (2)$$

The last equation corresponds to the very fast irreversible conformational transition in the ESn complex (*i.e.*, the search for the damaged nucleotide is completed). To solve the system of equations, we used the Laplace transform and the generating function formalism. By assuming that the enzyme starts from the kth site, the Laplace image of the initial system of equations takes the following form:

$$\begin{cases} sC_1 = -wC_1 + wC_2, \\ sC_i = -2wC_i + wC_{i-1} + wC_{i+1}, \\ sC_k - 1 = -2wC_k + wC_{k-1} + wC_{k+1}, \\ sC_{n-1} = -2wC_{n-1} + wC_{n-2}. \end{cases} \quad (3)$$

A set of Laplace images of the C_i values can be related to the generating function $f(z)$ of the complex variable z encoding information on all C_i values:

$$f(z) = C_1 + C_2z + C_3z + \ldots + C_{n-1}z^{n-2}. \quad (4)$$

The solution of system (3) for Laplace images gives the following analytical expression:

$$f(z) = -\frac{1}{w}\frac{1}{\lambda_2 - \lambda_1}\left(\frac{1}{z-\lambda_2} - \frac{1}{z-\lambda_1}\right) \cdot \left[w(z-1)C_1 + z^k - wz^nC_{n-1}\right], \quad (5)$$

where $\lambda_{1,2} = [(s + 2w) \pm s(s + 4w)]/(2w)$.

These expressions allow one to analytically calculate the C_{n-1} value and, consequently, the kinetic parameters of the repair reaction. For example, the reaction rate can be calculated as a "flux" of the enzyme $j(t)$ to the site n from the site $n - 1$:

$$j(t) = wC_{n-1}(t),$$

$$j(s) = \int_0^\infty j(t)e^{-st}\mathrm{d}t = wC_{n-1}(s) = \\ = \frac{\lambda_1^k(\lambda_2 - 1) - \lambda_2^k(\lambda_1 - 1)}{\lambda_1^n(\lambda_2 - 1) - \lambda_2^n(\lambda_1 - 1)}. \quad (6)$$

Based on this expression, both the average arrival time of the enzyme to the damaged site $\langle t\rangle$ and the dispersion of this time Δ^2 can be calculated:

$$\langle t\rangle = -\lim_{s\to 0}\partial_s j(s),\ \langle t^2\rangle = -\lim_{s\to 0}\partial_s(\partial_s j(s)),$$
$$\Delta^2 = \langle t^2\rangle - \langle t\rangle^2. \quad (7)$$

For these values, the following expressions were deduced:

$$\langle t(k)\rangle = \frac{(n+k-1)(n-k)}{2w},$$
$$\Delta^2(k) = (n+k-1)(n-k)\frac{n(n-1)+k(k-1)+1}{6w^2},$$
$$\Delta^2(0) = n(n-1)\frac{n(n-1)+1}{6w^2}. \quad (8)$$

If the enzyme comes from the bulk solvent to a random nucleotide, the averaging over k is required for obtaining expressions for all kinetic values. This gives:

$$j(s) = \frac{1}{\lambda_1^n(\lambda_2 - 1) - \lambda_2^n(\lambda_1 - 1)} \cdot \\ \cdot\left[\lambda_1(\lambda_2 - 1)\frac{\lambda_1^n - 1}{\lambda_1 - 1} - \lambda_2(\lambda_1 - 1)\frac{\lambda_2^n - 1}{\lambda_2 - 1}\right],$$
$$\langle t\rangle = \frac{(2n-3)n+1}{6w},\ \Delta^2 = (n-1)(2n-1)\frac{(14n-9)n+7}{180w^2}. \quad (9)$$

In the model, where the motion of the enzyme along the DNA fragment is considered as the one-dimensional continuous diffusion, an analytical expression for the average search time and its dispersion can also be obtained. The use of the continuous diffusion instead of random jumps is justified in the case of a large length (R) of DNA fragments, when jumps are efficiently averaged, resulting in the diffusive character of the sliding of the enzyme. In this model, the density of the distribution of the enzyme $C(r,t)$ obeys the equation

$$\partial_t C(r,t) = D\Delta C(r,t) \quad (10)$$

with the initial condition $C(r,t=0) = 1/R$ and the boundary condition

$$C(R,t) = 0,\ \left.\frac{\partial C(r,t)}{\partial r}\right|_{r=0} = 0. \quad (11)$$

This model was used to calculate the following kinetic parameters of the problem:

$$\langle t\rangle = R^2/(3D),\ \langle t^2\rangle = 4R^4/(15D^2),\ \Delta^2 = 7R^4/(45D^2). \quad (12)$$

Let us note that these results can also be obtained from Eq. (9) by assuming that $n >> 1$ and $n^2/w = R^2/D$.

The analytical results make it possible to quantify the kinetics of the enzymatic DNA repair by calculating the inverse Laplace transform. The kinetic curves for the repair of DNA fragments of different length are shown in Fig. 2. In this case, the character of the kinetics, which is generally considered as monoexponential or is approximated by a set of exponents, is an important question. As can be seen from Fig. 3, the exponential approximation for the kinetics is absolutely wrong, because the kinetics for each n has a contribution that drops according to a power law, and the exponential approximation is incorrect. This is associated with the diffusive character.

To conclude, we analytically calculated the distribution of the arrival times of the enzyme to the damaged DNA nucleotide and determined the dependence of this distribution on the length of the DNA fragment (the number of nucleotides) and the diffusion coefficients for the sliding of the enzyme along the DNA fragment. It was shown that the kinetics of the repair has an essentially non-exponential character due to diffusion effects. The results of the present study make it possible to determine the optimal conditions for investigation of the kinetics of the repair. A knowledge of the distribution of the arrival times will help in more precisely separating the repair steps and determining the kinetic parameters of the problem. In future, we plan to extend the results to the consideration of successive steps of one- and three-dimensional diffusion of enzymes, electrostatic interactions, and additional reversible conformational transitions in the system. The method developed in the present study makes it possible to complicate the problem.

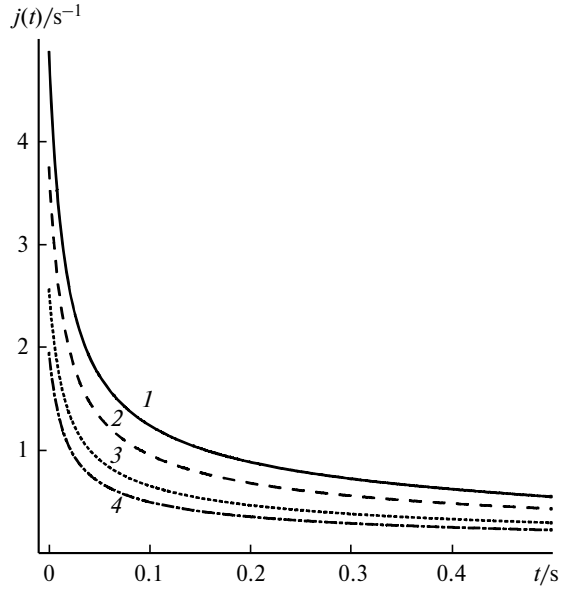

Fig. 2. Time plot of the repair rate $j(t)$ for DNA fragments of different length: n = 10 (*1*), 13 (*2*), 19 (*3*), and 25 (*4*).

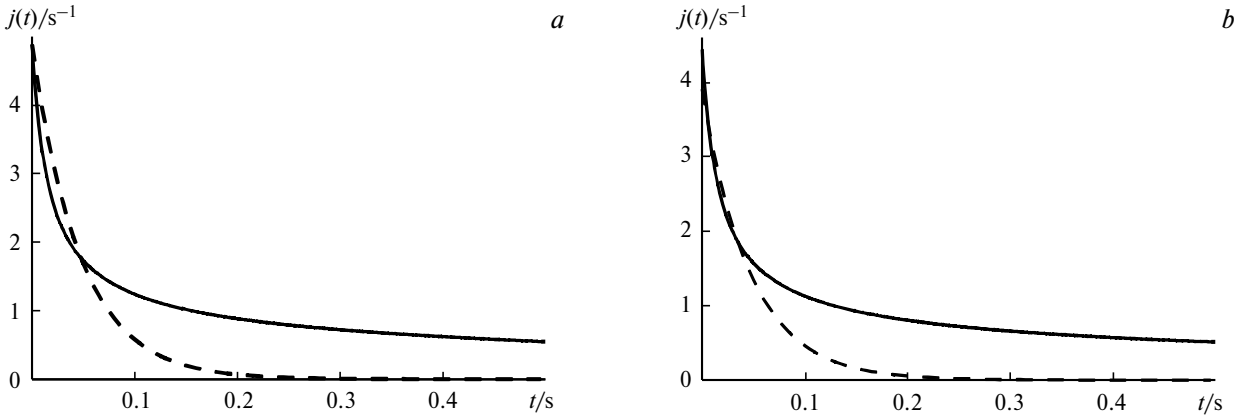

Fig. 3. Time plot of the repair rate $j(t)$ for DNA fragments of length $n = 10$ (*a*) and 25 (*b*) (solid lines); the dashed lines indicate the simulation of the kinetics by a monoexponential function.

This work was financially supported by the Russian Foundation for Basic Research (Project No. 11-03-00356), the Council on Grants at the President of the Russian Federation (Program for State Support of Leading Scientific Schools of the Russian Federation, Grant NSh-7643.2010.3), and the Siberian Branch of the Russian Academy of Sciences (Program for Support of Interdisciplinary Research Projects, Project No. 28).

References

1. G. Adam, M. Delbrück, in *Reduction of Dimensionality in Biological Diffusion Processes*, Eds A. Rich, N. Davidson, W. H. Freeman and Company, San Fransisco, 1968, p. 198.
2. A. Tafvizi, L. A. Mirny, A. M. van Oijen, *Chem. Phys. Chem.*, 2011, **12**, 1481.
3. M. Oda, H. Nakamura, *Genes Cells*, 2000, **5**, 319.
4. O. G. Berg, R. B. Winter, P. H. von Hippel, *Biochem.*, 1981, **20**, 6929.
5. R. B. Winter, P. H. von Hippel, *Biochem.*, 1981, **20**, 6961.
6. O. G. Berg, P. H. von Hippel, *Annu. Rev. Biophys. Biophys. Chem.*, 1985, **14**, 131.
7. S. E. Halford, *Biochem. Soc. Trans.*, 2009, **37**, 343.
8. A.-M. Florescu, M. Joyeux, *J. Phys. Chem. A*, 2010, **114**, 9662.
9. L. Mirny, M. Slutsky, Z. Wunderlich, A. Tafvizi, J. Leith, A. Kosmrlj, *J. Phys. A: Math. Theor.*, 2009, **42**, 434013.
10. G. Kolesov, Z. Wunderlich, O. N. Laikova, M. S. Gelfand, L. A. Mirny, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 13948.
11. R. K. Das, A. B. Kolomeisky, *Phys. Chem. Chem. Phys.*, 2010, **12**, 2999.
12. H.-X. Zhou, A. Szabo, *Phys. Rev. Lett.*, 2004, **93**, 178101.
13. N. A. Kuznetsov, D. O. Zharkov, V. V. Koval, M. Buckle, O. S. Fedorova, *Biochem.*, 2009, **48**, 11335.
14. L. Yu. Kanazhevskaya, V. V. Koval, D. O. Zharkov, P. R. Strauss, O. S. Fedorova, *Biochem.*, 2010, **49**, 6451.

Received July 22, 2011