Simulation of the kinetics of a sphere attached to a fluctuating polymer: Implications for target search by DNA-binding proteins

J. Chakrabarti*
S. N. Bose National Centre for Basic Sciences, Block-JD, Sector-III, Salt Lake, Calcutta 700091, India

S. Roy†
Department of Biophysics, Bose Institute, P-1/12 CIT Scheme VII M, Calcutta 700054, India

(Received 7 November 2002; revised manuscript received 30 September 2003; published 24 February 2004)

We simulate the dynamics of a sphere, bound with a certain probability onto the beads of a harmonic bead-spring polymer in a confined geometry. The sphere hops from ith to the spatially closest $i\pm k$th bead. There is a crossover between kinetics dominated by different mechanisms with increasing binding probability. We relate these observations to the context of a biologically important problem of target location by DNA-binding proteins within a genome.

DOI: 10.1103/PhysRevE.69.021904 PACS number(s): 87.15.Aa, 87.15.Kg, 87.15.He

I. INTRODUCTION

Gene expression, that is, flow of information from DNA to protein via mRNA, is regulated primarily at the level of transcription [1]. The initiation of transcription and its regulation are complex and finely tuned processes, requiring from a few protein factors in prokaryotes to many protein factors in eukaryotes. In general, important regulatory nucleotide sequences are located on the upstream of the coding region. This upstream contains specific binding region for RNA polymerase (promoter) and regulatory proteins (operator). Probably, the most intensively studied gene regulatory protein is the lac repressor of *E. coli* that binds to three specific operator sites. The surprisingly fast kinetics of the regulatory protein-operator site binding with an in-vitro rate 1000 times faster than a normal diffusion-controlled reaction is far from a clear understanding. A phenomenological kinetic model was proposed [5–7] to account for the enhancement of the rate of the DNA target sequence location: A regulatory protein is more likely to form a complex with a nonspecific segment of DNA, the nonspecific binding potential being assumed to be uniform over the DNA contour. The regulatory protein in the nonspecifically bound state undergoes two facilitating transfer mechanisms, namely, ''sliding'' and ''intersegment transfer.'' The sliding in this model can be viewed as one-dimensional diffusion of the protein along the contour of the DNA. The intersegment transfer involves the binding of the protein to two DNA sites that are quite far apart along the DNA contour but are brought together transiently by the segmental diffusion of the loops of the DNA. Evidences in favor of the facilitating mechanisms are becoming more compelling with the advent of experimental techniques to manipulate the individual biomolecules [8], although quantitative analysis of such experiments so as to understand the role of the facilitating mechanisms in the kinetics, are still lacking. On the other hand, the equilibrium protein-DNA binding potentials have been characterized in great details [1–4]. The binding potentials, unlike the assumption of the kinetic model, show sequence dependent variations over the DNA contour. With this backdrop we study a semiphenomenological model, mimicking the protein-DNA system, to understand the role of the binding potential and its variation in the kinetics of a protein bound to nonspecific sites which would aid the quantitative analysis of the experiments and thereby help immensely in building up microscopic understanding of the fast specific protein-DNA reaction pathway via the facilitating mechanisms.

II. THE MODEL

We represent the DNA strand by a polymer of N coarse-grained beads [9] of hard sphere diameter d with nearest neighbor harmonic interaction of spring constant K, performing overdamped motion in a solvent of viscosity η [2,10] in a confined cubic box of size L. The confined box mimics the small ''domain'' having just one finite DNA chain as in the usual experimental situations [6,7]. However, we take L large enough to ensure that the polymer would be random, and the wall effects are not significant. Instead of considering the detailed DNA-protein binding processes [1,2], we consider, for simplicity, a large hard sphere of diameter $d_p > d$, attached to a bead with a probability p_i to stick onto it, p_i being a phenomenological characterization of the binding at the ith site. The attached sphere has a slower time scale of motion than the typical diffusion time of the monomer beads in the solvent, given by $\tau = a_s^2/D$, where the mean separation between the beads, $a_s = 1/\rho^{1/3}$, ρ being the number density of the beads, and the monomer diffusion constant, $D = k_B T/6\pi\eta d$, k_B being the Boltzmann constant and T denotes the room temperature. Moreover, we consider strong binding, namely, large p_i, and the dissociation time slower than τ, as for the protein-DNA reactions in the low salt concentration limit [7]. While the sphere can move from one bead to another, the chain undergoes a number of changes in configuration via the thermal motion. The sphere position is updated as follows. The sphere detaches with a probability $1-p_i$ from the ith site every ten steps and is attached to another which comes spatially the closest to the ith site. This rule is referred to in the following as rule (A). Note that both small (size ~ 1) and large (size $\gg 1$) jumps are allowed.

*Email address: jaydeb@boson.bose.res.in
†Email address: sidroy@vsnl.com

Moreover, the sphere can be viewed as experiencing a bias p_i fluctuating with a time scale, given by its hopping time. The protein might detach from the DNA in protein-DNA binding processes. However, such detachment is difficult to envisage experimentally [11]. This leads us to ignore the detachment of the protein sphere for the simplification of the analysis. Our analysis shows clearly that the rate of association of the sphere with the selected target (50th) site, having $p_{50}=0.999$ and $p_i=p<0.999$ for other beads, crosses over with increasing p from bead diffusion dominated regime to one dominated by small jumps of the sphere. Further, having $p'<p$ at the sites adjacent to the target site enhances the rate of association.

III. SIMULATION DETAILS

We solve numerically [12] over a time step Δt, the overdamped stochastic Langevin equation without the hydrodynamic interactions as in the Rouse model [10], for the ith bead trajectory $\vec{r}_i(t)$ without the periodic boundary conditions:

$$\gamma_i \frac{d\vec{r}_i}{dt} = -\vec{\nabla}_{\vec{r}_i} \sum_{j\neq i} V(|\vec{r}_i - \vec{r}_j|) + \vec{F}_i^{(\mathrm{R})}(t), \qquad (1)$$

γ_i being the Stoke's friction on the ith bead. $\gamma_i = 6\pi\eta(d + d_p)$, if the sphere is attached to the ith bead, and $\gamma_i = 6\pi\eta d$, otherwise. The first term on the right-hand side of Eq. (1) is the force due to the nearest neighbor harmonic interactions. The terminal particles are having only one nearest neighbor. The random forces $\vec{F}_i^{(\mathrm{R})}$ are Gaussian processes with zero mean and variance $2(k_BT/\gamma_i)\Delta t$. The bead positions are checked against overlap due to the hardcores present in the system. The sphere is randomly placed over a bead initially and updated as per the rule (A), satisfying the hardcore constraints. We choose $N=100$, $d=500$ A° which is comparable to the persistence length of DNA and smaller than experimentally known sliding range of proteins over DNA [2], and $d_p=2d$. We treat a_s as the length scale, τ the time scale and k_BT the energy scale. We fix in the scaled units $L^*=5$, $\Delta t^*=0.01$, and $K^*=Ka_s^2/k_BT=7.5$ such that the elastic deformation energy is much higher than k_BT as for DNA [2]. Typically initial 20 000 steps are discarded to allow the system to reach the steady state, monitored by the energy, the pair correlation function [10] and the probability of finding the protein sphere at the ith bead. The pair correlation function shows an exponential decay for large distances between a pair of beads as in a random polymer [10]. After the steady state, the quantities of interest are calculated by averaging over $N_0=10\,000$ configurations sampled from each of $M=500$ independent runs, consisting of 25 different initial positions of the sphere, each repeated for 20 different realizations of the noise.

IV. NUMERICAL RESULTS

We consider first the cases of uniform $p_i=p$ for all i. Let $N^{\mu}_{i,|k|}$ be the frequency that the $i\pm k$th bead is spatially the closest to the ith bead onto which the sphere is attached, out of N_0 configurations in the μth ($\mu=1,2,\ldots,M$) run. Figure 1(a) shows $\langle N_{|k|}\rangle=(1/MN)\Sigma_{\mu,i}N^{\mu}_{i,|k|}$ as a function of $|k|$. $\langle N_{|k|}\rangle \sim |k|^{-\alpha}$, with $\alpha\simeq 1.5$ for $|k|\leq 10$, has a complicated exponential decay for larger $|k|$. The behavior of $\langle N_{|k|}\rangle$ can be understood qualitatively [13] by the excluded volume effect [10]. We also calculate the frequency $f^{\mu}_{i,|k|}$ that the sphere jumps from site i to $i\pm k$ for different $|k|$. Figure 1(b) shows $\langle f_{|k|}\rangle=(1/MN)\Sigma_{\mu,i}f^{\mu}_{i,|k|}$ as a function of $|k|$. We observe that $\langle f_{|k|}\rangle=(1-p)\langle N_{|k|}\rangle$ except for $|k|\sim N$. The jump from i to $i\pm k$ for the μth run is a binomial process with parameters $(1-p)$ and $N^{\mu}_{i,|k|}$, having a mean $f^{\mu}_{i,|k|}=(1-p)N^{\mu}_{i,|k|}$ [14]. We have thus MN binomial distributions for a given $|k|$ so that the mean number of jump of size $|k|$, averaged over M runs and N beads, $\langle f_{|k|}\rangle=((1-p)/MN)\Sigma_{\mu,i}N^{\mu}_{i,|k|}=(1-p)\langle N_{|k|}\rangle$, observed in the low $|k|$ regime where the fluctuations in $N^{\mu}_{i,|k|}$ are small.

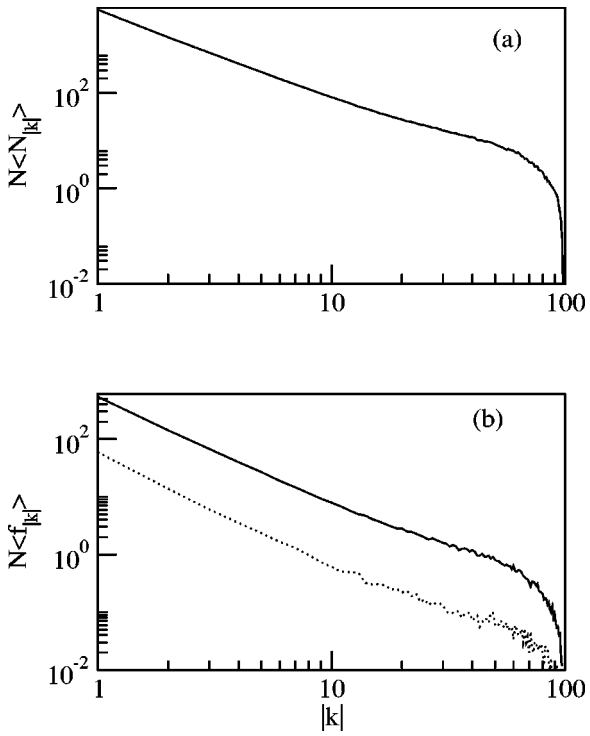

FIG. 1. (a) Log-log plot for $N\langle N_{|k|}\rangle$ vs $|k|$: $N\langle N_{|k|}\rangle \sim |k|^{-1.5}$ for $|k|\leq 10$ and $\log[N\langle N_k\rangle] \sim -1.5|k| + 0.18|k|^2 - 7.5\times 10^{-5}|k|^3$ for $|k| > 20$ (the fitted lines not shown). (b) Log-log plot for $N\langle f_{|k|}\rangle$ vs $|k|$: solid line for $p=0.9$ and the dotted line for $p=0.99$. The similar dependence as $\langle N_k\rangle$ in (a) is noteworthy except for $|k|\sim N$.

We further calculate the distribution $H^{\mu}_{i,|k|}(T^*_w)$ of waiting time T^*_w between two successive jumps of size $|k|$ out of $N^{\mu}_{i,|k|}$ observations. Figure 2(a) shows an exponential distribution of $H_{|k|}(T^{*'}_w)=(1/MN)\Sigma_{\mu,i}H^{\mu}_{i,|k|}(T^{*'}_w)$ over $T^{*'}_w = T^*_w/10\Delta t^*$ for $|k|=1$ with $p=0.9$. There is a substantial deviation from the exponential distribution for large $|k|$ ($=20$). If $N^{\mu}_{i,|k|}$ were large, as is the case for low $|k|$, the binomial process tends to the Poisson limit [14] where the waiting time distribution between two successive jumps is given by an exponential distribution [14]. The time spent by

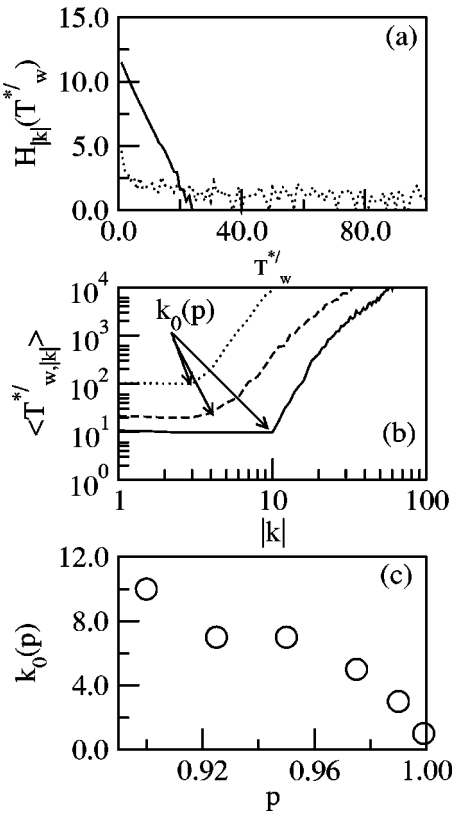

FIG. 2. (a) Log-normal plot for $MNH_{|k|}(T_w^{*'})$ as a function of $T_w^{*'}$ for $|k|=1$ (solid line) and $|k|=20$ (dotted line). Note the exponential distribution for $|k|=1$. (b) Log-normal plot for $\langle T_{w,|k|}^{*'}\rangle$ vs $|k|$ for three cases: (solid line) $p=0.9$; (dotted line) $p=0.99$; and (dashed line) target site at $i=50$, $p'=0.99$, and $p=0.9$ for the remaining beads where the contribution from the target site has been excluded. $\langle T_{w,|k|}^{*'}\rangle$ increases beyond $k_0(p)$, marked by arrows, up to which $\langle T_{w,|k|}^{*'}\rangle$ is $1/(1-p)$, independent of $|k|$. Note the slowing down of the jumps for $|k|<k_0(p)$ as $1/(1-p)$ with increasing p. (c) $k_0(p)$ vs p plot: Note the flat region and a sharp fall after $p=0.95$.

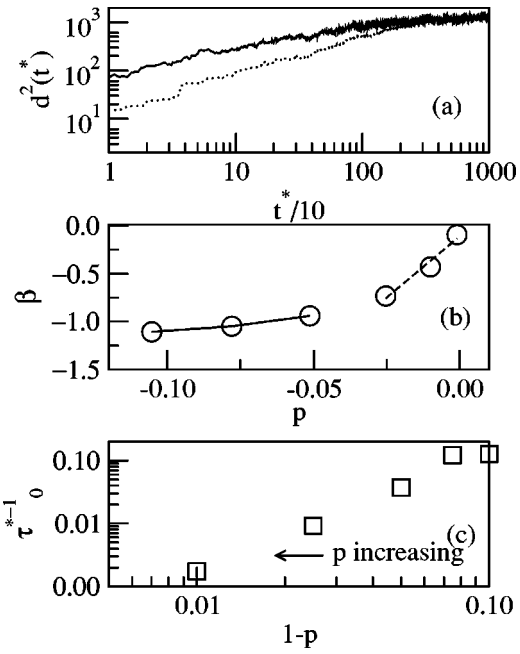

FIG. 3. (a) Log-log plot for $d^2(t^*)$ vs $t^*/10$: solid line for $p=0.9$ and dotted line for $p=0.99$. Note the saturation for large t^* and an initial subdiffusive growth. (b) β as a function of p in a log-log plot: Note a crossover from an initial slow dependence to a fast one, shown by the fitted solid and the dashed lines. (c) Log-log plot of $1/\tau_0^*$ as a function of $1-p$. Note the crossover in $1/\tau_0^*$.

the $i\pm k$th bead in the neighborhood of the ith bead to which the sphere is attached is $10N_{i,|k|}^{\mu}\Delta t^*$, and the number of jumps of size $|k|$ is $(1-p)N_{i,|k|}^{\mu}$, so that the number of jump per unit time is $(1-p)/10\Delta t^*$ which is the parameter of the exponential distribution, the mean being $10\Delta t^*/(1-p)$ [14], independent of μ,i, and $|k|$. The mean waiting time $\langle T_{w,|k|}^{*'}\rangle$, an indicative as well of the hopping time of the sphere over the chain, is thus $1/(1-p)$, independent of $|k|$, as shown in Fig. 2(b) for $|k|\leqslant k_0(p)$ for a given p. It may be worth to compare the hopping time to the mean approach time of a pair of beads. On the average the $i\pm k$th bead comes in the neighborhood of the ith bead $\langle N_{|k|}\rangle$ times over a period of $10N_0\Delta t^*$, the mean time of approach being $10N_0\Delta t^*/\langle N_{|k|}\rangle$ which is larger than the hopping time but not necessarily widely separated, especially for $|k|\leqslant k_0(p)$ in low p where the jumps are fast. This is in contrast to Ref. [15] that considers a particle hopping over a random heteropolymer where the hopping is much slower than mean approach time of the beads. Further, the rapid increase in $\langle T_{w,|k|}^{*'}\rangle$ beyond $|k|=k_0(p)$ indicates that these jumps are too slow to contribute significantly to the kinetics. Figure 2(b) also shows the $1/(1-p)$ dependence of $\langle T_{w,|k|}^{*}\rangle$ for $|k|\leqslant k_0(p)$ for different p. $k_0(p)$ falls off with increasing p, implying the suppression of large jumps. However, Fig. 2(c) shows that $k_0(p)$ falls off with p, albeit a discontinuity around $p=0.95$.

We also calculate the mean squared displacement of the sphere over the contour of the chain, given by $d^2(t^*)=\langle\{[i(t^*)-i(0)]b\}^2\rangle$, $i(t^*)$ being the location of the sphere at the ith bead at time t^*, $i(0)$ at the initial time, b the mean bond length and the average over $i(0)$. For a finite chain, $d^2(t^*)$ must saturate for sufficiently long time. Figure 3(a) indicates that the saturation approaches subdiffusively at $\sim At^{*\beta}$ with $\beta<1$, as observed for the motion of particles obeying fractional diffusion [16] in a complex geometry [17]. The subdiffusion here results from the compensating effects of the large jumps and the large waiting time needed for these jumps. β increases, while A decreases with increasing p ($\beta\simeq 1$ for $p=0.999$ as for a nearly diffusive motion), as the kinetics of the sphere gets dominated by slow and short jumps. We emphasize this by observing that the increase in β with p can be found as well from $[i(t^*)-i(0)]^2b^2P_i(t^*)$, $P_i(t^*)$ being the probability of finding the sphere at the ith bead, obtained by the numerical solution of the master equations:

$$\partial_{t^*}P_i(t^*)=\frac{1-p}{10\Delta t^*}\sum_j\left[\langle f_{|j-i|}\rangle P_j(t^*)-\langle f_{|i-j|}\rangle P_i(t^*)\right].$$

Here the first term describes the incoming of the sphere to, while the second one its loss from, the ith bead. The simu-

lated $\langle f_{|k|}\rangle$ has been used up to $|i-j|=|k|\leqslant k_0(p)$ and the prefactor accounts for the rate of these jumps. Figure 3(b) shows the simulation observation, confirmed from the solution of the master equations, of a crossover, namely, β increases slowly up to $p=0.95$ and then has a sharp increase.

Now we take the $i=50$ as a target site with $p_{50}=0.999$ and $p_i=p$ for other i. We calculate the mean association time τ_0^* of the sphere to the 50th bead, defined as the mean waiting time between two successive arrivals of the sphere at the 50th bead over M runs. $1/\tau_0^*$, the mean rate of association of the sphere to the target 50th site shown in Fig. 3(c), increases for large p, i.e., small $1-p$, and becomes independent of p for small p, or equivalently large $1-p$, the crossover being at $p=0.95$. Figures 2(b) and 2(c) show that large and fast jumps dominate for $p<0.95$. Here the mean time of approach of the $50\pm k$th bead to the 50th bead via bead diffusion, estimated by $10N_0N\Delta t^*/\langle N_{|k|}\rangle$, is slower than the hopping and becomes the limiting factor for the mean rate of association. Adding up all the contributions up to $|k|=k_0(p)$, we get the rate of approach $\Sigma_{|k|=1}^{k_0(p)}(\langle N_{|k|}\rangle/10N_0N\Delta t^*)\sim 0.09$ for $p=0.9$ and $p=0.925$ which is quite comparable to the saturated $1/\tau_0^*$. Thus the saturation of $1/\tau_0^*$ is governed by the rate of approach of the 50th and $50\pm k$th beads via thermal bead diffusion. This is quite analogous to the intersegment transfer process in the forward rate of transfer of protein from nonspecific to specific site [6,7]. For $p>0.95$ on the other hand, small jumps contribute to the kinetics such that the sphere spends a lot of time in traversing the chain before it can associate with the 50th bead, resulting in $1/\tau_0^*\sim(1-p)$ much smaller than the rate of approach of the beads via thermal diffusion. This picture is consistent with the sliding mechanism where the forward rate decreases as the dissociation rate from the nonspecific site decreases [6,7]. We have thus a crossover around $p=0.95$ in the target location rate kinetics dominated by intersegment transfer to the one dominated by sliding. Note in Fig. 3(b) that the crossover matches with that in β, even though $d^2(t^*)$ is subdiffusive around the crossover. The motion, dominated by sliding, becomes nearly diffusive only for large p.

We further consider inhomogeneity in the vicinity of the target site $i=50$, namely, $p_{49}=p_{51}=p'=0.99$ apart from the target site while $p_i=p=0.9$ for the remaining beads. $1/\tau_0^*$ ($\simeq 0.16$) enhances significantly compared to the corresponding case in Fig. 3(c) having the saturated $1/\tau_0^*$. The searching of the 50th bead is favored as the sphere is localized in the neighborhood at the 49th and 51st beads wherefrom the large jumps are suppressed due to $p<p'$. The suppression is reflected in Fig. 2(b) where $k_0(p)$ shifts to a lower value compared to the corresponding uniform case. However, $1/\tau_0^*$ ($\simeq 0.12$) decreases for $p>p'=0.75$ which enhances the large jumps at the 49th and the 51st sites. Further, $1/\tau_0^*$ ($\simeq 0.06$) decreases by having inhomogeneity at sites farther off the target site, in particular, $p_{40}=p_{60}=0.99$ along with the target site and $p_i=0.9$ for any other i. Thus the presence of sites with strong binding such that $p'>p$ in the neighborhood of the target site is significant in faster target location. We check the robustness of our results with a different updating rule: After detachment from the ith bead with probability $1-p_i$ every couple of steps, either the sphere slides to the $i\pm 1$st bead with a probability $q(=1/2$ in our case) or undergoes a large jump with a probability $1-q$ randomly to one of the beads $i\pm k$ with $|k|\neq 1$ within a radius of R_0^* ($=1$) around the sphere. In contrast to rule (A), the sliding to the nearest neighbor has a different probability from that of the large jumps depending on q.

V. CONCLUSION

In conclusion, our simple model captures qualitatively the intersegment transfer and sliding contributions to the kinetics of protein-DNA reaction. Thus our model would serve as a basis for the quantification of the kinetic contributions in the fast specific protein-DNA reaction pathway. In particular, we show a kinetic transition with p in the target location rate by the sphere from a regime dominated by bead thermal diffusion as in the intersegment transfer process to one dominated by short jumps of the sphere that could be identified with sliding process. Our analysis shows that the kinetic crossover takes place while the motion of the sphere over the backbone remains subdiffusive. However, the crossover is reflected in the kinetic quantities, like β and $k_0(p)$. In contrast to gross rate measurements [6,7], the kinetic quantities calculated in our analysis are amenable to direct measurements in single molecular biology experiments. We also show that the search process is favored by the presence of sites with $p<p'$ in the neighborhood of the target site, despite the subdiffusive motion. Our predictions can be verified by in-vitro experiments on a grafted polymer with a macromolecule having differential but strong binding affinity to the polymer. We hope to report on further improvements on our model, needed to be quantitavely comparable to the actual biological systems: The chain dynamics should be performed for a semiflexible model, the thermal dissociation of the sphere from the chain should be included and the sequence specific binding data should be taken into account through the p_i.

ACKNOWLEDGMENTS

J.C. thanks G. Shivashankar, A. Lahiri, and A. Chatterjee for helpful discussions and Srabani Chakrabarti for critically reading the manuscript.

[1] Bruce Alberts *et al.*, *Molecular Biology of The Cell*, 4th ed. (Garland Scince Taylor & Francis Group, New York, 2002).

[2] V.A. Bloomfield, D. Crothers, and I. Tinoco, *Nucleic Acids Structures, Properties and Functions* (USB, Sausalito, CA, 2000).

[3] Gary D. Stromo and Dana S. Fields, *Trends in Biochemical Sciemce* (Elsevier, New York, 1998), p. 109.

[4] B. Dubertret *et al.*, Phys. Rev. Lett. **86**, 6022 (2001).

[5] P.H. von Hippel and O.G. Berg, J. Biol. Chem. **264**, 675 (1989).
[6] O.G. Berg, R.B. Winter, and P.H. von Hippel, Biochemistry **20**, 6929 (1981).
[7] R.B. Winter, O.G. Berg, and P.H. von Hippel, Biochemistry **20**, 6961 (1981).
[8] D. Anselmetti *et al.*, Single Mol. **1**, 53 (2000); C. Bustamante *et al.*, J. Biol. Chem. **274**, 16 665 (1999).
[9] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, New York, 1979).
[10] M. Doi, *Introduction to Polymer Physics* (Clarendon Press, Oxford, 2001).
[11] N. Shimamoto, J. Biol. Chem. **274**, 15 293 (1999).
[12] M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids* (Clarendon Press, Oxford, 1989).
[13] The jth bead will be spatially nearest to the sphere at the ith bead, if their separation s lies between $d+d_p$ and b. The probability of this event is proportional to $(1/V)\int_{d+d_p}^{b} ds s^2 (1-v_c/s^3)^{n_{ij}(n_{ij}-1)/2} \exp(-3s^2/2n_{ij}b^2)$ [10], V being the system volume, v_c the excluded volume, and n_{ij} the number of beads between the ith and the jth bead. The integral, evaluated numerically, yields an algebraic dependence with an exponent 1.23 for low n_{ij} and an exponential tail for large n_{ij}. The simulation estimate of the exponent is slightly larger presumably due to the confinement.
[14] W. Feller, *An Introduction to Probablity Theory and its Applications* (Wiley Eastern, NewDelhi, 1991) Vol. 1, eighth Wiley Eastern reprint.
[15] D. Brockmann and T. Geisel, Phys. Rev. Lett. **91**, 048303 (2003).
[16] I.M. Sokolov *et al.*, e-print cond-mat/0107632; E. Barkai *et al.*, Phys. Rev. E **61**, 132 (2000).
[17] S. Jespersen *et al.*, Phys. Rev. E **59**, 2736 (1999); H.C. Fogedby, Phys. Rev. Lett. **73**, 2517 (1994); A. Maritan *et al.*, *ibid.* **71**, 1027 (1993).