Size Fluctuations Can Explain Anomalous Mobility in Field-Inversion Electrophoresis of DNA

Bruno H. Zimm
Department of Chemistry, University of California, San Diego, La Jolla, California 92093

(Received 8 September 1988)

We extend the reptation model of DNA gel electrophoresis by replacing the tube by a chain of narrow "straits" connecting open "lakes." The DNA chain can accumulate easily in the lakes, thus allowing large fluctuations in the length of the chain. Unlike the standard model, this model, simulated on a computer, shows a sharp minimum in mobility as the period is varied with periodically reversing fields, similar to experiments.

PACS numbers: 82.45.+z, 36.20.Ey, 87.15.−v

At present the reptation theory of the gel electrophoresis of DNA faces a fundamental difficulty. The steady-state theory of the electrophoretic mobility[1] depends only on even powers of the electric field, but experiments by Carle, Frank, and Olson,[2] and since repeated by many others, show that the mobility can be changed simply by periodically changing only the sign, not the magnitude, of the field; the current theory is unable to accomodate this.

We present here an extension of the reptation theory that predicts a minimum in the mobility as the period of field reversal is varied, similar to the experiments.

In contrast to the simple picture of a regular tube in which the migrating molecule, the "probe," finds itself, electron micrographs[3] show that agarose gels consist of bundles of fibers separated by open spaces up to 0.3 μm in width. We therefore assume that the probe lies in a series of such spaces, spaces that we call "lakes," connected by narrow gateways where the gel chains happen to tightly confine the probe chain, gateways that we call "straits." Segments of the probe diffuse and drift from one lake to another through the connecting straits, driven by the electric field and Brownian motion.

We save a great deal of computer time by assimilating the Brownian motion of the probe in the lakes into an average force, as in the theory of rubber elasticity. If one segment of the probe moves from one lake to another through the intervening strait, the entropy of the system changes, and this change produces an average force tending to pull the probe through the straight. The entropy associated with a piece of probe, modeled by a Gaussain chain of n segments lying in a lake between two straits a distance a apart, is

$$S = k_B \ln \left[C^n \left(\frac{3}{2\pi n b^2} \right)^{3/2} \exp\left(-\frac{3a^2}{2nb^2} \right) \right], \qquad (1)$$

and the average force associated with moving a length dl of probe through a strait is

$$f = T\frac{dS}{dl} = \frac{3k_BT}{2nb}\left(\frac{a^2}{nb^2} - 1 \right) + \frac{k_BT}{b}\ln C, \qquad (2)$$

where b is the length of a segment, $l = nb$, and C is a constant depending on the internal structure of the segment. We can think of this force as a tension exerted on the chain in the strait by the chain in the adjoining lake.

In the same way we find that the force in the one strait attached to a lake where the probe ends is $(k_BT/b)\ln C$, independent of the number of segments in the lake. In this work we will be interested only in differences of forces, so the $\ln C$ terms cancel.

The tension rises steeply when n is small but becomes almost independent of n when n exceeds $(a/b)^2$. Thus, the chain strongly resists excessive stretching but gives almost no resistance to the pooling of large amounts of chain in one lake, similar to the simulations by Deutsch.[4] We shall see that this asymmetry of the force function has important consequences.

Using these lakes, we make a simulation of the motion of the probe. The model, shown in Fig. 1, has the following characteristics: The probe lies in a chain of lakes connected by straits; the chain has a left end and a right end which remain so labeled regardless of orientation with respect to the field. The lakes all have the same length a, which is also the distance between centers with the assumption of negligible lengths for the straits, but the projections of this separation on the field direction, the x axis, differ from one strait to another. These projections divided by the strait length a generate the cosines of the angles between the axes of the straits and the field; these cosines multiplied by a parameter F, representing the field times the effective charge of the segment in the strait, give the effective electric forces tending to move the probe through the straits. The motion through a strait, expressed as the number of segments, dn/dt passing through the strait per unit time, is the sum of the effects of the field, of the tension, and of Brownian motion:

$$\frac{dn}{dt} = \frac{(x_l - x_r)F}{a\zeta} + \frac{f_l - f_r}{\zeta} + B(t). \qquad (3)$$

Here the subscripts l and r designate the lakes to the left and to the right of the strait; ζ is a resistance coefficient,

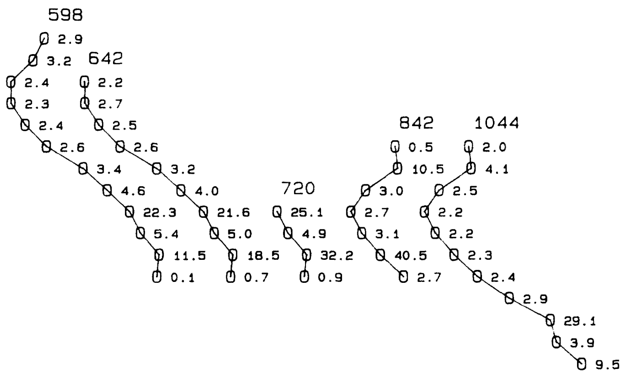

FIG. 1. A cycle of conformations of a 70-segment probe at successive times in steady field; lakes are represented by circles and the connecting straits by lines. The numbers to the right of the lakes are the numbers of segments occupying that lake at that time. For each conformation, the horizontal coordinate of each circle shows the x coordinate of the corresponding lake. The vertical coordinate shows the position of the lake along the probe axis, not the y or z coordinate of the lake, which is irrelevant to the motion; in effect, the conformation has been unfolded onto a plane. The large numerals at the top of each conformation show the time. Each conformation has been shifted to the right for clarity, but the same lakes are shown at the same vertical level in successive conformations. $F=0.5$ in the units of Fig. 2, and is directed toward the right; $n_{\max}=10$.

and $B(t)$ is a random number representing Brownian motion in the strait. At the ends of the chain we allow a new lake to be occupied when the number of segments in the preexisting end lake exceeds a present number $n_{\max}$. The x coordinate of this new lake is selected by addition of a random number uniformly distributed between $-a$ and a to the x coordinate of the preexisting end lake. Conversely, if the number of segments in an end lake falls below zero, we erase that lake. Thus the probe can migrate into new lakes and abandon old ones.

We have adopted a form for ζ that was suggested by a detailed computer simulation of the motion of a chain of n segments through a lake and its two adjoining straits in the presence of Brownian motion. If n was small so that the chain was under tension, then the field simply pulled the chain through the lake with a velocity proportional to the projection of the field on the line between the two straits. But if n was larger than $(a/b)^2$, then a strong field would extend loops of chain in the field direction so that the chain had to travel a circuitous path in traversing the lake. As a result of the longer path, the chain moved more slowly than when n was small. We take this effect into account by making ζ increase from a value of ζ_1 at $n=1$ to $5\zeta_1$ with increasing n and F according to

$$\zeta=\zeta_1[1+F^2(n-1)^2]/[1+\tfrac{1}{5}F^2(n-1)^2]\,. \qquad (4)$$

In this work the average value of ζ was about $3\zeta_1$.

In operation the program iterated a finite time step Δt and computed the changes in the number of segments in each lake produced by the Δn's calculated from Eq. (3). If the number in any of the interior lakes changed by more than 10% that time step was cut in half and tried again. End lakes were added or erased, as necessary, after each step. To avoid approaching the infinite tension at $n=0$, n was not allowed to drop below a preset number. This algorithm is not very accurate, but is simple and reliable. The total number of segments in the probe did not change during a run, though the number of lakes could, and did, vary widely. The probe was started with a set number of segments in an arbitrary conformation of lakes, and before any data were taken the program was allowed to iterate until the initial conformation was completely lost.

We have not taken explicit account of the possibility of long loops of the probe being pulled away from the main path by the field, as discussed by Deutsch.[5] To some extent, however, such effects are implicit in our lake tension and resistance functions f and ζ.

When the field F is zero, the model produces the expected random-walk conformation for the probe, with the mean n per lake equal to a^2/b^2. At low fields, that is, when the F term in Eq. (3) is consistently smaller than the f term, the motion produced is well described by the reptation picture,[1] a chain of fixed length moving in a rigid tunnel of random-walk conformation. But new effects appear when the field is large. Unlike the low-field case, now the number of segments in the lakes along the probe vary over a wide range. Figure 1 shows a typical series of conformations. Accumulations appear at kinks in the chain of lakes, where by Eq. (3) the flow into a lake is higher than the flow out because the x_l-x_r term is larger on the inflow strait than on the outflow strait; opposition to such accumulation is weak because of the asymmetry of the tension function.

Kinks originate by chance as new lakes with random-x increments are added to the chain ends, and the segment number of the lake at the kink builds up as the probe flows through. The segment number can greatly exceed the end-overflow number $n_{\max}$ if the kink is sharp and the probe is long. In this case an important event occurs when the trailing end of the probe reaches the lake at the kink and the lake upstream of the kink is abandoned; since n is greater than $n_{\max}$, and the lake is now an end, the program tries to occupy a new lake at a random position adjacent to the kink. This new lake is abandoned immediately if, by chance, it is upfield from the kink, but if it is down field it usually survives, although it can still possibly be erased by Brownian motion. If it survives, the field pulls chain into it from the over-full lake at the kink. Thus both ends of the probe are now down field from the former kink, which has become the apex of a λ-like conformation. The process is illustrated in Fig. 1.

From this point a race develops between the two arms of the λ, each driven by the field to empty the lake at the

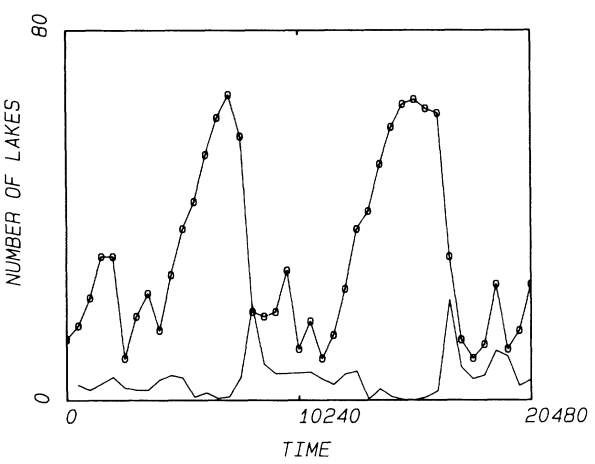

FIG. 2. Typical time series for a probe of 134 segments; number of occupied lakes (line with circles) and 20×center-of-mass velocity (simple line) as a function of time. The ratio $a/b=\sqrt{5}$; the electric force on a segment F has the value of 1 in units of $3k_BT/2b$, corresponding to segments 236 base pairs long in a field of 10 V/cm at 293 K. The velocity is expressed in units of $3k_BT/2\zeta_1 b$, and time in units of $2\zeta_1 ab/3k_BT$.

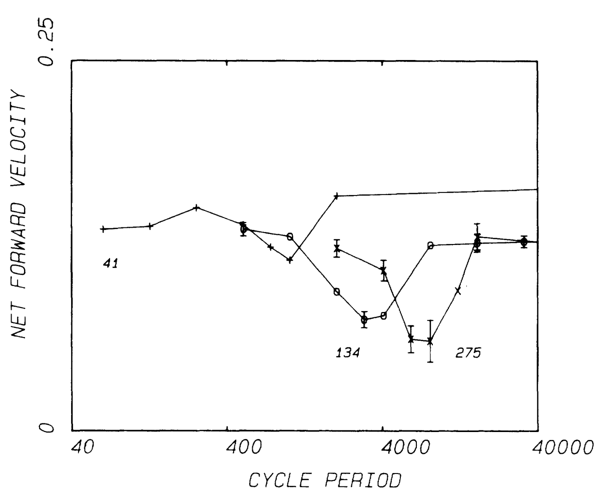

FIG. 3. Net forward velocity as a function of reversing-cycle period for probes with three different numbers of segments (numbers on plots). Limits at the right-hand ends of the plots are $\frac{1}{2}$ of the mean velocity in the steady field $F=1$. The reversing cycle had $\frac{3}{4}$ of the period with $F=1$ and $\frac{1}{4}$ with $F=-1$, and the velocity was averaged over an integral number of complete cycles. The total time in each case was about 500000. Units are the same as in Fig. 2. Where available, an estimate of the standard deviation of the net velocity is shown; this estimate required a separate calculation because of the finite autocorrelation time of the system.

apex until the segment number of that lake drops low enough for the tension at the apex to withstand the pull of the field on the shorter arm. The newly formed arm is usually the shorter, but in the random processes of adding new lakes it occasionally happens that the length of the newer arm overtakes, or comes close to overtaking, that of the older arm. When this happens a protracted tug of war ensues, during which the forward motion of the probe drops to a low value while the number of lakes, and the tension in the lakes near the apex, climb to a peak. Eventually one arm wins, pulling the other through the apex, and as the tension is relaxed the forward velocity of the center of mass reaches a temporary maximum which partially compensates the preceding minimum.

The number of occupied lakes cycles from high to low values as λ conformations form and dissipate; see Fig. 2. The recursion time varies from one cycle to the next, but is distributed with a standard deviation approximately half the mean. The mean recursion time is approximately proportional to the number of segments. The major peaks in the number of lakes are accompanied by valleys in the center-of-mass velocity. The cycling resembles the detailed simulations recently published by Deutsch,[4] in which alternation of compact and stretched conformations occurred. For another view of this situation, see Viovy.[6]

The cycling in and out of λ makes the system sensitive to periodic field reversals if the period of the reversal is close to the mean recursion time of the λ cycle. The phenomenon is shown in Fig. 3 for chains with three different numbers of segments, and the velocity variation averaged over many field reversals is shown in Fig. 4. In the long forward part of a cycle of length equal to the recursion time, a λ is likely to form and develop into a conformation with high tension. When the field is reversed, this high tension causes a large recoil (reminiscent of re-

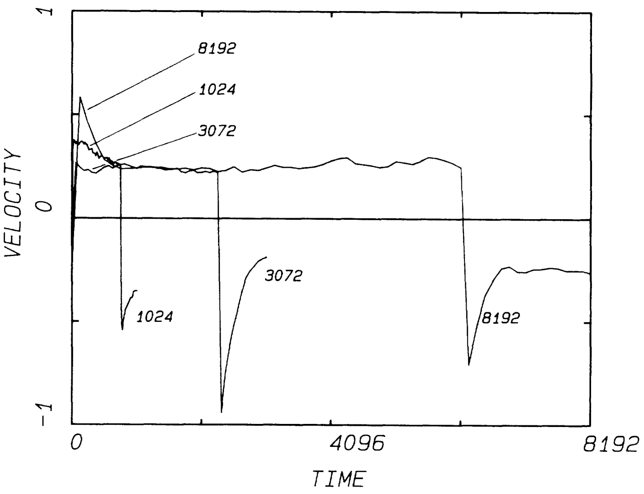

FIG. 4. Velocity, averaged over many cycles, as a function of the time from the beginning of a reversing-field cycle for a 134-segment probe. Values of the cycle period are shown by the numbers beside the plots. Units are the same as in Fig. 2, conditions are the same as in Fig. 3.

cent experimental linear-dichroism results[7]) which nullifies much of the forward progress of the forward part of the cycle. The backward part of the cycle is too short for a highly tensioned λ to form, so the recoil when the field is again switched forward is more modest. The result is that the average velocity in this cycle is less than would be expected from the steady-field velocity. Since the recursion time is larger for larger molecules, the period of minimum velocity is likewise larger. The effect is similar to that seen in Deutsch's simulations,[4] and also to that found experimentally by Carle, Frank, and Olson.[2]

---

[1] O. J. Lumpkin, P. Déjardin, and B. H. Zimm, Biopolymers **24**, 1573 (1985); G. W. Slater and J. Noolandi, Biopolymers **24**, 2181 (1985); J. Noolandi, J. Rousseau, G. W. Slater, C. Turmel, and M. Lalande, Phys. Rev. Lett. **58**, 2428 (1987).

[2] G. F. Carle, M. Frank, and M. V. Olson, Science **232**, 65 (1986).

[3] S. Waki, J. D. Harvey, and A. R. Bellamy, Biopolymers **21**, 1909 (1982); T. K. Attwood, B. J. Nelmes, and D. B. Sellen, Biopolymers **27**, 201 (1988).

[4] J. M. Deutsch, Science **240**, 922 (1988).

[5] J. M. Deutsch, Phys. Rev. Lett. **59**, 1255 (1987).

[6] J. L. Viovy, Phys. Rev. Lett. **60**, 855 (1988).

[7] G. Holzwarth, C. B. McKee, S. Steiger, and G. Crater, Nucleic Acids Res. **15**, 10031 (1987); G. Holzwarth, K. J. Platt, R. W. Whitcomb, and C. B. McKee, Biophys. J. **53**, 407a (1988); W. A. Baase, D. P. Moore, and J. A. Schellman, Biophs. J. **53**, 408a (1988).