General random walk model of ATP-driven helicase translocation along DNA

Y. Z. Chen,[1,2,3] Dong Mi,[3,4] He-Shang Song,[3] and Xian-Ju Wang[3]
[1]*ISIS Pharmaceuticals, 2292 Faraday Avenue, Carlsbad, California 92008*
[2]*Department of Physics, Purdue University, West Lafayette, Indiana 47907-1396*
[3]*Department of Physics, Dalian University of Technology, Dalian 116023, China*
[4]*Department of Physics, Graduate School of Academia Sinica, Beijing 100039, China*
(Received 11 March 1997)

A general random walk model is presented which can be used to statistically describe ATP-driven movement of a helicase (DNA unwinding enzyme) along a DNA chain with a nonuniform distribution of obstacles on the chain. These obstacles are representative of DNA-bound proteins, drugs, counterions, and DNA packing environment. We carried out a calculation on a DNA chain with an obstacle distribution that mimics DNA in chromatin (folded DNA-protein material in cells becomes chromosome in partially unfolded form). Our calculated helicase movement speed shows significant reduction with increasing obstacle strength. At the strong strength limit, the calculated speed is found to be consistent with the observed helicase unwinding rate for chromatin DNA. Therefore the model presented in this work is of potential application in the analysis of the effect of random obstacles on biomolecular translocation along DNA. The behavior of the helicase translocation under different obstacle strengths and along different lengths of DNA is discussed.
[S1063-651X(97)11207-7]

PACS number(s): 87.10.+e, 05.90.+m

INTRODUCTION

In the DNA replication process, a DNA polymerase complex translocates and pushes the replication fork along DNA chain [1]. A class of DNA unwinding enzymes, DNA helicases, is involved in the base pair separation at the fork [2,3]. This unwinding process is driven by the hydrolysis of ATP [3,4]. So far little is known regarding the mechanism by which a helicase unwinds a DNA double helix. It has been proposed that the helicase unwinding process can be facilitated by a thermal fluctuational base pair opening [5]. One can then hypothesize [6] that ATP-driven helicase unidirectional movement continues while thermal fluctuation opens up the DNA double helix in front of the helicase. The helicase in turn prevents reformation of the DNA base pairing by exerting particular DNA-protein interactions along the DNA single strand on which it is translocating.

In order to show whether a helicase can undergo the above-mentioned ATP-driven translocation, a variety of parameters that characterize this process needs to be determined [6]. Although the processivity (which measures the rate for the helicase to remain bound to the DNA chain) and directionality of such a translocation process have been measured in certain systems [7,8], the translocation process itself cannot generally be monitored by chemical means as there is no defined product associated with the translocation. Therefore theoretical models need to be developed such that the observed ATP consumption rates can be interpreted and the translocation process can be described.

In an attempt to derive the rate constant of an ATP-driven translocation process along a DNA chain, a biased random walk model was developed by Young *et al.* [6]. This model is based on Berg's random walk model of the diffusion process of a repressor along a strand of DNA [9]. A biased walk is introduced to statistically describe the translocation along the chain with uniformly distributed obstacles. However, in most situations the distribution of obstacles along the DNA chain is nonuniform. For instance, DNA in chromatin wraps around histone protein complexes at regular locations along the chain [1,10]. These DNA binding histone proteins act as periodic obstacles to the helicase translocation. DNA regulatory proteins and DNA binding drugs also act as obstacles at certain locations on the chain. In addition the varying counteria distribution and DNA packing environment can also affect the translocation nonuniformly. In order to take into consideration these nonuniform obstacles, it is necessary to develop a generalized model with a nonuniform biased walk. In this paper we present such a model and apply it to a system that mimics DNA in chromatin. We will examine whether our model can account for the observed slow-down of the replication rate in chromatin DNA. We will also analyze the effect of obstacle strength and DNA chain length on the translocation.

THEORETICAL MODEL

Following Berg [9] and Young *et al.* [6], we regard the ATP-driven translocation of a helicase as a biased random walk on a DNA chain with length L. We define a one-dimensional coordinate system along the chain with its origin at the left end. The helicase at position x is then assumed to walk to the right with a probability $\rho(x)$ and to the left with $1-\rho(x)$. $\rho(x)$ is position dependent because of the nonuniform distribution of obstacles along the chain. The walk is assumed to be biased towards the right of the chain and therefore $\frac{1}{2} \leqslant \rho(x) \leqslant 1$. Let Δt and Δx be the time and distance of each step, and we further assume that they are position independent; then, the mean time $T(x)$ for the helicase to walk from position x to L satisfies the equation

$$T(x)=\Delta t+\rho(x)T(x+\Delta x)+[1-\rho(x)]T(x-\Delta x). \quad (1)$$

In the case of $\rho(x)=\text{const}$, this equation becomes that given by Berg [9]. This equation can be rewritten as a second-order difference equation

$$[1-\rho(x)]\frac{[T(x+\Delta x)-2T(x)+T(x-\Delta x)]}{(\Delta x)^2}+\frac{[2\rho(x)-1]}{\Delta x}\frac{[T(x+\Delta x)-T(x)]}{\Delta x}+\frac{\Delta t}{(\Delta x)^2}=0. \quad (2)$$

For DNA the step Δx is no less than the distance between neighboring base pairs ($\sim$3.4 Å). Nonetheless, it is worth noting that, in the limit of small steps, $\Delta x\to 0$, Eq. (2) becomes the differential equation

$$[1-\rho(x)]\frac{d^2T(x)}{dx^2}+\frac{[2\rho(x)-1]}{\Delta x}\frac{dT(x)}{dx}+\frac{1}{D}=0, \quad (3)$$

where $D=(\Delta x)^2/\Delta t$.

This differential equation can be solved for a given set of boundary conditions. For a right biased walk $T(L)=0$. If we assume that $x=0$ is a reflecting boundary, at which the helicase always moves towards the right, then $T(0)-T(\Delta x)=\Delta t$, which gives $dT(0)/dx=-\Delta t/\Delta x$. These two conditions are used in this work.

Given the solution of Eq. (2) or (3), one can determine the speed $V(x)=(L-x)/T(x)$ and rate constant $K(x)=1/T(x)$, the inverse of the mean translocation time, of the helicase translocation from position x to the right end L. Replication in chromatin often involves multiple simultaneously replicating sections with varying lengths [1]. The experimentally measured replication speed is the average speed of these individual replication processes. To better compare with observations, the statistical average speed $\overline{V}$ will be calculated in this work:

$$\overline{V}=\frac{1}{L}\int_0^L\frac{L-x}{T(x)}dx. \quad (4)$$

Equation (3) has analytic solutions when $\rho(x)$ is position independent. For instance, when $\rho(x)=1$, which corresponds to a completely biased walk, Eq. (3) becomes

$$\frac{dT(x)}{dx}+\frac{\Delta t}{\Delta x}=0, \quad (5)$$

which has the solution

$$T(x)=\frac{\Delta t}{\Delta x}(L-x). \quad (6)$$

The average speed of translocation is then

$$\overline{V}=\frac{\Delta x}{\Delta t}. \quad (7)$$

As expected $\overline{V}$ is independent of the chain length.

In another case $\rho(x)=\frac{1}{2}$, Eq. (3) is reduced to

$$\frac{d^2T(x)}{dx^2}+\frac{2}{D}=0, \quad (8)$$

which is the equation derived by Berg to describe the diffusion process along a DNA chain [7]. Under our boundary conditions this equation has the solution

$$T(x)=\frac{1}{D}(L^2-x^2)+\frac{\Delta t}{\Delta x}(L-x), \quad (9)$$

from which one can derive the average translocation speed

$$\overline{V}=\frac{D}{L}\left[\ln\left(\frac{2L}{D}+\frac{\Delta t}{\Delta x}\right)-\ln\left(\frac{L}{D}+\frac{\Delta t}{\Delta x}\right)\right], \quad (10)$$

which is dependent on the chain length. For large L, $\overline{V}\to(D/L)\ln 2$.

If $\rho(x)$ is a complicated function of x, it is difficult to find analytic solutions for Eq. (3). In this case the translocation time can be derived by numerically solving Eq. (1) or (2). Equations (1) and (2) are also biologically more relevant as the minimum step Δx is no less than the distance between neighboring base pairs.

RESULTS AND DISCUSSIONS

To study the effect of chromatin packing on ATP-driven helicase translocation, we carried out a calculation to determine the average translocation speed and rate constant of a DNA chain with a distribution of obstacles that mimics a DNA in chromatin. For comparison a calculation on an obstacle-free DNA chain is also carried out.

The parameters for our model are chosen as follows: The step of the walk $\Delta x=3.4$ Å is the distance between neighboring DNA base pairs. To adequately describe a translocation on the DNA in chromatin a chain with at least 50 nucleosomes, each a unit section of chromatin, needs to be considered. Each nucleosome contains about 245 base pairs [1,10]. Therefore a DNA chain with 12 250 base pairs will be considered in this work. Such a chain has a length L of 41 650 Å. There is a lack of experimental data that can be used to determine the time for each walk Δt. Therefore the fastest observed DNA replication rate is used to tentatively determine Δt. The observed replication rate for bacteriophage T7 is 780 base pairs per second at 303 K [11]. This gives a Δt of 1.3×10^{-3} s. It is noted that the observed replication speed depends sensitively on temperature. The chromatin experiment to be compared in this work was carried out at 303 K [12]. Therefore the observed bacteriophage replication rate at the same temperature is used to define the parameter in our model.

The ATP-driven helicase translocation on an obstacle-free DNA chain can be considered as a completely biased random walk. Therefore the average translocation speed $\overline{V}$ is given by Eq. (7), which is 2.65×10^3 Å/s, the observed replication rate for bacteriophage T7.

In chromatin, DNA is packed into several levels of organization. The first level is the wrapping of the DNA onto histone octomers at regular locations on the DNA chain to form a beads-on-a-string protein-DNA chain. A unit containing a histone-DNA bead plus the linker DNA chain that connects to the neighboring beads is named nucleosome. The nucleosomes are further aligned and folded into several higher-level organizations [1,10]. These higher-level organi-

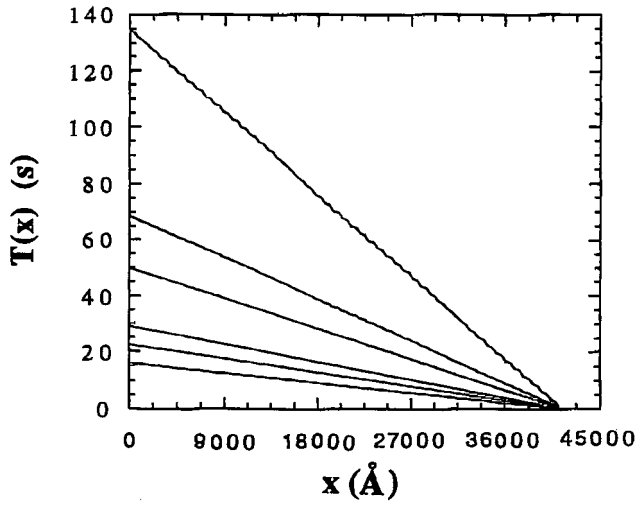

FIG. 1. The calculated mean time $T(x)$ of an ATP-driven helicase translocating from position x to L on a DNA chain with obstacle distribution $\rho(x)=1-\frac{1}{2}\epsilon\sin^2(\alpha x)$. The length of the DNA chain $L=41\,650$ Å. The lines from bottom to top correspond to $\epsilon=$ 0, 0.5, 0.7, 0.9, 0.95, and 1.0, respectively.

zations are unfolded before replication can take place and hence they have little effect on helicase translocation at the replication fork. On the other hand, biochemical tests indicated that, as the DNA polymerase complex (an enzyme complex with which the helicase is attached) passes, the nucleosomal DNA uncoils from the histone octomer without releasing it completely [1]. Therefore histone octomers act as obstacles to the helicase translocation. Because of the periodic distribution of these histone octomers, the probability for the biased walk changes periodically. It is reasonable to assume that the resistance to the helicase movement is much stronger in the central core region of a nucleosome, where the histone octomer is located, than in the regions at both ends of the nucleosome where the linker DNA is located. Therefore a sine periodic function can be used to simulate the effect of the histone octomers on the ATP-driven helicase translocation:

$$\rho(x)=1-\tfrac{1}{2}\epsilon\sin^2(\alpha x). \tag{11}$$

It is pointed out that other forms of periodic functions can also be used to describe the essential features of the histone resistance. Because of a lack of the knowledge on the details of histone-DNA interactions, it is unclear which periodic function better mimics the effect of histone octomers. Hence for simplicity a sine function is tentatively employed in this work.

The length of the DNA chain in a nucleosome is 833 Å, from which we obtain $\alpha=3.8\times10^{-3}$ Å^{-1}. ϵ describes the level of obstacle resistance. For a right-biased walk $0\leqslant\epsilon\leqslant1$. In the central core region the bias is likely to be reduced to a near-minimum value $\rho(x_c)\approx\frac{1}{2}$ due to the strong resistance of the histone octomer to the helicase movement. Therefore the likely value of ϵ in the chromatin is $\epsilon\approx1$.

To analyze the translocation at different levels of obstacle strength, several values of ϵ are considered. The calculated mean time $T(x)$'s for the helicase to travel from position x to L are shown in Fig. 1. All of the $T(x)$'s are found to follow

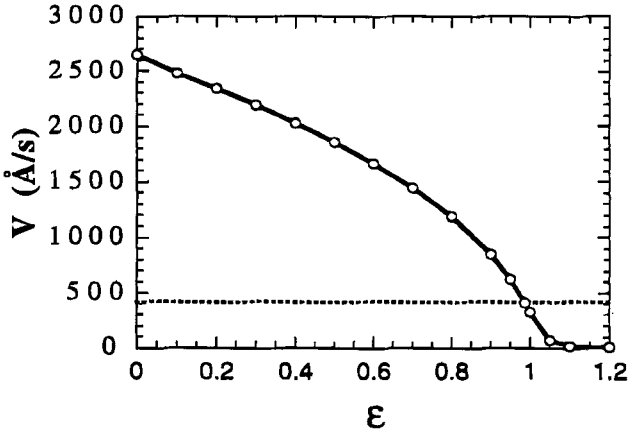

FIG. 2. The calculated average speed $\overline{V}$ of an ATP-driven helicase translocating along a DNA as a function of the obstacle strength ϵ (solid line). The obstacle distribution and chain length are the same as that in Fig. 1. The dashed line corresponds to the observed replication speed of 416.7 Å/s in the chromosomes of Chinese hamster cells [12].

approximately a linear relationship with $L-x$. In addition they show periodic zigzag features reflecting the effect of periodic obstacles in the chain. As expected both the slope and magnitude of the zigzag feature of $T(x)$ increase with increasing ϵ. The increase rate is relatively small for $0<\epsilon<0.5$ which corresponds to the weak obstacle strength. The rate increases more and more rapidly at higher values of ϵ. In the region of $0.95<\epsilon<1$ the rate is 3 times larger than that in the region of $0.7<\epsilon<0.9$.

The calculated average translocation speed $\overline{V}$ as a function of obstacle strength ϵ is shown in Fig. 2 (solid line). For comparison the observed replication rate for chromatin DNA is also included (dashed line). The observed rate falls into the region in $0.95<\epsilon<1$, which corresponds to the strong obstacle strength expected for chromatin. The best match is found at $\epsilon\approx0.99$ at which $\overline{V}\approx400$ Å/s. This is compared with the observed replication rate of 416.7 Å/s in the chromosomes of Chinese hamster cells [12].

The behavior of $\overline{V}$ at different obstacle strengths can be seen in Fig. 2. At low obstacle strength ($\epsilon<0.5$) $\overline{V}$ decreases

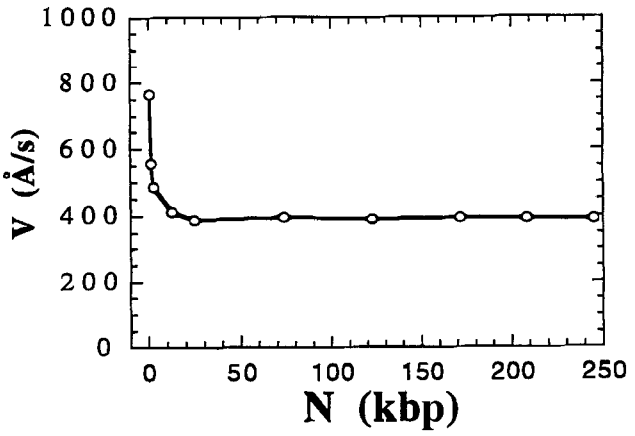

FIG. 3. The calculated average replication speed $\overline{V}$ in chromatin DNA as a function of the number of base pairs N in DNA chain. N is in units of kilo-base-pairs (kbp).

linearly as ϵ increases. The rate of decrease becomes larger and larger when ϵ further increases from 0.5 to 1. In the region of $\epsilon > 1$, where obstacle resistance is so strong that a left-biased walk occurs at the core region of histones, the decreasing rate gradually reduces. At $\epsilon > 1.1$, $\overline{V}$ levels off and approaches 0.

The effect of DNA chain length on the average speed $\overline{V}$ of translocation is shown in Fig. 3. From Fig. 3 one can see that $\overline{V}$ is very sensitive to the chain length for a chain with less than 12 250 base pairs. It changes from 765 Å/s for a 490-base-pair chain to 413 Å/s for a 12 250-base-pair chain. For a chain with more than 12 250 base pairs $\overline{V}$ becomes relatively insensitive to chain length.

CONCLUSION

A general random walk model is proposed which describes the essential features of the effect of random obstacles on ATP-driven helicase translocation along DNA. Our model is applied to the translocation along chromatin DNA. At the strong obstacle strength limit our results are consistent with experiments. The effect of the obstacle strength and the DNA chain length on the translocation is also analyzed. Our model has potential applications in other biomolecular translocation processes such as the translocation of RNA polymerase along DNA in transcription and the diffusion process of a protein or DNA binding drug along DNA.

[1] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J.D. Watson, *Molecular Biology of the Cell* (Garland, New York, 1992).
[2] M. Abdel-Monem and H. Hoffmann-Berling, Trends Biochem. Sci. **5**, 128 (1980).
[3] S. W. Matson and K. A. Kaiser-Rogers, Annu. Rev. Biochem. **59**, 289 (1990).
[4] S. W. Matson, S. Tabor, and C. C. Richardson, J. Biol. Chem. **258**, 14017 (1983).
[5] Y. Z. Chen, W. Zhuang, and E. W. Prohofsky, J. Biomol. Struct. Dyn. **10**, 415 (1992).
[6] M. C. Young, S. B. Kuhl, and P. H. von Hippel, J. Mol. Biol. **235**, 1436 (1994).
[7] T. Q. P. Uyeda, S. J. Kron, and J. A. Spudich, J. Mol. Biol. **214**, 699 (1990).
[8] M. E. Porter, J. M. Scholey, D. L. Stemple, G. P. A. Vigers, R. D. Vale, M. P. Sheetz, and J. R. McIntosh, J. Biol. Chem. **262**, 2794 (1987).
[9] H. C. Berg, *Random Walk in Biology* (Princeton University Press, Princeton, NJ, 1983).
[10] J. Widom, Annu. Rev. Biophys. Biophys. Chem. **18**, 365 (1989).
[11] S. D. Rabkin and C. C. Richardson, J. Mol. Biol. **204**, 903 (1988).
[12] J. A. Huberman and A. D. Riggs, J. Mol. Biol. **32**, 327 (1968).