Simulation of Megabase DNA Undergoing Gel Electrophoresis

T. A. J. Duke and J. L. Viovy

Groupe de Physico-Chimie Théorique, Ecole Supérieure de Physique et de Chimie Industrielles de Paris, 10 rue Vauquelin, 75231 Paris CEDEX 05, France

(Received 5 August 1991)

A novel type of simulation has been designed to study the dynamics of long DNA fragments, containing millions of base pairs, as they migrate during gel electrophoresis. It shows that in a continuous field, large molecules assume a branched configuration—a behavior qualitatively different from that previously described for smaller fragments. Simulations in crossed fields reveal a new kind of reorientation mechanism which explains the well-known but puzzling fact that the power of separation increases abruptly as the angle between fields is raised above 90° and remains approximately constant for a wide range of obtuse settings.

PACS numbers: 82.45.+z, 05.40.+j, 36.20.Ey

Despite the phenomenal success of the pulsed-field gel electrophoresis technique [1] for separating large DNA fragments, our knowledge of its mechanism remains far from complete. Theoretical concepts [2,3] based on the reptation model [4], in which the molecule snakes down a "tube" formed by the surrounding gel fibers, have indicated that the fragments are oriented by the field. An equally invaluable source of information has been the direct simulation of Deutsch [5] which revealed an instability in the dynamics that leads to the breakdown of the tube model at rather moderate field strengths. This is associated with the formation of "hernias," where a section of the molecule protrudes from the tube in a doubled-up loop. It causes the fragment to cycle between U-shaped and elongated configurations. A number of separate authors [6–12] have established a connection between the renewal time of molecular conformations and the switch time used in pulsed-field techniques, but a variety of mechanisms for renewal have been proposed and which of these apply to molecules in the megabase (Mb) range, where development of the technique is most desirable, is not settled. Clarification of this point requires a more efficient numerical method.

The Monte Carlo simulation technique is faster than directly solving equations of motion, but in the past, the use of *local* hopping rules based on detailed balance, directly borrowed from the Monte Carlo method used in equilibrium statistical mechanics, has limited its application to equally small molecules [7]. The problem is that electrophoresis is intrinsically a nonequilibrium situation and, while the approximate concept of "local detailed balance" has been shown to describe adequately the dynamics when no energy difference along the chain is much greater than kT (i.e., small fragments and low fields), it is insufficient for the more general case [13]. Consequently, we propose a fresh approach which authorizes *variable range* hopping with rules based on a *dynamical*, rather than equilibrium, argument. This provides a more comprehensive description since, on the one hand, a Boltzmann distribution is naturally approached whenever the overall motion of the molecule is slow enough to allow the equilibration of internal modes, while on the other, the rules reproduce the dynamics of an inextensible rope whenever the chain is stretched taut by the field.

The model is easily pictured. We think of the gel as a randomly connected three-dimensional network of pores with uniform diameter a. The DNA molecule is represented by a chain of N segments, each of contour length a and each permitted to be in one of two states (assumed equally probable at thermal equilibrium): either taut, in which case it stretches between two pores, or slack, residing coiled in a single pore. The sequence of pores occupied by the chain forms, in effect, a tube. In de Gennes' original conception of reptation [4], the overall translation of the molecule in the tube is a consequence of the diffusion of regions of excess chain or "length defects" along the tube axis. By analogy with this description, we regard each slack segment as a defect and aim to describe the movement of the molecule by the hopping motion of the defects along the chain

To derive the hopping rules, we focus our attention on a fully extended section of the chain immediately adjacent to a defect. Such a portion can slide along the tube contour by pulling tight the slack in the defect; simultaneously, excess length will accumulate at the other end of the portion, generating a new defect there. We thereby establish the equivalence of a fictitious construct—defect hopping of range n—and a physical process—the sliding of a string of n consecutive taut segments through the distance of one pore. This is analogous to the use of holes as a convenient description of conduction in semiconductors; just as in that case, where one has to determine the motion of the electrons in order to specify the dynamics of the holes, here we can deduce the rules for defect hopping from a consideration of the dynamics of an inextensible string.

It is straightforward to calculate the resolved force that a string experiences due to the external field and write the Smoluchowski equation for its motion along the tube contour, which includes a Brownian term and a friction coefficient proportional to the string's length. This can be solved to obtain a mean time scale for the motion through a distance of one pore (see also Ref. [14]), inversion of which yields an intrinsic rate r for a defect hop of range

© 1992 The American Physical Society

n:

$$r = \frac{\theta x/a}{n\{1-\exp(-\theta x/a)\}} r_0,$$

where $\theta = qEa/kT$ and x is the displacement of the hop measured in the field direction, q the effective charge carried by each segment, E the electric field, and r_0 a microscopic rate constant. At any instant, a defect typically has the opportunity to make a variety of hops of different range, landing anywhere between its present position and the location of the next defect on the chain. Each possible hop will have a different likelihood, reflected by its intrinsic rate. We propose that, in a fixed short time interval, the overall chance that the defect changes position is governed by its most likely motion; i.e., the probability of a move is proportional to the highest intrinsic rate. If a move is ordained by the draw of a random number, we then select the range of hop from among the various possibilities with a probability proportional to its intrinsic rate. The nature of the approximation involved here is first to average over the thermal noise to obtain a set of mean transition rates, then to reinsert randomness by choosing among the possible jumps. We have checked that this procedure yields the expected dynamical behavior in particular cases where an analytical argument is feasible (such as uniformly oriented molecules and fragments extended in U-shaped configurations) as well as the proper equilibrium statistics in low fields when the chain is closely Gaussian (the case $N\theta^2 < 1$) [2].

One further rule is added to the model to allow for the formation of hernias (a similar modification of the tube model has also been suggested by Zimm [15] and by Smith, Heller, and Bustamante [16]). Whenever two adjacent defects occupy the same pore, we authorize their conversion to a pair of taut segments, forming a loop that makes an excursion into a neighboring pore and returns. Thereby, subtubes branching off the main tube may be created and, since hernias can equally well grow out of a subtube, these branches may, in turn, subdivide. The reverse move that eliminates a loop is also permitted. The rule governing these events must take account of the entropic penalty of a loop, due to the constraint that one of the segments must retrace the path of the other instead of visiting a random neighboring pore. We therefore specify that the ratio of the probabilities of loop formation and annihilation is $1/z$, where z is the average coordination number of the gel pores. Assuming a random arrangement of uniformly sized pores, $z = 6$.

We have studied the behavior of this model for long chains in a variety of field conditions, both continuous and pulsed. To establish a correspondence with experimental data, we estimate that for double-stranded DNA in 1% agarose each segment corresponds to 1000 base pairs, $\theta = 1.0$ represents a field of order 10 $\mathrm{V\,cm^{-1}}$, and $a = 300$ nm.

In continuous fields, we observe three distinct types of behavior, depending on the molecular size and the field strength. For weak fields ($\theta \ll 1$) and short fragments ($N < 100$) the growth of hernias is entropically inhibited and the motion is consistent with the standard tube model [2,3]. The mobility varies inversely with molecular weight and falls to a plateau value as the fragment size increases. Chains of intermediate size ($100 < N < 1000$) display the "inchworm" dynamics discovered by Deutsch [5], cycling fairly regularly between extended and U-shaped conformations. In this range, the mobility is independent of molecular weight. Very long molecules ($N > 1000$) have a quite novel comportment, an example of which is shown in Fig. 1. They adopt ramified configurations in which the main tube is oriented along the field direction but hernias (also aligned with the field) branch off it at intervals; those which originate near the tail of the tube are longer than those at the head. As the molecule advances, all of the hernias grow simultaneously until the one most in the rear is released as the trailing end of the chain passes along it. The subsequent rapid retraction of the tail transfers slack to the head of the chain, initiating a new hernia there so that the overall form of the conformation is preserved. Thus, while there is a cycle where individual hernias grow and retract, the global motion is akin to biased reptation [2,3] with large tube length fluctuations. Inchworm dynamics [5] corresponds to the special case where the chain is not long

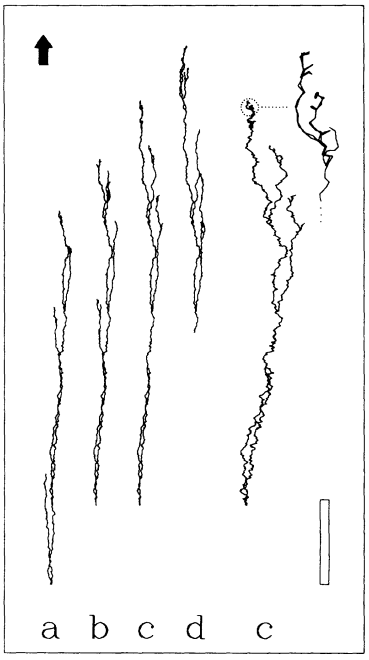

FIG. 1. Motion of a chain with $N = 2000$ segments (corresponding to 2 Mb) in a continuous field, strength $\theta = 1.0$, applied in the direction of the arrow. The time between each snapshot (a–d) is 2000 Monte Carlo units. The scale bar is the length of 100 pores (corresponding to 30 μm). One of the configurations is redrawn with an expanded horizontal scale to give a clearer indication of its branched structure, and its head section is magnified 10 times to show the leading hernias on the scale of the gel pores.

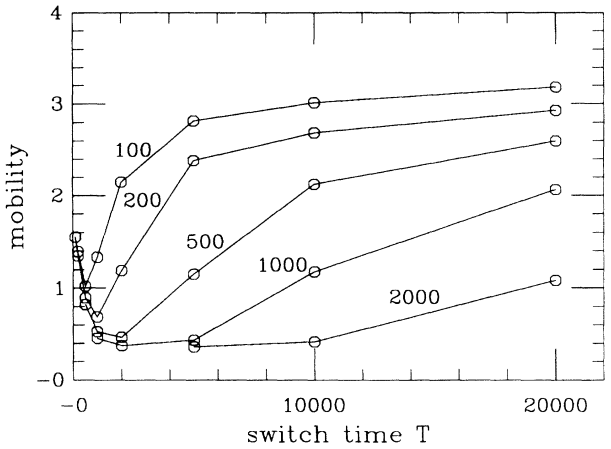

FIG. 2. Variation of mobility with pulse time T for chains of size $N=100$ to $N=2000$ in a crossed-field simulation; field angle $=120°$, field strength $\theta=1.0$. The error in the data is less than 5%.

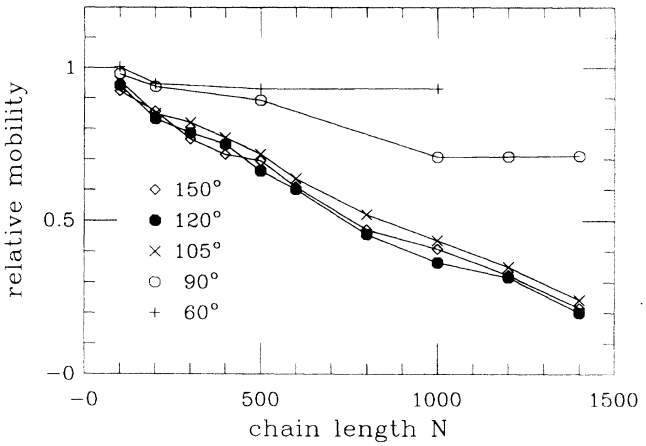

FIG. 3. Variation of mobility with chain length N as a function of field angle Φ. The pulse time is fixed at $T=10000$, and the field strength is $\theta=1.0$. The mobility is measured relative to its value in the limit of very long pulse times: $\mu_\infty = \mu_0 \cos(\Phi/2)$, where μ_0 is the continuous-field mobility (independent of chain length).

enough to support many hernias at the same time. It is interesting to note that the mobility remains at the same plateau value throughout.

We now turn our attention to pulsed-field situations. Figure 2 shows the mobility as a function of switch time and molecular size for a field angle of 120°. The agreement with the systematic experimental studies of Birren *et al.* [17] and Mathew, Smith, and Cantor [18] is very close. Note that the mobility curves display minima at a switch time proportional to the molecular weight so that a clear separation of fragments requires a pulse period that increases linearly with their size.

A number of different field configurations have been investigated. As indicated in Fig. 3, we find that a range of obtuse angles, from 105° to 150°, yield practically identical patterns of separation. An orthogonal field setting, on the other hand, differentiates very poorly between fragments. The abrupt improvement in separation above 90° is well known in practice [17] but has remained a puzzle since the inception of the technique. The reason is readily discerned by watching a film of the simulation at a pulse time that significantly reduces the mobility at 120° but not at 90°. While the motion is somewhat irregular and the types of configurations observed vary from cycle to cycle, the snapshots reproduced in Fig. 4 are representative of the most typical behavior during a single pulse. Immediately before the field is switched, the chain has a U-shaped conformation; for orthogonal fields, the U has a broad base formed during the corresponding pulse of the previous cycle, but at 120° it is usually very narrow. With the change of field, the chain ends turn to move in the new direction and simultaneously, hernias sprout from places where there is a buckle in the tube (it is at these points that a sudden drop in the field gradient mea-

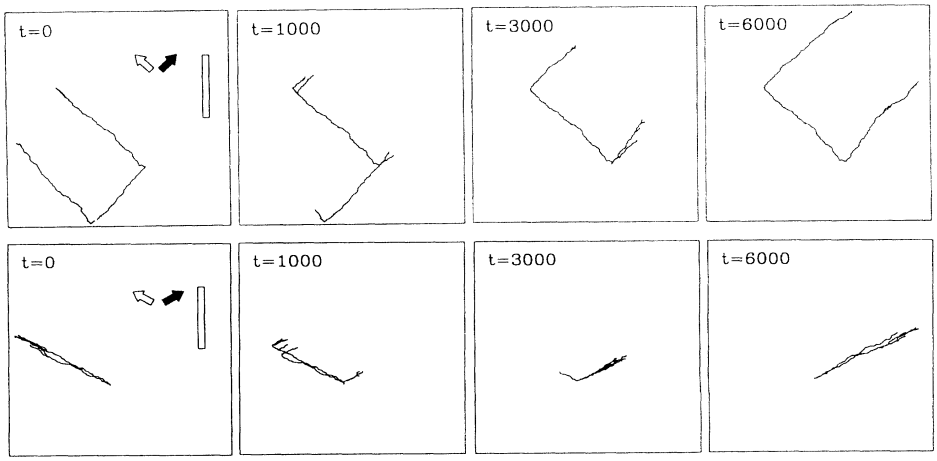

FIG. 4. Motion of a chain of size $N=1000$ during a single pulse in crossed-field electrophoresis with 90° (top row) and 120° (bottom row) field settings. Switch time $=6000$ Monte Carlo units, field strength $\theta=1.0$, and the field directions are indicated by the arrows. The time t is measured from the start of the pulse. The scale bar is the length of 100 pores.

sured along the tube axis causes an accumulation of chain, which can then seep out of the tube). Now a difference arises between the two cases: In the 90° setting there is a fairly even competition between the leading chain end and the foremost hernia and both continue to extend until the hernia unravels to produce, at the end of the pulse, another broad U, displaced further along the gel: In the 120° case, on the other hand, owing to the obtuse angle, a hernia created at the base of the U starts with an advantage over the chain end. It soon establishes superiority and pulls the extremity back into the arm of the U. Rapid retraction of the chain down the arm is then driven by (a component of) the field, leading to bunching near the base which immediately generates more hernias. The hernias grow, sucking up the rest of the molecule and finally unwinding so that, at the end of the pulse, the configuration is once again a narrow U shape. This behavior is entirely consistent with direct observations of DNA using fluorescent microscopy [19,20]. We conclude that after an orthogonal field switch the molecules continue to move forward as they reorient, at a speed independent of size so that no resolution can be obtained. By contrast, with the use of obtuse angles, the chains make little progress while they change direction (they retract backwards at first, then start to move forwards, but get held up while hernias resolve into a U shape); since longer chains take more time to reorient, their progress is hindered for a longer period and it is this feature that enables the segregation of fragments according to size.

The obtuse angle mechanism has some similarities with the "ratchet" motion [8,9] previously proposed on the basis of the biased reptation model [2,3]. However, in that case it was assumed that the molecules are initially uniformly aligned and reorient as the hindmost end slides into a new tube; instead, we find that the chains alternate between U shapes and the reorientation is mediated by hernia growth. This difference explains why the resolution does not vary gradually with the field angle but improves sharply above 90° and remains even for a wide range of obtuse settings. The reorientation in orthogonal pulses has some similarities with the competing-hernia mechanism discussed by Deutsch [21]. However, in common with experimental observations [22], we find that numerous hernias form and compete only if an oriented chain is allowed to relax with the field turned off before the field switch is made. In standard pulsed-field situations, by contrast, rather few hernias grow, preferentially from special locations where the tube is kinked, and the chain ends tend to dominate reorientation.

In conclusion, this model clearly delineates the different regimes of motion of DNA in a continuous field and accounts very well for the most straightforward pulsed-field technique. Our Monte Carlo method, suitably adapted, may prove to be useful in other problems concerning systems far from equilibrium, for which direct solution of the equations of motion is too slow.

We thank Michael Rubinstein for helpful discussions and Tony Maggs for assistance with computer graphics. This work was funded by the EEC SCIENCE programme, Grant No. SC1000475. Groupe de Physico-Chimie Théorique, Ecole Supérieure de Physique et de Chimie Industrielles de Paris, is Unité associée au CNRS No. 1382.

[1] D. C. Schwartz and C. R. Cantor, Cell **37**, 67 (1984).
[2] O. J. Lumpkin, P. Dejardin, and B. H. Zimm, Biopolymers **24**, 1573 (1985).
[3] G. W. Slater and J. Noolandi, Phys. Rev. Lett. **55**, 572 (1985).
[4] P. G. de Gennes, J. Chem. Phys. **55**, 572 (1971).
[5] J. M. Deutsch, Science **240**, 922 (1988).
[6] B. H. Zimm, Phys. Rev. Lett. **61**, 2965 (1988).
[7] T. A. J. Duke, Phys. Rev. Lett. **62**, 2877 (1989).
[8] J. L. Viovy, Electrophoresis **10**, 429 (1989).
[9] G. W. Slater and J. Noolandi, Electrophoresis **10**, 413 (1989).
[10] J. Noolandi, G. W. Slater, H. A. Lim, and J. L. Viovy, Science **243**, 1456 (1989).
[11] M. Olvera de la Cruz, D. Gilip, and E. O. Schaffer, Phys. Rev. Lett. **64**, 2324 (1990).
[12] T. L. Madden and J. M. Deutsch, J. Chem. Phys. **94**, 1584 (1991).
[13] M. Olvera de la Cruz, J. M. Deutsch, and S. F. Edwards, Phys. Rev. A **33**, 2047 (1986).
[14] J. Noolandi, J. Rousseau, G. W. Slater, C. Turmal, and M. Lalande, Phys. Rev. Lett. **55**, 572 (1987).
[15] B. H. Zimm, J. Chem. Phys. **94**, 2187 (1991).
[16] S. B. Smith, C. Heller, and C. Bustamante, Biochemistry **30**, 5264 (1991).
[17] B. W. Birren, E. Lai, S. M. Clark, L. Hood, and M. I. Simon, Nucl. Acids Res. **16**, 7563 (1988).
[18] M. K. Mathew, C. L. Smith, and C. R. Cantor, Biochemistry **27**, 9205 (1988).
[19] S. B. Smith, P. K. Aldridge, and J. B. Callis, Science **243**, 203 (1989).
[20] D. C. Schwartz and M. Koval, Nature (London) **338**, 520 (1989).
[21] J. M. Deutsch, Phys. Rev. Lett. **59**, 1255 (1987).
[22] S. Gurrieri, E. Rizzarelli, D. Beach, and C. Bustamante, Biochemistry **29**, 3396 (1990).