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_{off}^{ns}, and (re)adsorption after diffusion through the bulk. The dynamics of $n(x,t)$ is thus governed by the balance equation [1]

$$\begin{aligned}\frac{\partial n(x,t)}{\partial t} &= \left(D_{1d}\frac{\partial^2}{\partial x^2} - k_{off}^{ns}\right) n(x,t) \\ &+k_{off}^{ns}\int_{-\infty}^{\infty}\int_{0}^{t} W_{bulk}(x-x',t-t')n(x',t')dt'dx' \\ &+G(x,t) - j(t)\delta(x). \end{aligned} \quad (1)$$

Here, $j(t)$ denotes the flux into the target located at $x = 0$ under the condition that the target is fully absorbing, $n(x=0,t)=0$. $W_{bulk}(x-x',t-t')$ 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_{eq}^{ns}. The flux into the target site is then given by

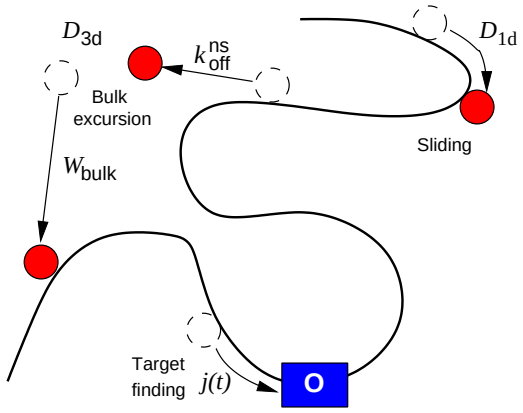

Figure S1: Schematic of the search mechanisms of Eq. (1): Enzymes can perform 1D motion along the DNA with diffusivity D_{1d}, unbind with rate k_{off}^{ns}, perform bulk diffusion with diffusivity D_{3d}, and rebind to the DNA. The rebinding is governed by the transfer kernel $W_{bulk}(x,t)$.

the steady state value [2]

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

where we have introduced a 1D rate constant given by

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

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

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

where K_{ns} is the binding constant for non-specific binding to DNA per length of DNA. The quantity $W_{bulk}(x,u=0)$ denotes the distribution of jump lengths along the DNA mediated by a single volume excursion. We calculate W_{bulk} in two limiting cases, assuming only a fraction x_{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_{int} (the range of the non-specific interaction) with non-specific reaction rate k_{on}^{ns} at the boundary such that the flux of enzymes per length of the DNA onto the DNA is k_{on}^{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 $K_{ns} = x_{act}k_{on}^{ns}/k_{off}^{ns}$, where x_{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_{3d} while the DNA is treated as fixed in space. Apart from the introduction of x_{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 Smoluchowsky to solve the problem of reaction with spheres [4].

Consider first an enzyme diffusing within a large volume V in which a straight piece of DNA with length L situated randomly. If we denote the probability that the enzyme has bound to the DNA before time t by $J_{\rm single}(t)$, then the probability that it has not bound to any of N pieces of DNA is (independence assumption)

$$P_{\rm survival}^{\rm foreign}(t) = (1 - J_{\rm single}(t))^N \ . \qquad (5)$$

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

$$P_{\rm survival}^{\rm foreign}(t) = \exp\left[-J_{\rm cap}(t)\right] . \qquad (6)$$

$J_{\rm cap}$ can be found by solving the 2D problem of enzymes reacting with a circular target with reaction rate constant $k_{\rm on}^{\rm ns}/(2\pi r_{\rm int})$ 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, $l_{\rm DNA}$, 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 $l_{\rm DNA}$ 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 $D_{\rm 3d}$ we will use the Einstein-Stokes formula for an object with radius 2.5 nm. This gives

$$D_{\rm 3d} = \frac{k_{\rm B}T}{6\pi\eta \times 2.5\,{\rm nm}} \sim 8 \times 10^8\,{\rm bp}^2/{\rm s} \, . \qquad (7)$$

The radius of non-specific interaction $r_{\rm int}$ will be chosen to be the radius of DNA (approx. 1 nm) plus the radius of EcoRV (roughly 2.5 nm), giving $r_{\rm int} = 3.5\,{\rm nm}$ (for conversion to bp (basepairs) we use $1\,{\rm bp} = 0.34{\rm nm}$). The sliding length $l_{\rm sl}$ we take to decrease logarithmicly with salt (Debye screening κ), $l_{\rm sl} = -l_{\rm sl,0}\log(b_{\rm sl}\kappa)$, with $l_{\rm sl,0} = 25\,{\rm bp}$ and $b_{\rm sl} = 0.5\,{\rm nm}$. The value of $k_{\rm on}^{\rm ns}$ has been calculated numerically by choosing the value that gives the experimentally measured $k_{\rm on}$ for the straight conformation. The choice of the two parameters $l_{\rm sl}$ and $r_{\rm int}$ 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 $k_{\rm on}^{\rm ns}$. When estimating the density $l_{\rm DNA}$ 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 $l_{\rm DNA}$ 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

$$j_{\rm M} = \left(\frac{3}{4\pi|s|\ell_{\rm p}}\right)^{3/2} \exp\left[-\frac{8\ell_{\rm p}^2}{|s|^2}\right] \ . \qquad (8)$$

where we take the persistence length $\ell_{\rm p}$ 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 ($L = 6538\,{\rm bp}$) we will have a density of DNA around the target:

$$l_{\rm DNA} = \int_{-L/2}^{L/2} ds \; j_{\rm M} \, . \qquad (9)$$

We do not know of any theory we can use to estimate $l_{\rm DNA}$ in the case of attraction between the DNA segments. Instead we have gauged roughly what likely values of $l_{\rm DNA}$ 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: $x_{\rm act} = 0.75\%$. The values of R obtained for these parameters are listed in Table 1 in the main text as $R_{\rm theory}$.

[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)).
[2] M. A. Lomholt et al, forthcoming paper
[3] O. G. Berg and M. Ehrenberg, Biophys. Chem. **15**, 41 (1982).
[4] M. von Smoluchowski, Z. physikalische Chemie **92**, 129 (1917).
[5] L. Ringrose et al., EMBO J. **18**, 6630 (1999).
[6] G. S. Manning, Biophys. J. **91**, 3607 (2006).

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.

1. Erskine, SG, Baldwin, GS & Halford, SE (1997) Rapid-reaction analysis of plasmid DNA cleavage by the EcoRV restriction endonuclease. *Biochemistry* 36:7567-76.
2. van den Broek, B, Noom, MC & Wuite, GJL (2005) DNA-tension dependence of restriction enzyme activity reveals mechanochemical properties of the reaction pathway. *Nucleic Acids Research* 33:2676-2684.