

Collision of a Field-Driven Polymer with a Post: Electrophoresis in Microlithographic Arrays

Edith M. Sevick

Department of Chemical Engineering, University of Colorado, Boulder, Colorado 80309-0424

David R. M. Williams

Institute of Advanced Studies, Research School of Physical Sciences and Engineering, The Australian National University, Canberra, ACT, Australia

(Received 19 October 1995)

We study the collision of a DNA or other polyelectrolyte chain with a point obstacle in the presence of an electric field via computer simulation. We find a very strong dependence of the average collision time, $\langle t_c \rangle$, and the average distance traveled during a collision, $\langle z_c \rangle$, upon the impact parameter, b . Despite the complexities of the chain-post interaction, $\langle t_c \rangle$ and $\langle z_c \rangle$ follow universal curves over a large range of molecular weights and field strengths. This result provides analytic formulas for the chain's mobility in an array of posts and yields insight into the effect of post spacing.

PACS numbers: 87.15.-v, 05.60.+w, 36.20.Ey, 82.45.+z

Electrophoresis is one of the most widely used techniques for size-separating charged chains such as nucleic acids or synthetic polyelectrolytes. The separation is achieved by driving charged chains through a random array of obstacles with an electric field. Traditionally, the obstacles are fibers of a gel; but, more recently, lithographically etched arrays of silicon have been proposed and demonstrated as a novel electrophoretic medium [1–3]. This medium has several advantages over traditional gel media; one of these is that nearly any array geometry can be fabricated. This freedom allows one to design an array geometry that provides a large variation of chain mobility with chain length, and, hence, enhances size separation. The crucial question is: What array geometry is optimal? The answer to this question requires a detailed knowledge of the interaction between a chain and an obstacle. Fluorescence imaging of single DNA chains during electrophoretic migration through gels [4–9] and a post array [1–3] demonstrates that this interaction is complex. The time and the downfield advance of the chain during its interaction with the post affect the overall mobility of the chain.

In this Letter, we present a computer simulation study of the wide range of interactions experienced by a chain and post, allowing for glancing as well as head-on impacts. Apart from the application to electrophoresis, this is a fundamental problem in polymer science which is analogous to the Rutherford scattering problem in atomic physics and has application to chains in any convective flow which impinge upon stationary obstacles [10]. Our simulations demonstrate that the average collision time, $\langle t_c \rangle$, and the average downfield distance advanced during a collision, $\langle z_c \rangle$, can be scaled to fall on universal curves which are independent of the molecular weight and field strength. Furthermore, these quantities depend strongly on the type of impact. These universal formulas are used to construct a general expression for the chain mobility in an array of posts in the single post approximation.

The polyelectrolyte chain is modeled as follows. Each chain has a degree of polymerization N , monomer size a , an effective charge per unit length λ , and is surrounded by fluid of viscosity η . All lengths are measured in terms of the monomer size. An electric field of strength E is applied in the z direction, and the x direction is perpendicular to the field. The polyelectrolyte is free draining and, in the absence of any obstacles, drifts along the field direction with a speed $v_0 = E\lambda/\eta A$, where A is a numerical constant. As shown previously [8,11] there are two distinct regimes, strong and weak stretching, delineated by $N\mathcal{E}$, where $\mathcal{E} \equiv E\lambda/kT$ is the dimensionless field strength. Here we consider only the high-field, strong-stretching regime $\mathcal{E}N \gg 1$, which describes usual experimental conditions. In this high-field regime, diffusion can be ignored and the motion of each chain can be described by a deterministic model.

Figure 1 represents events of the chain-post interaction. A chain impacts a post with impact parameter b , defined as the distance between the chain's center of mass and the post, measured perpendicular to the field. At short times, portions of the chain extend downfield on either side of the post, forming a multiply hooked conformation. The time from initial impact to hooking is assumed to be $t = z_c^0/v_0$, where z_c^0 is the chain's center of mass, measured in the z direction from the post; as such the hooking does not significantly slow the chain. However, subsequent motion of the chain may be seriously frustrated as different arms of the chain compete for length and exchange monomers across the pivot obstacle. The monomer exchange between arms is found from a simple one-dimensional equation of motion [8,11,12]:

$$\frac{\eta A}{kT} \frac{\partial z(n)}{\partial t} = \mathcal{E} \operatorname{sgn}[z(n)] + \frac{\partial}{\partial n} F\left(\frac{\partial z}{\partial n}\right),$$

where $z(n)$ is the position of the n th monomer, measured downfield of the post and assigned negative or positive sign, depending upon whether the monomer belongs to

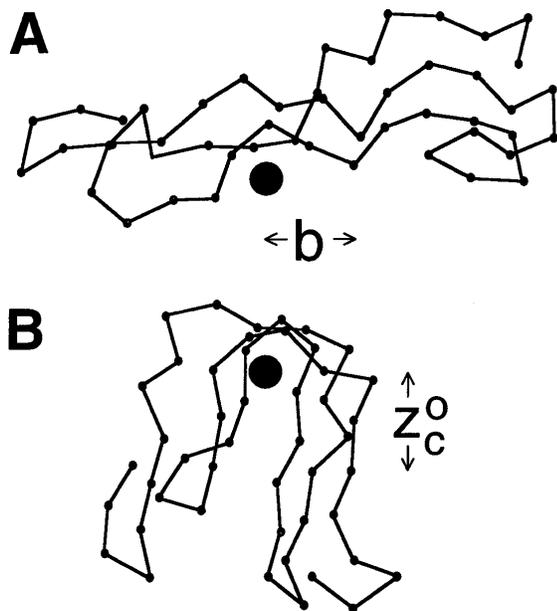


FIG. 1. (A) A schematic of an electric field driven polyelectrolyte prior to impact with a stationary post. The chain collision may be glancing or head-on and is characterized by the impact parameter, b , defined as the distance between the post and chain's center of mass at impact, measured perpendicular to the field direction. (B) A schematic of a chain multiply hooked upon the stationary obstacle with center of mass of the chain, measured from the post in the direction of the field, z_c^0 . The dynamics of unraveling of this hooked chain is modeled as arms which straddle the pivot and compete for arm growth in the field direction.

an arm hooked to the left or to the right of the obstacle. The spring force connecting beads is given by the inverse Langevin form $F(y) = \mathcal{L}^*(y)$ [13]. In practice, the detailed form of the force law does not matter provided it is linear at small extensions and diverges at the monomer size. The equation of motion, integrated forward in time until all monomers lie on one side of the post, provides the time of collision, t_c . The chain's center of mass at t_c is z_c .

The detailed dynamics of the chain interacting with the post is a complicated process [8,14]. However, we can postulate the existence of a characteristic time τ and length w and demonstrate that these are consistent with results of our numerical model. The characteristic time is likely to be related to the unwinding time for a single arm and should have the form $\tau \sim \eta N / \lambda E$. The relevant length scale in the x direction is the radius of gyration in that direction or $w \sim R_x \sim N^{1/2}$. Indeed, as shown in Fig. 2, the numerical data from the equation of motion for various molecular weights and field strengths follow a universal form $\langle t_c \rangle = \eta N / (\lambda E) f(b / N^{1/2})$, where f is a simple exponential provided b is less than a few R_x . We claim no special relevance for this exponential behavior, only that the function is well fitted by an exponential for small values of b . For large values of b the time

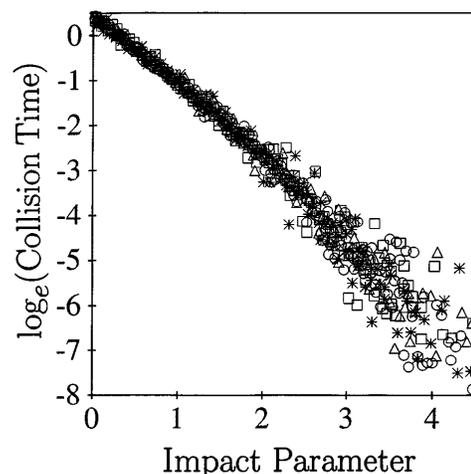


FIG. 2. A plot of the logarithm of the scaled average impact time, $\log_e(\langle t_c \rangle \lambda E / \eta N)$, versus the scaled impact parameter, $b / \langle R_x \rangle$. Here $\langle R_x \rangle$ is the average radius of gyration in the direction perpendicular to the field. The average is taken over a large number of chains with a prescribed degree of polymerization N and dimensionless field \mathcal{E} in the ranges $20 < N < 80$ and $10 < \mathcal{E} < 1000$, where different symbols have been used to indicate select (N, \mathcal{E}) . Chains which do not collide with the post, i.e., start with all their monomers on one side of the post, are assigned zero collision time, $t_c = 0$, and zero center of mass displacement, $z_c = 0$. A total of 3×10^5 chains are sampled in this and following graphs. The data from various (N, \mathcal{E}) collapse onto a single universal curve which appears exponential for $b / \langle R_x \rangle < 2$, i.e., the collision time is a very strong function of the impact parameter. At large b there is much scatter in the data as collisions are very rare.

decays more rapidly than this simple exponential: For $b > N/2$ the collision probability is strictly zero so that the average collision time at these extreme distances must be zero. The functional form, f , is a complex convolution between the probability of a collision and the various dynamical processes occurring during a collision. It is not, for example, a simple multiplication of the collision probability and a constant time, since, as Fig. 3 shows, the collision probability has a very different form and is virtually constant within one radius of gyration.

Figure 4 shows that the numerical data for the average center of mass at release, $\langle z_c \rangle$, versus the impact parameter, b , also falls on a universal curve. Here the relevant length scale in the z direction is N , as the chain is strongly stretched. Thus all of the data can be plotted as $\langle z_c \rangle = N g(-b / N^{1/2})$. The function g is not exponential, but is linear for about two radii of gyration.

Apart from the average collision time and distance, nonaveraged values are also of interest. Figure 5 plots the scaled center of mass distance traveled after impact, z_c / N , versus the scaled collision time, $t_c / (\eta N / \lambda E)$. Although there is some scatter in the plot, the numerical data all fall close to a simple universal curve: $z_c = N h(t_c / \tau)$. The function h is linear at small times, corresponding to chains which barely interact with a post. For long

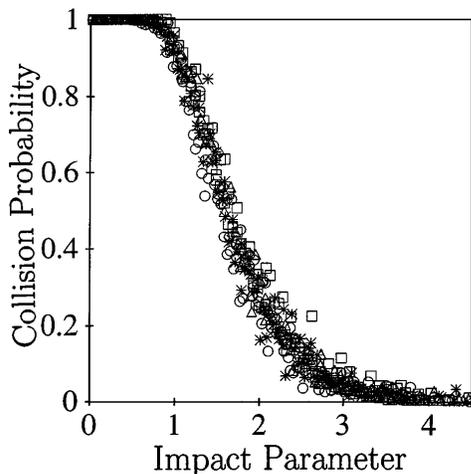


FIG. 3. A plot of the collision probability versus scaled impact parameter, $b/\langle R_x \rangle$. The data are the same as those in Fig. 2, where the average is taken over sets of chains of given N in dimensionless field strength \mathcal{E} .

collision times, the function h saturates at $\frac{1}{2}$. This is indicative of dynamics of a single hairpin where, in the strong field limit, release occurs when all monomers have been exchanged to a single arm of length $\sim N$ and center of mass $z_c \sim N/2$. Such hairpins have been seen experimentally via fluorescence microscopy of individual DNA molecules [4–7,15].

Although the universality of the numerical results is interesting in itself, these relations are useful in deriving a general formula for the mobility of a chain in an array of posts. Let the posts be arranged in rows separated dis-

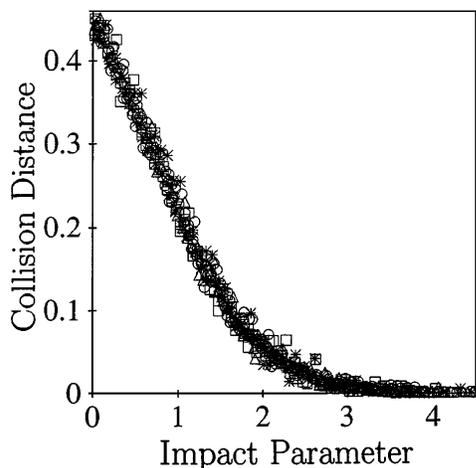


FIG. 4. A plot of the scaled average distance moved by the center of mass during a collision, $\langle z_c \rangle/N$, versus the scaled impact parameter $b/\langle R_x \rangle$. The data are the same as those in Fig. 2 where the average is taken over sets of chains of given N in dimensionless field \mathcal{E} . The data from various (N, \mathcal{E}) collapse onto a single universal curve which is linear for nearly direct impacts.

tance Z apart in the z direction and X apart in the perpendicular direction. The time taken for a chain to traverse a row of posts is the time for collision plus the time to drift to the next row: $t = t_c + (Z - z_c)/v_0$. The ratio of the mobility of the chain with obstacles, μ , to the mobility in the absence of obstacles, μ_0 , is simply the ratio of the chain velocities with and without obstacles: $\mu/\mu_0 = Z/(Z + t_c v_0 - z_c)$. The process of chain impact is stochastic, so that there is a distribution of times t_c and distances z_c which we characterize with the probability distribution P_b defined such that $P_b(t_c, z_c) db dt_c dz_c$ is the probability that a collision occurs with impact parameter between b and $b + db$, with collision time between t_c and $t_c + dt_c$, and center of mass advancement between z_c and $z_c + dz_c$. The average mobility is then given by $\langle \mu \rangle = \int db dz_c dt_c P_b(t_c, z_c) \mu$. This expression can be simplified by invoking the universal relation between z_c and t_c demonstrated in Fig. 4, $z_c = Nh(t_c/\tau)$. In addition, our data [16] for given $b < \sqrt{N}$ show that the standard deviation in t_c is less than $\langle t_c \rangle$, and, consequently, we approximate the distribution of t_c with a delta function centered at $\langle t_c \rangle$. Further simplification can be made by approximating the universal functions according to Figs. 2 and 5: Ignoring numerical prefactors, $f(x) = \exp(-x)$, $h(x) = x$ for $x < 1/2$, and $h(x) = 1/2$ for $x > 1/2$. In the limit $N \ll Z$, this expression reduces to

$$\frac{\langle \mu \rangle}{\mu_0} = 1 - BN^{3/2} X^{-1} Z^{-1},$$

where B is a numerical constant. The decrease in the mobility $\sim N^{3/2} X^{-1} Z^{-1}$ has a simple physical explanation.

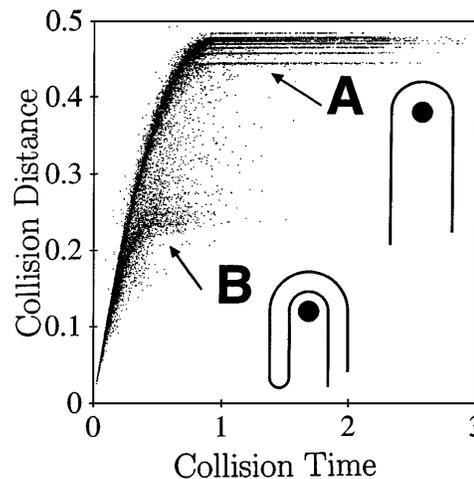


FIG. 5. A plot of the scaled center of mass advance, z_c/N , versus the collision time, t_c , for each of the 3×10^5 simulated chains. (A) At long times, single hairpins form with a competitive pair of arms [inset (a)]. In this frustrated conformation, the final center of mass, z_c , differs slightly among chains of different N and \mathcal{E} as indicated by the discrete levels. However, in the strong-stretching limit $N\mathcal{E} \rightarrow \infty$ the final displacement z_c approaches $\frac{1}{2}$. (B) At shorter times, the data show evidence for doubly hooked hairpins with $z_c \approx \frac{1}{4}$ [inset (b)].

It arises from a factor $N^{1/2}/X$ representing the probability of collision, multiplied by a factor N , representing the decrease in speed of the chain, and a factor Z^{-1} which is the number of obstacles per unit length in the Z direction. The mobility is a strongly decreasing function of N , and thus an array of posts should provide a reasonable means of chain separation.

The mobility formula given above is based upon the single-post approximation, where a chain interacts solely with one post at a time and the interaction is complete before impinging upon an additional post. In practice, however, the post-chain interaction will interfere with subsequent interactions. This interference can take two forms. First, if the radius of gyration is larger than the X spacing, then a chain can interact with more than one obstacle at one time [17]. The second effect deals with conformational relaxation of the newly released chain. After unhooking from a hairpin configuration a chain is strongly distorted. It first relaxes rapidly due to nonlinear elasticity, but then much more slowly due to linear elastic effects. In this latter stage the relaxation is controlled by the Rouse time $\tau_R \sim \eta N^2/kT$, which can be compared with the time taken to travel between rows of posts, $Z\eta/E\lambda$. Thus, for the chain to fully relax between rows requires $Z > \mathcal{E}N^2$. Row spacing less than this would violate the single post approximation and render the process non-Markovian, requiring a more sophisticated treatment.

E. M. S. acknowledges financial support for this work from Engineering Foundation initiation grant RI-A-94-11, the Colorado Institute for Research in Biotechnology (CIRB). D. R. M. W. was supported by a QEII fellowship.

- [1] W. D. Volkmuth and R. H. Austin, *Nature (London)* **358**, 600 (1992).
- [2] R. H. Austin and W. D. Volkmuth, *Analysis* **21**, 235 (1993).
- [3] W. D. Volkmuth, T. Duke, M. C. Wu, R. H. Austin, and A. Szabo, *Phys. Rev. Lett.* **72**, 2117 (1994).
- [4] S. Matsumoto, K. Morikawa, and M. Yanagida, *J. Mol. Biol.* **152**, 501 (1981).
- [5] S. B. Smith, L. Finzi, and C. Bustamante, *Science* **258**, 1122 (1992).
- [6] S. Gurrieri, E. Rizzarelli, D. Beach, and C. Bustamante, *Biochemistry* **29**, 3396 (1990).
- [7] L. Song and M. F. Maestre, *J. Biomol. Struct. Dyn.* **9**, 87 (1991).
- [8] Y. Masubuchi, H. Oana, T. Akiyama, M. Matsumoto, and M. Doi, *J. Phys. Soc. Jpn.* **64**, 1412 (1995).
- [9] H. Oana *et al.*, *Macromolecules* **27**, 6061 (1994).
- [10] Since our model is free draining, the results for hydrodynamic flows would usually be quantitatively different, but qualitatively similar.
- [11] E. M. Sevick and D. R. M. Williams, *Phys. Rev. E* **50**, R3357 (1994).
- [12] Although diffusion is generally unimportant, it can be important at very early times in long-lived hairpin states with equal arms [11]. Such configurations can get trapped, and we untrap them by slightly changing each monomer position.
- [13] P. J. Flory, *Principles of Polymer Chemistry* (Cornell, Ithaca, 1953).
- [14] G. I. Nixon and G. W. Slater, *Phys. Rev. E* **50**, 5033 (1994).
- [15] S. B. Smith, P. K. Aldridge, and J. B. Callis, *Science* **243**, 203 (1989).
- [16] E. M. Sevick and D. R. M. Williams (unpublished).
- [17] J. M. Deutsch, *Phys. Rev. Lett.* **59**, 1255 (1987).