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 specific 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

An 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 $k_{\text{on}}$ 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. 1 C). 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 pCco5 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 fluo-
would form a coil with a mean square end-to-end distance of 0.24\( \mu \text{m}^2 \). A relaxed pCco5 DNA molecule (6,538 bp) with free ends at smaller extensions, a decrease in the rate occurs. Why would the association rate on DNA held in an extended configuration be as slow as on coiled DNA. We define \( R \) as the ratio between the maximal association rate and the rate found on stretched DNA. The fastest association is measured at \( a = 0.24 \) (\( R = 1.7 \pm 0.3 \)). At smaller extensions, a decrease in the rate occurs. Why would the search rate acceleration peak at a fractional extension of 0.24? A relaxed pCco5 DNA molecule (6,538 bp) with free ends would form a coil with a mean square end-to-end distance (\( R_e^2 \)) \( \approx 2L_p = (0.49 \mu \text{m})^2 \) (29), corresponding to \( a = 0.22 \). 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 (Fig. 1A and B). At \( a \approx 0.2 \), 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 \( l_{DNA} \) 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, \( 1.8 \pm 0.2 \times 10^8 \text{M}^{-1}\text{s}^{-1} \), are in close agreement with values found earlier in bulk experiments in similar buffers: \( 1.2-2.2 \times 10^8 \text{M}^{-1}\text{s}^{-1} \) (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 3D DNA conformational freedom of the DNA at these extensions can be divided into three regimes: (entropically) stretched DNA (Top), globular DNA (relaxed coil) (Middle), and a DNA coil that is squeezed between the beads. (Bottom)

**Rate Acceleration on Coiled DNA.** Specific association rates \( k_{on} \) were computed from 385 individual cleavage reactions (see Materials and Methods). Variation of the DNA conformation includes three regimes (see Fig. 1C): (i) extended DNA, where jumping to different DNA segments cannot take place (\( \approx 1 \text{pN tension, } \approx 85\% \) extended); (ii) a DNA configuration that is close to a relaxed coil (fractional extension \( a \) between 0.2 and 0.3); and (iii) smaller fractional extensions, where the beads interfere with the coil and the DNA is slightly squeezed out of the narrow gap. Fig. 4 shows the association rates acquired for different DNA conformations. Strikingly, at 100 mM NaCl the specific association rate on DNA held in an extended configuration is almost twice as slow as on coiled DNA. We define \( R \) as the ratio between the maximal association rate and the rate found on stretched DNA. The fastest association is measured at \( a = 0.24 \) (\( R = 1.7 \pm 0.3 \)). At smaller extensions, a decrease in the rate occurs. Why would the search rate acceleration peak at a fractional extension of 0.24? A relaxed pCco5 DNA molecule (6,538 bp) with free ends would form a coil with a mean square end-to-end distance (\( R_e^2 \)) \( \approx 2L_p = (0.49 \mu \text{m})^2 \) (29), corresponding to \( a = 0.22 \). 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 (Fig. 1A and B). At \( a \approx 0.2 \), 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 \( l_{DNA} \) 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, \( 1.8 \pm 0.2 \times 10^8 \text{M}^{-1}\text{s}^{-1} \), are in close agreement with values found earlier in bulk experiments in similar buffers: \( 1.2-2.2 \times 10^8 \text{M}^{-1}\text{s}^{-1} \) (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 3D DNA conformational freedom of the DNA at these extensions can be divided into three regimes: (entropically) stretched DNA (Top), globular DNA (relaxed coil) (Middle), and a DNA coil that is squeezed between the beads. (Bottom)
explain the enzyme’s acceleration of target site association, we search rate would then be diffusive volume excursion in the bulk before rebinding to the recognition site. At lower salt conditions (0 and 25 mM) this effect disappears. (Inset) Association rates in 150 mM NaCl are ~10-fold lower for all DNA extensions. Each data point consists of a minimum of 30 measurements.

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 $l_{\text{eff}}$ (the average length of correlated motion along the contour of the DNA, including both hopping and sliding) is maximal. However, we could not probe the difference between the stretched and coiled DNA configuration at 60 mM NaCl. At that concentration the EcoRV association rate approaches the rate of phosphodiester DNA hydrolysis (0.4 s$^{-1}$), causing the reaction to be rate-limited by this process instead of by diffusion. Using a lower EcoRV concentration could diminish this problem, but as mentioned before we detected fluctuations in the actual amount of enzymes present in the flow chamber if we decreased the EcoRV concentration <1 nM. In reaction buffers containing 150 and 200 mM NaCl, target site location rates were dramatically lowered for all DNA configurations, probably attributable to the reduced affinity of EcoRV for both specific and nonspecific DNA in these buffers (35) (Figs. 4 Inset and 5). Here, the sliding length became so small that the enzyme effectively performed a 3D random walk and facilitation by nonspecific binding was absent.

**Theoretical Modeling of Intersegmental Jumping.** To quantitatively explain the enzyme’s acceleration of target site association, we consider EcoRV’s search for the specific cleavage site as a combination of (i) sliding diffusion along the DNA with diffusivity $D_{1d}$ until dissociation to the bulk (with rate $k_{\text{off}}^{\text{ns}}$) and (ii) 3D diffusion with diffusivity $D_{3d}$, where an enzyme eventually 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., $D_{3d} \gg k_{\text{on}}^{\text{ns}}$ ($k_{\text{on}}^{\text{ns}}$ has units of inverse concentration per length per time, i.e., length squared per time like $D_{3d}$), an enzyme would perform a long diffusive volume excursion in the bulk before rejoining to the DNA, losing its correlation to the previous dissociation site. The search rate would then be $k_{\text{on}} = 2k_{\text{on}}^{\text{ns}}l_{\text{sl}}$, with sliding length $l_{\text{sl}} = (D_{1d}l_{\text{sl}}^{\text{ns}})^{1/2}$ (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 $D_{3d} \ll k_{\text{on}}^{\text{ns}}$ an enzyme just dissociated from the DNA is likely to rebind immediately, $k_{\text{on}}$ will then be smaller than $2k_{\text{on}}^{\text{ns}}l_{\text{sl}}$, as the enzyme will often slide along previously visited DNA segments, an inefficient oversampling (Fig. 1A). The effective sliding length $l_{\text{eff}}^{\text{sl}} = [k_{\text{on}}^{\text{ns}}/(2\pi D_{3d})]^{1/2}$ (36) is a measure of the distance, including intrasegmental (microscopic) hops, that the protein explores before effectively leaving a DNA segment. Experiments suggest that $l_{\text{eff}}^{\text{sl}}$ 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 $K_{\text{sl}} = D_{1d}k_{\text{on}}^{\text{ns}}/(D_{3d}k_{\text{off}}^{\text{ns}})$ as $5.2 \times 10^3$ M$^{-1}$bp$^{-1}$. In terms of effective sliding this corresponds to $l_{\text{eff}}^{\text{sl}} = \sqrt{K_{\text{sl}}/(2\pi)} = 193$ bp (using 1 bp = 0.34 nm). However, for such a $l_{\text{eff}}^{\text{sl}}$ one would expect a search rate $D_{\text{sl}}l_{\text{eff}}^{\text{sl}}$ 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 $D_{\text{sl}} = 193$ 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 $k_{\text{on}} = 2 \times 10^{10}$ M$^{-1}$s$^{-1}$ (a result that is essentially independent of the detailed choice of $D_{1d}$, $k_{\text{on}}^{\text{ns}}$, 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 $k_{\text{on}} = x_{\text{act}}k_{\text{act}}$, $x_{\text{act}}$ being the fraction of open (and therefore active) EcoRV and $k_{\text{act}}$ its association rate constant. Assuming that $k_{\text{on}} = 1.5 \times 10^8$ M$^{-1}$s$^{-1}$ for stretched DNA when $l_{\text{eff}}^{\text{sl}} = 193$ bp and $k_{\text{act}} = 2 \times 10^8$ 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 segment, that is nearby (in 3D space) because of DNA looping (Fig. 1B). This conversion from an intrasegmental hop to an intersegmental jump reduces oversampling, because in chemical distance along the DNA the dissociation and rebinding points 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 $[\text{NaCl}]$ variation and with previously published values. The value of $R$ is found to rise above unity when $f_{DNAA}^{eff}$ is comparable with or larger than the typical distance $d_{DNAA}$ between DNA segments, where $d_{DNAA}$ 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 $d_{DNAA}$ 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 pC05 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](image)

<table>
<thead>
<tr>
<th>[NaCl], mM</th>
<th>$k_{on}^{straight}$, (M$^{-1}$s$^{-1}$)</th>
<th>$f_{DNAA}^{eff}$, bp</th>
<th>$1/V_{DNAA}$</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>$0.9 \times 10^8$</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 $f_{DNAA}^{eff}$ assuming that the estimate $x_{an} = 0.75\%$ holds for all $[\text{NaCl}]$.

### 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 configuration 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 rates $0.4 s^{-1}$), 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):

\[
\tau_R = \frac{2L^2\eta}{3\pi \ln(L/d)k_B T}.
\]

With a persistence length \( \xi = 50 \text{ nm} \), 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 \( k_B T \) 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 plasmons of \( \sim 10^8 \text{ M}^{-1} \text{ s}^{-1} \), this requires an enzyme concentration \( \lesssim 1 \text{ nM} \). 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 cleavage times differed by up to \( \pm 50\% \) between data sets, presumably caused by proteins sticking to and detaching from the walls of the chamber and tubing. At concentrations \( \lesssim 1 \text{ nM} \) such effects were never observed. Therefore, 1 nM EcoRV was used in the experiments.

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