How DNA coiling enhances target localization by proteins

B. van den Broek*, M. A. Lomholt†, S.-M. J. Kalisch*, R. Metzler‡, and G. J. L. Wuite*

*Department of Physics and Astronomy, Faculty of Sciences, Vrije Universiteit, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands; †MEMPHYS Center for Biomembrane Physics, Department of Physics and Chemistry, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark; and ‡Physik Department, Technical University of Munich, James Franck Strasse, 85747 Garching, Germany

Edited by Peter H. von Hippel, University of Oregon, Eugene, OR, and approved July 24, 2008 (received for review May 1, 2008)

Many genetic processes depend on proteins interacting with specific sequences on DNA. Despite the large excess of nonspecific DNA in the cell, proteins can locate their targets rapidly. After initial nonspecific binding, they are believed to find the target site by 1D diffusion (“sliding”) interspersed by 3D dissociation/reassociation, a process usually referred to as facilitated diffusion. The 3D events combine short intrasegmental “hops” along the DNA contour, intersegmental “jumps” between nearby DNA segments, and longer volume “excursions.” The impact of DNA conformation on the search pathway is, however, still unknown. Here, we show direct evidence that DNA coiling influences the association rate of EcoRV restriction enzymes. Using optical tweezers together with a fast buffer exchange system, we obtained association times of EcoRV on single DNA molecules as a function of DNA extension, separating intersegmental jumping from other search pathways. Depending on salt concentration, targeting rates almost double when the DNA conformation is changed from fully extended to a coiled configuration. Quantitative analysis by an extended facilitated diffusion model reveals that only a fraction of enzymes are ready to bind to DNA. Generalizing our results to the crowded environment of the cell we predict a major impact of intersegmental jumps on target localization speed on DNA.

DNA configuration | DNA–protein interaction | facilitated diffusion | intersegmental jumping | single-molecule

A n essential feature in biological processes on DNA is the ability of proteins to quickly locate specific DNA sequences in a vast surplus of nonspecific DNA (1, 2). A protein’s search for the target site is thought to be accelerated by facilitated diffusion along nonspecific DNA (3–6). Recent work has yielded considerable insight in the possible search strategies of site-specific proteins (7–15). Assisted by DNA looping some proteins can, for instance, intermittently bind to two DNA segments simultaneously. This way they can directly move from one to another chemically remote segment (16). This intersegmental transfer accelerates target finding on DNA because it assumes a constantly changing random configuration (11). Here, we demonstrate and quantify a similar mechanism, intersegmental jumping, for proteins with only one DNA-binding site.

Little experimental work on facilitated diffusion is available. To date, most studies have been investigating DNA cleavage by restriction enzymes in bulk assays, measuring association times as a function of DNA length, or monitoring processivity on DNA constructs with two sites (7, 17–21). Although in these biochemical assays valuable information can be obtained, association rates of proteins to specific sites are difficult to measure, and the underlying kinetics of the target search mechanism are often obscured. Furthermore, it remains experimentally challenging to distinguish 1D and 3D search pathways. In previous single-molecule assays only pure 1D protein search has been addressed (21–24). Here, we present single-molecule measurements of DNA cleavage by EcoRV on individual plasmid-size molecules (6,538 bp; one EcoRV site) having different degrees of conformational freedom. By tuning the DNA extension, the conformation of the DNA can be changed from a relaxed random configuration to an extended polymer. This procedure enables us to selectively “switch off” 3D intersegmental jumping, while leaving sliding, intrasegmental hopping, and long volume excursions intact. By acquiring specific association rates of EcoRV for these DNA conformations at different salt conditions, we can thus determine the relative impact of 3D intersegmental jumps (illustrated in Fig. 1 A and B) on the search process.

Results and Discussion

Experimental Approach. To determine the bimolecular association rate $k_{on}$ to a recognition site on linear DNA, for different polymer conformations, individual DNA molecules were tethered between two optically trapped beads in a multichannel flow chamber as described (25, 26). The degree of DNA coiling was set by changing the distance between the two beads (Fig. 1C). After the buffer flow was stopped DNA constructs were quickly (~0.5 s) transported into enzyme solution. The DNA molecule was briefly (~20 ms) stretched to a force of 5–10 pN every second, to check whether it was cut [supporting information (SI) Movie S1]. The transient stretching resulted in spikes in the force trace that disappeared when the DNA had been cleaved in the preceding second (Fig. 2). The cleavage time was defined as the time between moving the construct into the enzyme solution and scission of both DNA strands.

As we are interested in the influence of DNA configuration on the association rate, the complete cleavage reaction (comprising association, induced fit, DNA hydrolysis, and product release) should be limited by diffusion of the protein to the specific site. We have previously demonstrated that the induced-fit process at low DNA tension is very fast and that product release from the plasmid, in bulk experiments normally the slowest step (27, 28), is not rate-limiting in our optical tweezers experiments (25). This result leaves the actual strand cleavage reaction, the hydrolysis of two phosphodiester bonds in the DNA backbone, as a possible rate-limiting step. This latter rate was determined by measuring cleavage times of pCo5 DNA molecules in a stretched configuration at saturating enzyme conditions (500 nM) (Fig. 3). The rate of strand cleavage was found to be 0.60 s$^{-1}$. Consequently, the average time needed for the cleavage of both strands is 2.5 s. To satisfy the diffusion-limited requirement mentioned above, and to avoid flux...
would form a coil with a mean square end-to-end distance \(0.24\). A relaxed pCco5 DNA molecule (6,538 bp) with free ends the search rate acceleration peak at a fractional extension of

\[
\frac{R}{h} = 0.24\text{ (corresponding to } a = 0.22\text{.)}
\]

The maximal association rate is thus found at the extension where DNA is closest to its free random configuration. We attribute the rate difference with stretched DNA to the disappearance of intersegmental jumping in stretched DNA (Figs. 1A and B). At \(a \leq 0.2\), the presence of the two large beads causes a deformation of the DNA coil. In addition to possible obstruction of enzyme access routes, on average the local density of DNA segments around the specific site \(l_{\text{DNA}}\) will be lower than for the relaxed coil, resulting in the observed decrease in association rate. The second-order association rate constants deduced from the single-molecule cleavage events, \(1.8 \pm 0.2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}\) in the coiled configuration is in close agreement with values found earlier in bulk experiments in similar buffers: \(1.2–2.2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}\) (30–33).

**Salt Dependence of Association Rate.** Nonspecific protein–DNA interactions are largely of electrostatic nature and therefore depend on salt concentration (34). Changes in buffer conditions could thus induce a shift in the relative contributions of 1D and

![Typical data trace of a cleavage event](image)

Fig. 2. Typical data trace of a cleavage event. Displayed is the force on the DNA in the direction along the stretched molecule. Every second, the DNA is rapidly stretched to \(-5–10\) pN, resulting in sharp positive force spikes (negative spikes are caused by the fast movement of the beads and do not represent a real force on the DNA). Disappearance of these spikes implies that the DNA molecule has been cleaved in the preceding second. The best estimate for this event is exactly halfway between the last spike and the next stretching attempt. We thus define the reaction time as the time between transportation of the DNA construct into the enzyme-containing flow channel \((t = 0\) in the graph) and last upward spike plus 0.5 s. The statistical error that is introduced in the measured cleavage time \((-0.5\) s) averages out for the large number of data points.

**Determination of DNA hydrolysis rate, obtained with 500 nM EcoRV in reaction buffer with 100 mM NaCl.** Because in the optical tweezers experiments only the cleavage of the second strand is observed, the distribution of cleavage times shows a lag phase. The single exponential fit (dashed line) gives a strand cleavage rate \(k_1\) of \(0.60 \pm 0.03 \text{ s}^{-1}\) for the second strand. Because of degeneracy, cleavage of the first strand should in principle be twice as fast. As a result, the rate of hydrolysis for both DNA strands is \(0.40 \pm 0.03 \text{ s}^{-1}\). The average time required for DNA cleavage is the inverse, \(2.5 \pm 0.2\) s. The full distribution can also directly be fitted with such a two-step process \((k_1[\exp(-k_1t) - \exp(-2k_1t)])\). Doing so (solid line) yields a strand cleavage rate \(k_1\) of \(0.54 \pm 0.03 \text{ s}^{-1}\), comparable to the rate found above.

![Determination of DNA hydrolysis rate](image)

Fig. 3. Determination of DNA hydrolysis rate, obtained with 500 nM EcoRV in reaction buffer with 100 mM NaCl. Because in the optical tweezers experiments only the cleavage of the second strand is observed, the distribution of cleavage times shows a lag phase. The single exponential fit (dashed line) gives a strand cleavage rate \(k_1\) of \(0.60 \pm 0.03 \text{ s}^{-1}\) for the second strand. Because of degeneracy, cleavage of the first strand should in principle be twice as fast. As a result, the rate of hydrolysis for both DNA strands is \(0.40 \pm 0.03 \text{ s}^{-1}\). The average time required for DNA cleavage is the inverse, \(2.5 \pm 0.2\) s. The full distribution can also directly be fitted with such a two-step process \((k_1[\exp(-k_1t) - \exp(-2k_1t)])\). Doing so (solid line) yields a strand cleavage rate \(k_1\) of \(0.54 \pm 0.03 \text{ s}^{-1}\), comparable to the rate found above.
3D pathways in the search process. To examine this effect, we repeated the cleavage measurements on relaxed and extended DNA in buffers with different NaCl concentrations (Fig. 4). At 0 and 25 mM NaCl we observed almost invariable association rates for all DNA extensions, indicating a lowered probability of protein jumping. We found an optimal salt concentration for target finding at 60 mM, consistent with previous biochemical assays (18, 19). Apparently, at this salt concentration the effective sliding length \( l_{\text{eff}} \) approaches the rate of phosphodiester hydrolysis (0.4 s\(^{-1}\)) sliding diffusion along the DNA with diffusion-diffusivity \( D_{\text{sld}} \), until dissociation to the bulk (with rate \( k_{\text{off}}^{\text{ns}} \)) and (ii) 3D diffusion with diffusivity \( D_{\text{off}} \) where an enzyme eventually (re)binds nonspecifically to a DNA segment (with rate per length \( k_{\text{on}}^{\text{ns}} \)). Eventually, the target is found with rate \( k_{\text{on}} \).

Below we describe two limiting cases. If 3D diffusion were much faster than nonspecific binding, i.e., \( D_{\text{sld}} \gg k_{\text{on}}^{\text{ns}} \), an enzyme would perform a long diffusive volume excursion in the bulk before rebinding to the DNA, losing its correlation to the previous dissociation site. The search rate would then be \( k_{\text{on}} = 2 k_{\text{on}}^{\text{ns}} l_{\text{sl}} \), with sliding length \( l_{\text{sl}} = (D_{\text{sld}}/k_{\text{off}}^{\text{ns}})^{1/2} \). (11): typically only proteins binding nonspecifically within a distance \( l_{\text{sl}} \) from the target will actually find it. In this case the global conformation of the DNA would not influence the search dynamics, in contrast to what we observed for EcoRV.

Conversely, if \( D_{\text{sld}} \ll k_{\text{on}}^{\text{ns}} \) an enzyme just dissociated from the DNA is likely to rebind immediately, \( k_{\text{on}} \) will then be smaller than \( 2 k_{\text{on}}^{\text{ns}} l_{\text{sl}} \), as the enzyme will often slide along previously visited DNA segments, an inefficient oversampling (Fig. 1A). The effective sliding length \( l_{\text{eff}} \) is a measure of the distance, including intrasegmental (microscopic) hops, that the enzyme explores before effectively leaving a DNA segment. Experiments suggest that \( l_{\text{eff}} \) for EcoRV can be as large as hundreds of base pairs (18, 20).

Analysis by Zhou et al. (13) of these experimental data gives the maximal value of the ratio \( K_{\text{on}} = k_{\text{on}}^{\text{ns}}/(2\pi D_{\text{sld}})l_{\text{eff}}^{2} \) (36) as 5.2 \( \times \) 10\(^{3}\) M\(^{-1}\)bp\(^{-1}\). In terms of effective sliding this corresponds to \( l_{\text{eff}}^{2} = \sqrt{K_{\text{on}}(2\pi)} = 193 \) bp (using 1 bp = 0.34 nm). However, for such a \( l_{\text{eff}}^{2} \) one would expect a search rate \( \sim D_{\text{sld}} l_{\text{eff}}^{2} \) of order 10\(^{10}\) M\(^{-1}\)s\(^{-1}\), much faster than \( k_{\text{on}} = 10^{8} \) M\(^{-1}\)s\(^{-1}\) found experimentally. More precisely, using \( l_{\text{eff}}^{2} = 193 \) bp one can use the expression for \( k_{\text{on}} \) on infinitely long stretched DNA as discussed in ref. 36 to arrive at an expected \( k_{\text{on}} \approx 2 \times 10^{10} \) M\(^{-1}\)s\(^{-1}\) (a result that is essentially independent of the detailed choice of \( D_{\text{sld}}, k_{\text{on}}^{\text{ns}}, \) etc.). To explain the a priori unexpectedly low experimental rate we follow an observation by Erskine et al. (30): based on x-ray crystallography data, EcoRV may switch between an open state allowing for DNA binding, and a closed one that does not. The actual rate constant becomes \( k_{\text{on}} = x_{\text{act}} k_{\text{on}}^{\text{act}} x_{\text{off}} k_{\text{off}}^{\text{act}} \), where \( x_{\text{act}} \) and \( x_{\text{off}} \) being the fraction of open (and therefore active) EcoRV and \( k_{\text{on}}^{\text{act}} \) and \( k_{\text{off}}^{\text{act}} \) its association rate constant. Assuming that \( k_{\text{on}} = 1.5 \times 10^{8} \) M\(^{-1}\)s\(^{-1}\) for stretched DNA when \( l_{\text{eff}}^{2} = 193 \) bp and \( k_{\text{off}}^{\text{act}} = 2 \times 10^{10} \) M\(^{-1}\)s\(^{-1}\) this gives \( x_{\text{act}} \approx 0.75\% \). Physiologically, the advantage of this open/closed isomerization could be to provide the cell with a reservoir of EcoRV uniformly distributed in the cytoplasm, instead of being bound nonspecifically to the cell’s own DNA. Foreign DNA entering the cell would then immediately be surrounded by a higher concentration of EcoRV that, after switching to the open state, could readily attack the foreign DNA.

So why does coiled DNA enable faster target search than stretched DNA? An enzyme that would otherwise rebind quickly to the segment it just visited can instead be captured by another stretched DNA? An enzyme that would otherwise rebind quickly to the segment it just visited can instead be captured by another DNA segment. Experiments suggest that \( k_{\text{on}}^{\text{ns}} \) and \( k_{\text{off}}^{\text{ns}} \) are remote. Given an escape time from a segment of typical
order of 1 s (30), and with the estimate of 2 ms for the polymer configuration relaxation time (see Materials and Methods), one sees that the polymer configuration will change in between intersegmental jumps, such that successive intersegmental jumps are likely to occur at a point that also in 3D space is remote from the original segment. These considerations are quantified in SI Appendix, in which the work of Berg and Ehrenberg (36), where the DNA was treated as a straight cylinder, is generalized to the coiled configuration by including also foreign segments around the target site as randomly placed straight cylinders. This procedure allows us to estimate the probability of intrasegmental hops of different lengths being converted into intersegmental jumps, and how this affects the search rate $k_{on}$. In Table 1 we list values for the relative acceleration $R$ caused by coiling resulting from this theoretical modeling. The fitting parameters were chosen such that consistency is achieved simultaneously with both DNA extension and [NaCl] variation and with previously published values. The value of $R$ is found to rise above unity when $f_{DNAd}$ is comparable with or larger than the typical distance $f_{DNA}$ between DNA segments, where $f_{DNA}$ is the average density (length per volume) of foreign DNA segments around the specific target site. We interpret the high value of $R$ at 100 mM NaCl in terms of an increased $f_{DNA}$ because of a mutual attraction of DNA segments at these salt conditions. Such attraction was predicted by Lee et al. (37) from molecular dynamics simulations. Moreover, Qui et al. (38) experimentally demonstrated that such attraction exists between 25-bp DNA pieces above Mg$^{2+}$ concentrations of 16 and 10 mM in cases with, respectively, no monovalent salt and 20 mM NaCl present. In our experiments with 5 mM Mg$^{2+}$ we expected the onset of DNA attraction to occur between 20 and 100 mM NaCl. Table 1 reflects this expected increase of DNA density.

Previously, a 3-fold preference of cleaving supercoiled plasmids over relaxed plasmids was observed (17), in qualitative agreement with our results: whereas the density of DNA segments in supercoiled DNA is larger than in a relaxed coil, the density of DNA in the coiled state is larger than when it is held straight. Hence, the amount of reassociations to nearby DNA segments decreases from many to a few in the supercoiling experiment and from a few to zero in our experiment. Combining both studies thus demonstrates already a 6-fold enhancement of the searching rate of restriction enzymes caused by intrasegmental jumping. Considering that in the cell the density of DNA is even higher (because of crowding, condensing agents, and DNA-organizing proteins), we can conclude that in vivo the jumping pathway is an essential tool for efficiently targeting specific sites on DNA.

Materials and Methods

Experiments. EcoRV was recovered from ammonium sulfate precipitates as described (26). For the DNA cleavage experiments pCOS plasmid (6.5 kbp in length) was linearized by SpeI digestion (25). The linear DNA contains a single recognition site located almost in the middle. The degree of DNA coiling was controlled by changing the distance between the two attached beads.

Table 1. Fitting parameters and the corresponding theoretical value of $R$ (see SI Appendix) compared with the measurements

<table>
<thead>
<tr>
<th>[NaCl], mM</th>
<th>$k_{on}^{straight}$, (M$^{-1}$ s$^{-1}$)</th>
<th>$R_{DNAd}$, bp</th>
<th>$1/V_{DNAd}$, bp</th>
<th>$R_{theory}$</th>
<th>$R_{measured}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.8 $\times$ 10$^8$</td>
<td>97</td>
<td>518</td>
<td>1.11</td>
<td>1.3</td>
</tr>
<tr>
<td>25</td>
<td>1.0 $\times$ 10$^8$</td>
<td>123</td>
<td>485</td>
<td>1.15</td>
<td>1.1</td>
</tr>
<tr>
<td>100</td>
<td>1.0 $\times$ 10$^8$</td>
<td>122</td>
<td>120</td>
<td>1.63</td>
<td>1.7</td>
</tr>
<tr>
<td>150</td>
<td>0.09 $\times$ 10$^8$</td>
<td>9.1</td>
<td>80</td>
<td>1.10</td>
<td>1.3</td>
</tr>
</tbody>
</table>

The measured on rates for extended DNA, $k_{on}^{straight}$, are used as input for the theory to obtain the values of $f_{DNAd}$ assuming that the estimate $x_{RDNAd} = 0.75\%$ holds for all [NaCl].

Determination of Specific Association Rates. Specific association rates at the different DNA conformations were calculated from the average measured time required for cleavage on 30 DNA molecules at each DNA conformation for each NaCl concentration. For a one-step Poissonian process, the average of measured cleavage times equals the time constant of an exponential fit to the time-binned histogram. However, the cleavage of an individual DNA molecule in our experiments is effectively a multistep process, comprising both association to the specific site and DNA hydrolysis of both strands. At the used EcoRV concentration of 1 nM the association step is mostly rate-limiting (except at 60 mM NaCl). Yet, a histogram computed from the total of 385 events at 100 mM NaCl (shown in Fig. 6) still shows a lag phase for small cleavage times (with rate $0.4 \pm 0.1$ s$^{-1}$), owing to DNA hydrolysis (both strands), followed by an exponential drop-off. Consequently, by computing the rate simply from the averages of measured cleavage times, one would underestimate the actual association rate. We therefore corrected for DNA hydrolysis by subtracting the average hydrolysis time (2.5 s; see Fig. 3) from the statistical average of the measured cleavage time. The association rate is the inverse of this number (see SI Appendix for details).

Relaxation Time of DNA Fluctuations. The DNA between the beads undergoes constant thermal fluctuations. The transient stretching cycles could in principle lead to an enhanced mixing of DNA conformational states if the duration
of the stretch is shorter than the relaxation time $\tau_R$ of the coiled polymer. The latter can be calculated by using the Rouse model (40, 41):

$$\tau_R = \frac{2L^2\eta}{3\pi \ln(L/d)k_BT}.$$  

With a persistence length $\xi_0 = 50 \text{ nm}$, DNA diameter $d = 2 \text{ nm}$, contour length $L = 2.2 \mu\text{m}$, $\eta$ as the viscosity of the buffer ($\sim 10^{-3} \text{ Pa s}$), and $k_B$ as the thermal energy, we obtain $\tau_R = 2 \text{ ms}$. Compared with the stretching duration ($\sim 20 \text{ ms}$) this is fast enough to exclude any enhanced mixing.

**EcoRV Concentration.** We determined the average hydrolysis time for the cleavage of both strands to be 2.5 s (Fig. 3). To make association rate-limiting, a minimum cleavage time of $\sim 10 \text{ s}$ or more is required. With a specific association rate on plasmids of $10^8 \text{ M}^{-1} \text{ s}^{-1}$ (30), this requires an enzyme concentration $\leq 1 \text{ nM}$. Because at low [EcoRV] the specific association rate is proportional to the enzyme concentration, it is crucial that the number of enzymes in the flow chamber is constant in all measurements. To test which protein concentration still yields reliable results we performed cleaving experiments with progressively lower proteins concentrations. For each concentration we determined the average cleaving times for independent sets of trial experiments (i.e., a new protein dilution used in a cleaned or new flow chamber). These tests showed that for EcoRV concentrations $<0.7 \text{ nM}$ the average cleaving times differed by up to $\sim 50\%$ between data sets, presumably caused by proteins sticking to and detaching from the walls of the chamber and tubing. At concentrations $\geq 1 \text{ nM}$ such effects were never observed. Therefore, 1 nM EcoRV was used in the experiments.

**ACKNOWLEDGMENTS.** We thank Dave Hiller, John Perona, Tobias Ambjörnsen, Alexander Grosberg, and Roland Winkler for helpful discussions. This work was supported by a Fundamenteeel Onderzoek der Materie Projectruimte Grant, a Netherlands Organization for Scientific Research-Vernieuwings-impuls grant (to J.G.W.L.), and the Villum Kann Rasmussen Foundation (M.A.L.).
Supplementary Information

In this part we discuss the incorporation of 3D exchange events in an extended model for facilitated diffusion. The 3D exchange comprises intrasegmental hopping, intersegmental jumping, and long-ranging volume exchange. Our description of the target search process is based on the enzyme density $n(x, t)$ per DNA length, where $x$ is the distance along the DNA contour (chemical distance).

We include (compare Figure S1 in this note) 1D sliding along the DNA with diffusion constant $D_{1d}$, enzyme dissociation with rate $k_{\text{off}}$, and (re)adsorption after diffusion through the bulk. The dynamics of $n(x, t)$ is thus governed by the balance equation [1]

$$\frac{\partial n(x, t)}{\partial t} = \left( D_{1d} \frac{\partial^2}{\partial x^2} - k_{\text{off}}^{\text{ns}} \right) n(x, t) + k_{\text{on}}^{\text{ns}} \int_{-\infty}^{\infty} W_{\text{bulk}}(x-x', t-t') n(x', t') dt' dx' + G(x, t) - j(t) \delta(x). \quad (1)$$

Here, $j(t)$ denotes the flux into the target located at $x = 0$ under the condition that the target is fully absorbing, $n(x = 0, t) = 0$. $W_{\text{bulk}}(x-x', t-t')$ is the probability that an enzyme after a bulk excursion binds non-specifically to the point $x$ at time $t$ after unbinding from the DNA at position $x'$ at time $t'$. $G(x, t)$ is the flux onto the DNA of enzymes that have not previously been bound to the DNA (up to time $t$).

At sufficiently long times, a steady state exists, and the overall particle density on the DNA will reach a constant value $n_{\text{eq}}^{\text{ns}}$. The flux into the target site is then given by the steady state value [2]

$$j_{\text{stat}} \sim k_{1d} r_{\text{eq}}^{\text{ns}}, \quad (2)$$

where we have introduced a 1D rate constant given by

$$k_{1d}^{-1} = \int_{-\infty}^{\infty} dq \frac{1}{2\pi D_{1d} q^2 + k_{\text{off}}^{\text{ns}}[1 - W_{\text{bulk}}(q, u = 0)]}, \quad (3)$$

and $W_{\text{bulk}}(q, u)$ is the Fourier-Laplace transform of $W_{\text{bulk}}(x, t)$. The target association rate constant $k_{\text{on}}^{\text{ns}}$ is obtained in the steady state by dividing by the density of unbound enzymes in the bulk

$$k_{\text{on}} = j_{\text{stat}} / n_{\text{bulk}} = k_{1d} K_{\text{ns}}, \quad (4)$$

where $K_{\text{ns}}$ is the binding constant for non-specific binding to DNA per length of DNA. The quantity $W_{\text{bulk}}(x, u = 0)$ denotes the distribution of jump lengths along the DNA mediated by a single volume excursion. We calculate $W_{\text{bulk}}$ in two limiting cases, assuming only a fraction $x_{\text{act}}$ of enzymes is ready to bind, see main text:

(i) Straight DNA configuration

DNA in a straight configuration can be considered as a cylinder of radius $r_{\text{int}}$ (the range of the non-specific interaction) with non-specific reaction rate $k_{\text{on}}^{\text{ns}}$ at the boundary such that the flux of enzymes per length of the DNA onto the DNA is $k_{\text{on}}^{\text{ns}}$ times the bulk concentration of active (open) enzymes next to the DNA. The relation of the on-off rates with the binding constant is $K_{\text{ns}} = x_{\text{act}} k_{\text{on}}^{\text{ns}} / k_{\text{off}}^{\text{ns}}$, where $x_{\text{act}}$ is the fraction of active enzymes. We assume that the switching between the active and dormant (closed) configuration is sufficiently slow not to alter the dynamic considerations below. When in the bulk the enzymes diffuse with diffusion constant $D_{1d}$ while the DNA is treated as fixed in space. Apart from the introduction of $x_{\text{act}}$ then our model reduces for the straight configuration to the classical result from Ref. [3]. However, our approach allows for generalization to a coiled DNA configuration.

(ii) Random DNA configuration with persistence length

To describe enzyme diffusion in the bulk when the DNA assumes a random configuration we use a simple model where the local segment the enzyme has recently unbound from is locally viewed as a straight cylinder. The other segments are then included using a superposition technique by considering separately the problem of capture by each foreign segment. The processes are combined by choosing the realization where the binding happens first. To do this we make the approximation that the capture processes are independent. This method is equivalent to the method used by von Smoluchowski to solve the problem of reaction with spheres [4].

![Figure S1: Schematic of the search mechanisms of Eq. (1): Enzymes can perform 1D motion along the DNA with diffusivity $D_{1d}$, unbind with rate $k_{\text{off}}^{\text{ns}}$, perform bulk diffusion with diffusivity $D_{3d}$, and rebind to the DNA. The rebinding is governed by the transfer kernel $W_{\text{bulk}}(x, t).$](image)
Consider first an enzyme diffusing within a large volume \( V \) in which a straight piece of DNA with length \( L \) is situated randomly. If we denote the probability that the enzyme has bound to the DNA before time \( t \) by \( J_{\text{single}}(t) \), then the probability that it has not bound to any of \( N \) pieces of DNA is (independence assumption)

\[
P_{\text{foreign}}(t) = (1 - J_{\text{single}}(t))^N \tag{5}.
\]

Taking the limit of \( V, N, L \to \infty \) and fixing \( l_{\text{DNA}} = NL/V \) (\( l_{\text{DNA}} \) is the density (length per volume) of DNA locally around the original segment) we obtain

\[
P_{\text{foreign}}(t) = \exp[-J_{\text{cap}}(t)]. \tag{6}
\]

\( J_{\text{cap}} \) can be found by solving the 2D problem of enzymes reacting with a circular target with reaction rate constant \( k_{\text{on}}/(2\pi l_{\text{int}}) \) per length at its boundary. In this problem the initial density of enzymes in the surrounding infinite space is taken to be uniform and equal to the density of length, \( l_{\text{DNA}} \), of foreign DNA segments in the original problem. The resulting equations can be evaluated numerically. An enzyme captured by a foreign segment is assumed to be sufficiently far away in chemical distance along the DNA contour after this intersegmental jump (due to the large persistence length) to avoid overlap in the 1D motion before and after the jump [2].

Note that we assume that the density of ‘other segments’ is uniform everywhere. Although this is not strictly true, and \( l_{\text{DNA}} \) actually fluctuates, on average this should be a fair approximation, since an enzyme diffusing further away than a distance where the non-uniformity of DNA segments comes into play will be likely to perform a ‘very long jump’ anyway. Again it doesn’t matter how exactly we bookkeep such long jumps, since the important distinction is with regards to the overlap between the 1D motions before and after a 3D exchange.

EcoRV

We are now ready to fit the model to our data obtained for EcoRV. To do this we will have to estimate some of the parameters. For the bulk diffusion constant \( D_{3d} \) we will use the Einstein-Stokes formula for an object with radius 2.5 nm. This gives

\[
D_{3d} = \frac{k_B T}{6\pi \eta \times 2.5 \text{nm}} \approx 8 \times 10^8 \text{bp}^2/\text{s}. \tag{7}
\]

The radius of non-specific interaction \( r_{\text{int}} \) will be chosen to be the radius of DNA (approx. 1 nm) plus the radius of EcoRV (roughly 2.5 nm), giving \( r_{\text{int}} = 3.5 \text{nm} \) (for conversion to bp (basepairs) we use 1 bp = 0.34 nm).

The sliding length \( l_d \) we take to decrease logarithmically with salt (Debye screening \( \kappa \)), \( l_d = -l_{d,0} \log(b_{\text{salt}} \kappa) \), with \( l_{d,0} = 25 \text{bp} \) and \( b_{\text{salt}} = 0.5 \text{nm} \). The value of \( k_{\text{on}}^{\text{na}} \) has been calculated numerically by choosing the value that gives the experimentally measured \( k_{\text{on}} \) for the straight conformation.

The choice of the two parameters \( l_d \) and \( r_{\text{int}} \) does not influence the value \( R \) of the relative rate increase in the coiled configuration much. The dependence on these two parameters mostly comes in through the effective sliding length, which is adjusted through fitting of \( k_{\text{on}}^{\text{na}} \). When estimating the density \( l_{\text{DNA}} \) it is important to distinguish between two cases, whether the divalent salt induces a short range attraction between DNA segments or not. Following the arguments in the main text we assume that no short range attraction occurs at 0 mM and 25 mM salt. The estimate of \( l_{\text{DNA}} \) is then achieved through the probability density within the Worm Like Chain model that a point a contour distance \( s \) away has looped back. According to [5] this cyclization probability is well approximated by

\[
j_M = \left( \frac{3}{4\pi s l_p} \right)^{3/2} \exp \left[ -\frac{8 l_p^2}{|s|^2} \right]. \tag{8}
\]

where we take the persistence length \( l_p \) to be the same function of ionic strength as found in [6] for monovalent salt. If the target is situated in the middle of a chain with length \( L = 6538 \text{ bp} \) we will have a density of DNA around the target:

\[
l_{\text{DNA}} = \int_{-L/2}^{L/2} ds j_M. \tag{9}
\]

We do not know of any theory we can use to estimate \( l_{\text{DNA}} \) in the case of attraction between the DNA segments. Instead we have gauged roughly what likely values of \( l_{\text{DNA}} \) are based on the present theory and the experimentally found values of \( R \) at 100 and 150 mM NaCl. Finally we have used the estimate of the main text: \( x_{\text{act}} = 0.75\% \). The values of \( R \) obtained for these parameters are listed in Table 1 in the main text as \( R_{\text{theory}} \).

[1] For identical enzymes, their mutual avoidance is actually included in Eq. (1), as on encounter it does not matter whether they deflect each other or swap identities (I. M. Sokolov et al., Phys. Rev. E 72, 041102 (2005)).
Supplementary note on the different timescales involved in the experiment

In the cleavage experiments several processes are taking place simultaneously or sequentially. Here we elaborate on how these processes are connected and how we can separate out the association rate.

A DNA molecule, tethered between two optically trapped beads, undergoes constant conformational changes due to Brownian motion. The relaxation time of the polymer is on the order of a millisecond (see Materials and Methods). The sliding interaction time of an EcoRV protein associated to a non-specific DNA segment typically is 0.1 to 1 second (1). When it rebinds to another nearby segment (in coiled DNA), the polymer has thus completely rearranged itself in the meantime, which reduces the possibility that this segment has been visited before. This process speeds up target search.

The enzymatic cleavage reaction itself consists of consecutive reaction steps: association, induced-fit reaction, chemical breakage of two phosphodiester bonds in the two backbones of the DNA, and release of the cleaved DNA parts. In our experiments we are interested in the association rate, i.e. the (inverse of the) mean time it takes to localize the single specific target sequence on our DNA molecule. At low enzyme concentration this process is diffusion-limited and depends linearly on enzyme concentration. At the used concentration of 1 nM the association time typically is 6-10 seconds. The induced-fit reaction is much faster and takes place at the millisecond timescale. Phosphodiester hydrolysis takes another 1.7 seconds per DNA strand (0.6 s\(^{-1}\)). As the EcoRV dimer contains two active sites that act independently, the two hydrolysis reactions occur simultaneously, leading to an average DNA cutting time (both DNA strands) of 2.5 seconds (0.4 s\(^{-1}\)). Finally, the dissociation of the enzyme from the cleaved DNA is fast in our setup. (In biochemical experiments this is often the rate-limiting step, but we have shown previously that even small forces (~ 1 pN) are sufficient to pull the complex apart (2). The periodic stretching every second ensures a quick product release.)

In our experiments we measure the time it takes to complete all the reaction steps described above. Overall the cutting time of a single DNA molecule is rate-limited by diffusion to the specific site, i.e. the association step. However, the second slowest step, DNA backbone hydrolysis, still contributes to the total reaction time. Because the latter is independent of DNA configuration, and only occurs after association, on average the total cleavage time will be 2.5 seconds longer than the association time. We can thus obtain the association time by subtracting this number from the measured average cleavage time for each DNA configuration.

We can only detect cleavage of a single DNA molecule by stretching it and measuring its force response. In our setup we periodically stretch the DNA every second. An intact DNA results in a force spike; when the DNA is cut, the force remains zero. Therefore the precision of measuring the cleavage time of a single DNA molecule cannot be greater than 1 second. However, because the frequency of stretching is (much) faster than the cleavage time, errors for individual molecules average out for large number of measurements. The cleavage time can thus still be determined with higher accuracy.


Supporting Information

van den Broek et al. 10.1073/pnas.0804248105

Movie S1. Rapid periodic stretching of a lambda phage DNA molecule between two optically trapped beads, as used for cleavage detection of DNA in nonstretched configurations.

Other Supporting Information Files

SI Appendix (PDF)