Band Inversion in Gel Electrophoresis of DNA

M. Doi and T. Kobayashi
Department of Physics, Faculty of Science, Tokyo Metropolitan University, Setagaya, Tokyo 158, Japan

Y. Makino and M. Ogawa
Fuji Film Asaka Laboratory, Sensui, Asaka, Saitama 351, Japan

G. W. Slater and J. Noolandi
Xerox Research Center of Canada, Mississauga, Ontario, Canada L5K 2L1
(Received 2 May 1988)

Band-inversion phenomena, found in a recent computer simulation and a real experiment of steady-state gel electrophoresis, is theoretically analyzed by a simple model that describes the conformational change of DNA reptating through a gel. The model predicts the distribution function of the end-to-end vector and the steady-state mobility, which agrees quite well with the result of a computer simulation without the use of any adjustable parameters.

PACS numbers: 87.15.He, 82.45.+z, 82.70.Gg

Gel electrophoresis is widely used to separate DNA fragments with different molecular weight. If the molecular weight of DNA is low, or if the electric field E is sufficiently small, the migration velocity V_G of DNA is proportional to the applied electric field, and the mobility $\mu = V_G/E$ decreases with the molecular weight M according to $\mu \propto 1/M$. This behavior has been interpreted[1] by the reptation model.[2] However, for high molecular weights or for strong electric fields, a nonlinear behavior is observed: The mobility increases with the electric field, and becomes rather insensitive to the molecular weight.

Lumpkin, Dejardin, and Zimm (LDZ)[3] gave a first theory for such nonlinear behavior. Assuming that DNA moves towards the field, with its head segments (the segments pointing to the field) being aligned by the field, they predicted that the mobility monotonically decreases with the increase of M according to

$$\mu = C_1(E)/M + C_2(E) , \tag{1}$$

where C_1 and C_2 are positive constants which depend on the magnitude of the applied field.

Noolandi and co-workers[4-6] formulated this idea into a mathematical model which they call the biased reptation model. By computer simulation, they found a rather unexpected result: With the increase of M, the mobility first decreases but then starts to increase again before it approaches the limiting value. In the actual experiments of gel electrophoresis, this means that the band moving faster does not always correspond to smaller DNA: It may be a band of larger DNA. Noolandi *et al.*[6] showed that this band inversion phenomena is indeed observed experimentally.

A phenomenological explanation for the band inversion was given by Noolandi *et al.*[6] They noticed that although the DNA in their computer simulation migrates through gels mostly in an elongated conformation, it occasionally takes a compact conformation which has small end-to-end vectors. Once it takes a compact conformation, it is trapped in the conformation for a long time, during which it moves quite slowly. The band inversion occurs since very long chains take compact conformation rarely, thus moving faster than the chains with intermediate lengths.

In this paper we shall give a quantitative theory for the band inversion. We shall take the biased reptation model as the model of DNA motion. This model does not account for certain aspects of DNA motion such as contour length fluctuation and tube leakage,[7] which, as it was pointed out by Deutsch,[8,9] might become important for strong electric field or for very large DNA molecules. However, for a certain range of experimental conditions that is discussed in Ref. 6, the use of the reptation model would be justified at least in the first approximation.

Starting from the model formulated by Noolandi *et al.*, we shall derive a simple equation which approximately describes the time evolution of chain conformation. Though the theory is simple, it can reproduce the result of the computer simulation quite well without using any adjustable parameters.

In the biased reptation model, it is assumed that the chain is made up of N segments of length a (the average pore size), and can move only along its own contour. If there is an electric field E in the x direction, the chain moves along itself with the mean curvilinear velocity[2]

$$v_1 = QEh_x/L\xi , \tag{2}$$

where Q is the total charge of the chain, h_x is the x component of the end-to-end vector $\mathbf{h}$, $L = Na$ is the total contour length, and ξ is the friction constant of the chain ($\xi \propto M$). The average velocity is superimposed on the Brownian motion of the chain which is characterized by

the curvillinear diffusion constant $D_l = k_B T/\xi$.

The average velocity of the center of mass is related to the curvillinear velocity by[2]

$$V_{Gx} = \langle h_x v_l \rangle / L . \tag{3}$$

Thus the mobility $\mu = V_{Gx}/E$ is given by

$$\mu = Q\langle h_x^2 \rangle / L^2 \xi = 3\mu_0 \langle h_x^2 \rangle / L^2 , \tag{4}$$

where $\mu_0 \equiv Q/3\xi$ is a segmental reptation mobility, which is independent of molecular weight.

Equation (4) indicates that the nonlinear behavior of the mobility is caused by the change of $\langle h_x^2 \rangle$ by the electric field. If the field is very weak, $\langle h_x^2 \rangle$ is equal to the equilibrium value $Na^2/3$, so that $\mu \propto 1/N \propto 1/M$. With the increase of the field, $\langle h_x^2 \rangle$ increases, and so does the mobility. Thus we shall now consider how h_x changes with time.

As the chain moves along the tube, the head segment leaves the tube and creates new tube segments. The biased reptation model assumes that the new tube segments are aligned toward the field: The distribution of the unit vector **u** a long the new segment is given by

$$P(u) \propto \exp\left(\frac{QaE}{2Nk_B T} u_x \right) = \exp(\Theta u_x) , \tag{5}$$

where $\Theta = QaE/2Nk_B T$ is the scaled electric field. From Eq. (5), it follows that

$$\langle u_x \rangle = \coth(\Theta) - 1/\Theta \equiv P(\Theta) \tag{6}$$

and

$$\langle u_x^2 \rangle - \langle u_x \rangle^2 = 1/\Theta^2 - 1/\sinh^2(\Theta) \equiv Q(\Theta) . \tag{7}$$

We shall now set up an approximate equation which describes the time evolution of $h_x(t)$. First we consider the case of strong electric field and disregard the effect of Brownian motion. In a time interval Δt, the chain moves the distance $v_l \Delta t$ along the tube, so that the head creates $\Delta n \equiv v_l \Delta t/a$ tube segments:

$$\Delta n = \frac{v_l \Delta t}{a} = 3\mu_0 \frac{h_x}{L} E \Delta t . \tag{8}$$

Thus the change of $h_x(t)$ may be written as

$$h_x(t+\Delta t) - h_x(t) = aP(\Theta)\Delta n - [h_x(t)/N]\Delta n + \Delta R(t) . \tag{9}$$

The first term on the right-hand side is the mean end-to-end vector of the new tube segments created by the head. The second term represents the mean end-to-end vector of the segments which have disappeared in the time interval Δt. [Here we have assumed that if the end-to-end vector of a chain is **h**, the mean end-to-end vector of a segment is **h**/N. This approximation will be good as far as the orientational distribution of segments is homogeneous along the chain.] The last term in Eq. (9) represents statistical fluctuations of the vectors which have been created or destroyed in the time interval Δt: $\Delta R(t)$ is a Gaussian stochastic variable characterized by the moments

$$\langle \Delta R(t) \rangle = 0 ,$$
$$\langle \Delta R(t) \Delta R(t') \rangle = \begin{cases} 0, & \text{if } |t-t'| > \Delta t, \\ a^2 Q(\Theta)\Delta n, & \text{if } |t-t'| < \Delta t. \end{cases} \tag{10}$$

The second moment of $\Delta R(t)$ is determined by the fact that the mean square end-to-end vector of the new tube segments is $(aP\Delta n)^2 + a^2 Q \Delta n$.

Equation (9), supplemented by Eqs. (8) and (10) is a stochastic difference equation. Unfortunately, this equation becomes invalid for weak electric fields where the effect of Brownian motion becomes important. Particularly, it fails completely for zero field: If $E = 0$, Δn and $\Delta R(t)$ are identically zero according to Eqs. (8) and (10), so that $h_x(t)$ does not change with time. To remove this failure, we shall modify Eq. (8).

It is known[7] that for $E = 0$, the time evolution of $h_x(t)$ is approximately described by the stochastic equation

$$h_x(t+\Delta t) - h_x(t) = -\frac{h_x(t)}{\tau_{rep}} \Delta t + \Delta R(t) , \tag{11}$$

where τ_{rep} is the reptation time:

$$\tau_{rep} = \xi N^2 a^2 / \pi^2 k_B T . \tag{12}$$

We modify Eq. (8) such that Eq. (9) reduces to Eq. (11) in the case of $E = 0$. A simple choice is

$$\Delta n = (3\mu_0 h_x E/L + N/\tau_{rep})\Delta t = \tau_{rep}^{-1}[(2/\pi^2) N \Theta h_x/a + 1]\Delta t. \tag{13}$$

Obviously there are other choices, but the essential feature of our conclusion [such as Eq. (1)] is insensitive to the actual form of Δn.

Equations (7), (10), and (13) are the stochastic equations which approximately describe the time evolution of $h_x(t)$ under an arbitrary electric field. The equations can be rewritten in a standard form of stochastic difference equation:

$$h_x(t+\Delta t) = h_x(t) + g\Delta t + \sigma^{1/2} \Delta r(t) , \tag{14}$$

where

$$g(h_x) = \left(aP(\Theta) - \frac{h_x}{N} \right) \frac{\Delta n}{\Delta t}, \quad \sigma = a^2 Q(\Theta) \frac{\Delta n}{\Delta t} , \tag{15}$$

and $r(t)$ is a Gaussian random number satisfying

$$\langle \Delta r(t) \rangle = 0,$$
$$\langle \Delta r(t) \Delta r(t') \rangle = \begin{cases} 0, & \text{if } |t-t'| > \Delta t, \\ 2, & \text{if } |t-t'| < \Delta t. \end{cases} \tag{16}$$

It is known that if $h_x(t)$ obeys the stochastic equation

(14), its distribution function $\Psi(h_x,t)$ satisfies[7]

$$\frac{\partial\Psi}{\partial t}=\frac{\partial}{\partial h_x}\left[\sigma\left(\frac{\partial\Psi}{\partial h_x}+\frac{\partial U}{\partial h_x}\Psi\right)\right], \tag{17}$$

where

$$\frac{\partial U}{\partial h_x}=-\frac{1}{\sigma}\left(g-\frac{\partial\sigma}{\partial h_x}\right). \tag{18}$$

Equation (17) is a diffusion equation for a Brownian particle moving in a potential U, with a diffusion constant σ. From Eqs. (13), (15), and (17), we have

$$\sigma(h_x)=\frac{a^2Q(\Theta)}{\tau_{\rm rep}}[(2/\pi^2)N\Theta|h_x/a+1], \tag{19}$$

$$U(h_x)=\ln(\sigma(h_x))-[2Na^2Q(\Theta)]^{-1}[|h_x|-NaP(\Theta)]^2, \tag{20}$$

where h_x has been replaced by $|h_x|$ since $\sigma(h_x)$ and $U(h_x)$ must be even functions of h_x. The steady-state solution of Eq. (17) is

$$\Psi_{\rm st}(h_x)\propto\exp[-U(h_x)]\propto[(2/\pi^2)N\Theta|h_x|/a+1]^{-1}\exp[-(2Na^2Q)^{-1}(|h_x|-NaP)^2]. \tag{21}$$

For $E=0$, P and Q are equal to 0 and $\frac{1}{3}$, respectively, so that Eq. (21) reduces to the correct Gaussian distribution function. For nonzero E, Eq. (21) consists of two factors: The exponential factor represents the Gaussian distribution whose peak position and variance are given by the LDZ theory. The front factor, on the other hand, has a peak at $h_x=0$. This weighs the compact conformation, and gives rise to the self-trapping in the compact conformation. Figure 1 shows typical distribution functions. The parameters here correspond to Fig. 2 in Ref. 6. Self-trapping is most conspicuous for $N=70$, which is in accordance with the qualitative observation made in Ref. 6.

Given the distribution function, the mean square end-to-end distance can be easily calculated. If $NP^2/Q\gg1$, $\langle h_x^2\rangle$ can be evaluated analytically, and the result is

$$\langle h_x^2\rangle=Na^2[NP(\Theta)^2-Q(\Theta)]. \tag{22}$$

Notice that the sign in front of $Q(\Theta)$ is minus, not plus as it is assumed in the LDZ theory. The minus sign comes from the fact that the weight of the compact state is large. From Eqs. (4) and (22), it follows that the mobility increases with N at very large N:

$$\mu/\mu_0=3(P^2-Q/N),\quad \text{for } NP^2/Q\gg1. \tag{23}$$

In the general case, the mobility can be calculated from Eqs. (4) and (20) by numerical integration. The result is shown in Fig. 2. The calculated mobility agrees well with the result of the computer simulation especially when both Θ and N are large. (Notice that no adjustable parameter has been used in this comparison.) The agreement becomes less satisfactory at small N (because the discrete nature of the simulation model becomes important), and at low field (presumably because the effect of Brownian motion, which we accounted for rather arbitrarily, becomes important). On the other hand, the result of the LDZ theory, shown by the dot-dashed lines, is considerably different from the result of the simulation.

In Fig. 3, the value of N which gives the minimum mobility is plotted against Θ. The experimental value for N^* in the figure is approximately described by

$$N^*(\Theta)=14.4/\Theta^2. \tag{24}$$

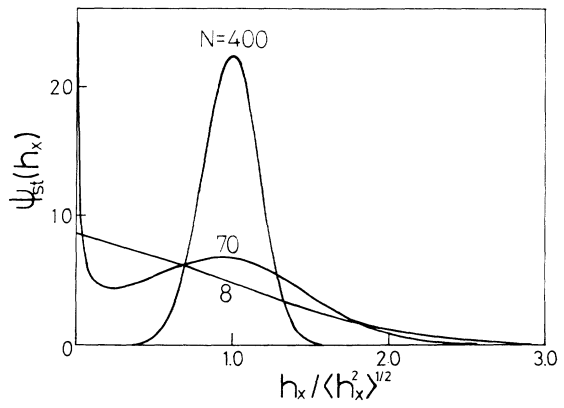

FIG. 1. Distribution of the end-to-end vector in the steady-state electrophoresis. The scaled electric field is $\Theta=0.5$ and $N=8$, 70, and 400. These values correspond to Fig. 2 in Ref. 6. Notice that the distribution function for $N=70$ has double peaks.

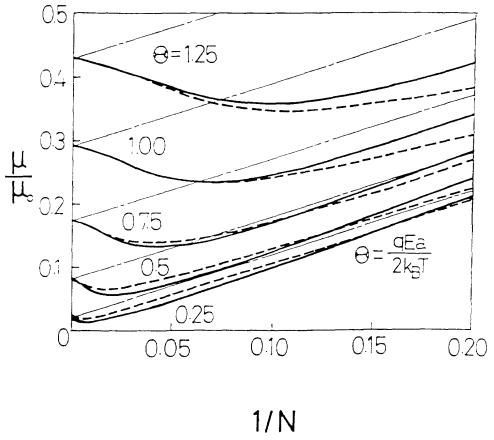

FIG. 2. Relative mobility μ/μ_0 plotted against $1/N$ for various values of scaled electric field. Dashed lines, results of the computer simulation; dash-long-dash lines, results of LDZ theory; solid lines, results of the present theory.

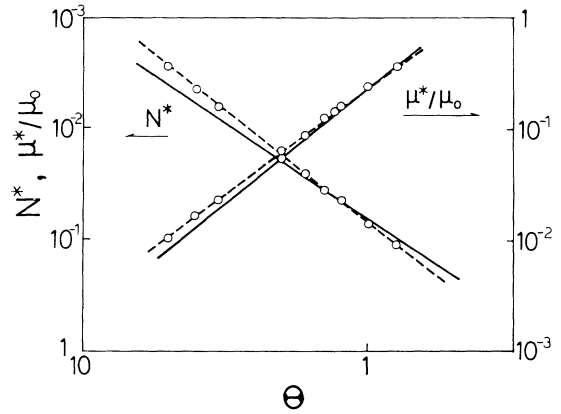

FIG. 3. Minimum mobility μ^*/μ_0 and the corresponding value of N^* at the minimum is plotted against Θ. Circles denote the value obtained by a computer simulation. Solid lines are the result of the present theory.

The band inversion takes place if N exceeds this value.

We have proposed an approximate kinetic equation [Eq. (17)] for the distribution function of the end-to-end vector. The equation, together with Eq. (4), can predict the nonlinear behavior of gel electrophoresis for arbitrary electric field. (Generalization of the equation for a three-dimensional field is trivial.) It works quite well for a steady field. However, perhaps a more interesting application of the equation can be found in the case of an alternating field, which has now become technologically quite important. Research toward this direction is now in progress.

[1]O. J. Lumpkin and B. H. Zimm, Biopolymers **21**, 995 (1982).

[2]P. G. de Gennes, J. Chem. Phys. **55**, 572 (1971).

[3]O. J. Lumpkin, P. Dejardin, and B. H. Zimm, Biopolymers **24**, 1573 (1985).

[4]G. W. Slater and J. Noolandi, Phys. Rev. Lett. **55**, 1579 (1985).

[5]G. W. Slater, J. Rousseau, and J. Noolandi, Biopolymers **26**, 863 (1987).

[6]J. Noolandi, J. Rousseau, G. W. Slater, C. Turmel, and M. Lalande, Phys. Rev. Lett. **58**, 2428 (1987).

[7]M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford. Univ. Press, London, 1986).

[8]J. M. Deutsch, Phys. Rev. Lett. **59**, 1255 (1987), and Science **240**, 922 (1988).

[9]E. O. Shaffer and M. Olvera de la Cruz, to be published.