Multiple Regimes of Collision of an Electrophoretically Translating Polymer Chain against a Thin Post

Sean P. Holleran* and Ronald G. Larson

Chemical and Biochemical Engineering, University of Pennsylvania, 220 South 33rd Street, Philadelphia, PA 19104

Received June 6, 2007; In Final Form March 11, 2008; Revised Manuscript Received March 10, 2008

ABSTRACT: We use a previously developed bead-spring Brownian dynamics model for simulating the topological interactions between polymers and thin obstacles to study electrophoretically translating DNA strands colliding directly with an immovable post over a wide range of chain lengths ($25 \leq N_K \leq 1500$) and field strengths or velocities ($10^{-4} \leq Pe_{Kuhn} \leq 10$ °). Here N_K is the number of Kuhn steps in the chain and Pe_{Kuhn} is the Peclet number based on the Kuhn length. This Peclet number is the ratio of the field-induced polymer motion to the Brownian motion. We find that the mean distance $\langle \Delta \bar{x} \rangle$ by which the chain migration is delayed by the entanglement interaction increases with higher fields, encompassing four distinct regimes. The two fastest regimes exhibit variations of the classic rope-and-pulley dynamics, in which the chain is draped around the entanglement and the longer of the two dangling arms pulls the shorter arm around the obstacle. In one of these regimes, occurring at the highest field strength, the dimensionless delay distance reaches its theoretical upper limit at $\langle \Delta \bar{x} \rangle / N_K = 0.5$, and in the other, at moderately high field strength, the incomplete extension of each arm is modeled as a terminal relaxed "ball" attached to a fully extended "chain." In the two slower regimes, the polymer retains a coil-like shape as it diffuses laterally and eventually clears the post without deforming. We develop theoretical expressions that describe both the average delay and the distribution of delays for all regimes, except the slowest one, which is distinguished by a peculiar fractional power law relationship $\langle \Delta \bar{x} \rangle / N_K \propto Pe_{Kuhn}^{0.27}$.

INTRODUCTION

Long polymer chains, such as DNA, can be separated by length by driving them electrophoretically through a mesh of obstacles, such as a polymer gel matrix or a microfabricated array of posts.[1–4] In all such processes, the basic interaction that leads to a polymer-length-dependent delay is an entanglement of the polymer with the post. Hence, there is considerable interest in understanding how this delay is produced and in modeling its dependence on chain length and field strength. The simplest model that allows the basic entanglement interaction to be studied with the fewest distracting complications is that of a single long polymer molecule encountering a single thin post. This "simple" case has been studied experimentally,[5,6] computationally,[7–11] and theoretically[12,13] and has been shown to display a remarkably rich range of behavior. Many studies have focused on the "high field" regime, in which the polymer, when encountering the post, often forms a "rope-and-pulley" structure with two extended arms, one on either side of the obstacle, and the interaction is governed by the rate in which the shorter of the two arms is pulled back and around the post by the force exerted on the longer arm.[14,15]

Randall and Doyle[6] recently carried out an extensive and careful experimental examination of the interactions of long DNA molecules driven by an electric field past a single thin post and cataloged four different kinds of interactions, some of which had previously been seen in simulations.[7–11,14,15] In "U" interactions, the DNA forms two nearly equal-length arms, one of which eventually migrates downfield, dragging the other arm (usually the slightly shorter one) against the field and around the post. The "J" interactions are same as the "U" interactions, except one arm is quite a bit shorter than the other and is dragged more quickly against the field and around the post. In "X" interactions, the end of the long arm continues to unravel even as the short arm is pulled up the field. In "W" interactions, which are fairly rare, the molecule is draped more than once over the pole. Extending a model for U/J interactions developed by Sevick and Williams,[16] Randall and Doyle were able to develop models for U, J, and X modes that described the time for both "hooking" and "unhooking," as well as the delay in migration produced by the interaction with the post, and they tested their models against data for DNA molecules of two different lengths at a couple of (rather high) field strengths. Randall and Doyle found that, while in detail the results were rather complex, the delay in migration introduced by the post could in most cases be described rather simply using a formula we present shortly. However, the range of field strengths and chain lengths examined in this study was rather narrow and did not consider "low field" regimes, where the molecule is hardly deformed by its interaction with the pole, nor the short chains, presumably because they would be difficult to image.

A different range of conditions has been reached by Brownian dynamics or Monte Carlo simulations of pearl necklace or bead-rod models. These models are limited to rather short chain lengths and fast electrophoretic velocities because they are coarse-grained at the level of a Kuhn-step length (which is the length of a rod in a bead-rod simulation and can be thought of as the effective random-walk step size of the polymer chain). Many 100s or 1000s of rods would therefore be required to simulate a long polymer using a bead-rod simulation, which is computationally overwhelming.

To overcome both the limitations on chain length and electrophoretic velocity, we have developed a bead-spring model in which each spring can represent many Kuhn steps and yet preserve topological "noncrossability" constraints, and so we are able to evaluate longer chains in slower fields than has been previously possible. We showed elsewhere[17] that our new Brownian dynamic method with spring noncrossability can span an unlimited range of chain lengths, exemplified by simulating chains with lengths in the range $25 \leq N_K \leq 1500$ (N_K is the number of Kuhn steps in the chain), and a wide range of field strengths, covering the range $10^{-4} \leq Pe_{Kuhn} \leq 10^0$ in example simulations (Pe_{Kuhn} is the Peclet number based on the Kuhn-step length, which we will define more precisely shortly). Thus

* Corresponding author. E-mail: seanholl@seas.upenn.edu.

10.1021/ma0712596 CCC: $40.75
Published on Web 05/21/2008

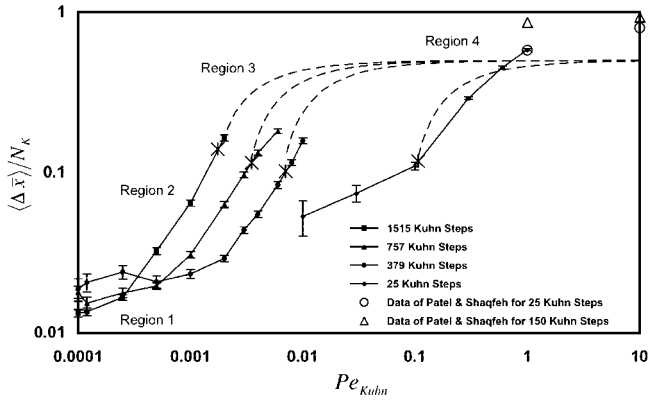

Figure 1. Master plot of normalized delay distances $\langle \Delta \bar{x} \rangle / N_K$ as functions of Pe_{Kuhn} for chains of four different lengths. The results for the three longest length chains ($N_K = 379$, 757, and 1515) were obtained using coarse-grained bead-spring simulations in which each spring represents many Kuhn steps, while for the shortest chain ($N_K = 25$), 25 stiff Fraenkel springs were used, which resembled rods. The four regions are discussed in the text. The dashed lines represent the theoretical predictions for R3 for each of the four chain lengths. The Patel and Shaqfeh data are for bead-rod simulations at high Pe_{Kuhn}. The large asterisks show the values of Pe_{Kuhn}, where σ^2/μ^2 transitions from R2 into R3, as shown in Figure 9.

we are able to explore a much wider range of conditions than has been accessed before either experimentally or computationally.

Figure 1 contains our previous simulation results (shown as symbols) for polymers initially positioned just upstream of the thin, infinitely long, post and with the center of mass lined up with the post to ensure a direct collision with it. Also included are two data points for each of two chain lengths from a previously published bead-rod simulation from another group.[10] In this paper, we will explore these results in detail and will propose that the interactions between the chain and the thin post can be grouped into four distinct regions (also shown in Figure 1), distinguished by the field strength and chain length, that will be explained in what follows. We will develop predictive models for of all of the regions, except the slowest region (R1), which we explain only qualitatively. We have explicit simulation data in the two regions of lowest field strength (R1 and R2); we borrow from the literature to describe the region in the highest field strength (R4); and we will develop an approximation to predict the data in the transition region (R3) between R2 and R4.

We measure the effect of the interaction between the chain and the post using the steady-state ensemble-averaged collision delay distance ($\langle \Delta \bar{x} \rangle$). This measure has been made nondimensional by division by the Kuhn-step length (b_K). $\Delta \bar{x}$ is the dimensionless difference between the x location (where x is the field direction) of the center of mass of a chain in a simulation containing the obstacle and the one translating at the average speed the chain would move without the obstacle. The location of the center of mass of a chain without the obstacle is easily calculated from the known uniform velocity and the elapsed time. A positive $\Delta \bar{x}$ represents a distance penalty that the chain suffers as a result of its entanglement interaction. A large collision distance implies that the chain is greatly delayed by the post. We normalize this measure by N_K, the number of Kuhn steps in the chain. (We note that to relate our results to those of fine-scale bead-rod models, we use N_K to characterize chain length even though our simulations are carried out with a coarse-grained bead-spring model in which each spring represents $N_{K,S} = N_K/N_S$ Kuhn steps, where N_S is the number of springs used to represent the entire molecule.)

Randall and Doyle[6] have determined experimentally and theoretically the value of around 0.5 for the upper limit of $\langle \Delta \bar{x} \rangle / N_K$ in R4, assuming the "unhooking" time dominates the delay distance. We will explain our approximation method for R3 in the discussion section of this work. We will also describe the physics that govern the dynamics in R2 and R3. We suggest a possible mechanism dominating R1 but without making quantitative predictions for this regime.

METHODS

We detailed in our previous paper[17] a method that allows us to impart spring-post repulsions into Brownian dynamics bead-spring simulations of arbitrarily long chains. This method reaches convergence with respect to changes in the time step (Δt), the number of springs (N_S), and the parameters used in the repulsive spring potential (namely, the range and strength of the potential[17]), so that the predictions are free of dependence on these nonphysical simulation parameters and should be virtually identical to those that would be obtained from a fine-grained bead-rod model of a polymer impacting with an asymptotically thin post that extends infinitely in a direction perpendicular to the direction of the electrophoretic field.

Our method is restricted in that the minimum number of springs necessary to achieve convergence is a function of the field strength (Pe_{Kuhn}) because the configurations become more distorted at higher fields and they require more and shorter springs to resolve the tight bends produced in the chain at high fields. There is thus a limit to how large a field strength we can simulate for each chain length, a limit that is set by the computational resources available. Moreover, our method provides no advantage over simpler bead-rod simulations if the number of springs we require to achieve convergence approaches the number of Kuhn steps in the chain. This limit is approached at high fields. Hence, we are able to simulate long chains at low Pe_{Kuhn} using the coarse-grained bead-spring model and short chains at high Pe_{Kuhn} using a fine-grained bead-Fraenkel-spring model in which each spring is effectively a rod but not long chains at high Pe_{Kuhn}. This is why some parts of Figure 1 are missing and will be filled in with predictions that we will describe below.

Note that the Peclet number is a ratio between the rates of electrophoretic transport and diffusive transport. A high Peclet number implies that the drag dominates, while a low Peclet number implies that the Brownian motion is dominant. We define a microscopic Peclet number that balances convection and diffusion at the level of a Kuhn-step length as:

$$Pe_{Kuhn} = \frac{V \zeta_{Kuhn} b_K}{k_B T} \tag{1}$$

where V is the velocity of the coil in the absence of the post and is directly proportional to the field strength. For our bead-spring model, we interpret ζ_{Kuhn} in eq 1 as the drag per Kuhn step, $\zeta_{Kuhn} \equiv \zeta_{tot}/N_K$, which we obtain from the total drag coefficient summed over all beads, $\zeta_{tot} \equiv \zeta_b(N_s + 1)$, where ζ_b is the bead drag coefficient divided by the number of Kuhn steps, N_K, in the entire chain. The number of Kuhn steps in the entire chain is given by the ratio of the extended chain length to the Kuhn-step length. We take ζ_{Kuhn} to be the same for chains of all lengths. Longer chains therefore have a higher total drag and also have correspondingly more Kuhn steps.

We model polymer chains of four different lengths: 1515 Kuhn steps (200 μm), 757 Kuhn steps (100 μm), 379 Kuhn steps (50 μm), and 25 Kuhn steps (3.3 μm). The lengths in the parentheses are the approximate lengths of double-stranded DNA molecules having the corresponding numbers of Kuhn steps. We take the Kuhn-step length to be $b_K = 0.132\,\mu$m, corresponding to that for optically stained double-stranded DNA molecules.[18] The three longer chains, which have too many Kuhn steps to be modeled by a bead-rod model, are modeled

with 20-bead simulations in which each of the 19 springs represents many Kuhn steps. We use the worm-like chain spring law for these long chains for which the number of Kuhn steps per spring is greater than 15. The shorter chain cannot be modeled in this way but is short enough to be modeled with a 26-bead simulation with each of the 25 springs corresponding to a single Kuhn step. To accomplish this, we use the "FENE-Fraenkel" spring law because it results in a stiff rod-like spring, and thus these chains mimic bead-rod chains, which are appropriate and are computationally tractable for shorter chains.[17,19]

Each simulation begins with a random coil (each run has an independent random coil configuration) that is initially placed a small distance upstream of the post (with the bead at the most downstream of the coil at a distance of 0.1 R_g away from the post, and the results are insensitive to exactly how close this is). The coil is also placed such that its center of mass is in line with the center of the post. A drag force equal to $\zeta_b V$ is imposed on each bead of the chain to drive it downfield. The simulation continues until the last bead to pass the post (the bead farthest upstream) reaches a distance of more than seven-and-a-half Kuhn lengths downstream of the post, which is far enough to be sure that it is no longer entangled with the post. The time t required to reach this position is then converted into a delay distance Δx by subtracting the distance traveled (the difference between the original location of the center of mass and the location of the center of mass at the end of the run) from the distance Vt a chain would have traveled in the same time t in the absence of the post and without diffusion. The value Δx is made dimensionless, $\Delta \bar{x} \equiv \Delta x/b_K$, as described above. Each data point in Figure 1, representing a single chain length and value of Pe_{Kuhn}, is the mean normalized delay distance $\langle \Delta \bar{x} \rangle / N_K$ averaged over 800 trials (although some data points are averaged over 3200 trials, which is discussed later). Further details of the simulation can be found in Holleran and Larson.[17]

RESULTS AND DISCUSSION

Region 4 (R4): $Pe_{Kuhn}N_K \geq 20$. Convective transport is the dominant driving force in R4. It is in this region that the rope-and-pulley formation is most prevalent and the arms of the chains become nearly fully extended. This has been discussed elsewhere by Randall and Doyle[6] and Patel and Shaqfeh,[10] and we will not devote much attention to it here. If the delay is dominated by the "unhooking" process and the hook is equally likely to occur anywhere along the chain, Randall and Doyle find that at high enough Peclet number chains of all lengths will reach an asymptotic delay of around $\langle \Delta \bar{x} \rangle / N_K = 0.5$. (Consideration by Randall and Doyle of the time to form the hook and the drift of the center of mass downstream while the chain is unhooking leads to values of $\langle \Delta \bar{x} \rangle / N_K$ ranging from 0.4 to 0.7). If we choose the lower boundary of R4 as the point at which $\langle \Delta \bar{x} \rangle / N_K$ reaches around 90% of the "asymptotic" value of 0.5, we find that that this boundary corresponds roughly to $Pe_{Kuhn}N_K \approx 20$ (see Figure 1). Hence, longer chains enter R4 at lower values of Pe_{Kuhn}.

Most chains in R4 form the "rope-and-pulley" interaction (i.e., the U, J, or X configuration of Randall and Doyle) because the convective force is too great to allow diffusion to help move the chain around the post. The random Brownian motions have a negligible effect at this Pe_{Kuhn}. The coil usually forms two extended arms (one on either side of the post, although in the "X" configuration the end of one of the arms continues to unravel even while the "rope-like" portion of the chain is moving over the pulley) and escape is controlled by how long it takes for the shorter arm to be pulled back (by the spring force) around the post. One key variable is the location along the length of the chain where the intersection with the post will occur (the location of the pulley). For coils whose center of mass is lined up with the post, there is expected to be a somewhat greater likelihood that the collision will be nearer the center of the chain than to the ends of the chain, especially at high fields, where the chain has little chance to diffuse before the collision occurs. However, for chains not lined up with the post, or for low fields, where the chain can drift laterally before the collision, a uniform distribution is most likely. A uniform distribution was assumed by Randall and Doyle in analyzing their DNA experiments. For simplicity, we will follow the Randall−Doyle approximation here and take the length of the shorter arm to be evenly distributed between zero and one-half of the total length of the chain but will note later that at high field this approximation is not quite adequate to explain our simulation results. The longest entanglement interaction thus occurs when the polymer intersects the post at a location near the midpoint of the chain.

We attempted to reach R4 with our simulations, but we run up against the limitation that an increase in Pe_{Kuhn} requires an increase in the number of springs needed to obtain converged results that are insensitive to the number of springs used. We came closest to reaching R4 for our work with the shortest chain at high field strengths. Recall that the each data point on Figure 1 is an average over 800 trials. Even though we do not reach a high enough Pe_{Kuhn} to find a regime where $\langle \Delta \bar{x} \rangle / N_K$ is independent of Pe_{Kuhn}, we are able to see some individual trials that clearly show the rope-and-pulley dynamics. Figure 2a contains frames from a movie of the 25 Kuhn-step chain at $Pe_{Kuhn} = 1$. It is clear in this series of images that the chain has formed two separate and distinct arms.

We must note that we actually find a case where $\langle \Delta \bar{x} \rangle / N_K > 0.5$ and this appears to contradict the theory of Randall and Doyle.[6] We can explain this by noting that there is a bias in our setup, and as noted above, we are unfairly favoring longer lasting interactions at high Pe_{Kuhn} because we begin with the chain lined up with the post and therefore with a somewhat greater likelihood of forming the collision near the center of the chain than near the ends. This is a negligible effect at low Pe_{Kuhn} because there is ample opportunity for the diffusion to eliminate the bias, but at high Pe_{Kuhn}, where convection dominates diffusion, this effect can produce an increase in long-lasting interactions.

We evaluated the effect of the bias that led to too many long-time interactions that resulted in $\langle \Delta \bar{x} \rangle / N_K > 0.5$ by rerunning the simulations for the two highest Pe_{Kuhn} with the initial location of the center of the coil varying laterally uniformly over the range $\pm R_g$ away from a head-on impact with the post. This allowed some chains that were not lined up with the post to simply brush past the post without having a rope-and-pulley interaction. With this distribution of initial positions, we found that $\langle \Delta \bar{x} \rangle / N_K$ dropped by a factor of 2 for both $Pe_{Kuhn} = 0.60$ and $Pe_{Kuhn} = 1.00$ and fell well below the $\langle \Delta \bar{x} \rangle = 0.5$ limit from Randall and Doyle. We note that the points from Shaqfeh and Patel (on Figure 1) are also above the $\langle \Delta \bar{x} \rangle / N_K = 0.5$ limit by roughly a factor of 2, presumably because their chains were also lined up exactly with the post, with no offset, as in ours. We believe that if Shaqfeh and Patel had removed their bias in the location of chain-post their results would also have been beneath the limit.

Randall and Doyle relate the time required for the chain to unhook from the post to the location along the chain where the interaction begins. Thus, in R4, they showed that[6]

$$t_{unhook} = -\frac{L}{2\mu E} \ln\left(1 - \frac{2x_1}{L}\right) \quad (2)$$

where L is the length of the chain, x_1 is the length of the shorter arm, and μE (the product of mobility and field strength) is the polymer velocity V created by the electric field. While this

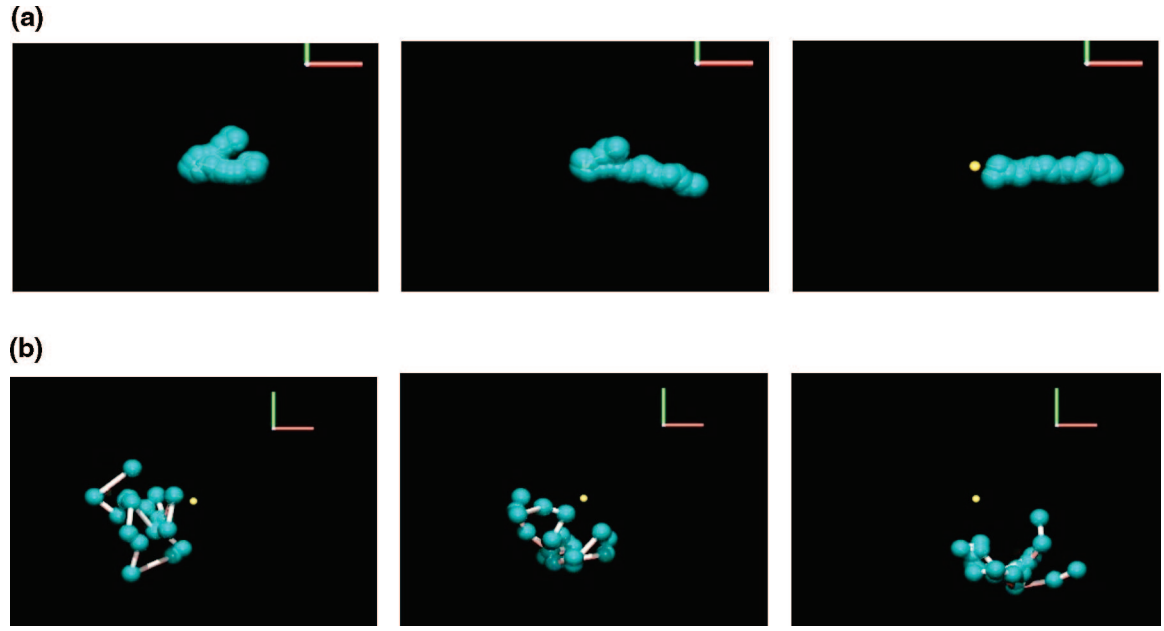

Figure 2. Images from simulations of polymers escaping from entanglements with a post, where the post is shown by a dot. (a) A 25-Kuhn-step chain under strong convection ($Pe_{\mathrm{Kuhn}} = 1$). The rope-and-pulley formation (common for R3 and R4) is evident. The "camera" (or view) is in close in because this is a short chain. (b) A 757-Kuhn-step chain under low field ($Pe_{\mathrm{Kuhn}} = 0.001$). The chain remains in its coiled conformation (common for R1 and R2). The "camera" is farther back because this is a long chain.

formula is strictly only valid for the unhooking time of the "U" and "J" configurations, Randall and Doyle showed that the above formula also gives a good approximation for the delay (or "hold-up") time for U, J, and X configurations, encompassing most of the configurations they observed. We can recast this equation to measure dimensionless delay distance so that it corresponds with our method of analysis. We first define $y \equiv x_1/L$, and we also note $\langle \Delta \bar{x} \rangle / N_{\mathrm{K}} = t_{\mathrm{unhook}}/(L/\mu E)$. We rewrite eq 2 for R4.

$$\frac{\Delta \bar{x}}{N_{\mathrm{K}}} = -\frac{1}{2} \ln (1 - 2y) \tag{3}$$

This has the appropriate behavior. The distance penalty is zero when $y = 0$, and this is the case when the interaction occurs at the end point of the chain. The distance penalty is infinitely large when $y = 0.50$, and this is the case when the interaction occurs at the exact center of the chain.

In R4, according to our definition, the convection is strong enough to "nearly completely" unravel both arms of the chain. In reality, the tips of each arm will almost never be unraveled, because typically the tension in the arm increases from zero at the tip of the arm to a maximum at the "pulley," since the effect of the drag along the arm adds up, as in a tug-of-war, with maximum tension at the "pulley." We can say, however, that if a chain hooks at its midpoint, then each arm can become "nearly unraveled" if the tension F along each arm remains high enough to stretch at least 90% of the arm by 50% or more. This will mean that the tension generated by the drag on a chain end encompassing 5% of the whole chain (i.e., 10% of the arm) is sufficient to overcome Brownian motion on the Kuhn segment connecting that chain end to the rest of the chain. Defining ξ as a coordinate that runs from 0 at the tip to 0.5 at the pulley, the transitional condition marking the boundary of R4 is therefore defined by $F(\xi = 0.05)\, b_{\mathrm{K}}/k_{\mathrm{B}}T = 1$. The number of Kuhn steps in the end segment is $0.05N_{\mathrm{K}}$, which experiences a total drag of $0.05\, V\, N_{\mathrm{K}}\, \zeta_{\mathrm{Kuhn}}$, which is then equal to the tension at position $\xi = 0.05$. Hence, the lower boundary of R4 occurs at $0.05\, V\, \zeta_{\mathrm{Kuhn}}\, b_K N_{\mathrm{K}}/k_B T = 1$, or $Pe_{\mathrm{Kuhn}} N_{\mathrm{K}} = 20$.

Region 3 (R3): $2 \leq Pe_{\mathrm{Kuhn}} N_{\mathrm{K}} < 20$. R3 is really a broad cross-over between R4, where the convection is strong enough to nearly fully extend the molecule, and R2, where the molecule is barely deformed from a coil. In R3, there will be a continuous variation of stretch from nearly fully extended near the "pulley" to completely unextended at the tip of the arm. Moreover, the relatively unstretched portion of the arm will experience large fluctuations in extension due to Brownian motion, acting like a "Brownian yo-yo," which will influence the time required to pull the short arm around the pulley.

Thus, accurate modeling of the polymer dynamics in R3 will be complex. To develop simplified analytical expressions for the delay distance will require drastic simplifications. For the case of convection that is not quite strong enough to fully extend the chains, Randall and Doyle assumed that the chain is uniformly stretched and neglected fluctuations in stretch. This allows them to continue to use eq 2, even when the drag on the chain is not strong enough to fully extend it, simply by replacing the chain contour length L by the average chain stretch length. While this approach proved successful in modeling unhooking times for the DNA chains examined by Randall and Doyle, it does not describe our simulation results, as we show at the end of this section.

Hence, we will simplify this case by taking essentially the opposite tack. Rather than assuming that the stretch is uniform, we will break the arm into two regions: the inner portion near the pulley which we take to be fully extended and the outer portion is taken to be "unextended". By unextended, we mean that the stretch is sufficiently weak that fluctuations allow this portion of the arm to instantly wrap around the pulley as soon as it reaches the pulley. This assumption allows us to apply eq 2 to R3, just as Randall and Doyle did, but in our model we shorten the arm not by assuming uniform incomplete stretch but instead by assuming that a portion of the chain is completely stretched and neglect entirely the stretch of the rest of the chain. We note that versions of this idea of a chain that is fully stretched over part of its contour but coiled at the end can also be found in the "X" configuration of Randall and Doyle[6] and in the "stem-flower" configuration of Dorfman.[12,13]

In our simple model, we can derive the complete dependence of the distribution of delay distance on velocity and chain length

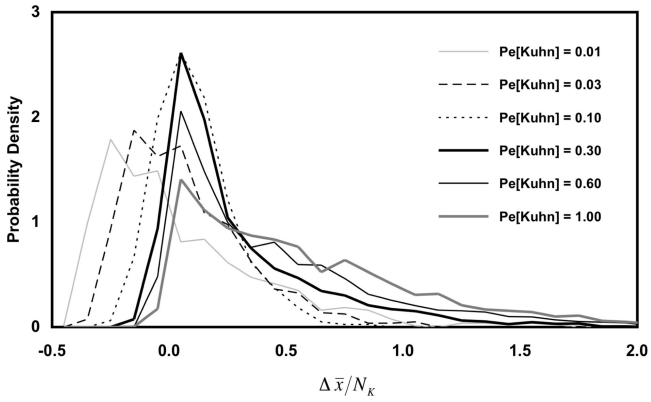

Figure 3. Probability distribution of $\Delta \bar{x}/N_K$ for various values of Pe_{Kuhn} for simulations of 25 Kuhn-step chains. Notice that at $Pe_{Kuhn} = 0.10$ there is a significant change in the shape of the curves.

without any additional assumptions about average stretch. After introducing our model and comparing its predictions to our simulations, we will discuss its relationship to the model of Randall and Doyle applied to R3. As mentioned, in our model, the chain is fully stretched except at the ends of the arms, where some of the monomers exist in the form of a coil or clump. The size of the clump (i.e., the number of monomers in the clump) is a function of Pe_{Kuhn}; as the field strength is lowered, more monomers and a greater fraction of the chain length are found in the end coils. This "ball and chain" concept reduces the $\langle \Delta \bar{x} \rangle / N_K$ values in R3 for two reasons. First, there is a shorter length of chain that needs to be pulled back around the pulley, since we assume that the "ball" instantaneously escapes the pulley once it reaches it. In addition, in R3, there are some cases when the interaction will have a zero delay because the intersection point is inside of the ball, and this chain will get past the post without any delay.

We define b as the fraction of the chain in each ball. The dimensionless length of the short arm (y) can vary between zero and one-half (as was discussed in R4); thus, b is also allowed to vary between zero and one-half. We have $b = 0$ when the short arm is fully extended (case R4) and $b = y$ when the entire amount of the shorter arm is at the end coil (and this case will result in a zero time interaction).

Now consider our data for the 25 Kuhn-step chain; this is the only length chain that we were able to simulate at a high enough Pe_{Kuhn} to obtain data in the R3 range. For a chain with only 25 Kuhn steps there are rather few monomers available to distribute between the extended and coiled portions of the chain. The ball and chain idea should work much better for the longer chains because for them it is easier to have a sizable number of monomers in both the ball and the chain portions. However, we have no simulation data in R3 for long chains and we must content ourselves with testing the ball and chain model using the results for the 25 Kuhn-step chain.

We show in Figure 3 histograms for all of the $\Delta \bar{x}/N_K$ values for the 25-Kuhn-step chain. There are six curves here, one for each of the six Pe_{Kuhn}. The peak of each curve shows the most common value for $\Delta \bar{x}/N_K$ at that Pe_{Kuhn}. Notice that, starting from the lowest value of Pe_{Kuhn}, the peak initially moves to the right as Pe_{Kuhn} is increased. However, for $Pe_{Kuhn} \geq 0.10$, the peak remains at the same $\Delta \bar{x}/N_K$ but begins to decrease in height. The contribution of the right-side tail becomes more significant at high Pe_{Kuhn}. The peak location remains at $\Delta \bar{x}/N_K = 0.2$ for $Pe_{Kuhn} \geq 0.10$, but with increasing Pe_{Kuhn}, there are more long-lasting interactions. These long-lasting interactions lead to an increasing $\langle \Delta \bar{x} \rangle / N_K$ with increasing Pe_{Kuhn} up until a saturation point, which is expected when Pe_{Kuhn} reaches unity and R4 is entered. We note here that the curves for the three highest Pe_{Kuhn} are based on 3200 trials (while the curves for the three lowest Pe_{Kuhn} are based on 800 trials); this was done to improve the statistics for our future analysis of the high Pe_{Kuhn} results.

We now simply add a term to eq 3 to account for the fact that only the portion $(y - b)$ of the short arm needs to be pulled over the pulley before the chain is released. For the purpose of this simplified analysis, the ball is assumed to be of essentially zero size and to roll instantaneously over the pulley. However, we still assume that the ball exerts the same drag as it would if it were fully extended. Thus, we are here neglecting conformation-dependent hydrodynamic interactions and essentially assuming a free-draining chain, as is also assumed in our simulations. The delay induced by the last fraction b of the chain must therefore be subtracted from eq 3. This delay is $-1/2 \ln(1-2b)$, which when subtracted from the right side of eq 3 gives

$$\frac{\Delta \bar{x}}{N_K} = -\frac{1}{2} \ln (1 - 2y) + \frac{1}{2} \ln (1 - 2b) \tag{4}$$

This extra term allows us to adjust $\Delta \bar{x}/N_K$ for various ball sizes. When $b = 0$, eq 4 reduces to eq 3, which is valid in R4.

We use this expression for the distribution of delay distances to find the probability distribution for $\Delta \bar{x}/N_K$, assuming that y is evenly distributed between zero and one-half. A delta function at zero is introduced to represent the cases in which there is a zero-delay interaction because the intersection point is located within the ball. The factor of 2 in the second term below is required so that the integral of the probability will sum to unity even in the case when $b = 0$:

$$P\left(\frac{\Delta \bar{x}}{N_K}\right) = 2b\, \delta(0) + 2\,(1 - 2b) \exp\left(-2\,\frac{\Delta \bar{x}}{N_K}\right) \tag{5}$$

We note that when $Pe_{Kuhn} = 1$ the ball should be fully unraveled because the drag induced by the field matches the Brownian force on a single Kuhn step. The end monomers begin to ball up for smaller Pe_{Kuhn} because there is then not enough drag force on the end Kuhn segment to overwhelm the Brownian force. However, for Pe_{Kuhn} slightly below unity, the drag force on the final few Kuhn segments will accumulate, producing enough tension to fully extend all but the last few Kuhn steps which form the "ball" at the end of the chain. Since, in the free-draining limit, the drag on the ball is proportional to the number of Kuhn segments in the ball, the number of Kuhn steps in the ball should be inversely proportional to Pe_{Kuhn}. Since b is the fraction of the Kuhn steps in the ball, we must have

$$b \approx \frac{1}{(Pe_{Kuhn} N_K)} \tag{6}$$

We are now able to use eqs 5 and 6 to predict the probability distribution of $\Delta \bar{x}/N_K$ as a function of Pe_{Kuhn} in R3. Panels a−c in Figure 4 show our predictions and our simulation data for the three values of Pe_{Kuhn} that are in R3 for the 25 Kuhn-step chain. We evidently do a good job of predicting the tail portion of the distribution. There is some discrepancy at small $\Delta \bar{x}/N_K$, where our simple ball-and-chain model assumes that there is a delta function at $\Delta \bar{x}/N_K = 0$ corresponding to cases where the ball makes contact with the post and immediately escapes with zero delay. Obviously, in the simulations, these cases require a small but finite delay, and thus produce a peak at small $\Delta \bar{x}/N_K$ that is not captured in our oversimplified model.

We have no R3 data from the three longer chains to compare with theory. However, we can predict how these chains would behave in R3 by averaging over the probability distribution in eq 5:

$$\frac{\langle \Delta \bar{x} \rangle}{N_K} = \int_0^\infty P\left(\frac{\Delta \bar{x}}{N_K}\right)\left(\frac{\Delta \bar{x}}{N_K}\right) \mathrm{d}\left(\frac{\Delta \bar{x}}{N_K}\right) \quad (7)$$

which, after a little mathematics, yields

$$\frac{\langle \Delta \bar{x} \rangle}{N_K} = \frac{1}{2}(1-2b) = \frac{1}{2}\left(1 - \frac{2}{Pe_{\mathrm{Kuhn}} N_K}\right) \quad (8)$$

Since panels a−c in Figures 4 show that our probability prediction is in good agreement with the limited delay distribution data we were able to acquire for R3 for the shortest chain, we use this probability distribution to predict the curves of $\langle \Delta \bar{x} \rangle / N_K$ vs Pe_{Kuhn} in R3 for chains of any length. Such predictions for chains of length 1515, 757, and 379 Kuhn steps are shown as dashed lines in Figure 1. We calculate the curves down to the regions where they come into near overlap with the corresponding data that we have for these long chains. The reasonable match up of the curves predicted for R3 with the high Pe_{Kuhn} end of the R2 data for the long chains indicates that our theory for R3 is at least consistent with R2 simulation data. Note also that the predicted result in R3 for the shortest chain (25 Kuhn steps) is in reasonable agreement with data for this chain length, especially considering that the ball-and-chain model is not expected to work very well when the ball and the chain each contain only a few Kuhn segments, and so there is really no sharp distinction between the ball and the chain. We also notice that eq 8 reaches Doyle−Randall limit of $\langle \Delta \bar{x} \rangle / N_K = 0.5$ as the field strength is increased (in the limit of high Pe_{Kuhn}).

Our model is an alternative to that of Randall and Doyle discussed above, who simply rescale the chain length by its average stretch and apply eq 2. Since, according to eq 2, the average delay is proportional to the chain length, we can test this idea using our simulation data by simply rescaling the x axis of Figure 3 by dividing the delay distance by the average delay distance. To keep the area under the histogram equal to unity, we must also rescale the y axis by multiplication by the average delay distance. If the Randall−Doyle idea applies to our simulation data, then all the curves in Figure 5 should superpose after this rescaling. Obviously they do not, and this outcome shows that this model is not applicable to our simulation results in R3.

However, Randall and Doyle's model does fit well the data they obtained for DNA electrophoresis around a small but finite-sized post, where, as input to the model, they took the measured average extension of the chain. The good agreement with the DNA data seems to conflict with the poor agreement the Randall−Doyle model gives to our simulation results. Possible sources for the differences might include the facts that our simulations neglect hydrodynamic interactions and the post used in the experiments had a diameter comparable to the DNA radius

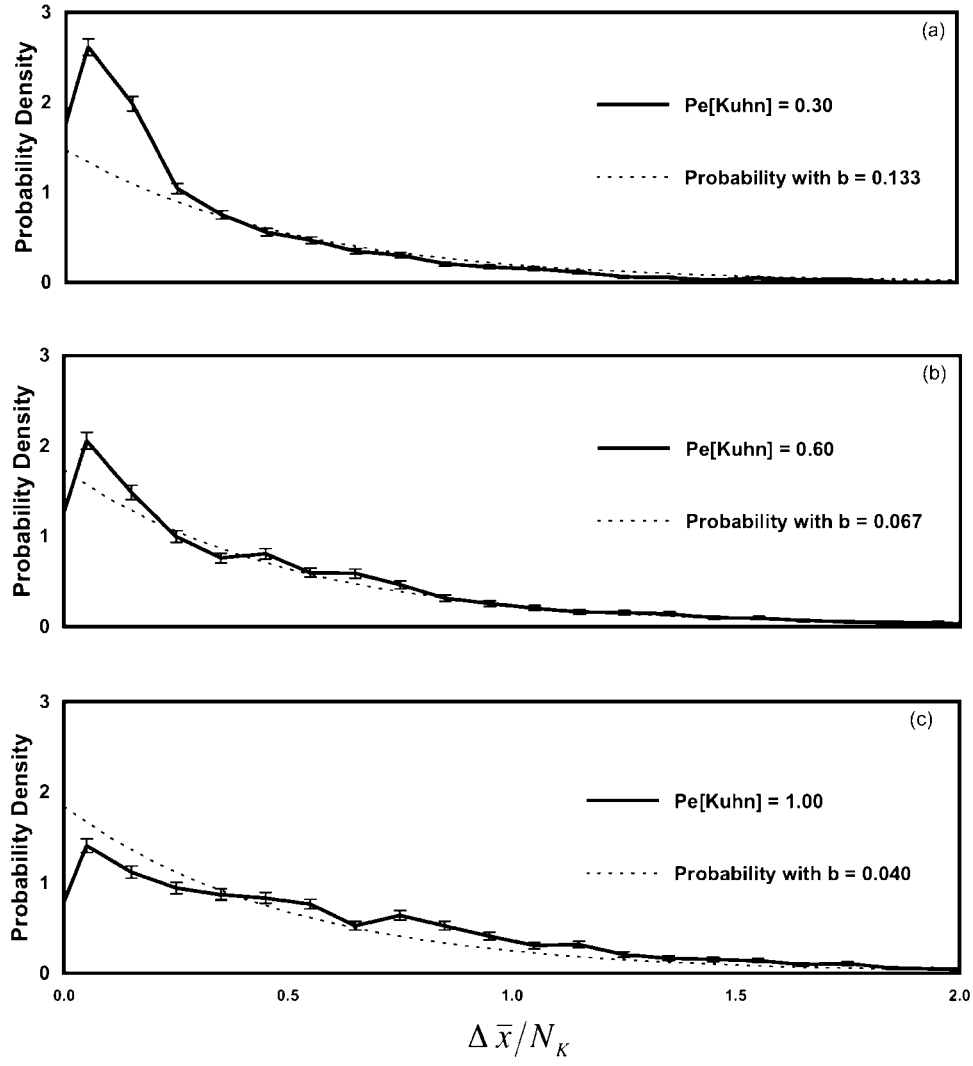

Figure 4. The solid line shows the probability distribution for the 25-Kuhn-step chain obtained from the simulation. The dashed line shows the prediction of the probability distribution with b given by eq 7, $b = 1/(Pe_{\mathrm{Kuhn}} N_K)$. (a) $Pe_{\mathrm{Kuhn}} = 0.30$ and $b = 0.133$. (b) $Pe_{\mathrm{Kuhn}} = 0.60$ and $b = 0.067$. (c) $Pe_{\mathrm{Kuhn}} = 1$ and $b = 0.04$.

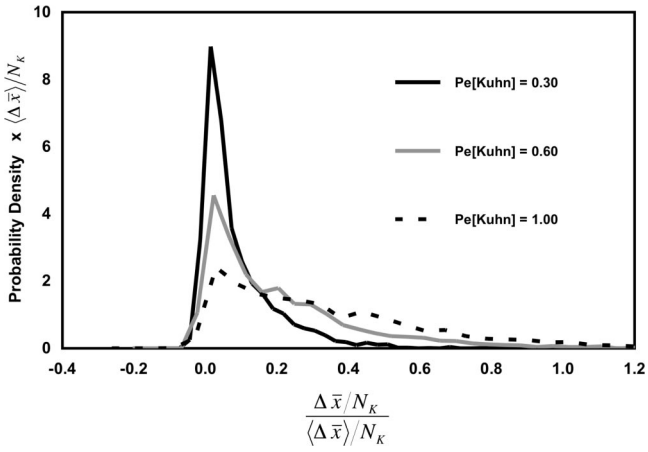

Figure 5. The data of Figure 3 for Pe_{Kuhn} in region 3 ($Pe_{\mathrm{Kuhn}} = 0.30$, 0.60, and 1.00) replotted against delay normalized by the average delay. This is done to test the Randall−Doyle theory.

of gyration, and so there might have been some effect of the finite diameter of the post on their results. Also, their experimental data only encompassed two different chain lengths and a couple of different Peclet numbers, not adequate for a complete test. Our model is perhaps even more inadequately tested, since we are only able to apply it to the case $N_K = 25$, which is too short a chain length to give a strong test of the model. Nevertheless, for this chain length, our histograms of delay distances are in good agreement with the model, except for small delay distances where the predicted delta function is, in the simulations, smeared into a peak of finite height. In truth, both our model and that of Randall and Doyle are rather crude for R3, and deciding which one is best will require more extensive simulations (and experiments) encompassing a wider range of chain lengths and field strengths.

Note in the Randall−Doyle model, as the Peclet number decreases, the average chain stretch shrinks until the chain becomes a random coil at Pe_{Kuhn} of around unity. In our model the ball grows to consume the entire chain, when Pe_{Kuhn} drops to around unity. In either model, this condition marks the end of R3. From simulation data for the three longest chains Figure 1, we observe that the condition for entering the new regime is closer to $Pe_{\mathrm{Kuhn}}N_K \approx 2$ than to unity.

Finally, we note that Dorfman[12,13] developed a "stem-flower" model[20] for a regime equivalent to our R3. Dorfman's model is similar to our "ball-and-chain" model, where in Dorfman the "stem" is our "chain" and the "flower" our "ball," except that in Dorfman's version of our eq 4 (i.e., eq 25 in Dorfman[12,13]) the drag force from the "flower" is taken to be proportional to the length of the flower region rather than proportional to the number of monomers in the flower (or "ball"), as we assume. Fundamentally, this difference is due to our neglect of hydrodynamic interaction (consistent with our simulations), which is included in the stem-flower model through the proportionality of drag to length of the flower rather than to the number of monomers in the chain. In addition, Dorfman leaves out the collisions of the "flower" with the post, since such "null collisions" cause no delay, while we include the effect of this as a delta function added to the distribution of delay times (or delay distances). This difference between the two models is essentially a difference in the definition of what constitutes as "collision" and could be incorporated into the probability of a "collision" occurring in models for electrophoresis in arrays of posts, such as that discussed below.

Region 2 (R2): $10/\sqrt{N_K} \le Pe_{\mathrm{Kuhn}}N_K < 2$. In R2, the convective force is no longer strong enough to extend the arms of the chain. There are no more rope-and-pulley formations. There is no "chain" at all over any part of the polymer but only a single coil attempting to move beyond the post. We see in Figure 2b images of a 757 Kuhn-step chain at $Pe_{\mathrm{Kuhn}} = 0.001$. The strand never forms a rope-and-pulley. Instead, it eventually translates past the post as a coil.

We believe that in R2 when the polymer coil collides with the post it can no longer deform and be dragged around the post. Instead, the collision will merely block the motion of the coil. Thus, the coil acts roughly like a globular object blocked by a thin post. Electrophoresis of rigid objects in arrays of obstacles has been considered in a series of papers by Slater and co-workers.[21] However, the case of a single such object colliding with, and maneuvering around, a thin post does not seem to have been analyzed yet, and, in any event, the diffusion of a "soft object" like a polymer will be somewhat different from that of a completely rigid particle. For the case of a polymer, the chain can move past the post only if it first diffuses a distance on the order of the radius of gyration away from the plane defined by the post and the migration direction so that the chain can clear the post without a significant collision with it. Patel and Shaqfeh[10] studied the relationship between the offset from this plane passing through the post (the so-called "y offset") and the resulting delay. Their simulation data not only show a gradual decrease in delay with increasing offset but also show that the delay penalty is reduced by a factor of 2 or so, when the offset is only about one-half the radius of gyration. This large drop in delay for an offset of only $R_g/2$ is reasonable, since any offset greater than $R_g/2$ will often result in weak interactions with the post and many "grazing incidences" that do not significantly slow down the coil. Also, the coil may deform somewhat in R2 (although not enough to form a "rope and pulley"), and this will enable it to squeeze past the post even if it has not diffused a full radius of gyration out of the ($y = 0$) plane of incidence with the post. Finally, when there is an offset, in the time it takes for the coil to translate to a position that it forms a distinct entanglement contact with the post, it may diffuse farther away and completely clear the post with no delay. Hence, for simplicity, we will here approximate the results of Patel and Shaqfeh by a sharp cutoff at $\frac{1}{2}R_g$ and assume that the chain cannot pass by the post until it diffuses laterally a distance $\frac{1}{2}R_g$, and when it has done so, it is free to translate beyond the post without further delay.

We use this simple picture to develop the following scaling argument and plot our data on a universal curve to show evidence of the correctness of this picture. We make two changes in the scaling of the axis from our original Figure 1 to introduce the diffusion idea. The first change takes account of the assumption that it is the diffusion of the chain as a whole that controls the delay. We therefore use a Peclet number based on the diffusivity of the whole chain, which is given by $D = k_BT/\zeta_{\mathrm{total}}$, and change the x axis of Figure 1 from Pe_{Kuhn} to Pe_{chain}, where

$$Pe_{\mathrm{chain}} = \frac{V\,\zeta_{\mathrm{tot}}\,R_g}{k_BT} = Pe_{\mathrm{Kuhn}}\left[\frac{\zeta_{\mathrm{tot}}\,R_g}{\zeta_{\mathrm{Kuhn}}\,b_K}\right] = Pe_{\mathrm{Kuhn}}\left[\frac{N_K^{3/2}}{\sqrt{6}}\right] \tag{9}$$

This changes the characteristic length used in the definition of the Peclet number from the Kuhn step to the radius of gyration and the drag coefficient from that for a single Kuhn step to one for the entire chain ($\zeta_{\mathrm{tot}} = N_K\zeta_{\mathrm{Kuhn}}$).

The original measure of average delay, $\langle\Delta\bar{x}\rangle/N_K$, is a dimensionless length. Since in R2 we believe that the delay is controlled by the time it takes for the whole chain to diffuse laterally by $\frac{1}{2}R_g$, we consider two measures of time:

$$\text{time} = \frac{\langle \Delta x \rangle}{V} \tag{10}$$

$$\text{time}_{\text{diff}} = \frac{\left(\frac{1}{2} R_{\text{g}}\right)^2}{D} = \frac{{R_{\text{g}}}^2}{4\,D} \tag{11}$$

Equation 10 is the average translation distance penalty divided by the fluid velocity. This is a measure of the time spent during the interaction. Equation 11 is a measure of the time required to diffuse a distance ½R_{g} sideways from the post. A ratio of these two times is our dimensionless time.

$$\begin{aligned} \frac{\text{time}}{\text{time}_{\text{diff}}} &= \frac{\langle \Delta x \rangle}{V} \frac{4\,D}{{R_{\text{g}}}^2} \\ &= \frac{\langle \Delta x \rangle}{V} \frac{4\,k_{\text{B}}T}{\zeta_{\text{tot}}} \frac{6}{N_{\text{K}}\,{b_{\text{K}}}^2} \\ &= \left(\frac{\langle \Delta \bar{x} \rangle}{N_{\text{K}}}\right)\left(\frac{24}{Pe_{\text{Kuhn}}\,N_{\text{K}}}\right) = z \end{aligned} \tag{12}$$

Thus, the original measure of dimensionless distance, $\langle \Delta \bar{x} \rangle / N_{\text{K}}$, can be scaled with the inverse of Pe_{Kuhn} (a dimensionless velocity) and the ratio $24/N_{\text{K}}$ to give an appropriate measure of dimensionless delay time.

We plot in Figure 6 our results for all chain lengths with this dimensionless time as the *y* axis. We see that all of the data in R2 collapse onto the same curve. We note moreover that there is a plateau in the curve when the dimensionless time has a value of unity. The data for the 25 Kuhn chain deviate from the universal behavior at higher Pe_{chain}. This is because the data for this chain length reach the R3/R4 regions at higher Pe_{chain}, where the diffusion is no longer the key factor. We take this universal agreement among all chain lengths as strong evidence that our diffusive scaling is correct for R2.

We further plot in Figure 7 our delay distribution data from the longer 379-Kuhn-step chain but using dimensionless time to determine the bins. We see from Figure 1 that R2 for this chain length begins around $Pe_{\text{Kuhn}} = 0.003$, and we see in Figure 7 that there is a clear distinction in the behavior of the data for Pe_{Kuhn} above and below this value of 0.003. The R2 results all peak at the same value of the delay time, and all reach roughly the same peak probability. The data collapse onto a single curve for data in R2; the data for $Pe_{\text{Kuhn}} \geq 0.003$. An effect exactly like this is also seen for the 757-Kuhn-step and 1515-Kuhn-step chains (although the critical Pe_{Kuhn} varies for different chain lengths).

We show in Figure 8a the results for all three chain lengths and for all Pe_{Kuhn} that are in R2. Note again that the *x*-axis is a measure of the dimensionless time. There are five data sets

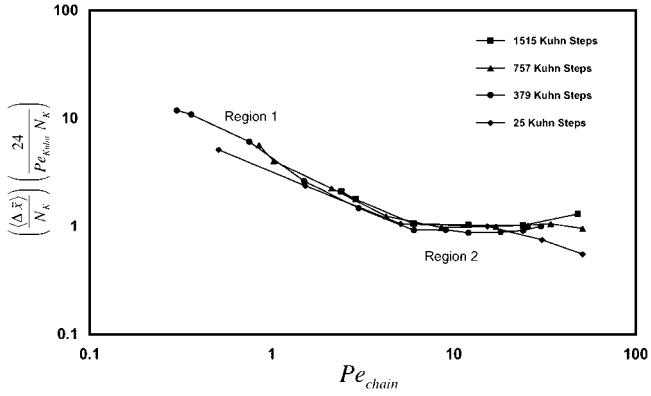

Figure 6. The data from Figure 1 are plotted with rescaled *x* and *y* axes. Note that the results for all four lengths fall on a universal curve. Also note that the R2 plateau occurs at 1.0. This scaling is based on the diffusion-dominated physics in R1 and R2.

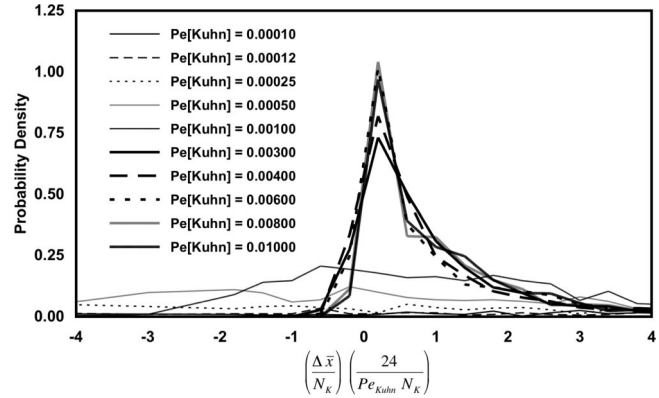

Figure 7. Rescaled probability distribution of $(\Delta \bar{x}/N_{\text{K}})(24/Pe_{\text{Kuhn}}N_{\text{K}})$ for various Pe_{Kuhn} for the 379-Kuhn-step chain simulations. Notice that the rescaled curves for $Pe_{\text{Kuhn}} > 0.001$ are nearly identical.

from R2 for the 379-Kuhn-step chain, four data sets from R2 for the 757-Kuhn-step chain, and three data sets from R2 for the 1515-Kuhn-step chain. We are pleased to see that all these results collapse (roughly) into a single curve for all Pe_{Kuhn} and N_{K} within R2. This is additional strong evidence that our diffusion scaling is sound. The individual peaks occur at the same dimensionless time, but their magnitude varies somewhat, in a manner that does not seem to correlate with either Pe_{Kuhn} or N_{K}.

We note that our picture for the dynamics in R2, in which we assume that the delay is entirely due to the time required for the coil to diffuse laterally a distance of $R_{\text{g}}/2$, is equivalent to a classical "first-passage-time" problem. In this problem, one seeks to determine the time required for a random diffuser to reach a fixed distance from its starting point. There is a distribution of times, and this distribution is a function of the diffusivity of the object and the distance it is required to travel. We expect to have a "double-sided first-passage-time" problem because the coil can pass the post on either side. That is, the coil can get around the post equally easily by drifting to the right or to the left by an amount comparable to the radius of gyration of the coil. Nagar and Pradhan present this distribution of double-sided first-passage times for a one dimensionally diffusing object[22]

$$P(t, L) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{D(2n-1)\pi}{L^2} \exp\left[-\frac{(2n-1)^2\pi^2 D}{4L^2} t\right] \tag{13}$$

where *L* is the distance that must be traveled, *D* is the diffusivity, and *t* is the required time. We recast eq 13 in dimensionless variables to align this formula with our nomenclature for R2. We choose $L = \alpha(R_{\text{g}}/2)$ as the distance the polymer coil must travel (based on our previous comment about the results from Patel and Shaqfeh), where α is a parameter that will allow us to adjust the distance to best fit our results. We therefore make the following substitutions.

$$L = \alpha \frac{1}{2} R_{\text{g}} = \alpha \frac{1}{2} \left(\frac{N_{\text{K}}\,{b_{\text{K}}}^2}{6}\right)^{1/2} \tag{14}$$

$$D = \left(\frac{k_{\text{B}}T}{N_{\text{K}}\,\zeta_{\text{K}}}\right) \tag{15}$$

$$t = \frac{\Delta x}{V} \tag{16}$$

This transforms eq 13 into

$$P(t,L)\,dt = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{\left[\frac{k_BT}{N_K\zeta_K}\right](2n-1)\pi}{\left(\frac{1}{4}\right)\left[\frac{N_K b_K^2}{6}\right]\alpha^2}\times \exp\left[\frac{(2n-1)^2\pi^2\left[\frac{k_BT}{N_K\zeta_K}\right]\left[\frac{\Delta x}{V}\right]}{-4\left(\frac{1}{4}\right)\left[\frac{N_K b_K^2}{6}\right]\alpha^2}\right]\frac{d(\Delta x)}{V} \quad (17)$$

Using the original definition of Pe_{Kuhn} (equation 1) and the expression for the dimensionless time (equation 12), eq 17 will simplify to

$$P(t,L)\,dt = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{(2n-1)\pi\left(\frac{24}{Pe_K N_K}\right)}{N_K\, b_K\,\alpha^2}\times \exp\left[-\frac{(2n-1)^2\pi^2 z}{4\,\alpha^2}\right]d(\Delta x) \quad (18)$$

The final step is to replace the $d(\Delta x)$ with dz (see eq 12).

$$d(\Delta x) = dz\left[\frac{Pe_K\, b_K\, N_K^2}{24}\right] \quad (19)$$

The result is

$$P(z)\,dz = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{(2n-1)\pi}{\alpha^2}\exp\left[-\frac{(2n-1)^2\pi^2 z}{4\,\alpha^2}\right]dz \quad (20)$$

We plot eq 20 on Figure 8b to compare the predictions of this simple double-sided first-passage-time theory with our simulation results. This approximation is successful in predicting the scaling, but it does a poor job of matching the results. We find $\alpha = 1.35$ to give the closest agreement with our data. This would imply that the actual distance the coil must diffuse (on average) out of the plane of the post is 0.675 R_g (close to the ½R_g that we had discussed previously).

We find, however, that the solution to the single-sided first-passage-time, where the coil must reach a certain distance in, say, the direction to the left of the pole and cannot escape by going to the right, actually does a good job describing our data (and this does not require the α parameter to scale the length measure). This might suggest that once the coil has drifted in one direction or the other it is convected slightly downstream and so has "committed" itself to passing by the pole on one particular side. This might help explain why the one-sided first-passage-time distribution fits the simulations better than does the two-sided distribution. This cannot be a complete explanation, however, since the theoretical one-sided distribution does not interpose any barrier that would prevent the coil from first moving a short distance in the "right" direction and then passing over the post and drifting arbitrarily far in the "wrong" direction. In the simulations, the pole would

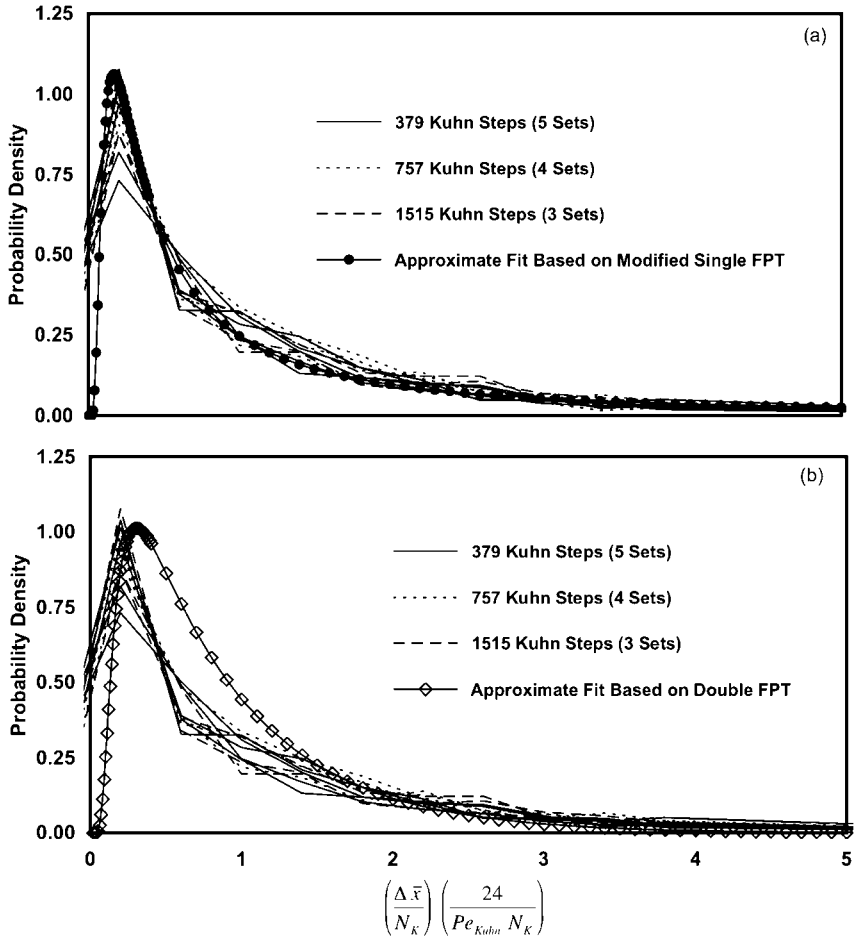

Figure 8. Rescaled probability distribution of $\Delta\bar{x}/N_K$ $(24/Pe_{Kuhn}\,N_K)$ restricted to Pe_{Kuhn} values in R2 for the chains with 379, 757, and 1515 Kuhn steps. These are independent of Pe_{Kuhn} or N_K. In (a), the solid circles are a phenomenological fit to the simulation data based on a modified single-sided first-passage-time, eq 25 (with $\beta = 0.02$). In (b), the open diamonds show the prediction based on the double-sided first-passage-time, eq 22 (with $\alpha = 1.35$).

either interfere with such motion or, if it did not interfere, the coil would pass by the pole on the either side and hence should be described by the two-sided distribution. Thus, the true situation is not necessarily described well by either the one- or two-sided distribution. Empirically, we observe that the one-sided distribution fits much better the simulation data than does the two-sided distribution. The solution to the single-sided problem is presented by Mazo[23]

$$P(\xi, t)\,\mathrm{d}t = \frac{\xi}{\sqrt{4\pi D t^3}} \exp\left[\frac{-\xi^2}{4 D t}\right] \mathrm{d}t \tag{21}$$

where ξ is the measure of distance, D is the diffusivity, and t is the required time. We repeat the scaling procedure described above to recast eq 21 as a function of z.

$$P(z) = \frac{1}{\sqrt{4\pi z^3}} \exp\left[-\frac{1}{4z}\right] \tag{22}$$

We recognize that eq 22, while fitting the distribution of delay times well, has an infinite mean, because once the coil drifts in the "wrong" direction, as noted above, there is a chance it will wander arbitrarily far before drifting back in the "right" direction. Obviously, this problem will not occur with a "two-sided" first-passage-time distribution for which the coil is allowed to pass the pole on either side. (The actual mean in the simulations is set by the extreme tail of the distribution function, and the infinite mean does not prevent the above one-sided distribution function from providing a good fit to the simulation data over the range of delay times considered.) The infinite mean will be a problem later as we seek to use this approximation to determine the variance. We therefore need to make one additional small correction to eq 22 so that it will have a finite mean. We multiply the function by a rapidly decreasing exponential ($\exp[-\beta z]$). We also must divide by ($\exp[-\sqrt{\beta}]$) to normalize the probability to unity.

$$P^*(z) = \frac{1}{\sqrt{4\pi z^3}} \exp\left[-\frac{1}{4z}\right] \frac{\exp[-\beta z]}{\exp[-\sqrt{\beta}]} \tag{23}$$

We plot eq 23 on Figure 8a to compare the predictions of this empirical function (based on the single-sided first-passage-time theory) with our simulation results. We find $\beta = 0.02$ to give the closest agreement with our data. We note that the mean value of eq 23 is around 3.5. We would expect, from Figure 6, that the mean value of z in R2 would be close to unity. This discrepancy can be attributed to the arbitrary manner in which we elected to alter eq 22 to result in a finite mean. A mean of 3.5 is much closer to unity than a mean of infinity. A more complicated extra term may have accomplished both tasks (getting an finite mean and getting a mean of unity). However, we accept eq 23 as a sufficient approximation for our later use.

We note that the plateau in Figure 6 has a lower bound at $Pe_{\mathrm{chain}} \approx 5$, corresponding to $Pe_{\mathrm{Kuhn}} N_{\mathrm{K}}^{3/2} \approx 10$. This marks the lower bound of R2.

Region (R1): $Pe_{\mathrm{Kuhn}} N_{\mathrm{K}}^{3/2} < 10$. R1 is where the convective force is very weak compared to the diffusive force. This is the very-slow-convection regime. Note in Figure 1 that for all four chain lengths we have data in this region and that the data all fall along the same line (this line can be fit to a power law that scales as $Pe_{\mathrm{Kuhn}}^{0.27}$).

As with R2, the polymer in R1 does not reach the rope-and-pulley formation because there is insufficient force to extend the chain. The polymer must translate beyond the post by randomly diffusing out of the plane containing the post and the field direction and then be moved downstream of the post by convective motion.

However, in R1, it is possible for a chain to diffuse far enough laterally to be able to clear the post but is convected so slowly that it might frequently diffuse back into line with the post before clearing it. It might therefore need to clear the post multiple times before the coil is able to move beyond the post. We note that in Figure 6 the transition from R2 to R1 happens at the same Pe_{chain} for all lengths of chain.

Note that some negative times appear for the R1 curves on Figure 7. In this very low Pe_{Kuhn} regime, where diffusion dominates the motion, it is possible for a coil to have moved downstream faster than it would have if it were carried by the convection alone. To explore this, we ran the two sets of simulations, $N_{\mathrm{K}} = 379/Pe_{\mathrm{Kuhn}} = 0.003$ (this is in R2) and $N_{\mathrm{K}} = 379/Pe_{\mathrm{Kuhn}} = 0.001$ (this is in R1), both with and without the post in the field. Figure 9a shows the comparison in R2. Note that the run without the post yields a distribution of delay distances centered and symmetric about zero, as expected, and the average delay time is zero. The run with the post in R2 has a peak to the positive side of zero and has a lengthy tail on the positive side, with relatively few molecules having a negative delay, indicating that while lateral diffusion determines the time required for the chain to clear the post (as discussed earlier) longitudinal diffusion (i.e., in the field direction) is less important than convection in determining the distribution of delay distances, especially in the tail region. We note that the relative importance of longitudinal diffusion compared to convection is influenced by the distance (about seven-Kuhn-step lengths) that the chain must travel downfield to reach the "finish line" where the simulation stops, and the time to reach this "finish line" is recorded. If we had put the "finish line" farther downfield, there would have been more time for both convection and diffusion to operate, but the importance of longitudinal diffusion, relative to convection, would have been diminished, and the fraction of chains with a negative "delay" time would have been smaller. The average delay time, which is the time to reach the "finish line" minus the time required for a chain purely convected in the absence of the post, for this run was greater than zero, indicating that the post had an influence on the chain.

Figure 9b shows that in R1 there is only a small difference between the runs with and without the post. The average delay time for the run without the post is zero, while it is slightly greater than zero when the post is present. In R1, the delay time is dominated by longitudinal diffusion and the presence of the post merely perturbs that diffusion. The asymmetry of the distribution even without the post is due to the fact that the largest negative delay possible occurs when the coil jumps immediately from the starting point to the finish line, while, because of diffusion, the positive delay can be arbitrarily large. Hence, the distribution of delay times is asymmetric about zero, even in the absence of the post.

Since R1 is controlled by the subtle perturbation to the distribution of delay times that the post introduces, and the distance to the finish line likely influences significantly the behavior in R1, predicting even semiquantitatively the behavior in this regime would be rather difficult. In addition, since chain motion is so slow, this regime is probably also of limited importance to electrophoretic separation. Hence we content ourselves here with the qualitative description of the behavior in R1 that we have just given.

Variance. We have developed in this work formulas that predict the probability distributions in R4/R3 (equation 5) and R2 (equation 23). Note that R4 is defined by eq 5 when $b = 0$. We now evaluate the mean and variance in each of these regions. We first must rewrite eq 23 in terms of $\Delta \bar{x}/N_{\mathrm{K}}$.

$$P^*\left(\frac{\Delta\bar{x}}{N_K}\right)=\frac{\left(\frac{24}{Pe_KN_K}\right)}{\sqrt{4\pi\left(\frac{24}{Pe_KN_K}\right)^3\left(\frac{\Delta\bar{x}}{N_K}\right)^3}}\times \exp\left[\frac{-\left(\frac{Pe_KN_K}{24}\right)}{4\left(\frac{\Delta\bar{x}}{N_K}\right)}\right]\frac{\exp\left[-\beta\left(\frac{24}{Pe_KN_K}\right)\left(\frac{\Delta\bar{x}}{N_K}\right)\right]}{\exp[-\sqrt{\beta}]} \quad (24)$$

Using eqs 5 and 24, we deduce the data shown in Table 1. Recall that $b = f(Pe_K, N_K)$ and that $\beta = 0.02$. We plot in Figure 10 the dimensionless ratio σ^2/μ^2 for chains with various Pe_K and N_K. We note that all chains reach the same value $(\sigma^2/\mu^2) = 1$ at high Pe_K and all chains have the same value $(\sigma^2/\mu^2) = 7.07$ at low Pe_K. It is only in R3 where σ^2/μ^2 depends on Pe_K for chains of different lengths.

The value of Pe_K at which the ratio σ^2/μ^2 for R2 intersects the value of $\sigma^2/\mu^2 = 7.07$ for R3 is marked on Figure 1 with the large asterisk mark. This depends on N_K. Note that this mark is a good measure of where the mean reduced delay $\langle\Delta\bar{x}\rangle/N_K$ for each chain shows a transition between R3 and R2 behavior. The agreement between the crossover as measured by the intersection of the values of σ^2/μ^2 and as measured by the crossover in the value of $\langle\Delta\bar{x}\rangle/N_K$ is better with the three longer chains, but we have commented earlier that the R3 approximation is less likely to be successful for chains with a small number of Kuhn steps.

Our simulations have dealt only with the interaction of a single chain with a single post. However, we can apply our findings qualitatively to the more useful case of a chain interacting with a sparse array of posts by first considering a row of posts arrayed perpendicular to the migration distance (with a spacing which we will discuss momentarily) and then considering n columns of these rows of posts in succession, allowing the chain to interact successively with one row after the next. This will be a simplified but useful extrapolation from the work presented in this paper to the case of a "sparse" array of posts. This can be combined with some of the work of Dorfman[12,13] to develop a full model for an array of sparse posts.

We have been assuming in each simulation that each chain is aligned with the post. We can imagine a situation, where there is a row of posts with a well-defined spacing between each post. A chain may approach this row and, depending on the spacing, either have an entanglement interaction with a post or pass through between two posts unaffected. For simplicity, we shall take a "step function" approach to the interactions and assume that the collision is either effectively a "head on" collision of the type we have analyzed here or a "clean miss". That is, the off-center collisions are treated as either close enough to being head on that they can treated as such or are lumped into those that experienced a clean miss. We will define f to be the probability of an effective head-on collision and 1-f the probability of an effective clean miss. From the discussion above, f will be approximately given as the ratio between R_g and the distance separating two posts in the same row. The percentage of clean misses will affect both σ^2 and μ for the average interaction of the chain with the first row. The "total" mean, incorporating the misses, is easily calculated as $\mu_T = f\mu$.

It is more complicated to determine the average total variance for the interaction with the row of posts. Consider the probability of delay to be a combination of two terms, the first term being the probability P that we discussed earlier and the second term being a delta function at zero (representing a miss). The contribution from each term is weighted by f.

$$P_T\left(\frac{\Delta\bar{x}}{N_K}\right) = fP\left(\frac{\Delta\bar{x}}{N_K}\right)+(1-f)\,\delta\left(\frac{\Delta\bar{x}}{N_K}\right) \quad (25)$$

The variance of eq 25

$$\sigma_T^{\,2} = \int_0^\infty\left(\frac{\Delta\bar{x}}{N_K}-f\mu\right)^2 fP\left(\frac{\Delta\bar{x}}{N_K}\right)\mathrm{d}\left(\frac{\Delta\bar{x}}{N_K}\right)+\int_0^\infty\left(\frac{\Delta\bar{x}}{N_K}-f\mu\right)^2(1-f)\,\delta\left(\frac{\Delta\bar{x}}{N_K}\right)\mathrm{d}\left(\frac{\Delta\bar{x}}{N_K}\right) \quad (26)$$

is found, after some mathematics, to be

$$\sigma_T^{\,2} = f\sigma^2+2\mu^2f(1-f)-\mu^2f(1-f^2)+\mu^2f^2(1-f) \quad (27)$$

Here σ^2 and μ have already been determined for each region. Equation 27 provides a simple expression for the adjusted variance as a function of σ^2, μ, and f for a row of equally spaced posts with any post density. The ratio $\sigma_T^{\,2}/\mu_T^{\,2}$ is then determined as

$$\frac{\sigma_T^{\,2}}{\mu_T^{\,2}} = \frac{1}{f}\frac{\sigma^2}{\mu^2}+\frac{1}{f}2(1-f)-\frac{1}{f}(1-f^2)+(1-f) \quad (28)$$

Note that, for $f = 1.0$, this returns the original σ^2/μ^2.

We now consider an array of n columns of these post rows. Both the variance and the mean will scale with n. If the rows of posts are spaced widely enough apart that a DNA molecule can completely relax after any encounter it experiences in the first row of posts before it encounters a post in the second array, we can take each encounter to be independent of the others. Then, according to the central-limit theory, for large n, a set of DNA molecules of given length passing through the array of posts will emerge distributed as a Gaussian with mean delay $n\mu_T$ and standard deviation $\sqrt{n}\sigma_T$. The relative width of the peak, which is the ratio of the standard deviation to the mean, scales inversely with the square root of n.

$$\textit{relative peak width} = \frac{1}{\sqrt{n}}\frac{\sigma_T}{\mu_T} \quad (29)$$

Since the above formula allows one to determine the bandwidths and resolution of different bands of migrating DNA, it could be used to design an array that can electrophoretically separate DNA strands by size. It is important to note that this is for a dilute DNA solution in a sparse array of posts. The posts must be sufficiently sparse that the chain is able to reform the random coils between interactions. The simulation method we have used here could also be used to determine the interactions of a DNA molecule with multiple, more densely arrayed, posts, where the interactions of the molecule with one post are coupled to the interactions with other posts. Such coupled interactions could take place either because the spacing in a single row is tight enough for the chain to interact simultaneously with more than one post or because the column spacing is small enough that the molecule does not recover completely from one interaction before it has another. One can imagine combining the results of such simulations with the formulas we have derived above by using "interaction coefficients" to adjust the dilute-post formulas for the effects of more densely arrayed posts, thus producing empirically useful formulas for designing such arrays for DNA separations. However, such additional work is beyond the scope of this study. A theory that converts single entanglement encounters into expressions predicting separation parameters has been put forward by Dorfman, who, however, did not consider the two low-field cases (R1 and R2) that we have explored here.

SUMMARY

We have used our previously developed bead-spring Brownian dynamics model to simulate the interactions between a

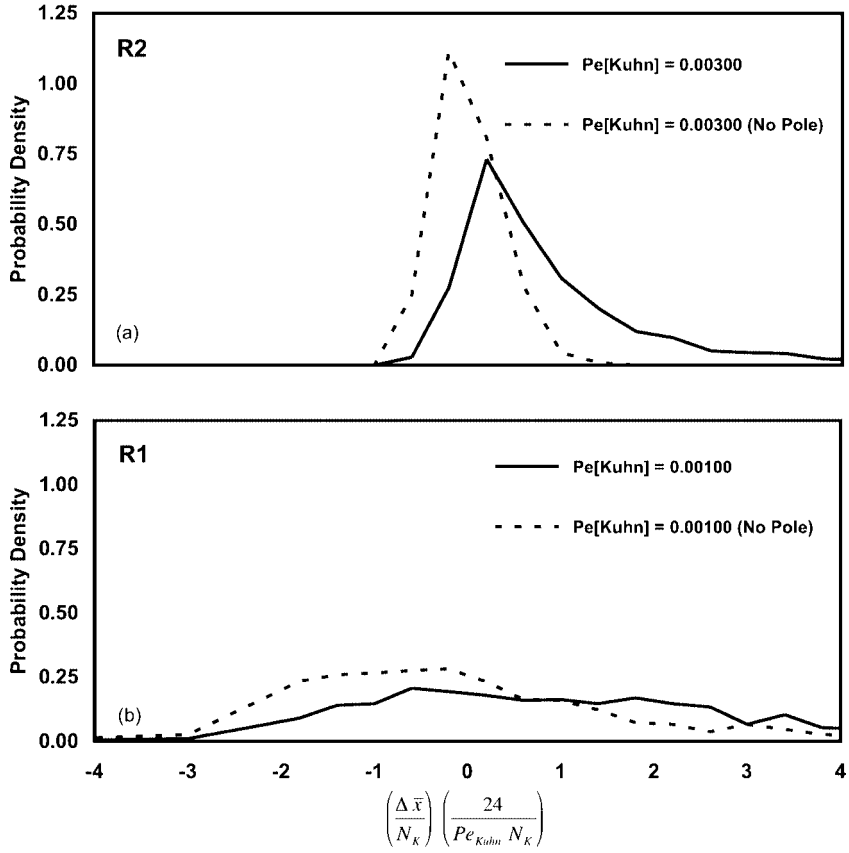

Figure 9. Rescaled probability distribution of $(\Delta \bar{x}/N_K)(24/Pe_{Kuhn}N_K)$ with and without the post for R1 and R2. (a) R2: $N_K = 379$ and $Pe_{Kuhn} = 0.003$. (b) R1: $N_K = 379$ and $Pe_{Kuhn} = 0.001$.

Table 1

	mean (μ)	variance (σ^2)	σ^2/μ^2
R4	0.5	0.25	1
R3	$0.5 - b$	$0.25 - b^2$	$(0.25 - b^2)/(1 - 2b + b^2)$
R2	$0.5(24/Pe_K N_K)^{-1}\beta^{-1/2}$	$0.25(24/Pe_K N_K)^2\beta^{3/2}$	$\beta^{-1/2}$

polymer chain and a thin obstacle. We studied electrophoretically translating DNA strands entangling with an immovable post, sampling a wide range of chain lengths ($25 \leq N_K \leq 1500$) and spanning 4 orders of magnitude in field strength ($10^{-4} \leq Pe_{Kuhn} \leq 10°$). We found that the delay in chain migration distance created by the entanglement is greater at higher fields and that there are four distinct regimes that describe these entanglements, two of which (R3 and R4) are dominated by convection, and two (R1 and R2) by diffusion. We discussed the physics that govern the four regions, and for all but the first region, (lowest field strength) we presented analytic approximations for the distribution of delay distances, including the mean and standard deviation of the delay.

The high field strength region (R4) is characterized by rope-and-pulley interactions (or variants of these), in which there is sufficient convective force for the polymer always to reach a fully extended configuration. Previous work[6] has shown that all chains (regardless of length) reach a limit of around $\langle \Delta \bar{x} \rangle / N_K = 0.5$ in this region. In a region with a less dominant convective force (R3), the chain is not able to fully extend its arms on other side of the obstacle. Hence, the polymer does not form the rope-and-pulley configuration; instead we model it as a ball-and-chain configuration in which the size (the number of monomers) of the ball is inversely related to the field strength. This results in shorter-lived interactions than those present in R4. While limitations of computer time prevented us from generating simulation data for R3 for the longer chains, we were

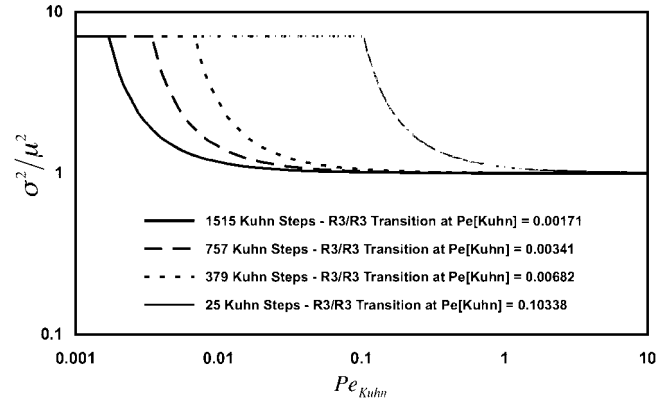

Figure 10. A plot of σ^2/μ^2 for four different length chains. $\sigma^2/\mu^2 = 1$ for all chains in the high-field regime, R4. $\sigma^2/\mu^2 = 7.07$ for all chains in the low-field regime, R2. σ^2/μ^2 is a function of Pe_{Kuhn} and N_K in the intermediate-field regime R3. The values of Pe_{Kuhn} for the transitions between R2 and R3 are shown with large asterisks on Figure 1.

able to develop a model to predict the behavior in R3 including the variation of $\langle \Delta \bar{x} \rangle / N_K$ with N_K and Pe_{Kuhn}, and this model agreed well with the simulation data that we were able to obtain. We note, however, that Randall and Doyle developed an alternative theory for this region that assumed uniform stretch rather than a ball-and-chain. While not describing our simulation data, the Randall−Doyle model does seem to describe well experimental data on DNA interactions with isolated posts.

In the two slower regimes (R2 and R1), the convective force is not able to alter significantly the shape of the polymer, and the chain retains a coil-like shape as it diffuses laterally and eventually passes the post largely without deforming. We argue that the controlling effect in R2 is the time required for the coil to diffuse a distance $\frac{1}{2}R_g$ out of the plane of the post. We

rescaled our data based on this argument and found universal behavior of all chain lengths in this regime. We also developed an approximation for this region based on the solution to the first-passage time problem. While lateral diffusion controls the delay distance in R2, with negligible effect of longitudinal diffusion in the slowest region (R1), convection is so slow that the distribution of delay times is dominated by longitudinal diffusion that is perturbed subtly by interactions with the post. Predicting even semiquantitatively the behavior in this regime would be rather difficult. We also note that R1 features a peculiar fractional power law relationship $\langle \Delta \bar{x} \rangle / N_K \propto Pe_{Kuhn}^{0.27}$.

Our results are unique in addressing a comprehensive range of electrophoretic strengths and chain lengths and in providing accurate analytical expressions for the distribution of delay distances induced by a single post. We find approximate expressions for the mean and variance of the probability distribution for each of the three fastest regimes and use these along with the central-limit theorem to generalize our results from a single chain interacting with a single post to a dilute solution of chains interacting with a sparse array of posts. Our methods can readily be extended to consider the effect of nonsparse arrays, with potential applications in the field of size-dependent separations.

References and Notes

(1) Barron, A. E.; Soane, D. S.; Blanch, H. W. *J. Chromatogr., A* **1993**, *652*, 3.
(2) Doyle, P. S.; Bibette, J.; Bancaud, A.; Viovy, J. L. *Science* **2002**, *295*, 2237.
(3) Volkmuth, W. D.; Austin, R. H. *Nature* **1992**, *358*, 600.
(4) Volkmuth, W. D.; Duke, T.; Wu, M. C.; Austin, R. H. *Phys. Rev. Lett.* **1994**, *72*, 2117.
(5) Randall, G. C.; Doyle, P. S. *Phys. Rev. Lett.* **2004**, *93*, 058102.
(6) Randall, G. C.; Doyle, P. S. *Macromolecules* **2006**, *39*, 7734.
(7) Starkweather, M. E.; Muthukumar, M.; Hoagland, D. A. *Macromolecules* **1998**, *31*, 5495.
(8) Nixon, G. I.; Slater, G. W. *Phys. Rev. E* **1994**, *50*, 5033.
(9) Saville, P. M.; Sevick, E. M. *Macromolecules* **1999**, *32*, 892.
(10) Patel, P. D.; Shaqfeh, E. S. G. *J. Chem. Phys.* **2003**, *118*, 2941.
(11) Padding, J. T.; Briels, W. J. *J. Chem. Phys.* **2002**, *117*, 925.
(12) Dorfman, K. D. *Phys. Rev. E* **2006**, *73*, 061922. (corrigendum to be submitted).
(13) Dorfman, K. D. *Phys. Rev. E* **2008**, *77*, 019901(E).
(14) Sevick, E. M.; Williams, D. R. M. *Phys. Rev. Lett.* **1996**, *76*, 2595.
(15) Kenward, M.; Slater, G. W. *Eur. Phys. J. E* **2006**, *20*, 125.
(16) Sevick, E. M.; Williams, D. R. M. *Phys. Rev. E* **1994**, *50*, 3357.
(17) Holleran, S. P.; Larson, R. G. *Rheol. Acta* **2007**, in press.
(18) Li, L.; Larson, R. G. *Macromolecules* **2000**, *33*, 1411.
(19) Hsieh, C.; Jain, S.; Larson, R. G. *J. Chem. Phys.* **2006**, *124*, 044911.
(20) Brochard-Wyart, F. *Europhys. Lett.* **1995**, *30*, 387.
(21) Mercier, J. F.; Slater, G. W. *Macromolecules* **2001**, *34*, 3437.
(22) Nagar, A.; Pradhan, P. *Physica A* **2003**, *320*, 141.
(23) Mazo, R. M. *Brownian Motion: Fluctuations, Dynamics, and Applications*; Clarendon Press: Oxford, 2002.

MA0712596