Effect of external fluctuations on the affinity-specificity negative correlation in DNA-probe interactions

R. Murugan*

Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai, India, 400005

(Received 8 November 2005; revised manuscript received 25 January 2006; published 23 May 2006)

We show that the site-specific interaction of a probe DNA with the template DNA can be well modeled as an unbiased random jump process, where the probe molecule first nonspecifically binds to the template DNA and then searches for the specific site via unbiased random jump motion on the template DNA. By investigating the effects of increasing the jump size, and the fluctuations in the position of the specific site and the fluctuations in the specific site interval on the affinity-specificity negative correlation, we show that (1) increasing the jump size will in turn increase the affinity of the probe toward its target site on the template DNA, however, with a limiting value—the maximum affinity condition; (2) the degree of supercoiling or condensation of the template DNA as well as the electrostatic interactions between the probe and the template in turn control the jump size associated with the dynamics of the probe on the template DNA; (3) under a maximum specificity condition (therefore with minimum affinity), by introducing an external fluctuation in the relative position of the target site on the template DNA with respect to the probe, one can still improve the affinity rate; (4) on the other hand, one can improve the specificity of the probe toward the target site on the template DNA by introducing external fluctuations in the target-site interval. Finally, we propose the design strategies and optimum experimental conditions to simultaneously enhance the affinity as well as the specificity of probe toward its target site on the template DNA.

DOI: 10.1103/PhysRevE.73.051915 PACS number(s): 87.15.Kg

I. INTRODUCTION

Recognition of a specific site on the DNA lattice in the presence of an enormous amount of nonspecific sites by a small stretch of another DNA or RNA is an important and fundamental process in molecular biology [1]. In the biotechnological applications point of view, the small stretch of DNA that interacts with a specific-site on the large size genomic DNA is called as the probe. Here the probe DNA is often a radio-labeled one and it is mainly used to detect the presence of a particular sequence, e.g., the gene of interest, on the large size genomic DNA that is the target or the template in the present context. This detection technique is called Southern blotting (Ref. [1]). Apart from this, many other molecular biological techniques such as PCR and DNA fingerprinting are all based on the site-specific interactions between the probe-DNA and the template-DNA (Ref. [1]). In all these techniques, the process of recognition of the target site on the template DNA by the probe DNA is known as annealing. Here the temperature associated with the annealing will be maintained in such a way that both the template DNA as well as the probe DNA will be in a single stranded form. Therefore, throughout this article it is assumed that the template DNA and the probe DNA are in a single stranded form. The site-specific interaction of a probe-DNA with the template-DNA is an unusual phenomenon that differs from most of the biomolecular processes such as enzyme-substrate and drug-protein interactions in a way that in the former case the recognition process is actually a one-dimensional nucleation-zipping type interaction (Ref. [2]), whereas in the later case the recognition process is a three dimensional lock-and-key type (Ref. [3]) interaction.

Here we should note that the detection efficiency of the probe DNA depends on two factors, viz., (1) its affinity toward the specific site on the target DNA; and (2) its specificity, i.e., the ability to discriminate the specific site from other nonspecific sites on the target DNA. Therefore the probe DNA that is used to detect the presence of a particular sequence on the target DNA should be designed in such a way that it should possess higher specificity as well as a higher affinity toward the target site. Unfortunately, the sequence specificity and the binding affinity associated with the interactions of the probe DNA with the target DNA correlate negatively with each other (Ref. [4]), except for some specially designed nucleic acid probes such as oligonucleotide aptamers (Refs. [5,6]). Here the term negative correlation is used to describe the decrease in the affinity of the probe toward the target site on the template DNA upon a concurrent increase in the specificity of the probe toward its target site, and vice versa.

Since using the probes is the cheapest as well as the fastest method of detecting the presence of a particular sequence on the large size genomic DNA (otherwise one needs to carry out the sequencing of entire DNA, which is of course expensive and time consuming), there is always an urge and demand to develop optimizing tools to design nucleic acid probes with an enhanced specificity as well as the affinity. In this context, a new type of probes are actively being designed and synthesized. One of such example is the PNAs (Protein Nucleic Acids) (Ref. [7]) that has a protein like backbone with ordinary DNA bases. Site-specific binding of PNAs to DNA has shown a remarkable enhancement of both

*Present address: Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20740. Electronic address: muruga@tifr.res.in or rmurugan@gmail.com. Fax: 011 +9122-2280-4610.

the specificity as well as the affinity. On the other hand, it is necessary to understand the origin of this affinity-specificity negative correlation by theoretical or experimental means so that one can develop simple selection rules for better designing of the probe or one can optimize the experimental conditions for better usage of the nucleic acid probes. Earlier studies have shown that a two-step model (Ref. [4]) could explain the negative correlation between the affinity and the specificity (Ref. [8]) in case of PNA-DNA interactions. Nevertheless, this is a deterministic model and it does not include the underlying microscopic processes that is actually stochastic in nature. The recently developed random jump model (Ref. [9]) suggested that the negative correlation between the specificity and affinity of the probe toward the specific site on the target DNA originates as a consequence when the probe DNA searches for its target site on the template DNA by unbiased random jump motion and this model also predicted that the affinity-specificity anticorrelation would diminish upon increasing the temperature or decreasing the viscosity of the medium. However, all the previously mentioned models describe only the dynamics of the probe on the DNA lattice, and none of the models takes the dynamics of the template-DNA into account, which also introduces fluctuations in the system, i.e., all the previously mentioned models assumed a fixed target site, whereas in the real situations the relative position of the target site with respect to the probe DNA is a fluctuating quantity. In this paper we investigate (1) the effect of increasing the jump sizes associated with the dynamics of the probe DNA and (2) the effect of external fluctuations or noises in the relative position of the target site as well as the target site interval on the template DNA with respect to the probe DNA on the affinity-specificity anticorrelation in DNA-probe interactions. This paper is organized as follows. First we present the origin of the random jump model, assuming a fixed target site on the template DNA and the consequences of increasing the jump size associated with the dynamics of the probe on the affinity-specific anticorrelation. Subsequently, we discuss the effects of introducing fluctuations in the relative position of the target site as well as the target site interval on the template DNA on the affinity-specificity anti correlation.

II. RANDOM JUMP MODEL ON DNA PROBE INTERACTIONS

In this section we develop the random jump model on DNA-probe interactions in a step by step manner and derive the basic results that are in turn necessary for the later sections of this paper. Hereafter we simply call the probe DNA as the probe and the template DNA as the template, and it is assumed that both of them are in single stranded form. According to our model, the probe interacts with the specific site on the template via two steps (Refs. [10,11]) viz. the probe first nonspecifically binds to the template and then searches for the specific site by performing unbiased random jumps on the template. Since we insist the two-step condition, whenever the probe encounters the helical ends of the template, we assume that it returns back to template rather than detaching from it. Our arguments for the existence of unbiased random jump motion of the probe on the template in the process of searching for its target site are as follows.

First, one should note that the interaction between the probe and the template is a special case of the DNA renaturation process (Refs. [12,13]). The main difference is that in the later case, the interaction between the two complementary strands of the template is considered, whereas in the former case we consider the interaction between the template and the probe that is much smaller in size than that of the template. The process of renaturation of DNA has been extensively studied. Earlier studies on DNA renaturation have shown that the reagents those are capable of inducing the supercoiling or condensation of the template, such as divalent cations and an organic cosolvent like phenol, in turn, increase the renaturation rate (Refs. [14–17]). In other words, the rate at which a probe is targeting the specific site on the template is positively correlated with the supercoiled or condensed nature of the template.

Second, we should note that the jump size associated with the dynamics of the probe on the template can be positively correlated with the degree of supercoiling or condensation of the template, i.e., upon condensation or supercoiling, the probe can freely jump from the present position to a distal site on the template since two distal sites of the template can be brought closer together by a ring closure event (Refs. [10,11]). However, the distance associated with the occurrence of such ring closure events is a function of the degree of condensation or supercoiling of the template. For example, in case of a linear template, only the sliding motion (here the step size is a unit base pair) of the probe is favored. Here one should note that under solution conditions the template is mostly in a condensed or supercoiled state (since the buffer solution that contains the template will be always having divalent cations such as Mg^{2+}). Moreover, under solution conditions, the conformation of the template itself is a fluctuating quantity as we have stated in the Introduction, and therefore the occurrence of such ring closure events as well as the locations of such occurrences are equally probable and uncorrelated or unbiased random quantities. Apart from this, the template is also prone to the formation of local hairpin structures, which, in turn, impede the sliding motion of the probe on the template. In other words, the probe needs to simultaneously perform more than one type of facilitating movements (Refs. [10,11]) such as sliding (the jump size is the unit base), hopping (the jump size is a few bases), and intersegmental transfers (here the jump size is a few hundred to a thousand bases), as in the case of DNA-protein interactions to locate the target site on the template in a reasonable time. That is to say, the dynamics of probe on the template can be well modeled as an unbiased random jump process where the jump size is controlled by the degree of supercoiling or the degree of condensation of the template. Here the unbiased jump condition is due to the fact that the ring closure event can occur between any two distal sites with equal probabilities, where the distance of such occurrences are controlled by the degree of supercoiling or condensation of the template. These are all the basic ideas upon which we develop our random jump model on DNA probe interactions.

A. Forces affecting the affinity and the specificity of the probes

Here we should note that the hydrophobic forces originating from the base staking (Ref. [1]) between the probe and the template are the main forces associated with the nonspecific interactions, whereas in case of specific interactions apart from the hydrophobic forces arising from the base staking, hydrogen bonding between the complementary bases is also involved. On the other hand, there is also an electrostatic repulsive force due to the negatively charged phosphate groups present in the backbones of the template and the probe that counteracts the hydrophobic forces those are arising from the base staking interaction. Due to this reason the melting temperature of the site specifically bound template-probe complex will increase upon raising the ionic strength of the medium (Refs. [18]) contrasting from the site specifically bound DNA-protein complexes (Refs. [10,11]) where there are electrostatic attractive forces present in between the negatively charged DNA template and the positively charged protein molecule. This is mainly because increasing the ionic strength of the medium will weaken the electrostatic repulsion between the probe and the template, which in turn stabilizes the hydrophobic forces arising from the base staking leading to an increase in the melting temperature.

However, in the present context, this electrostatic repulsive force can also enhance the jump size associated with the dynamics of the probe on the template, i.e., the kinetic affinity toward the target site increases, however, with a decrease in the specificity associated with the probe toward its target site on the template. This can be explained as follows. As we have discussed in the earlier sections, to make jumps with larger jump sizes, the nonspecific interactions between the probe and the template should be a minimum apart from the requirement of higher degree of supercoiling of the template. However, the electrostatic repulsive forces also counteract with the hydrogen bonding interactions of the probe at the specific site that in turn reduces the specificity of the probe. Moreover, recent studies indicate that increasing the jump size k beyond certain critical k_c level, i.e., $k \geqslant k_c \approx 2N^{2/3}$, where N is the size of the template under consideration will not enhance the rate of the site-specific association of a protein/probe molecule with the DNA (Ref. [19]), i.e., via manipulating the degree of supercoiling of the template by the addition of either divalent cations or organics cosolvents, one cannot enhance the kinetic affinity of the probe toward its target site on the template beyond a certain level. We call this the maximum affinity condition. Under these maximum affinity conditions, the specificity associated with the probe toward its target site on the template DNA can be increased only by reducing the repulsive forces between the charged backbones of the template and the probe. This is the idea that is exploited in case of protein nucleic acid (PNA) probes, which is described in the following section.

B. Specialty of protein nucleic acid (PNA) probes

The PNA possesses a positively charged protein backbone rather than the negatively charged sugar-phosphate backbone, as in case of ssDNA, which in turn introduces an electrostatic attraction between the probe and the template and hence enhances the specificity of the probe toward its target site on the template, as we have discussed in an earlier section. For the same reason the melting temperature associated with the site specifically bound PNA-ssDNA complex slightly decreases upon increasing the ionic strength of the medium [18,20,21] contrasting from the site specifically bound DNA-DNA complex. Since the electrostatic attractive force is weaker than that of the protein-DNA interactions (because the distance between the peptide backbone of the PNA and the sugar-phosphate backbone of the DNA is much larger than that of the distance in the DNA-protein case), the jump size associated with the dynamics of the PNA-probe on the template is not much affected, i.e., the affinity is not retarded by this attractive force, however, the specificity is significantly enhanced. This is the reason why the PNAs show a remarkable enhancement of both the affinity as well as the specificity toward the target site on the template. In short, one can conclude that the PNA-probes take the advantages of DNA-DNA interactions as well as the DNA-protein interactions to enhance both the affinity and the specificity toward their target sites on the template. Now in the following sections we develop our random jump model in a much more rigorous manner. First we derive the expressions for the case where the target site on the template is fixed.

III. DNA-PROBE INTERACTIONS WITH FIXED TARGET SITE

Let us consider a DNA lattice of N (bases) bps in length, containing the specific site at the lattice position a such that $0<a<N$, where the set of lattice points $\{0,N\}$ constitutes (these are the helical ends) the reflecting boundaries and the lattice point $x=a$ is the only absorbing boundary (i.e., the specific site). Here the reflecting boundaries at the helical ends are introduced to ensure that the interaction between the probe and the template DNA occurs in two steps, as we have seen in the earlier section, and we assume that the absorbing point $x=a$ is fixed. Now let us assume that the probe DNA nonspecifically bound at the lattice position $x=x_0$ at time $t=0$, and currently searching for the specific site by unbiased random jump motion with a jump size of k bps. Here the jump size k indicates that starting from a position x, in the next step the probe can be found anywhere in the interval $x\pm k$ with equal probabilities (which is equal to $1/2k$). Here one should note that this equal probability assumption is again controlled by the degree of supercoiling/condensation of the template DNA. It is obvious to note that in case of linear DNA this assumption is clearly not valid since the probe molecule cannot make jumps with different jump sizes with equal probabilities on the linear template DNA. As we have discussed in the earlier sections, under usual experimental conditions the template DNA will be in a supercoiled or condensed state. Therefore our equal probability assumption is valid both under *in vitro* conditions. Now the probability of finding the probe DNA on the template DNA can be described by the following birth-death master equation:

$$\partial_t P_{x,t} = \sum_{i=1}^{k} [P_{x-i,t} + P_{x+i,t} - 2P_{x,t}]. \quad (1)$$

Here $P_{x,t}$ is the probability of finding the probe DNA at the position x of the template at time t. The Fokker-Plank equation (FPE) associated [22,23] with the master equation (1) is simply given as

$$\partial_t P_{x,t} = \frac{D_k}{2} \partial_x^2 P_{x,t}, \quad (2)$$

where $D_k = k^{-1}\Sigma_{i=1}^{k} i^2 = (k+1)(2k+1)/6$ is the one-dimensional phenomenological diffusion coefficient in the dimensionless form (Refs. [22,23]). The mean first passage time (T_x) associated with escape of the probe DNA through the specific site a can be easily derived from the backward FPE, i.e.,

$$d_x^2 T_x = -\frac{2}{D_k}. \quad (3)$$

When the jump size $k=1$ (therefore $D_k=1$), the mean first passage time (MFPT) associated with the site-specific association of the probe at the target site can be derived as follows. If the initial position x_0 is such that $0 \leqslant x_0 \leqslant a$, then the MFPT is given as $T_{L,x_0} = a^2 - {x_0}^2$. Whereas if the initial position x_0 is such that $a \leqslant x_0 \leqslant N$, then the MFPT is given as $T_{R,x_0} = (a^2 - {x_0}^2) + 2N(x_0 - a)$. Here the reflecting boundary conditions are $d_x T_x|_{x=0} = d_x T_x|_{x=N} = 0$ and the absorbing boundary condition is $T_x|_{x=a} = 0$. However, when $k > 1$, though the initial position of the probe on the template DNA is in the interval $0 \leqslant x_0 \leqslant a$, there is a definite probability associated with the probe DNA to escape from the interval $[0, a-1]$ into the interval $[a+1, N]$ without actually getting absorbed at the target-site lattice position a. Since we have three boundary conditions, Eq. (3) cannot be solved analytically. However, this problem can be solved in a different way as follows [24].

First let us consider only the interval $[0, a-1]$ and let us compute the MFPT associated with the probe to escape only through the point $x=a$. Suppose if we consider an M number of trajectories starting from the position $x=x_0$, where $0 < x_0 < a$, an M/k number of trajectories will end at the position $x=a$ and therefore removed from the system with a MFPT of $T_{L,x_0,0} = D_k^{-1}[a^2 - {x_0}^2]$ and an Mi/k number of trajectories will end in the interval $[a+1, a+i]$ with MFPTs $T_{L,x_0,i} = D_k^{-1}[(a+i)^2 - {x_0}^2]$ and therefore put back again into the interval $[0, a-1]$. However, the MFPT that is taken by the trajectories those are hitting the interval $[a+1, a+i]$ will simply add up to $T_{L,x_0,0}$ with an appropriate weighting factor. Therefore the MFPT associated with the escape of the probe DNA only through the position $x=a$ is given by the weighted sum $\vec{T}_a = T_{L,x_0,0} + D_k^{-1}\Sigma_{i=1}^{k} \mu_i [i^2 + 2ai^2]$, where the weighting factor is $\mu_i = i/k$,

$$\vec{T}_a = T_{L,x_0,0} + 2a + f(k). \quad (4)$$

Here $f(k) = 3k(k+1)/2(2k+1)$, and one should note that $\vec{T}_a$ is the mean time for which the probe molecule stays in the interval $[0, a-1]$ before it gets absorbed at $x=a$. When we consider the situation where the absorbing point is situated such that $a=N-1$, then from Eq. (4) one can easily show that there exists a minimum value of MFPT ($\vec{T}_a$) at the critical jump size equal to $k \geqslant k_c \approx 2N^{2/3}$. In other words, increasing the jump size beyond the critical value $k \geqslant k_c \approx 2N^{2/3}$ will not enhance the affinity of the probe DNA toward the specific site on the target DNA at all (Ref. [19]).

Similar to the derivation of Eq. (4), one obtains the expression for mean time for which the probe DNA stays in the interval $[a+1, N]$ before it gets absorbed at the lattice point a as follows:

$$\overleftarrow{T}_a = T_{L,x_0,0} + 2(N-a) - f(k) \quad (5)$$

Here the term $T_{L,x_0,0}$ in Eq. (5) is to account for the time that is taken by the probe DNA to enter into the interval $[a+1, N]$ from the interval $[0, a-1]$. Now let us assume that the current position of the probe on the template DNA is $x = a-1$. Now the probability associated with the escape of the probe in to the interval $[a+1, N]$ from the lattice point $x = a-1$ is simply $\overleftarrow{p}_a = (k-1)/2k$, from which we obtain the probability associated with escape of the probe through the target site from the interval $[0, a-1]$ as $\vec{p}_a = (k+1)/2k$. Now using these splitting probabilities one can write the expression for the overall MFPT taken by the probe DNA to get absorbed at the position $x=a$ starting from the interval $0 < x_0 < a$ as follows (Ref. [24]),

$$\overleftrightarrow{T}_a = \vec{p}_a \vec{T}_a + \overleftarrow{p}_a \overleftarrow{T}_a = T_{L,x_0,0} + N - \frac{N}{k} + \frac{2a}{k} + \frac{3(k+1)}{2(2k+1)}. \quad (6)$$

Here we should note that $\lim_{k\to\infty} \overleftrightarrow{T}_a = N$ and $\lim_{k\to 1} \overleftrightarrow{T}_a \approx T_{L,x_0}$. Since we insist on the two-step assumptions in the probe-template DNA interactions, the limit $k \to \infty$ in Eq. (6) is meaningless, i.e., to be consistent with the two-step model the jump size has to satisfy the inequality $0 < k \leqslant N$. However, from Eqs. (4) and (5) for $x_0 = 0$, one can easily show that

$$\lim_{k \geqslant k_c} \overleftrightarrow{T}_a = T(x_0 = 0 | k \geqslant k_c) = N. \quad (7)$$

A. DNA probe interactions under reversible conditions

So far we assumed that the specific site at $x=a$ on the template DNA is a pure-absorbing boundary which is not actually realistic due to fact that there is always a probability p_e such that $p_e > 0$ associated with the site specifically bound probe DNA to escape into the intervals $[0, a-1]$ and $[a+1, N]$. The escape of the specifically-bound probe DNA from the specific site on the template DNA can also be a result of probe secondary structure (e.g., hairpin looping) or single-strand structure of the template or probe DNA (Refs. [25–27]). Since at any time the probability associated with the nonspecifically bound probe to find the specific site is $1/2k$, one can write the escape probability p_e as $p_e = \alpha/2k$, where α is the proportionality constant. Now the modified splitting probabilities $\vec{p}_{a,m}$ as well as $\overleftarrow{p}_{a,m}$ can be written as $\overleftarrow{p}_{a,m} = \overleftarrow{p}_a + \alpha/4k$ and $\vec{p}_{a,m} = \vec{p}_a + \alpha/4k$ from which we can con-

clude that $\lim_{k\geq k_c} \tilde{p}_{a,m}=\tilde{p}_a$ and $\lim_{k\geq k_c} \vec{p}_{a,m}=\vec{p}_a$, i.e., when the jump size k associated with the probe DNA is equal to or beyond the critical value $k\geq k_c\approx 2N^{2/3}$, one can ignore the probability associated with the escape of the site-specifically bound probe from the specific site on the template DNA and therefore Eq. (7) is still valid under the reversible binding condition too.

B. Increasing the probe affinity beyond the critical jump size

Equation (7) states that under maximum specificity condition (i.e., putting the condition that the point $x=a$ is the only absorbing point), by manipulating the jump size k, the maximum achievable target finding rate is $r_{k>k_v}\approx 1/N$. In other words, by manipulating the degree of supercoiling or condensation of the template, one can achieve a maximum target finding rate of $r_{k>k_v}\approx 1/N$ with maximum specificity, i.e., without any mismatch base pairing. If one wants to increase the target finding rate further, then either the specificity of the target site with respect to the probe DNA has to be reduced or the electrostatic repulsive force between the backbones of the template and probe needs to be reduced, as in the case of PNAs. This can be demonstrated as follows. Now let us assume that the set of lattice points $x=a\pm\delta_s$ are absorbing points (therefore the specificity of the target site is reduced to an order proportional to $1/\delta_s$). Then the MFPT associated with the escape of the probe through the interval $a-\delta_s<x<a+\delta_s$ at the jump size $k>k_c\approx 2N^{2/3}$ is simply given as follows:

$$T_{\delta_s}(x_0=0|k\geq k_c)=\left(\sum_{i=0}^{2\delta_s}\frac{1}{N}\right)^{-1}=\frac{N}{2\delta_s+1} \tag{8}$$

Therefore the escape rate that is the inverse of MFPT is given as $r_{k>k_v}\approx(2\delta_s+1)/N$ (one should note that specificity $\propto 1/\delta_s$), indicating the existence of affinity-specificity anticorrelation, i.e., $r_{k>k_v}\propto[1/\delta_s]^{-1}$, even at or above the critical jump size $k>k_c$. Eq. (8) can be derived as follows. Since each point in the interval $a-\delta_s<x<a+\delta_s$ is an absorbing point, when $k>k_c$, the overall escape rate is simply the sum of escape rates associated with each absorbing point in the interval $a-\delta_s<x<a+\delta_s$ including the point $x=a$, i.e., $r_{k>k_v}\approx\frac{2\delta_s+1}{N}$ which in turn gives Eq. (8) since $T_{\delta_s}(x_0=0|k\geq k_c)=1/r_{k>k_c}$.

Here one should note that the phenomenological one dimensional diffusion coefficient associated with the probe DNA is directly proportional to the jump size k. On the other hand, the observed diffusion coefficient associated with the probe DNA is directly proportional to the temperature and inversely proportional to the viscosity of the medium, i.e., the one dimensional jump size k can be increased by increasing the temperature or by reducing the viscosity of the medium. In our earlier study (Ref. [9]) we have shown that when the temperature of the medium becomes extremely large, the affinity-specificity anticorrelation would vanish. In this context, in the present study we show that raising the temperature of the medium reduces the affinity-specificity anticorrelation associated with the probe-template DNA interactions, but the anticorrelation never reduces to zero until the probe DNA is nonspecifically attached to the template DNA. However, one should note that raising the temperature of the medium beyond a certain limit will in turn raise the jump size k beyond N (here this is size of the template DNA), where our two-step assumption of DNA probe interactions breaks down leading to the disappearance of affinity-specificity anticorrelation. Here one should recall the fact that the jump size k can also be positively correlated with the degree of supercoiling of the template DNA, i.e., a higher degree of supercoiling facilitates higher jump sizes. Therefore one can easily conclude that increasing the degree of supercoiling of the template DNA will reduce the affinity-specificity negative correlation in the DNA-probe interactions to a minimum value, but not to zero.

IV. DNA PROBE INTERACTIONS WITH FLUCTUATING TARGET SITES

So far, we have assumed that the relative position of the target site on the template DNA with respect to the probe is a fixed quantity that is an oversimplification of the real process, since under solution conditions the relative position of the target site on the template is a fluctuating quantity. In this section we assume that the fluctuations in the relative position of the target site are external and delta-correlated Gaussian white noises, and we investigate the effect of such fluctuations on the rate associated with the escape of the probe in to the specific site a. First, we consider the DNA-probe interaction with maximum specificity, i.e., $\delta_s=0$, and with the jump size $k>k_c$.

A. Effect of fluctuations in the position of the target site

Here the external white noise introduces fluctuations in the relative positions of the probe and the specific site on the template DNA, i.e., before the probe reaches the specific site a, the relative position of the specific site "a" with respect to the probe will be changed in a random manner. Here the fluctuation in the position of the specific site can be positively correlated with the translational-dynamics of the template DNA, which in turn depends on the temperature as well as the viscosity of the medium. Suppose if we denote the noise strength as δ_n base pairs, we can say that when the probe reaches the position a, due to the external noise or fluctuations, the position a might have shifted to any one of the lattice positions in the interval $a-\delta_n<x<a+\delta_n$. In other words, the specific site (i.e., absorbing point) fluctuates in the interval $a-\delta_n<x<a+\delta_n$ with equal probabilities. The argument for the validity of this equal-probability assumption is similar to that we have discussed in the earlier sections. Let us assume that the rate of fluctuations in the absorbing point (a) is almost equal to the rate of dynamics of the probe on the template DNA. Then the MFPT can be shown to be

$$T_{\delta_n}(x_0=0|k\geq k_c)=\left(\frac{1}{\delta_n+1}\sum_{i=0}^{2\delta_n}\frac{1}{N}\right)^{-1}=\frac{N}{2}\left(1+\frac{1}{2\delta_n+1}\right). \tag{9}$$

Here the escape rate associated with each possible lattice point in the interval $a-\delta_n<x<a+\delta_n$ is weighted with

the total noise strength δ_n+1. Now one can easily derive the limit

$$\lim_{\delta_n\to\infty} T_{\delta_n}(x_0=0|k\geqslant k_c)=N/2. \tag{10}$$

The limit given by Eq. (10) clearly states that even in the presence of maximum DNA-probe specificity (i.e., $\delta_s=0$) and the jump size $k>k_c$, one can still increase the affinity rate by introducing random fluctuations in the relative position of the target site.

B. Effect of fluctuations in the target site interval

Now let us assume that the length of the specific-site stretch (i.e., absorbing interval) $\delta_s>0$, and an external noise is introduced on $\delta_s=\delta_r$ such that the absorbing interval $[a-\delta_r,a+\delta_r]$ itself randomly fluctuates with equal probabilities. Here the fluctuation in the absorbing interval can be positively correlated with the bending-stretching dynamics of the template DNA. Again, we assume that the rate of fluctuations in the absorbing interval is almost equal to the rate of dynamics of the probe on the template DNA. Under this condition when $k>k_c$, the MFPT associated with the escape of the probe into the fluctuating absorbing interval is given as follows:

$$T_{\delta_r}(x_0=0|k\geqslant k_c)=\left(\frac{1}{\delta_r+1}\sum_{i=0}^{\delta_r}\frac{2i+1}{N}\right)^{-1}=\frac{N}{\delta_r+1}. \tag{11}$$

Here the escape rate associated with each possible absorbing interval is weighted with the total strength of fluctuations δ_r+1. Equation (11) clearly states that under fixed specificity conditions, the kinetic affinity associated with the probe toward its target site on the template DNA can still be increased by introducing fluctuations in the target site interval.

C. Generalizing to many specific sites on the template DNA

Now we generalize our concepts to the case where there are many specific sites on the template DNA to which the probe can interact simultaneously. Let us assume that the probe is situated at the position $x=0$ of the template DNA at time $t=0$, and there are m numbers of specific sites a_j where $j=1,2,3,\ldots,m$, such that $0<a_1<a_2<\cdots<a_m<N$, and each site spans an interval of $[a_j-\delta_{sj},a_j+\delta_{sj}]$, i.e., specificity associated with the jth target site is directly proportional to $1/\delta_{sj}$. Now, in the absence of external fluctuations, and in the presence of jump size $k>k_c$, it is easy to verify that the overall MFPT associated with the escape of the probe DNA into any one of the interval $[a_j-\delta_{sj},a_j+\delta_{sj}]$ is simply given as follows:

$$T_{\delta_s,m}(x_0=0|k\geqslant k_c)=\left(\sum_{j=1}^{m}\sum_{i=0}^{2\delta_{sj}}\frac{1}{N}\right)^{-1}=\frac{N}{m+\sum_{j=1}^{m}2\delta_{sj}}. \tag{12}$$

It is obvious to note that when the specificity associated with each target site is maximum, i.e., $\delta_{s1}=\ldots..=\delta_{sj}=\ldots=\delta_{sm}=0$, then $T_{\delta_s=0,m}(x_0=0|k\geqslant k_c)=N/m$. Suppose when there are fluctuations in the relative position of each specific site a_j with strength of δ_{nj} base pairs, one can easily derive the following general relationship:

$$T_{\delta_n,m}(x_0=0|k\geqslant k_c)=\left(\sum_{j=1}^{m}\frac{1}{\delta_{nj}+1}\sum_{i=0}^{2\delta_{nj}}\frac{1}{N}\right)^{-1}=\frac{N}{m+\sum_{j=1}^{m}\frac{\delta_{nj}}{\delta_{nj}+1}}. \tag{13}$$

Similarly, when there are fluctuations in the absorbing interval $[a_j-\delta_{rj},a_j-\delta_{rj}]$ with a fluctuation strength of δ_{rj}, one obtains the expression for the overall MFPT as follows:

$$T_{\delta_{s,m}}(x_0=0|k\geqslant k_c)=\left(\sum_{j=1}^{m}\frac{1}{\delta_{sj}+1}\sum_{i=0}^{\delta_{sj}}\frac{2i+1}{N}\right)^{-1}=\frac{N}{m+\sum_{j=1}^{m}\delta_{sj}}. \tag{14}$$

Here one should note the limit $\lim_{\delta_n\to\infty} T_{\delta_{s,m}}(x_0=0|k\geqslant k_c)=N/m$. The main conclusion derived from Eqs. (12)–(14) is that when large amount random fluctuations are introduced in the relative position (a) of the target site on the template DNA with respect to the probe DNA, the maximum achievable target finding rate in the presence of maximum specificity is doubled, i.e., $r_{k>k_c}\propto 2m$.

V. SIMULATION RESULTS AND DISCUSSION

To check the validity of Eqs. (8)–(11), simple random walk simulations on the DNA lattice are carried out. Here the settings are as follows. The total number of DNA lattice points $N=1015$; the specific site is at the $a=500$ bps position, the probe DNA was at $x=0$ at time $t=0$; the set of points $\{0,N\}$ constitutes the reflecting boundaries (i.e. whenever the probe hits these point, it will be reflected back into the interval $[0,N]$), only the point $x=a$ constitutes the absorbing boundary (therefore here $m=1$) in the case of maximum specificity. The specificity was relaxed by introducing the absorbing interval $[a-\delta_s,a+\delta_s]$, where the specificity factor δ_s is varied in the range of $0<\delta_s<10$. Here whenever the probe hit any one of the points in the interval $[a-\delta_s,a+\delta_s]$, it will be removed from the system with equal probabilities. The fluctuations in the position $x=a$ was introduced in the interval $[a-\delta_n,a+\delta_n]$ with equal probabilities, where δ_n was varied in the range of $0<\delta_n<10$. Similarly, the fluctuations in the absorbing interval $[a-\delta_r,a+\delta_r]$ was introduced with equal probabilities where δ_r was varied in the range of $0<\delta_r<10$. A total of 10^6 trajectories were averaged to obtain the MFPTs. The simulation results are shown in Fig. 1. The simulated MFPTs at the jump size $k>k_c$, at various values of the specificity factor and external noise are in accordance with the predictions given by Eqs. (8)–(11).

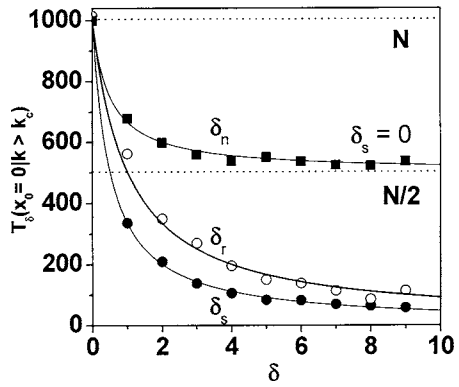

FIG. 1. Variation of (mean first passage time) MFPTs with the specificity factor δ_s (hollow circles), strength of fluctuation in the target-site interval δ_r (filled circles), and strength of fluctuation in the relative position of the specific-site δ_n (filled squares). Here the jump size k is chosen such that $k > k_c$. Here the total length of the DNA is 1015 (N) base pairs and the solid lines are the predictions by Eqs. (8), (9), and (11).

To check the validity of Eqs. (12)–(14) aforementioned simulation experiments were repeated with $m=2$. In this case two similar target sites (i.e., with equal δ values) are introduced at the lattice points $x=\{500, 1000\}$. The overall MFPT at various δ values are shown in Fig. 2 and the corresponding splitting probabilities (i.e., the probability associated with the escape of the probe DNA in each of the target sites on the template DNA as a function of the jump size k) associated with each target site at various δ values are shown in Fig. 3(a)–3(c). The MFPT values obtained from simulation agrees well with Eqs. (7)–(9) with $m=2$. When the jump size $k > k_c$ and $\delta_s=0$, in the presence of fluctuations in the relative positions of the target sites in the range of $0 < \delta_n < 10$, the splitting probabilities associated with each target-site are almost equal i.e., $p_{500}=p_{1000}=1/2$ (Fig. 3(a)). However, when absorbing interval $\delta_s > 0$, even at $k \gg k_c$, $p_{500} \neq p_{1000}$, and we

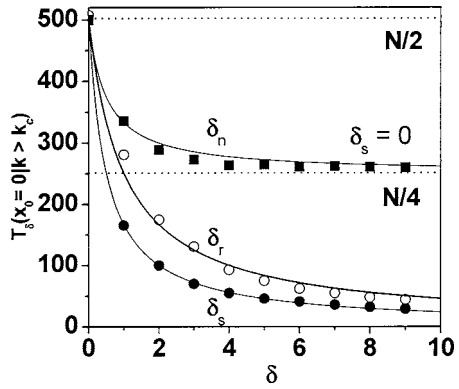

FIG. 2. Variation of MFPTs with the specificity factor δ_s (filled circles), strength of fluctuation in the target-site interval δ_r (hollow circles), and strength of fluctuation in the relative position of the specific site δ_n (filled squares) in the presence of similar target sites at the lattice positions $x=\{500, 1000\}$. Here the jump size is chosen such that $k > k_c$. Here the total length of the DNA is 1010 base pairs and the solid lines are the predictions by Eqs. (12)–(14) with $m=2$.

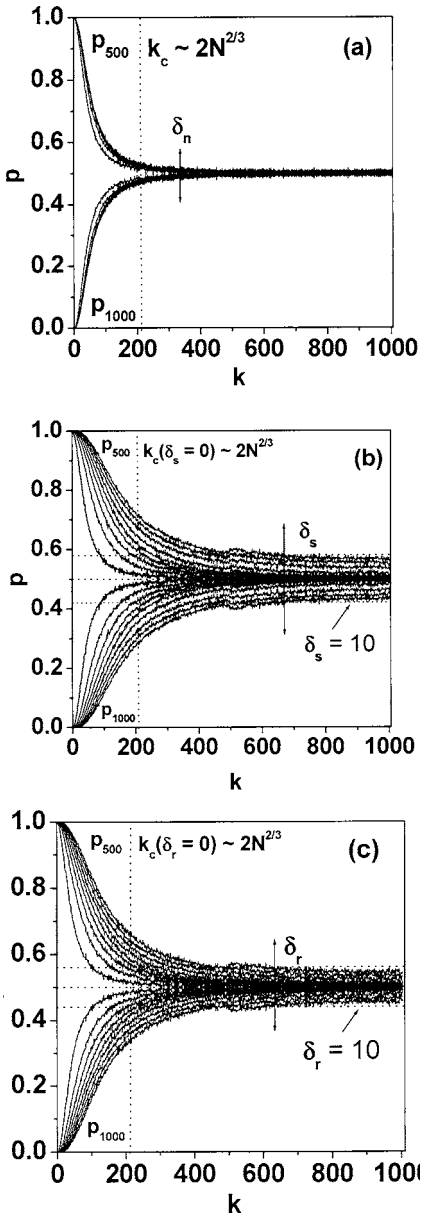

FIG. 3. Splitting probabilities (p) associated with the lattice points $x=\{500, 1000\}$ as a function of the jump size k. (a) In the presence of external fluctuations in the target site in the range of $0 < \delta_n < 10$. (b) Upon decreasing the specificity factor in the target-site interval with fluctuation strength in the range of $0 < \delta_s < 10$. (c) In the presence of external fluctuations in the target-site interval with fluctuation strength in the range of $0 < \delta_r < 10$.

observe the inequalities $p_{500} > 1/2$ and $p_{1000} < 1/2$ (Figs. 3(b) and 3(c). The total difference in the splitting probabilities, i.e., $\Delta P_s = |p_{500} - p_{1000}|_{k > k_c'}$ is a linear function of δ_s with a definite slope η. This behavior can be explained as follows.

Let us consider a case where the probe DNA was at the position $x=0$ at time $t=0$, and there are two target intervals at the positions $x=\{500 \pm 10, 1000 \pm 10\}$. When the jump size $k < 20$, then it is obvious to note that $p_{500}=1$ and $p_{1000}=0$, i.e., there is a gain in the splitting probability associated with the position $x=500$, with concurrent loss associated with the position at $x=1000$. Upon the introduction of the external

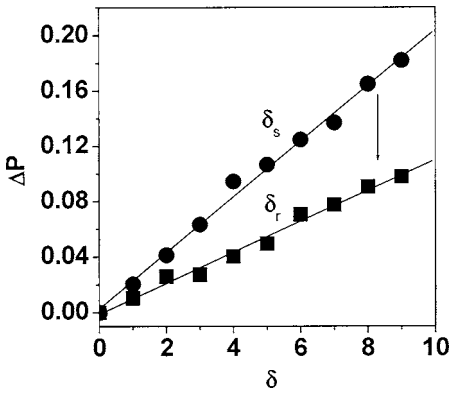

FIG. 4. Variation of the difference in the splitting probabilities associated with the lattice positions $x=\{500,1000\}$ as a function of the specificity factor in the range of $0<\delta_s<10$ (filled circles) and as a function of the strength of fluctuations associated with the target-site intervals in the range of $0<\delta_r<10$ (filled squares).

fluctuations in the absorbing intervals $[500-\delta_r,500+\delta_r]$ and $[1000-\delta_r,1000+\delta_r]$, the magnitude of the slope η significantly decreased and $\Delta P_r<\Delta P_s$ (Fig. 4). This is an important result to be considered especially in the PCR amplification of multiple target sites with a single primer set (e.g., in case of RAPD, random amplified polymorphic DNA, or DNA-fingerprinting), especially when there are fluctuations in the segment of DNA, where the target site is present. Our results clearly indicate that when the specificity is maximum, i.e., $\delta_s=0$, the rate of escape of the probe into the target site is directly proportional to the strength of the external fluctuations in the relative position of the target site. Since the helical ends are prone to higher fluctuations, the target-sites that are present near the helical ends of the template DNA possess higher opportunity to get amplified rather than the target sites that are present in the interior regions of the template, even under identical experimental conditions. This may produce undesired results in PCR dependent diagnostic methods such as RAPD and DNA fingerprinting. However, this effect can be compensated by designing a primer set such that it possesses the specificity factor $\delta_{s,\mathrm{int}}>\delta_{s,\mathrm{ext}}$, where $\delta_{s,\mathrm{int}}$ is the specificity factor associated with the interior target sites and $\delta_{s,\mathrm{ext}}$ is the specificity factor associated with the exterior target sites (here we should note that the specificity $\propto 1/\delta_s$).

On the other hand, when the specificity associated with the target sites are much smaller than by introducing external fluctuations in the target-site intervals, one can improve the specificity of the target sites without affecting the affinity of the probe toward the target site on the template DNA. Here one should note that in the case of PCR amplification of less specific sites, the specificity of amplification can be improved by simply raising the annealing temperature, since raising the temperature introduces more fluctuations in the relative position of the target site as well as in the target-site intervals. Finally, we present a strategy to reduce the affinity-specificity anticorrelation in DNA probe interactions as follows. First, the probe should be designed in such a way that the electrostatic attractive repulsive force between the backbones of the template and the probe should be a minimum, as in the case of PNAs. If it is not possible, then one can always achieve this by manipulating the ionic strength of the medium. Second, we can include the reagents that can induce the supercoiling or condensation of the template such as divalent cations or organic cosolvents such as phenol into the reaction mixture. Finally, as we have shown earlier, to decrease the affinity-specificity anticorrelation, the temperature of the reaction mixture can be increased or the viscosity of the reaction mixture can be decreased.

VI. SUMMARY

The affinity and the specificity associated with the interaction of a probe DNA with a specific site on the template DNA negatively correlates with each other, i.e., an increase in affinity leads to a decrease in the specificity of the probe DNA toward the specific site on the template DNA and, vice versa. DNA-probe interactions differ from most of the biomolecular processes such as enzyme-substrate and drug-protein interactions in a way that in the former case the recognition process is actually a one-dimensional nucleation-zipping type interaction, whereas in the later case the recognition process is a three dimensional lock-and-key type interaction. Noting the fact that none of the earlier models on DNA-probe interactions considered the dynamics of the template DNA, in this paper we have investigated the effect of external fluctuations or noises on this affinity-specificity negative correlation associated with the DNA-probe or DNA-protein interactions. We discuss our model in the context of efficient PNA probes which show a remarkable enhancement of both the affinity as well as the specificity toward the target site on the template DNA. Results have shown that (1) the DNA-probe interactions can be well modeled as a random jump motion, where the probe molecule first nonspecifically binds to the template DNA and then searches for the specific site via unbiased random jumps on the template DNA; (2) increasing the jump size will in turn increase the affinity of the probe toward its target site on the template DNA, however, with a limiting value—the maximum affinity condition; (3) the degree of supercoiling of the template DNA as well as the electrostatic interactions between the probe and the template in turn control the jump size associated with the dynamics of the probe on the template DNA; (4) under a maximum specificity condition (therefore with minimum affinity), by introducing an external fluctuation in the relative position of the target site on the template DNA with respect to the probe, one can still improve the affinity rate; (5) on the other hand, one can improve the specificity of the probe toward the target site on the template DNA by introducing external fluctuations in the target-site interval.

ACKNOWLEDGMENT

This work is supported by TIFR, Mumbai.

[1] B. Lewin in *Genes VIII* (Oxford University Press, Oxford, 2003); B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, in *Molecular Biology of the Cell* (Garland Science, New York, 2002).
[2] P. H. von Hippel and O. G. Berg, Proc. Natl. Acad. Sci. U.S.A. **83**, 1608 (1986).
[3] D. E. Koshland, Angew. Chem., Int. Ed. Engl. **33**, 2375 (1994).
[4] A. Lomakin and M. D. Frank-Kamenetskii, J. Mol. Biol. **276**, 57 (1998).
[5] M. Famulok, G. Mayer, and M. Blind, Acc. Chem. Res. **33**, 591 (2000).
[6] E. N. Brody and L. J. Gold, Biotechnology **74**, 5 (2000).
[7] P. E. Nielsen, M. Egholm, R. H. Berg, and O. Buchardt, Science **254**, 1497 (1991).
[8] V. V. Demidov and M. D. Frank-Kamenetskii, TIBS **29**, 62 (2004).
[9] R. Murugan, Biophys. Chem. **116**, 105 (2005).
[10] O. G. Berg, R. B. Winter, and P. V. Hippel, Biochemistry **20**, 6929 (1981).
[11] O. G. Berg, R. B. Winter, and P. V. Hippel, Biochemistry **20**, 6961 (1981).
[12] J. G. Wetmur and N. Davidson, J. Mol. Spectrosc. **31**, 349 (1968).
[13] J. G. Wetmur, Annu. Rev. Biophys. Bioeng. **5**, 337 (1976).
[14] B. W. Pontius and P. Berg, Proc. Natl. Acad. Sci. U.S.A. **88**, 8237 (1991).
[15] A. Goldar and J. L. Sikorav, Eur. Phys. J.: Appl. Phys. **14**, 211 (2004).
[16] I. Chaperon and J. L. Sikorav, Biopolymers **46**, 195 (1998).
[17] J. L. Sikorav and G. M. Church, J. Mol. Spectrosc. **222**, 1085 (1991).
[18] S. Tomac, M. Sarkar, T. Ratilainen, P. Wittung, P. E. Nielsen, B. Norden, and A. Graslund, J. Am. Chem. Soc. **24**, 5544 (1996).
[19] R. Murugan, Biophys. Chem. **120**, 143 (2006).
[20] M. D. Frank-Kamenetskii, V. V. Anshelevich, and A. V. Lukashin, Sov. Phys. Usp. **30**, 317 (1987).
[21] J. P. Bond, C. F. Anderson, and M. T. Record, Biophys. J. **67**, 825 (1994).
[22] C. W. Gardiner, in *Handbook of Stochastic Methods*, edited by H. Haken (Springer-Verlag, Berlin, 1983), p. 260.
[23] R. Murugan, Phys. Rev. E **69**, 011911 (2004).
[24] R. Murugan, J. Phys. A **39**, 1575 (2006).
[25] J. Jr SantaLucia and D. Hicks, Annu. Rev. Biophys. Biomol. Struct. **33**, 415 (2004).
[26] V. A. Bloomfield, Biochem. Biophys. Res. Commun. **34**, 765 (1969).
[27] G. Vesnaver and K. J. Breslauer, Proc. Natl. Acad. Sci. U.S.A. **88**, 3569 (1991).