DNA electrophoresis in microfluidic post arrays under moderate electric fields

Kevin D. Dorfman*

Department of Chemical Engineering and Materials Science, University of Minnesota, 421 Washington Avenue SE, Minneapolis, Minnesota 55455, USA

(Received 24 March 2006; published 29 June 2006)

The dynamics of long DNA moving through microfluidic arrays of micron-sized posts under a moderate electric field are modeled by a Scher-Lax continuous-time random walk. The microscale model consists of a repetitive sequence of three steps: (i) collision with the post and extension into two arms, (ii) electric-field-driven unhooking from the post, and (iii) uniform translation until the next collision. The model features two random variables: the initial offset between the two arms of the polymer during a given collision and the distance traveled between collisions. For experimentally realistic values of the electric field strength and DNA molecular weight, scaling laws indicate that the chain will generally be in a stem-flower conformation when unhooking from the post. Compared to a taut-chain model at the same field strength, the stem-flower conformation reduces the time engaged with the post and increases the collision frequency. Analytical expressions for the mean velocity and dispersivity are derived as a function of the post density, post spacing, free-solution mobility, Kuhn length, and sequence length. The incomplete extension of the chain does not strongly affect the mean velocity, but tends to increase the dispersivity relative to a taut chain. As a result, the separation resolution decreases as the field decreases for a moderate field, in agreement with experiments. The quantitative agreement between the model and experimental data is satisfactory, especially considering that the model contains no adjustable parameters.

DOI: 10.1103/PhysRevE.73.061922 PACS number(s): 87.15.Tt, 05.40.Fb

I. INTRODUCTION

DNA electrophoresis is one of the most commonly used tools in molecular biology and genetics. The typical medium for separating double-stranded DNA is an agarose gel. The mean pore size of agarose (several hundred nanometers) is much smaller than the radius of gyration of large DNA (microns). As a consequence, the DNA must unwind and then reptate through the gel. Long DNA cannot be separated in this manner because the reptation tube becomes oriented with the electric field [1]. As a rule of thumb, the upper limit for a gel separation in a constant field is on the order of several tens of kilobase pairs (bp).

One way to circumvent the upper size limit for dc gel separations is to use an artificial array of posts, which can be microfabricated in silicon [2,3], quartz [4], or PDMS [5,6]. Alternatively, a quasihexagonal post array can be formed by the self-assembly of small superparamagnetic beads in a magnetic field [7]. The post diameter in the array is on the order of the radius of gyration of the DNA, whereas the spacing between the posts is large enough to allow the DNA to move freely through the interstices in its coiled conformation. Although the free electrophoretic velocity is independent of molecular weight, from time to time the DNA collides with one of the posts. The nominal time for a collision scales linearly with molecular weight, giving rise to a separation. The power of this separation technique is now well established—long DNA that are impossible to separate in agarose under a dc field can be readily separated in minutes or less in a post array [4,8].

In developing these separation devices, it is essential to know how the mean velocity $\bar{U}^*$ and dispersivity (effective diffusivity) $\bar{D}^*$ depend on molecular weight and the device parameters. The difference in mean velocities between two different species furnishes a measure of the separation between them, whereas their dispersivities quantify the sharpness of the separation. Recently, Minc *et al.* [9] introduced a continuous-time random walk (CTRW) model based on geometration [10] to calculate these quantities. Geometration models the dynamics as the repetitive three-step cycle indicated in Fig. 1. In the first step, the DNA collides with a post and extends into two arms, with the length distribution between the arms being a random variable. Once the two arms are extended, the stronger electrical force acting on the longer arm induces a "rope-over-pulley" disengagement, where the longer arm "eats" the shorter one. Importantly, these models [9,10] assume that the chain is completely extended while unhooking, making the rope-over-pulley analogy a literal one. After unhooking from the post, the DNA

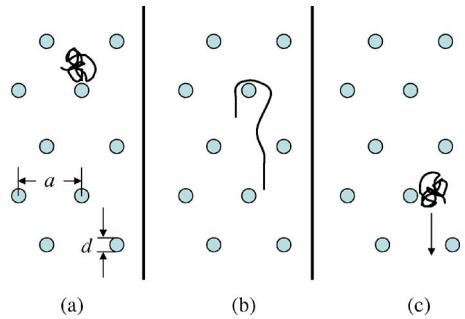

FIG. 1. (Color online) Schematic of the geometrative motion of DNA through a post array. The post diameter d is on the order of the radius of gyration of the DNA, while the post spacing a is much greater. The cyclic motion consists of three steps: (a) collision with the post, (b) unhooking, and (c) translation to the next collision.

*Electronic address: dorfman@cems.umn.edu

translates at its free solution mobility until it collides with the next post. In gel geometration models, the distance between collisions is fixed at the DNA contour length [10] or some deterministic distance farther downstream [11], whereas a model more appropriate for post arrays [9] allows the distance between collisions to be a second random variable that depends on the post density. The cyclic process with a random distance between collisions is equivalent to a Scher-Lax CTRW [12] on a lattice defined by the posts. The walk is characterized by a transition probability density $\psi(n,t)$ that, during a given cycle, the DNA will translate over n rows of the post array during a time t. The net effect of many such collision-translation cycles ultimately gives rise to a mean velocity $\bar{U}^*$, and the fluctuations in the two random variables cause a dispersivity $\bar{D}^*$.

The latter model [9] explains many of the overall trends in the separation process and provides a reasonable prediction of the separation resolution without any adjustable parameters. However, it fails to explain why the separation resolution is lower at lower fields. Explicitly, in the model proposed by Minc *et al.* [9], $\bar{U}^*$ and $\bar{D}^*$ both scale linearly with the electric field E, analogous to a simpler model [13]. If the mean velocity and dispersivity are linear in E, it follows that the separation resolution is independent of the electric field [cf. Eq. (57)]. However, experiments have shown that the separation resolution actually increases with electric field, eventually reaching a plateau at an electric field strength around 20 V/cm [8]. In the experimental work [8], it was suggested that the reduced resolution was a result of incomplete extension during collisions with the posts.

The goal of the present contribution is to quantify the effect of incomplete extension, using the CTRW framework presented by Minc *et al.* [9] and relaxing the requirement that the chain be taut during the unhooking process. The resulting theory retains the positive attributes of the taut-chain model [9] (most importantly, an analytical solution with no adjustable parameters) and limits to the latter when the electrical energy is large compared to the thermal energy. We will see that, for many realistic experimental parameters, incomplete extension due to a weak field plays an important role in determining the dispersivity by reducing the unhooking time and increasing the collision frequency. Moreover, the present model qualitatively accounts for the increase in separation resolution with electric field, furnishing satisfactory quantitative agreement with the experimental data without any adjustable parameters. Finally, as an incidental benefit, some mathematical details omitted for brevity in Ref. [9] will be presented here in greater detail.

The paper is organized as follows. In the next section, we compute the unhooking speed for two different chain conformations [15] and provide a working definition for what is meant by a "moderate" field. Section 3 presents a modification of the CTRW model [9] to account for incomplete stretching, ultimately arriving at analytical expressions for $\bar{U}^*$ in Eq. (48) and $\bar{D}^*$ in Eq. (50). The new results are compared with those for a taut chain [9] and experimental data [8] in Sec. IV, and Sec. V concludes with some possible improvements to the present theory.

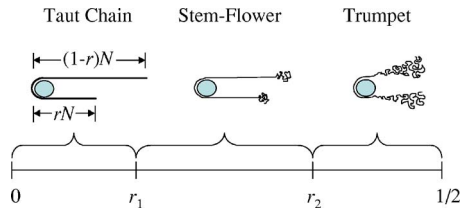

FIG. 2. (Color online) Schematic of the chain conformation at different points during the unhooking process. The values of r_1 and r_2 correspond to critical values of the offset r between the two arms that lead to a change in conformation.

II. COLLISION WITH A POST IN A MODERATE FIELD

Let us first consider what happens in the geometration model when a chain possessing N Kuhn steps of size l_k collides with the post. As depicted in Fig. 1, we assume that, upon colliding, the chain extends into two arms. As indicated in the left-side schematic of Fig. 2, the instantaneous number of Kuhn steps in the longer arm (1) and shorter arm (2) can be computed from $N_1(t)=[1-r(t)]N$ and $N_2(t)=r(t)N$, where $r(t)$ is the current fraction of the weight in the shorter arm. Immediately after unraveling, the length distribution is given by $r(0)=r_0\in[0,1/2]$. At the end of the unhooking process, $r(t)=0$. The extension of each arm is denoted by L_i. In contrast to previous models [9,10], we do not require that the chain be extended to its contour length, $L_i=N_i l_k$. Indeed, as depicted schematically in Fig. 2, the extension of the chain should depend on the tension in the arms, which itself depends on the current value of r. Only for a taut chain is $L_i=N_i l_k$.

The curvilinear velocity v during unhooking is computed from the balance between the electrical forces F_i acting on each of the two arms and the friction resisting the curvilinear motion,

$$F_1 = F_2 + \zeta v. \tag{1}$$

The electrical forces are modeled without hydrodynamic interactions [16],

$$F_i = N_i q_k E, \tag{2}$$

where q_k is the effective charge of a Kuhn step [17] and E is the field strength. Approximating the friction on arm i by $\zeta_i \sim \eta L_i$ in a fluid of viscosity η, the force balance (1) becomes

$$N_1 q_k E = N_2 q_k E + \eta[L_1(v) + L_2(v)]v. \tag{3}$$

To simplify the subsequent analysis, we define the dimensionless velocity and arm length as

$$\bar{v} \equiv v/\mu_0 E, \quad \bar{L}_i \equiv L_i/l_k. \tag{4}$$

Noting that the free-solution mobility is approximately equal to the effective charge on a Kuhn step divided by the friction, $\mu_0 \approx q_k/\eta l_k$, Eq. (3) adopts the form

$$N_1 = N_2 + [\bar{L}_1(\bar{v}) + \bar{L}_2(\bar{v})]\bar{v}. \tag{5}$$

The curvilinear velocity quantifies the rate of mass transfer from the short arm to the longer one. It is thereby equal to

rate that the length of the long arm increases,

$$\bar{v} = \frac{d\bar{L}_1}{d\tau}, \tag{6}$$

where

$$\tau \equiv t\left(\frac{\mu_0 E}{l_k}\right) \tag{7}$$

is a dimensionless time scale.

A. Unhooking velocities

To compute the total unhooking time, it is necessary to first determine how the arm length depends on the curvilinear velocity. To do so, we will make use of classic results for the conformation of a tethered polymer chain in a uniform flow field [14,15]. There are two qualitative differences between the canonical tethered-chain problem [14,15] and the current one: namely, (i) the chain is not physically tethered to the post and (ii) the tension in the chain is the result of an electric field, not a hydrodynamic flow. The first restriction is ameliorated by assuming that the retarding force of the post is strong enough to keep the pivot point of the rope-over-pulley constant, which is equivalent to modeling the instantaneous conformation by two chains tethered on either side of the post, with the relative weights of these chains computed from the offset value $r(t)$. The second concern is resolved by the electrohydrodynamic equivalence principle [18], which states that chain stretching results in a hydrodynamic flow field are equivalent to stretching results with an electrophoretic velocity $v=\mu_0 E$. Indeed, this equivalence was implicitly recognized in a previous study of polyelectrolyte collisions in weak fields [16].

In what follows, it will prove useful to define the dimensionless quantity

$$\lambda_f \equiv \frac{kT}{\eta\mu_0 E l_k^2}. \tag{8}$$

The logic behind this choice will become more apparent shortly. For the moment, though, we can interpret λ_f as the balance between electrohydrodynamic forces acting on the scale of a Kuhn step and the thermal energy. To translate this into realistic experimental values, note that $kT \sim 10^{-21}$ J, the viscosity of a buffer is close to that of water ($\eta \sim 1$ cp), typical free electrophoretic mobilities are around $\mu_0 \sim 10^{-4}$ cm^2/(V s) [8], and the Kuhn length of DNA is $l_k \sim 100$ nm. With the electric field E given in V/cm,

$$\lambda_f \sim \frac{100}{E(\mathrm{V/cm})}. \tag{9}$$

Thus, λ_f is around unity for a "strong" field of 100 V/cm and around a hundred for a "weak" field of 1 V/cm.

Once the velocity exceeds $\bar{v}_1 \approx \lambda_f/N$ [15], the chain adopts the stem-flower conformation indicated in the middle schematic of Fig. 2. In this conformation, the stem portion is taut while the flower remains relaxed like a trumpet. The number of monomers in the flower is given by [15]

$$n_f = 3\lambda_f/\bar{v}, \tag{10}$$

and the size of the flower is [15]

$$X_f = n_f l_k/3. \tag{11}$$

As in Ref. [15], the numerical prefactors are retained to emphasize the incomplete extension of the chain. The choice of λ_f in Eq. (8) is now apparent—it represents the characteristic length scale for the flower.

The extension of the chain is the sum of the stem length, $(N-n_f)l_k$, and the flower, X_f,

$$\bar{L} = N - 2\lambda_f/\bar{v}. \tag{12}$$

The flower regime ends at a critical velocity $\bar{v}_2$, where the flower contains only one Kuhn step. Substituting $n_f=1$ into Eq. (10),

$$\bar{v}_2 = 3\lambda_f. \tag{13}$$

By substituting Eq. (12) into Eq. (5) and making use of the definition of r, the stem-flower curvilinear velocity adopts the form

$$\bar{v} = 1 - 2r + 4\lambda_f/N \quad \text{for } \max(0, r_1) < r < r_2. \tag{14}$$

The lower limit is

$$r_1 = \frac{1}{2}\left(1 - 3\lambda_f + 4\frac{\lambda_f}{N}\right) \approx \frac{1}{2}(1 - 3\lambda_f), \tag{15}$$

where the approximation follows from the fact that $N \gg 1$. The upper limit

$$r_2 = \frac{1}{2} - \frac{N}{3\lambda_f} + \frac{1}{6}\sqrt{\left(\frac{2N}{\lambda_f}\right)^2 - 3} \tag{16}$$

is computed from the analogous calculation for a trumpet conformation [14].

The stem-flower regime disappears when $\lambda_f \to 0$. Thus, we can identify the model presented by Minc *et al.* [9] as corresponding to the infinite-field limit, where the chain is taut,

$$\bar{L} = N, \tag{17}$$

for any value of $r(t)$. In this limit, the velocity decouples from $\bar{L}$ and the taut-chain velocity adopts the simple form

$$\bar{v} = 1 - 2r \quad \text{for} \quad r < r_1. \tag{18}$$

Before proceeding to show the validity of Eq. (14) in a moderate field, a few remarks are in order. First, note that the velocity expressions are discontinuous as r passes through r_1. The source of the discontinuity is the nature of the stem-flower model, which was derived from a scaling argument [15]. In reality, we would expect the chain to evolve smoothly between the two states, with the curvilinear velocity undergoing a continuous transition between regimes. A more detailed model that accounts for the local tension [19,20] would capture this effect. Indeed, the entire analysis thus far is predicated on the assumption that the chain conformation evolves much more quickly than the chain accelerates. This assumption must break down at high fields.

However, changes in the conformation should be small at high fields, since the chain will be in a taut conformation for all values of r (except when $r \approx 1/2$).

B. Unhooking in a moderate field

Having established that the taut-chain model only dominates the chain conformation in the infinite-field limit, we now address the question of what happens in a finite field and how to define a "moderate" field. The taut regime will disappear entirely if r_1 in Eq. (15) becomes negative. In the realistic limit of long chains, where $N \gg 1$, this occurs when

$$\lambda_f \geqslant 1/3. \tag{19}$$

Using Eq. (9) in Eq. (19), the taut regime thus only appears in very strong electric fields, somewhere on the order of several hundred V/cm. As typical experiments [2–4,7,8,21] use tens of V/cm, Eq. (19) predicts that the chains will never actually be taut at all. Videomicroscopy [3,21] confirms this qualitative prediction, as there always appears to be a bright spot at the end of the chain, indicating the presence of a flower.

For very small offsets between the arms ($r \sim 1/2$), the trumpet regime [14,16] must play a role because the velocity is very small. However, the range of r values that accommodate a trumpet collision can be extremely limited. For the large values of N realized in experiments, extremely low fields are needed for the trumpet regime to make a significant contribution. Indeed, André *et al.* [16] noted previously that the range of electric fields that permit a "fully trumpet" unhooking process should be very limited for DNA.

Given that the taut-chain regime disappears for reasonable field strengths and the trumpet regime has a very limited range of validity (most of which is located around a singularity of dubious physicality at $r=1/2$), in what follows we will assume that the chain is always in a stem-flower conformation when unhooking from the post. The latter provides an operational definition for what we mean by a "moderate" electric field—the electric field must be small enough to ignore both the trumpet and taut chain regimes. Neglecting the trumpet corresponds to a lower bound of $N/\lambda_f \approx O(10)$, whereas neglecting the taut chain requires that λ_f be at least $O(1)$. Inasmuch as N is typically several hundred, both conditions should be satisfied in experiments with electric fields on the order of 10 V/cm.

One important consequence in ignoring the trumpet conformation is that it removes a singularity at $r=1/2$. The singularity is a result of neglecting diffusion in the model, since, at very low velocities, diffusion should break the metastable state [3]. In previous analyses of this type [9,10,22], the singularity played a small role, since the divergence is logarithmic and the probability distribution for r_0 is not strongly weighted at $r=1/2$ [22]. The $r=1/2$ singularity disappears in the stem-flower model of the collision process, since the flower "adds" a fixed speed $4\lambda_f/N$ to the taut-chain velocity. We can thus treat the small velocity in the stem-flower state near $r \approx 1/2$ as a crude approximation of a "diffusive escape velocity" from the metastable state.

There is one other minor point to consider in a fully stem-flower model. When the chain comes very close to unhooking from the post, the stem-flower configuration should begin to break down on the short arm as the number of monomers decreases. Moreover, the velocity of the two arms may differ due to changes in the local tension [19]. However, the chain will be moving very quickly at this point (due to the strong electrical force acting on the longer arm), so we would expect that whatever conformation arises in the short arm proximate to unhooking from the post will be short lived and thus contribute little to the total unhooking time.

III. CONTINUOUS-TIME RANDOM WALK MODEL

In the present section, we modify the taut-chain CTRW model [9] to account for a stem-flower configuration during the unhooking step. Note that the unhooking velocity $\bar{v}$ for the stem-flower in Eq. (14) reduces to that of a taut chain [Eq. (18)] in the limit $\lambda_f \to 0$. As a result, the analysis which follows should also reduce to the taut-chain results [9] in the infinite-field limit.

A. Cycle time

When the polymer collides with the post, it is in a coiled conformation and needs to first unravel into two arms. The process is quite complex (see, for example, Ref. [16]), but to make progress analytically [23] we will adopt the preaveraged extension model [10]. In the latter, the speed of the unraveling of each arm is assumed to be proportional to the electrical force acting on the arm. Thus, the total time for unraveling is

$$\tau_1 = N. \tag{20}$$

After forming two arms, the mass distribution between the arms is given by r_0. As in prior work [9,10], r_0 is chosen to be a uniformly distributed random variable between 0 and 1/2. A previous study of post collisions [22] demonstrated that more complicated distributions furnish qualitatively similar results, so we continue to invoke the uniform distribution for simplicity.

Given an initial condition $r(0)=r_0$, we now need to derive the time for the second step—i.e., the time τ_2 where $r(\tau_2)=0$. The differential equation for r is obtained by substituting Eq. (12) for the longer arm 1 into Eq. (6) and then eliminate $\bar{v}$ with Eq. (14), ultimately arriving at

$$-\left[\frac{N}{1-2r+4\lambda_f/N} + \frac{4\lambda_f}{(1-2r+4\lambda_f/N)^3}\right]\frac{dr}{d\tau} = 1. \tag{21}$$

Integrating this equation from $r=r_0$ to $r=0$, the unhooking time is found to be

$$\tau_2(r_0) = \tau_{2_a}(r_0) + \tau_{2_b}(r_0), \tag{22}$$

where the two contributions to the total time are

$$\tau_{2_a} = \frac{N}{2}\ln\left(\frac{1+4\lambda_f/N}{1-2r_0+4\lambda_f/N}\right), \tag{23}$$

$$\tau_{2_b} = \frac{\lambda_f}{(1-2r_0+4\lambda_f/N)^2} - \frac{\lambda_f}{(1+4\lambda_f/N)^2}. \tag{24}$$

In general, the logarithmic term dominates the total unhooking time. Approximating the unhooking time by

$$\tau_2(r_0) \approx \frac{N}{2}\ln\left(\frac{1+4\lambda_f/N}{1-2r_0+4\lambda_f/N}\right) \tag{25}$$

thus only introduces small errors in the final numerical prefactors, while still permitting an analytical solution. Note that $\tau_2(r_0) \sim N/2\ln[1/(1-2r_0)]$ as $\lambda_f \to 0$, which is the unhooking time for the fully extended chain [9,10].

At the instant before the chain disengages from the post, the curvilinear velocity is $\bar{v}=1+4\lambda_f/N$. Using Eq. (12), this corresponds to an extension

$$\bar{L}_f = \frac{N+2\lambda_f}{1+4\lambda_f/N}. \tag{26}$$

In the infinite-field limit, $\bar{L}_f \to N$, which is the extension of a taut chain (17).

After unhooking, the DNA is assumed to translate at a uniform speed $\mu_0 E$ until it hits the next post. The time required to move the rear of the chain (the part that last leaves the post) to the next collision point a distance na downstream would then be $na/\mu_0 E$ or, in dimensionless notation,

$$\tau_3 = n\alpha, \tag{27}$$

where $\alpha=a/l_k$ is the number of Kuhn steps that would fit along the center-to-center post spacing. As in the taut-chain model [9], we assume again that the collision probability is proportional to the areal post density, $\rho=d/a$, for a post diameter d and center-to-center spacing a, as indicated in Fig. 1. Collisions are forbidden along the extended chain backbone, so the first possible collision can occur at row

$$n^* = \bar{L}_f/\alpha. \tag{28}$$

The probability $h(n)$ of colliding in the nth row is given by

$$h(n) = \begin{cases} \rho(1-\rho)^{n-n^*} & \text{for } n \geqslant n^*, \\ 0 & \text{otherwise.} \end{cases} \tag{29}$$

B. Transition probability density for the CTRW

Knowledge of the cycle times τ_i and the collision probability $h(n)$ allows us to compute the transition probability density $\psi(n,t)$ that, during a given cycle, the DNA moves past n posts in a time t. The transition probability density is partially separable into the probability $h(n)$ of making a movement over n rows and the conditional probability $g(\tau|n)$ of making such a movement in time τ,

$$\psi(n,\tau) = g(\tau|n)h(n). \tag{30}$$

The derivation of the conditional probability $g(\tau|n)$ begins with the total time for a cycle through n rows,

$$\tau = N + \frac{N}{2}\ln\left(\frac{1+4\lambda_f/N}{1-2r_0+4\lambda_f/N}\right) + n\alpha. \tag{31}$$

The minimum cycle time is

$$\tau_{\min} = N + n\alpha, \tag{32}$$

and the maximum cycle time is

$$\tau_{\max} = \tau_{\min} + \frac{N}{2}\ln\left(\frac{N+4\lambda_f}{4\lambda_f}\right). \tag{33}$$

The probability $P(\tau \leqslant \tau'|n)$ that the cycle time over n rows is less than some time $\tau' \in [\tau_{\min}, \tau_{\max}]$ is given by

$$P(\tau \leqslant \tau'|n) = \int_0^{\tau'} g(\tau|n)d\tau. \tag{34}$$

Since r is uniformly distributed between 0 and 1/2,

$$P(\tau \leqslant \tau'|n) = 2r_0(\tau'|n), \tag{35}$$

where $r_0(\tau|n)$ is the inverse of Eq. (31),

$$r_0 = \frac{1}{2} + \frac{2\lambda_f}{N} - \frac{1}{2}\left(1+\frac{4\lambda_f}{N}\right)\exp\left\{-\frac{2[\tau-(N+n\alpha)]}{N}\right\}. \tag{36}$$

The equation for $g(\tau|n)$ is then obtained by differentiating Eq. (34) with respect to τ and using Eq. (35),

$$g(\tau|n) = 2\frac{dr_0}{d\tau}, \tag{37}$$

where we have used the fact that $g(0|n)=0$. Substituting Eq. (36) into Eq. (37) and differentiating yields

$$g(\tau|n) = \frac{2(1+4\lambda_f/N)}{N}\exp\left\{-\frac{2[\tau-(N+n\alpha)]}{N}\right\}. \tag{38}$$

The latter density is valid for $\tau_{\min} \leqslant \tau \leqslant \tau_{\max}$, and $g(\tau|n)$ vanishes otherwise. In the infinite-field limit, $g(\tau|n)$ reduces to the result given in Ref. [9].

C. Moments of the walk

The transition probability density $\psi(n,\tau)$ defines a Scher-Lax CTRW on a lattice defined by the posts. The Fourier-Laplace transform of the probability density for the walk is given by [12]

$$p(k,s) = \frac{1-\Lambda(0,s)}{s[1-\Lambda(k,s)]}, \tag{39}$$

where $\Lambda(k,s)$ is the Fourier-Laplace transform of $\psi(n,\tau)$,

$$\Lambda(k,s) = \sum_{n \geqslant n^*} e^{-ikn}\int_{\tau_{\min}}^{\tau_{\max}} e^{-s\tau}\psi(n,\tau)d\tau. \tag{40}$$

The Fourier variable k is made dimensionless with a^{-1}, and the Laplace variable s is made dimensionless with $\mu_0 E/l_k$. Use Eqs. (29) and (38) in Eq. (30), substitute the result into Eq. (40), and then perform the integration and summation to yield

$$\Lambda(k,s) = \frac{\mathcal{L}(s)e^{-s(N+n^*\alpha)}e^{-ikn^*}}{1-(1-\rho)e^{-ik-s\alpha}}, \tag{41}$$

where

$$\mathcal{L}(s) = \frac{2\rho(1+4\lambda_f/N)}{sN+2}\left[1-\left(\frac{4\lambda_f}{N+4\lambda_f}\right)^{(sN+2)/2}\right]. \tag{42}$$

The mth spatial moment of the probability $p(k,s)$ is computed from

$$M_m(s) = i^m\left(\frac{d^m p}{dk^m}\right)_{k=0}. \tag{43}$$

Using $m=0$ in the latter and inverting furnishes $M_0(t)=1$, reflecting the conservation of probability density. The next two moments are given by

$$M_1(s) = \frac{i}{s}\left(\frac{\Lambda_0'}{1-\Lambda_0}\right), \tag{44}$$

$$M_2(s) = \frac{-1}{s}\left[\frac{\Lambda_0''}{1-\Lambda_0}+2\left(\frac{\Lambda_0'}{1-\Lambda_0}\right)^2\right], \tag{45}$$

where the primes indicate differentiation with respect to k and the subscript indicates evaluating the resulting function at $k=0$. In the long-time limit ($s\to 0$), the dimensional moments have the asymptotic forms

$$M_1 \sim \frac{1}{s^2}\left(\frac{\bar{U}^*}{\mu_0 E\alpha}\right)+\frac{B}{s}, \tag{46}$$

$$M_2 \sim \frac{2}{s^3}\left(\frac{\bar{U}^*}{\mu_0 E\alpha}\right)^2+\frac{2}{s^2}\left(\frac{\bar{D}^*+\bar{U}^*Ba}{\mu_0 Ea\alpha}\right), \tag{47}$$

which serve as definitions of the mean velocity $\bar{U}^*$ and dispersivity $\bar{D}^*$. The quantity B is a physically irrelevant constant, which only arises in the calculation of $\bar{D}^*$.

D. Calculation of the mean velocity and dispersivity

The calculations leading up to the final expressions for $\bar{U}^*$ and $\bar{D}^*$ involve a reasonable amount of algebra, which is suppressed here for brevity. By first substituting Eqs. (41) and (42) into Eq. (44), performing the derivative, expanding for small s, and then comparing the leading-order term with Eq. (46), the mean velocity is seen to adopt the form

$$\frac{\bar{U}^*}{\mu_0 E} = \frac{2\alpha(\rho n^*+1-\rho)}{3\rho N+2\rho n^*\alpha+2\alpha(1-\rho)-4\rho\lambda_f\phi}, \tag{48}$$

where ϕ is defined as

$$\phi = \ln\left(1+\frac{N}{4\lambda_f}\right). \tag{49}$$

Note that the dimensionless quantity $\bar{U}^*/\mu_0 E$ is equivalent to the net electrophoretic mobility in the array made dimensionless with the free solution mobility.

Performing the same operations with Eqs. (45) and (47) yields the dispersivity

$$\frac{\bar{D}^*}{\mu_0 Ea} = \frac{\sigma_t+\sigma_{sf}}{[3\rho N+2\rho n^*\alpha+2\alpha(1-\rho)-4\rho\lambda_f\phi]^3}. \tag{50}$$

The first contribution

$$\sigma_t = \rho\alpha N^2[10+\rho(2n^*-11)+(1-n^*)^2\rho^2] \tag{51}$$

is analogous to the result for a taut chain. The second contribution

$$\sigma_{sf} = 8f_1\rho(1-\rho)\alpha\lambda_f\phi-2f_2\rho^2\alpha\lambda_f\phi^2 \tag{52}$$

represents the change in the dispersivity due to incomplete stretching in the stem-flower state. In the latter, the functions f_1 and f_2 are

$$f_1 = 2\alpha n^*+\rho\beta(6N+4n^*\alpha)+(\rho\beta)^2(3N+2n^*\alpha), \tag{53}$$

$$f_2 = N[1+(\rho\beta)^2]+8\lambda_f(2n^*-1)+\rho\beta(2N+8\lambda_f\beta), \tag{54}$$

where $\beta=n^*-1$.

It is straightforward to show that this new result reduces to that derived for a taut chain [9]. In the limit $\lambda_f\to 0$, $n^*=N/\alpha$. Although ϕ diverges logarithmically as $\lambda_f\to 0$, it always appears in combination with positive powers of λ_f. Consequently, all terms resulting from the incomplete extension of the chain vanish. The velocity and dispersivity then reduce to

$$\frac{\bar{U}^*}{\mu_0 E} = \frac{2(\rho n^*+1-\rho)}{5\rho n^*+2(1-\rho)} \tag{55}$$

and

$$\frac{\bar{D}^*}{\mu_0 Ea} = \frac{\rho(n^*)^2[10+\rho(2n^*-11)+(1-n^*)^2\rho^2]}{[5\rho n^*+2(1-\rho)]^3}, \tag{56}$$

which are the results derived in Ref. [9].

Knowledge of $\bar{U}^*$ and $\bar{D}^*$ also allows us to compute the separation resolution—i.e., the ability to separate two species of different size N. While our physical insights will come mostly from an interpretation of Eqs. (48) and (50), the separation resolution is the quantity of interest in experiments and the most easily observed test of the model. For a separation length L_s, the separation resolution is defined as

$$R_s = \frac{\Delta\bar{U}^*}{\langle\bar{U}^*\rangle}\sqrt{\frac{\langle N_{TP}\rangle}{16}}, \tag{57}$$

where $\Delta\bar{U}^*$ is the magnitude of the velocity difference and

$$N_{TP} = \frac{\bar{U}^*L_s}{2\bar{D}^*} \tag{58}$$

is the number of theoretical plates for a given species with average velocity $\bar{U}^*$ and $\bar{D}^*$.

IV. DISCUSSION

A. Average collision time

Let us first consider how the incomplete extension modifies the average time for a collision—i.e., the time to complete the first two steps of the geometration cycle. The probability density $g_c(\tau_c)$ for the total collision time,

$$g_c(\tau_c) = \frac{2(1+4\lambda_f/N)}{N} \exp\left\{-\frac{2(\tau_c - N)}{N}\right\}, \tag{59}$$

is derived in a similar manner to $g(\tau|n)$, except that the contribution from the third step of the geometration cycle is ignored. Following Eqs. (32) and (33), the latter expression for $g_c(\tau_c)$ is only valid for

$$N \leqslant \tau \leqslant N\left(1+\frac{\phi}{2}\right), \tag{60}$$

with $g_c(\tau_c)=0$ otherwise. The first moment of $g_c(\tau_c)$ is the average collision time,

$$\langle \tau_c \rangle = \frac{3}{2}N - 2\lambda_f \phi. \tag{61}$$

Since $\lambda_f \phi$ is always small compared to N, then the average collision time is dominated by the taut-chain contribution, $3N/2$.

Although incomplete stretching slightly reduces the average collision time, the extent of the decrease is insufficient to completely explain the experimental results. In single-molecule experiments on T4 DNA (169 kbp, $N=560$) moving through self-assembled arrays of magnetic posts [21], the average collision time was in the range $0.39N < \langle \tau_c \rangle < 0.53N$. The largest field correction only reduces the taut chain result, $1.5N$, by approximately $0.15N$.

There are at least two possible explanations for the continued discrepancy between the theoretical prediction of $\langle \tau_c \rangle$ and what is observed in experiments. As mentioned in the context of the taut-chain theory [9], the electrophoretic mobility during the unhooking process is assumed to be the same as the free solution velocity. One might expect that the arms are in a hydrodynamically favorable conformation during unhooking, whereupon the mobility appearing in the dimensionless curvilinear velocity equation (14) should be replaced by some effective mobility $\mu' > \mu_0$. Exactly determining μ' for a stem flower is a formidable task, since the stem and flower possess different conformations and thus contribute differently to the total friction. A second explanation is that the decoupling of the unraveling and unhooking processes is an oversimplification. Indeed, there is no reason why the chains should have to extend completely before beginning the rope-over-pulley disengagement, since the electrical force will be out of balance during the unraveling process. Moreover, it follows from Eq. (60) that this model will never predict an average trapping time less than N. The potential of a simultaneous unraveling-unhooking model is reinforced by considering what would happen if the rope-over-pulley time was the only relevant time scale for the total collision process. A naive calculation of the total collision time from the taut-chain model gives $\langle \tau_c \rangle \sim N/2$, which would be in very good agreement with experiments. Although clearly not a proof, this result does suggest that the coupling between unraveling and unhooking should be explored in more detail. A good starting point would be a model that considers the local tension [19,20] during the entire unraveling and unhooking process.

B. Mean velocity

Comparing the stem-flower unhooking velocity in Eq. (14) to its taut-chain counterpart in Eq. (18), it is clear that unhooking is aided by the stem-flower conformation because the relaxation in the flower reduces the total extension of the chain. Since the friction scales linearly with the length of the chain, the drag on a long stem flower is equivalent to a smaller taut chain without reducing the electrical force [24]. However, our result for $\bar{U}^*$ in Eq. (48) indicates that the incomplete stretching only introduces a small logarithmic correction into the retardation caused by the posts [13].

The final correction to $\bar{U}^*/\mu_0 E$ is the result of the competition between two effects. As seen in Eq. (61), the average dimensionless collision time decreases relative to a taut chain as the field is decreased. At the same time, the reduced extension of the chain as it disengages from the post [Eq. (26)] decreases the average distance to the next collision. To be more quantitative, the value of n^* computed from Eq. (28) with Eq. (26) will be less than in the taut-chain case, where $n^* = N/\alpha$. The average number of posts between collisions, given by the first moment of the distribution $h(n)$ [9], is

$$\langle n \rangle = n^* + \left(\frac{a}{d} - 1\right), \tag{62}$$

which decreases with the field [25]. Since the CTRW calculation includes not only the individual contributions from these effects, but also the coupling between them, it is not entirely straightforward to say that the decrease in the collision time dominates the increased collision frequency in determining the effect of the field on the mean electrophoretic mobility. Moreover, the simplified (nonsingular) model adopted for collisions around $r_0 = 1/2$ will also serve to increase the net mobility of a stem flower relative to the model adopted for a taut chain [9].

C. Effective diffusivity

While the electrophoretic mobility only depends weakly on the field, the dispersivity $\bar{D}^*$ [Eq. (50)] can change dramatically with field strength. Before delving into the details, let us first recall the key contributions to the dispersivity. There are two random variables in the model that give rise to dispersion: (i) the initial offset r_0 between the two arms and (ii) the distance between collisions, n. The role of fluctuations in r_0 is straightforward—each collision event can contribute to the dispersion if the offset between the arms at the start of the collision differs from the average offset. On the other hand, the number of rows, n, between collisions makes two different contributions to the dispersion. Similar to r_0, fluctuations in n are an intrinsic source of dispersion. How-

ever, the average distance between collisions also plays a role in determining $\bar{D}^*$ by changing the number of times the distribution for r_0 is sampled. The dual role of n is a natural consequence of the fact that the CTRW is only partially separable; the value of n appears both in the collision probability $h(n)$ and the conditional probability density $g(\tau|n)$, while the value of r_0 only serves to determine τ.

The fluctuations in n should be independent of the field because the probability of colliding in row $n \geqslant n^*$ is assumed to be purely geometric, increasing with the post density ρ. This hypothesis is confirmed by computing the variance in n,

$$\langle n^2 \rangle - \langle n \rangle^2 = \frac{1-\rho}{\rho^2}, \tag{63}$$

which is indeed independent of n^* and thus of the field.

Although the fluctuations in n are constant as the field changes, we have already seen that the dimensionless average collision time $\langle \tau_c \rangle$ decreases, while the frequency of collisions, $\langle n \rangle^{-1}$, increases. Although the two effects cannot be rigorously separated (hence the need for the full CTRW calculation above), we can gain some physical insight by considering each effect individually. The amount of dispersion caused by each collision is approximated by the variance of $g_c(\tau)$,

$$\langle \tau_c^2 \rangle - \langle \tau_c \rangle^2 = \frac{N^2}{4} - \lambda_f N \phi^2 \left(1 + \frac{4\lambda_f}{N} \right). \tag{64}$$

The correction for a finite field can be significant for very weak fields (i.e., when $\lambda_f \sim N$). However, weak fields also result in strong increases in the collision frequency by decreasing n^*. Thus, weak fields lead to many low-dispersion collisions, whereas strong fields lead to fewer high-dispersion collisions. The question then reduces to determining the more important effect, the reduced dispersion per collision [Eq. (64)] or the increased collision frequency [Eq. (62)]. Unfortunately, this is a coupled effect, so continuing to think of the two factors individually will not answer the question.

Rather, it is necessary to consider the full result in Eq. (50). As seen in Fig. 3, the dispersivity increases monotonically with $\lambda_f \sim E^{-1}$ at a fixed value of a. Thus, we can conclude that the increased frequency of collisions dominates the reduction in the dispersion per collision, leading to an increase in the dispersion at weaker fields for the experimentally relevant parameter range considered here. The dispersivity also increases with increasing molecular weight N. The latter is a result of the fact that the dispersion per collision scales like N^2, whereas N-dependent changes to the frequency of the collisions scale like N.

A monotonic decrease in the dimensionless dispersivity with the field is also observed in Fig. 4, where $\bar{D}^*/\mu_0 Ea$ is plotted as a function of λ_f at various values of the post spacing, a. In this figure, the post spacing is fixed at $d=1$ μm and the molecular weight $N=560$ corresponds to T4-DNA. The dependence on a is readily understood by the balance between the two terms: n^* and $(a/d-1)$, appearing in Eq. (62). When a is large and the posts are widely spaced, then changes in n^* do not have a strong effect on $\langle n \rangle$. As a result, the dimensionless quantity $\bar{D}^*/\mu_0 Ea$ varies weakly with the field. When a is not large, then changes in n^* strongly impact $\langle n \rangle$. In this second case, $\bar{D}^*/\mu_0 Ea$ increases sharply as the field decreases, as seen in Fig. 4. The results are less dramatic for shorter chains because n^* plays an ever lessening role in determining $\langle n \rangle$, leading to weaker variations in $\bar{D}^*/\mu_0 Ea$ with the field.

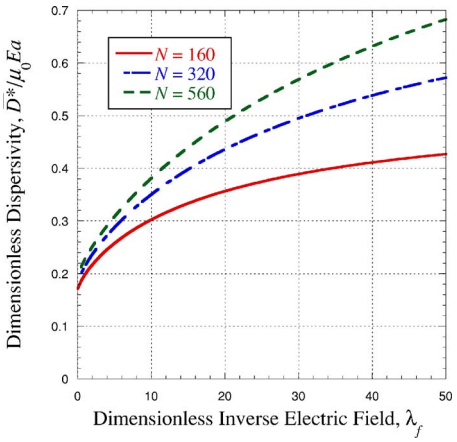

FIG. 3. (Color online) Plot of the dimensionless dispersivity [Eq. (50)] as a function of λ_f for $\rho=0.2$ and $\alpha=50$. In dimensional terms, this corresponds to a post diameter $d=1$ μm and center-to-center spacing $a=5$ μm. The different curves correspond to approximate values for λ-DNA ($N=160$), 2λ-DNA ($N=320$), and T4-DNA ($N=560$).

D. Separation resolution and comparison with experiments

From a practical standpoint, the most important result of the calculation presented here is the separation resolution in Eq. (57), which quantifies the ability to separate two different species. Figure 5 compares the theoretical results for the separation resolution with experimental data on separations

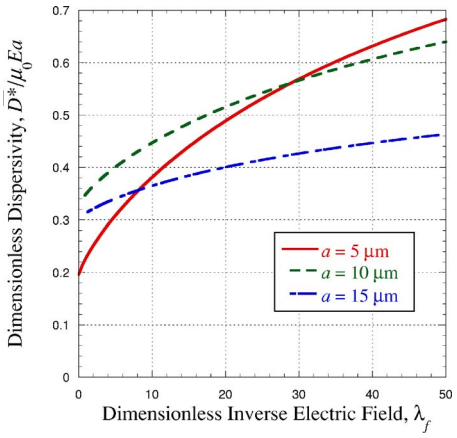

FIG. 4. (Color online) Plot of the dimensionless dispersivity [Eq. (50)] as a function of λ_f for a post diameter $d=1$ μm and T4-DNA ($N=560$). The different curves correspond to increased spacing between posts.

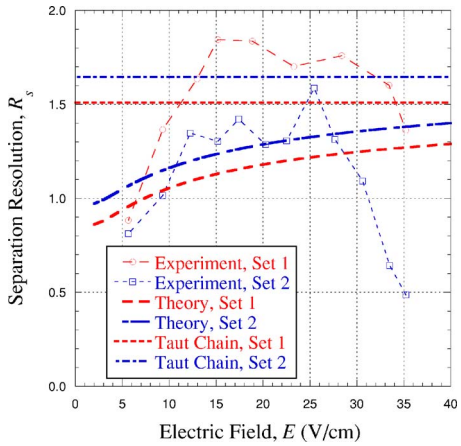

FIG. 5. (Color online) Comparison of experiments [8] and theory for separations of λ-DNA and 2λ-DNA in self-assembled post arrays of magnetic beads. Data set 1 corresponds to $d=1.4$ μm and $a=4.1$ μm. Data set 2 corresponds to $d=1.0$ μm and $a=3.8$ μm. The free mobility is 2.8×10^{-4} cm^2/(V s) [8]. The data were collected a distance $L_s=7.5$ mm downstream from the injection point. The curves labeled "Theory" are computed from Eqs. (48), (50), and (57). The lines labeled "Taut Chain" are computed from Eqs. (55)–(57) and correspond to the model presented in [9]. The dashed lines on the experimental data are added to guide the eye.

between λ-DNA and 2λ-DNA in self-assembled magnetic arrays [8]. The two curves are theoretical results for a moderate field, and the two lines are theoretical results for a taut chain.

The moderate field theory presented here qualitatively captures the key trend in the experiments when the field is less than 20 V/cm: namely, the increase in the separation resolution with increasing electric field. The preceding discussion of $\bar{U}^*$ and $\bar{D}^*$ provides us with the reason behind the increase. The effective species electrophoretic mobilities only weakly depend on the electric field, but the dispersion increases dramatically at lower fields. As a result, the spacing between the bands is relatively insensitive to the electric field, whereas the bands get wider as the field is decreased. The predicted rate of increase slows down as the field increases, mimicking the plateau in the experimental data.

Qualitatively, the results achieved here are a marked improvement over existing models of this particular separation [9,13], which predict that the resolution should be independent of the field. From a quantitative standpoint, the agreement between the present theory and experiment is satisfactory, especially in light of the fact that the theory contains no free fitting parameters. Indeed, all of the parameters appearing in the theory (molecular weight, post size, post spacing, free solution mobility, and electric field strength) are readily accessible from experiments. Although the taut-chain theory appears to be in better quantitative agreement with the experimental data, uncertainties in the experimental data and the physical parameters appearing in the models make such a claim tenuous.

There are two major shortcomings with both CTRW theories (stem flower and taut chain): (i) the order of the theory curves with array density is reversed compared to the experimental data, and (ii) the theories do not predict the loss of resolution at higher electric fields. Both of these deficits are the result of the overly simplistic collision model for $h(n)$ given by Eq. (29). The latter model assumes that the location of the DNA in the unit cell of the post array is unbiased. The probability of the DNA colliding with the post is then estimated by the likelihood that its center of mass will overlap with some part of the post as the DNA passes through the unit cell. In order for the DNA location to be unbiased, the electric field must be uniform. This is a reasonable model for a dilute array, since the perturbation to the electric field caused by an insulating post is localized around that post [5]. The model breaks down in a more dense array, as the electric field is strongly deformed by the nearby posts. Moreover, the curved field lines tend to drive the DNA towards a vertical post-to-post equipotential line that is centered in the gap between two posts (with the direction of motion indicated in Fig. 1), which greatly increases the probability of colliding with the next row in a hexagonal array. As a result, the collision probability increases more quickly than ρ in dense arrays and the resolution thereby increases. This effect is absent in the present collision model, leading to the discrepancy between the theory and experiment.

As the taut-chain model [9] fails to predict the decline in the resolution at strong fields, it is unsurprising that the present model fails as well, since the stem-flower effect is only important at weak fields. The failure of both models at high fields is related to the assumption in Eq. (29) that the chain can collide at row n^*. In order to collide, the chain must be at least partially relaxed. While the flower provides some relaxation at low fields, the chain will be almost fully extended in a strong field. As a result, the chain must relax during its translation through the array before it is able to collide, whereupon the first possible collision is located at some row $n>n^*$. Both of the issues identified here (the effect of the curved field lines and the relaxation at high fields) are important topics for further research in this area. The present contribution provides a framework for incorporating more complex collision models, and the results presented here demonstrate the potential of the CTRW approach to model electrophoretic separations in microfluidic post arrays.

V. CONCLUSIONS

In the present contribution, the CTRW model of DNA electrophoresis in microfluidic post arrays [9] was modified to account for incomplete extension of the chain during unhooking from the post. It was shown that the dominant conformation during the unhooking process is a stem flower, leading to a decrease in the collision time and increase in the collision frequency when compared to a taut chain. While the stem-flower conformation does not have a strong effect on the effective electrophoretic mobility, it results in a significant increase in the dispersivity over what would be expected from a taut chain. The increased dispersivity leads to a reduction in the separation resolution at weaker fields, in agreement with experiments.

Although the theory developed here constitutes an important step towards understanding DNA electrophoresis in mi-

crofluidic post arrays, the problem is far from resolved. A particular shortcoming of the present model is the excessively long average unhooking time. Indeed, the underestimate of the separation resolution can be directly attributed to an overly long unhooking time, since long unhooking times increase both the dispersion caused by a single post collision and the total time required to reach the detector. While some potential solutions to the problem were noted in the Discussion, understanding the post collision process remains an area of fundamental research. A better model for the post collision process, when combined with the CTRW approach developed here, will ultimately provide a powerful analytical tool for understanding DNA electrophoresis in microfluidic post arrays.

ACKNOWLEDGMENT

I am grateful to Nabil Laachi at the University of Minnesota for a critical reading of this manuscript.

[1] J. L. Viovy, Rev. Mod. Phys. **72**, 813 (2000).
[2] W. D. Volkmuth and R. H. Austin, Nature (London) **358**, 600 (1992).
[3] W. D. Volkmuth, T. Duke, M. C. Wu, R. H. Austin, and A. Szabo Phys. Rev. Lett. **72**, 2117 (1994).
[4] N. Kaji, Y. Tezuka, Y. Takamura, M. Ueda, T. Nishimoto, H. Nakanishi, Y. Horiike, and Y. Baba, Anal. Chem. **76**, 15 (2004).
[5] G. C. Randall and P. S. Doyle, Phys. Rev. Lett. **93**, 058102 (2004).
[6] G. C. Randall and P. S. Doyle, Macromolecules **38**, 2410 (2005).
[7] P. S. Doyle, J. Bibette, A. Bancaud, and J. L. Viovy, Science **295**, 2237 (2002).
[8] N. Minc, C. Futterer, K. Dorfman, A. Bancaud, C. Gosse, C. Goubault, and J. L. Viovy, Anal. Chem. **76**, 3770 (2004).
[9] N. Minc, J. L. Viovy, and K. D. Dorfman, Phys. Rev. Lett. **94**, 198105 (2005a).
[10] S. Popelka, Z. Kabatek, J. L. Viovy, and B. Gas, J. Chromatogr., A **838**, 45 (1999).
[11] M. Krawczyk, J. Dulak, and K. Kulakowski, Electrophoresis **23**, 182 (2002).
[12] H. Scher and M. Lax, Phys. Rev. B **7**, 4491 (1973).
[13] K. D. Dorfman and J. L. Viovy, Phys. Rev. E **69**, 011901 (2004).
[14] F. Brochard-Wyart, Europhys. Lett. **23**, 105 (1993).
[15] F. Brochard-Wyart, Europhys. Lett. **30**, 387 (1995).
[16] P. André, D. Long, and A. Ajdari, Eur. Phys. J. B **4**, 307 (1998).
[17] The effective charge represents the linear reponse between the applied electric field and the force acting on the Kuhn step. It differs from the intrinsic charge of the Kuhn step due to the action of the electric field on the Debye layer.
[18] D. Long, J. L. Viovy, and A. Ajdari, Phys. Rev. Lett. **76**, 3858 (1996).
[19] E. M. Sevick and D. R. M. Williams, Phys. Rev. E **50**, R3357 (1994).
[20] Y. Masubuchi, H. Oana, T. Akiyama, M. Matsumoto, and M. Doi, J. Phys. Soc. Jpn. **64**, 1412 (1995).
[21] N. Minc, P. Bokov, K. B. Zeldovich, C. Futterer, J. L. Viovy, and K. D. Dorfman, Electrophoresis **26**, 362 (2005b).
[22] G. I. Nixon and G. W. Slater, Phys. Rev. E **50**, 5033 (1994).
[23] In order to arrive at an analytical result, it is necessary to eventually invert the total cycle time for r_0. The full trapping time (22) cannot be inverted analytically. This limitation also motivates the approximation in Eq. (25).
[24] It is tempting to think that the net electrical force animating the stem flower should be reduced as well when compared to a taut chain, since the Kuhn steps contained in the flower are not necessarily aligned with the electric field. However, the size of the flower is independent of the length of the stem (at least until very close to disengagement, where the stem disappears). As a result, the reduced electrical force arising from misalignment of the bases with the field will be the same, on average, for the flowers on each side of the post because they have the same size. Inasmuch as the electrical force is a result of the balance between the force acting on the long arm and that acting on the short arm, then the net force only depends on the difference in the stem length, with the electrical forces due to the flowers canceling out.
[25] Although we have not explicitly included any relaxation effects in the microscale model, they would only serve to increase the frequency of collisions in a stem flower relative to a taut chain.