Molecular search in gene regulation

— Houston, 5th December 2019 —

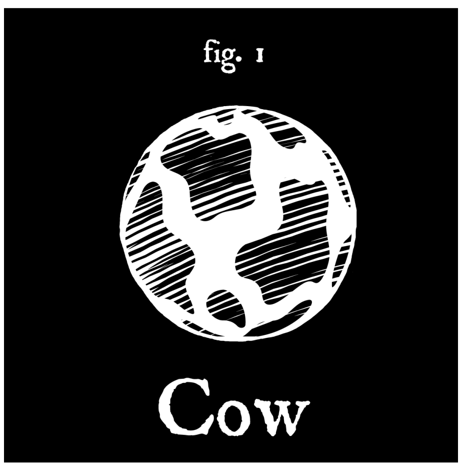
Main protagonist: bacteria cells such as E.coli

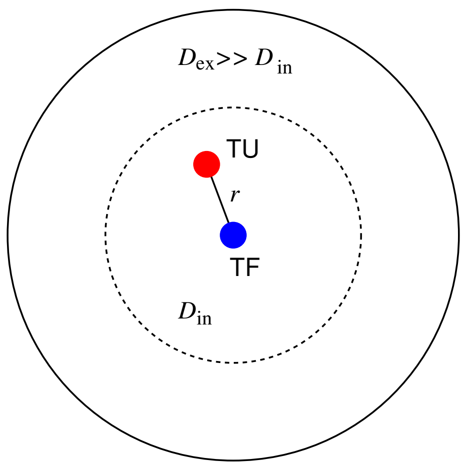
Cell size: roughly $2\mu\text{m} \times 1/2\mu\text{m}$
Cell volume: $\approx 1\mu\text{m}^3$

DNA length: 4.7×10^6 base pairs or $\approx 1.6\text{mm}$

Number of proteins in cell: $\approx 2.4 \times 10^6$
Different proteins (# genes): 4,300

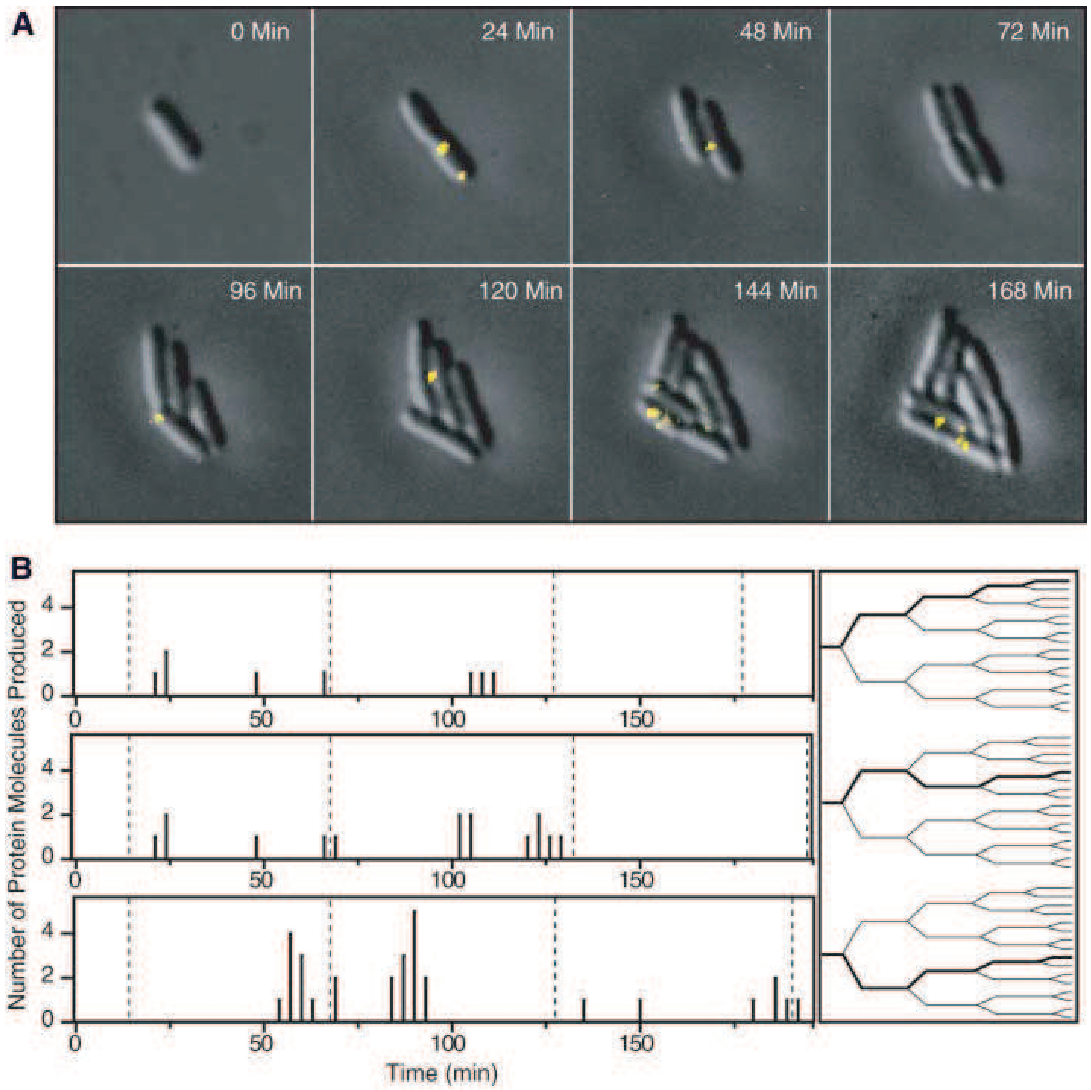
Some proteins occur only as few or few tens of copies/cell (nM concentrations)





RW Corbin et al, PNAS (2003)

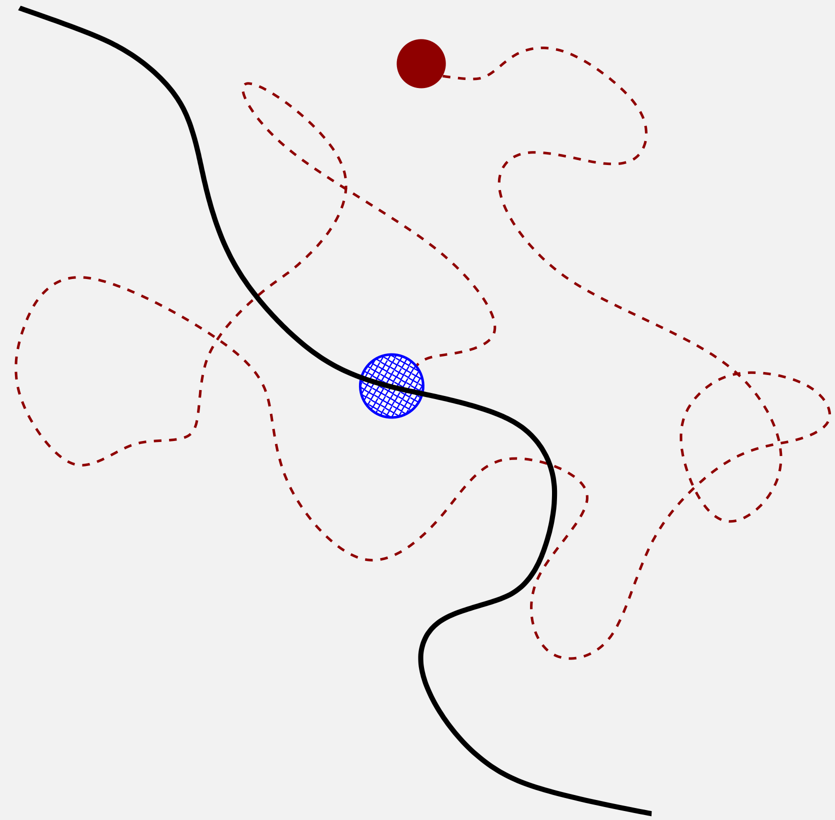
Gene expression one molecule at a time



synthesised proteins (bursty) along three cell lineages, dashed lines marking cell divisions

Yu et al, Science (2006); I Golding et al, Cell (2005)

Smoluchowski search picture



Search rate for a particle with diffusivity D_{3d} to find an immobile target of radius a (assuming immediate binding):

$$k_{\mathrm{on}}^{S} = 4\pi D_{\mathrm{3d}} a$$

Protein-DNA interaction: $a \approx \{\text{few bp}\} \approx 1\text{nm}$
$D_{\mathrm{3d}} \approx 10\mu\mathrm{m}^2/\mathrm{sec}$ (typically $\varnothing_{\mathrm{TF}} \approx 5\text{nm}$):

$$k_{\mathrm{on}}^{S} \approx \frac{10^8}{(\mathrm{mol/l}) \times \mathrm{sec}}$$

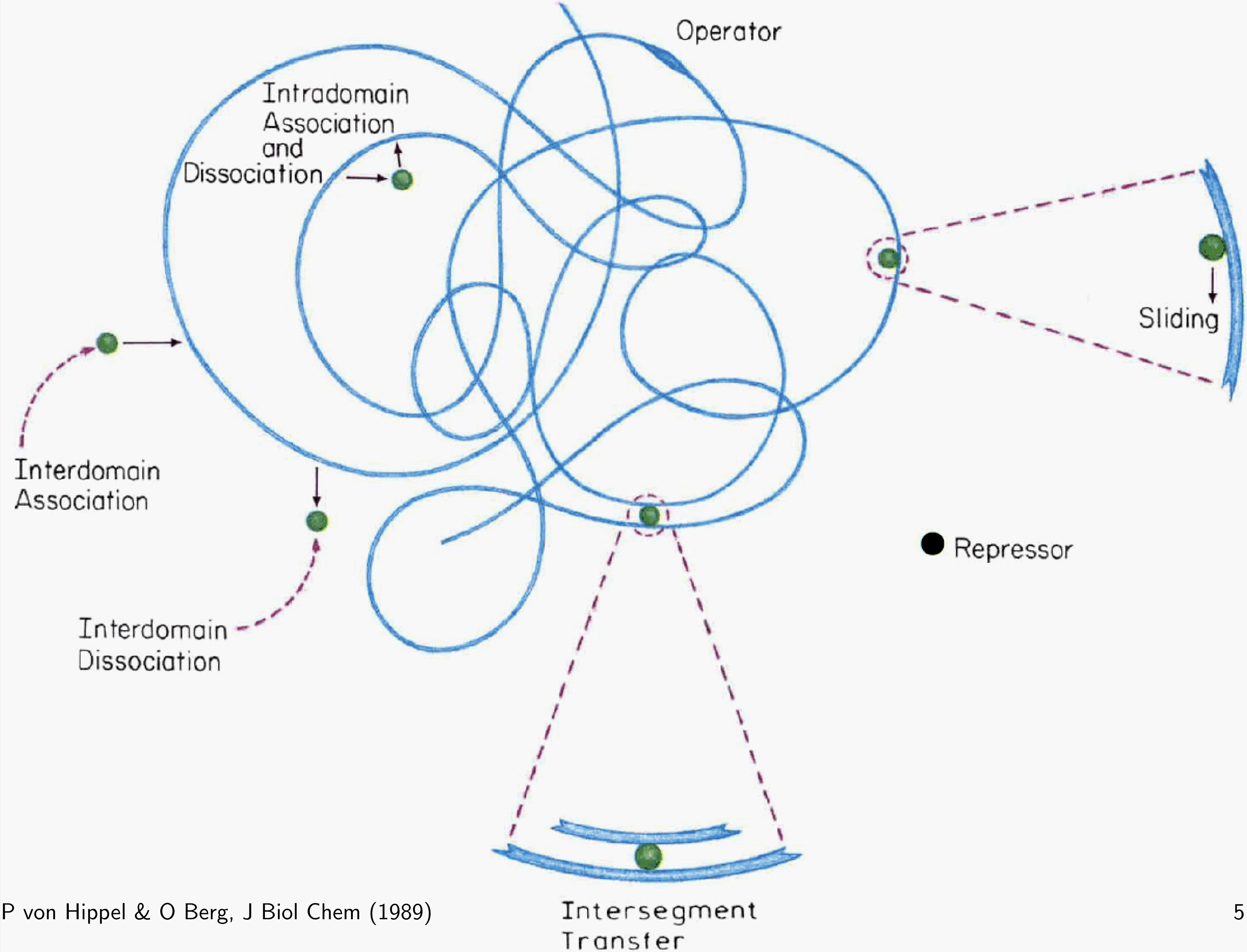
Lac repressor [AD Riggs, S Bourgeois, M Cohn, J Mol Biol 53, 401 (1970)]:

$$k_{\mathrm{on}} \approx \frac{10^{10}}{(\mathrm{mol}/l) \times \mathrm{sec}}$$

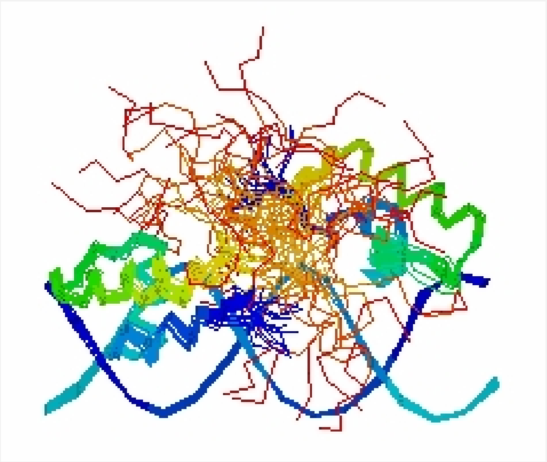
→ Facilitated diffusion picture

M v Smoluchowski, Physikal. Zeitschr. (1916); P von Hippel and O Berg, J Biol Chem (1989)

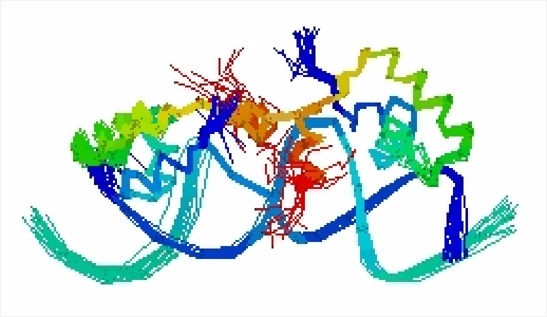
Facilitated diffusion: the Berg-von Hippel model



Non-specific binding energy based on *in vivo* data



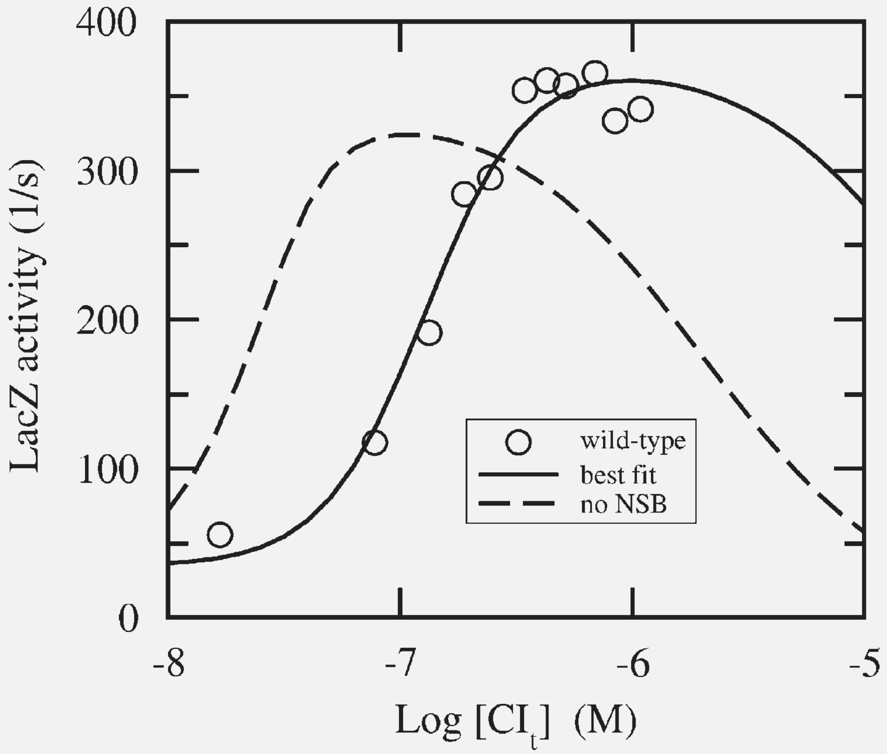
Lac repressor, nonspecific binding



Lac repressor, specific binding

$$[X] = [X_{free}] + \left[X_{@O_P}\right] + [X_{NSB}]$$

$$\Delta G_{NSB}(CI) = -4.1 \pm 0.9\,\text{kcal/mol},$$
$$\Delta G_{NSB}(Cro) = -4.2 \pm 0.8\,\text{kcal/mol}$$

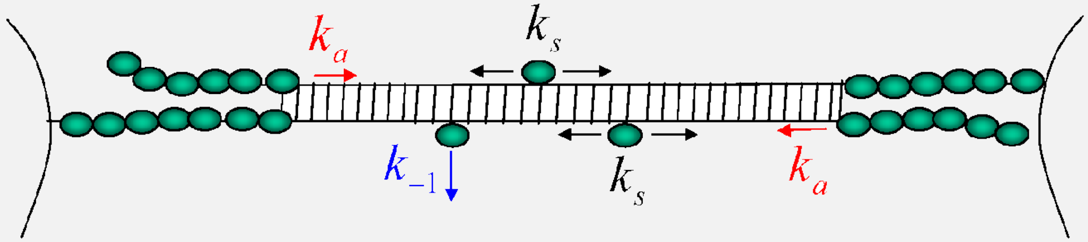


A Bakk & RM, FEBS Lett (2004); J Theoret Biol (2004)

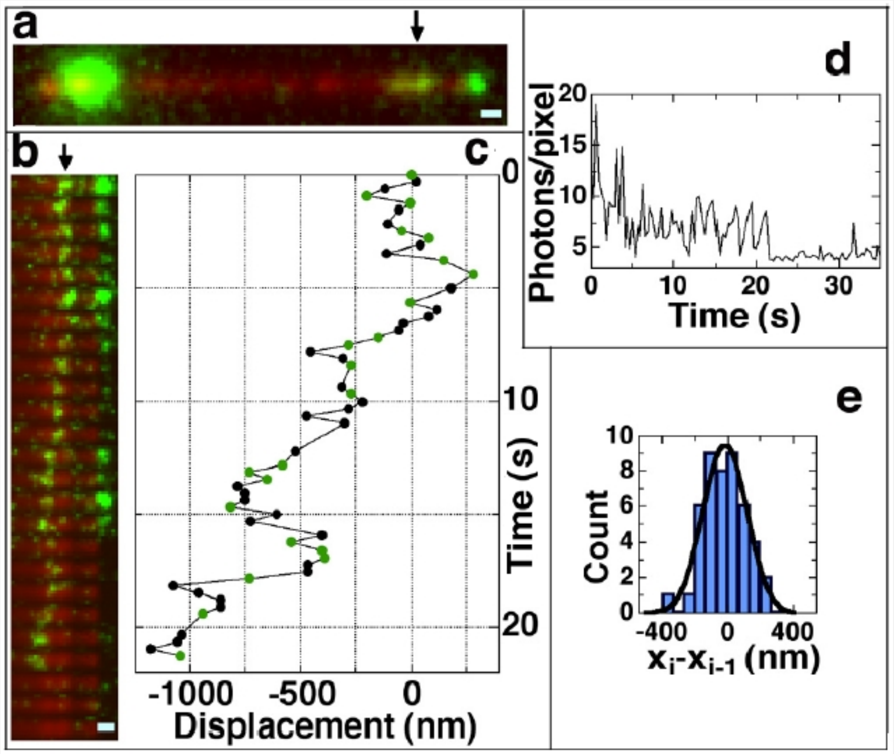
Proof of 1D search mode

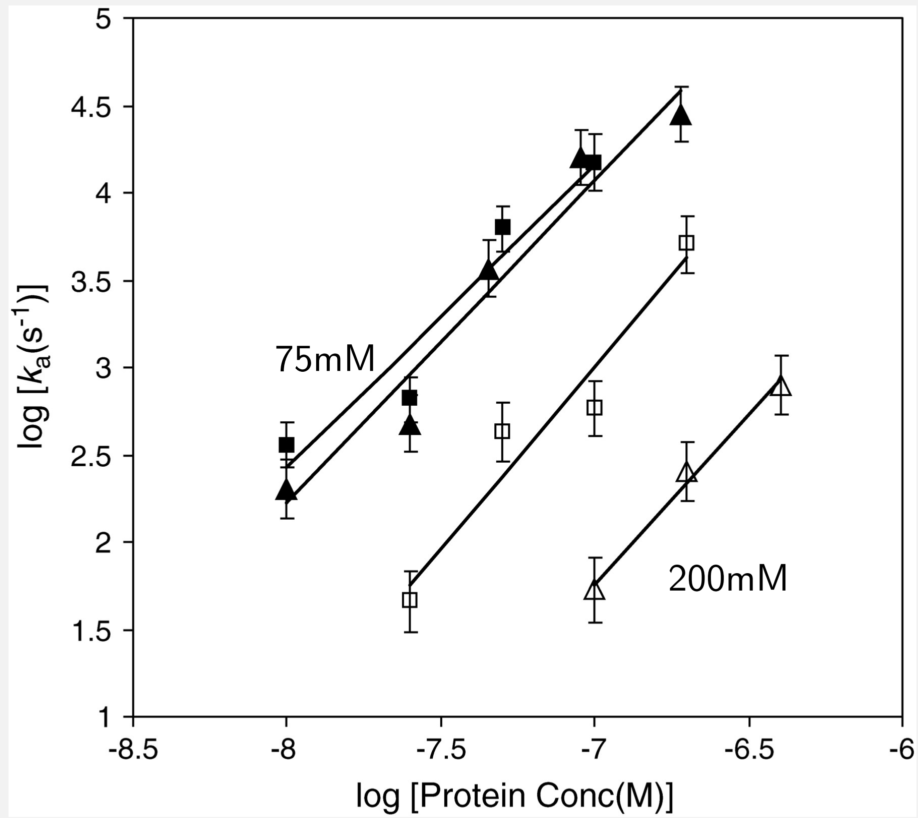
McGhee & von Hippel isotherm

$$f = \frac{N\lambda}{L} \simeq K_{\rm ns}\lambda C, \;\; f \ll 1$$



$$k_a \simeq \begin{cases} C, & \text{1D/3D Berg \& von Hippel} \\ C^2, & \text{Pure 1D search} \end{cases}$$





$\Delta = 1.74 \pm 0.35,\ 1.85 \pm 0.24,\ 2.08 \pm 0.39,\ 1.95 \pm 0.17$

YM Wang, RH Austin & EC Cox, PRL (2006); IM Sokolov, RM, K Pant & MC Williams, Biophys J (2005)

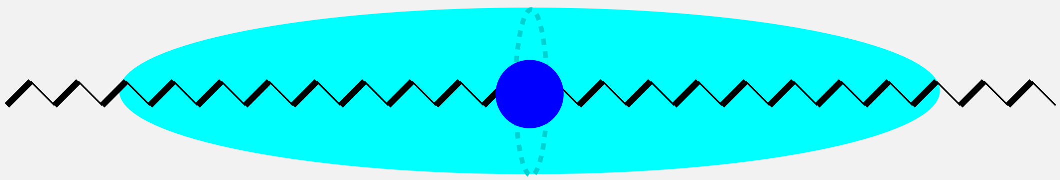
The antenna effect

Target search rate for cylindrical DNA model:

$$k_{\text{on}} \sim 4\pi D_{3d} \ell_{\text{sl}}^{\text{eff}} \times \frac{1}{\sqrt{\ln(\ell_{\text{sl}}^{\text{eff}}/r_{\text{int}})}}$$

Sliding length:

$$\ell_{\text{sl}} = \sqrt{\frac{D_{1d}}{k_{\text{off}}}}$$



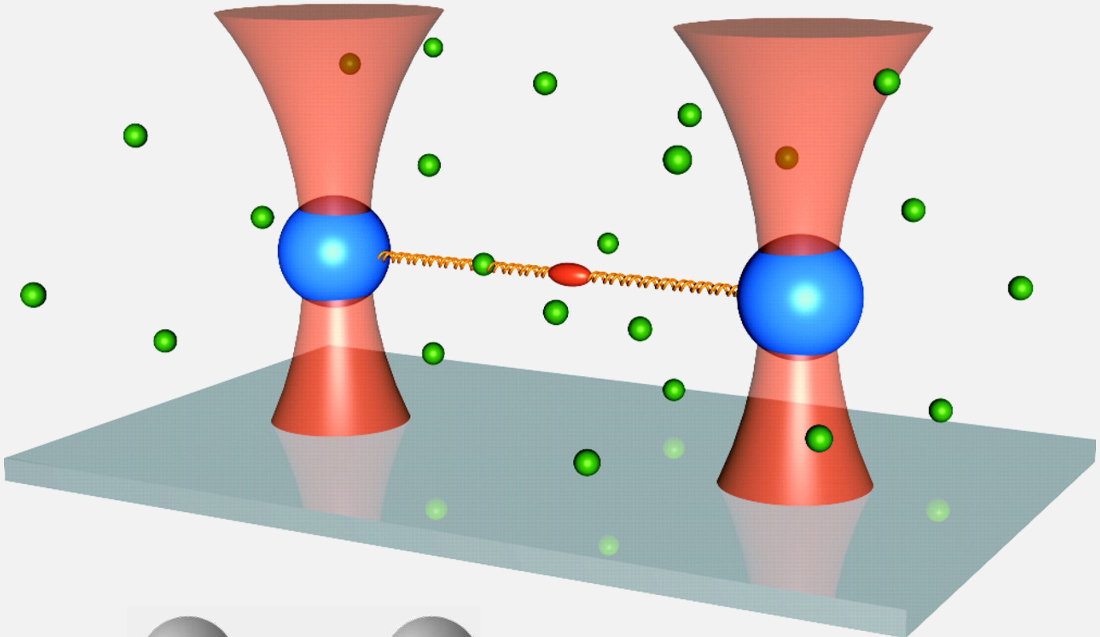
Effective sliding length:

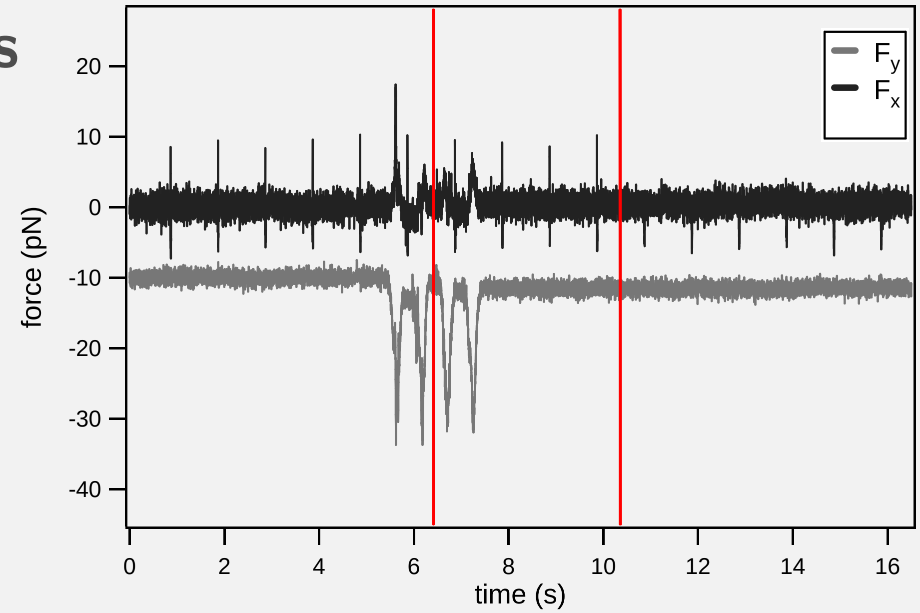
$$\ell_{\text{sl}}^{\text{eff}} = \sqrt{\frac{k_{\text{on}}}{2\pi D_{3d}}} \times \ell_{\text{sl}}$$

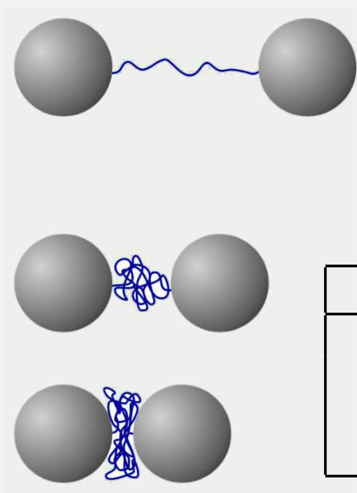
microhop correction: $\sqrt{\frac{k_{\text{on}}}{2\pi D_{3d}}}$

MA Lomholt, T Ambjörnsson & RM, PRL (2005)

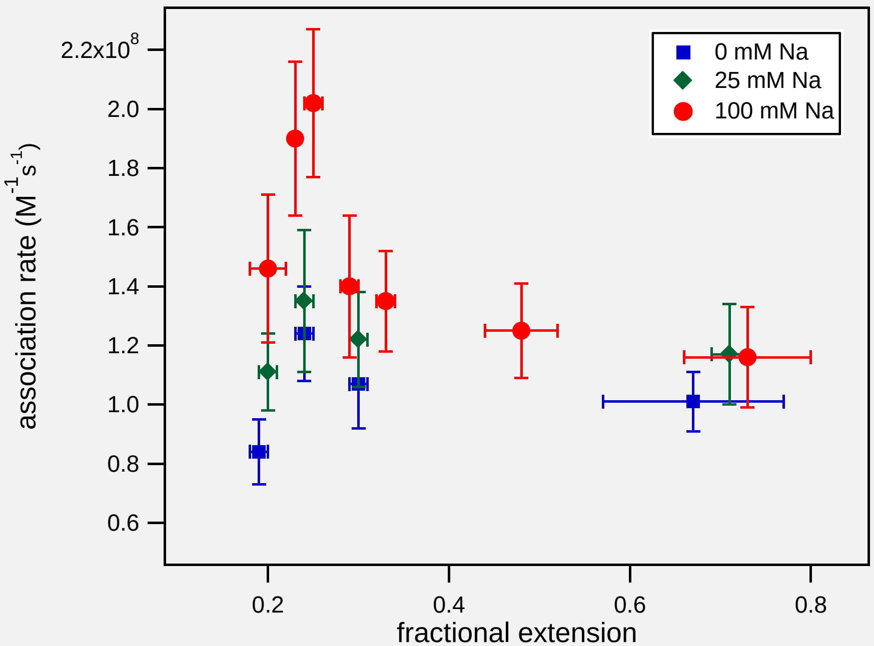
The rôle of DNA conformations





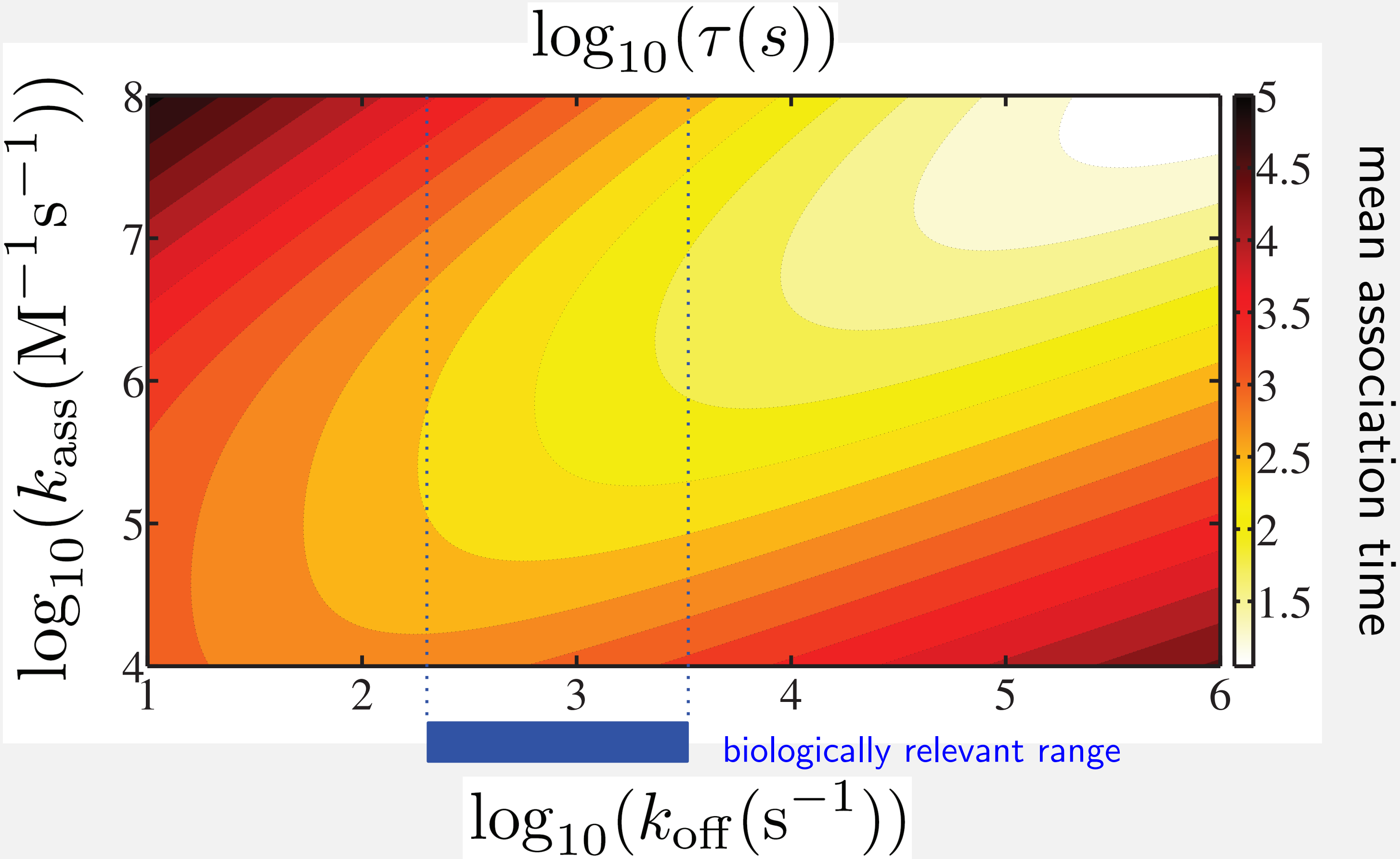


[NaCl]	R_{theory}	R_{measured}
0 mM	1.18	1.3 ± 0.2
25 mM	1.23	1.1 ± 0.2
100 mM	1.67	1.7 ± 0.3
150 mM	1.15	1.3 ± 0.4



B vd Broek, SM Kalisch, MA Lomholt, RM & G Wuite, PNAS (2008); MA Lomholt, B vd Broek, SM Kalisch, G Wuite & RM, PNAS (2009)

In vivo gene regulation consistent with facilitated diffusion



@ optimum the target association time is $\tau \approx 311$sec (no fit parameter)

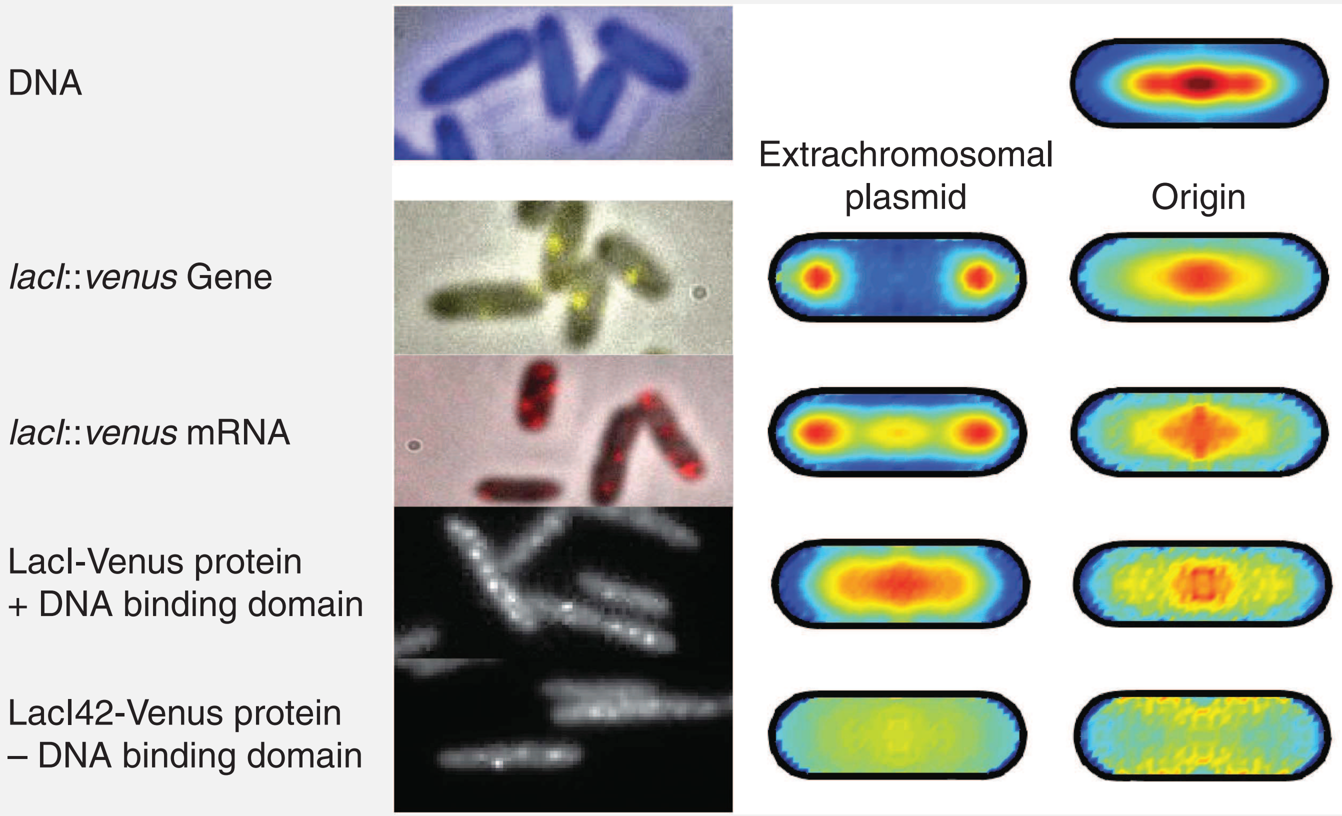
single molecule experiment: $\tau_{\text{exp}} = 354$sec [Elf et al, Science (2007)]

M Bauer & RM, PLoS ONE (2013)

Spatial aspects: do gene locations matter?

Képès: TF targets are typically located next to or at regular distances from the TF gene
$\rightarrow$ TF gene-target pairs close in 3D

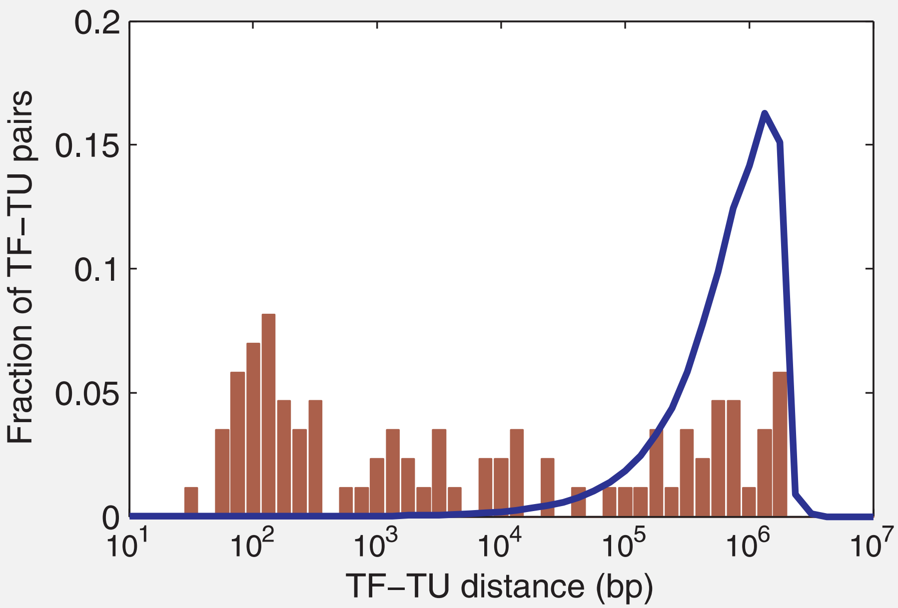
Kuhlman & Cox: • TF distribution highly heterogeneous • TF gene influences distribution: localisation of TF near TF gene

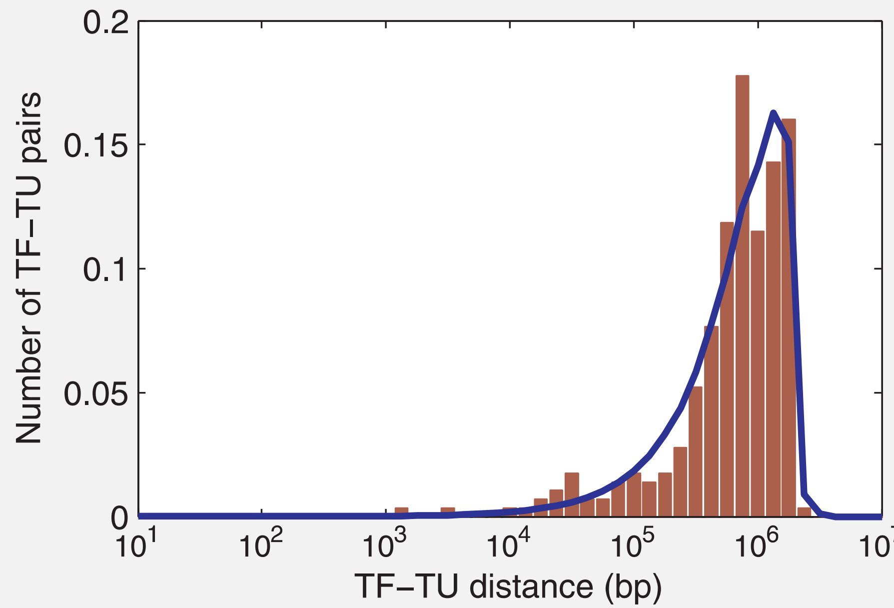


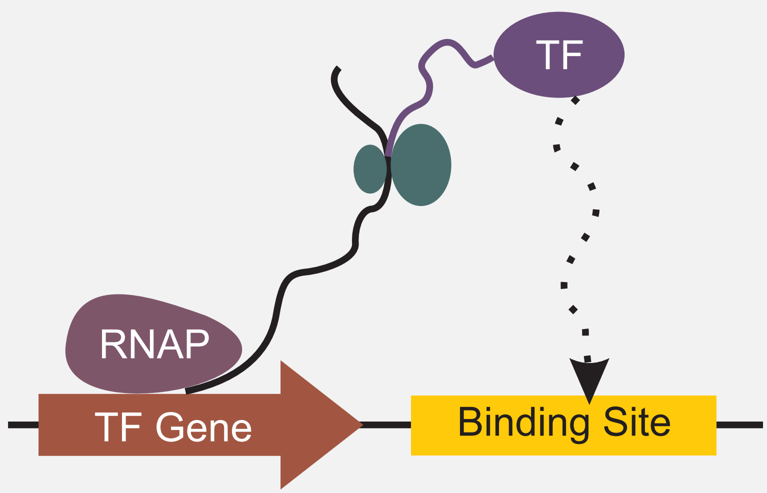
Kepes et al, J Mol Biol (2004); Kuhlman & Cox, Mol Syst Biol (2012)

TF-regulation effects gene-proximity: rapid-search-hypothesis

Gene-gene distance distribution for local TFs (regulate < 4 operons, left) and global (regulate ≥ 4 operons, right). Blue line: random location of genes

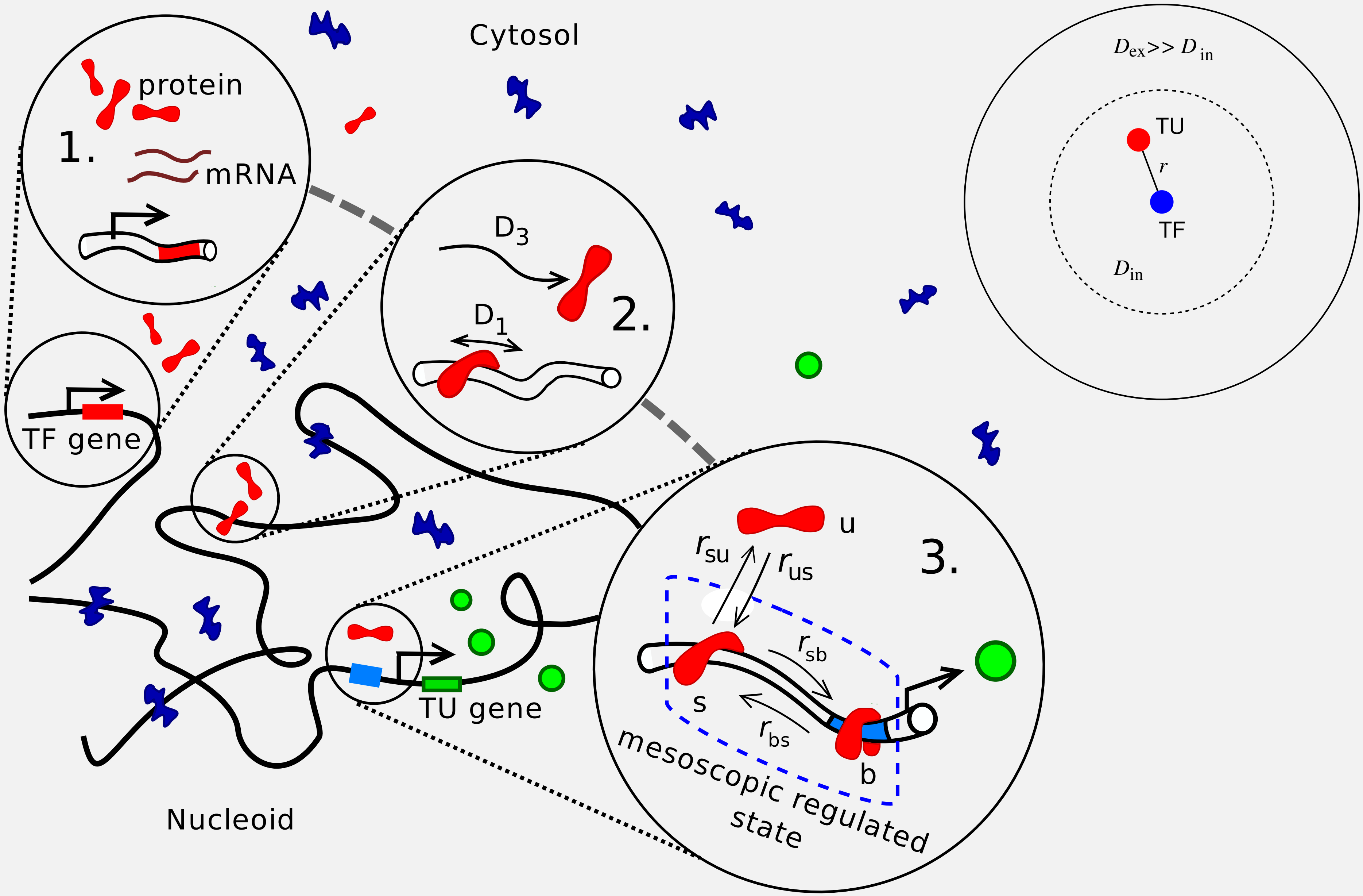






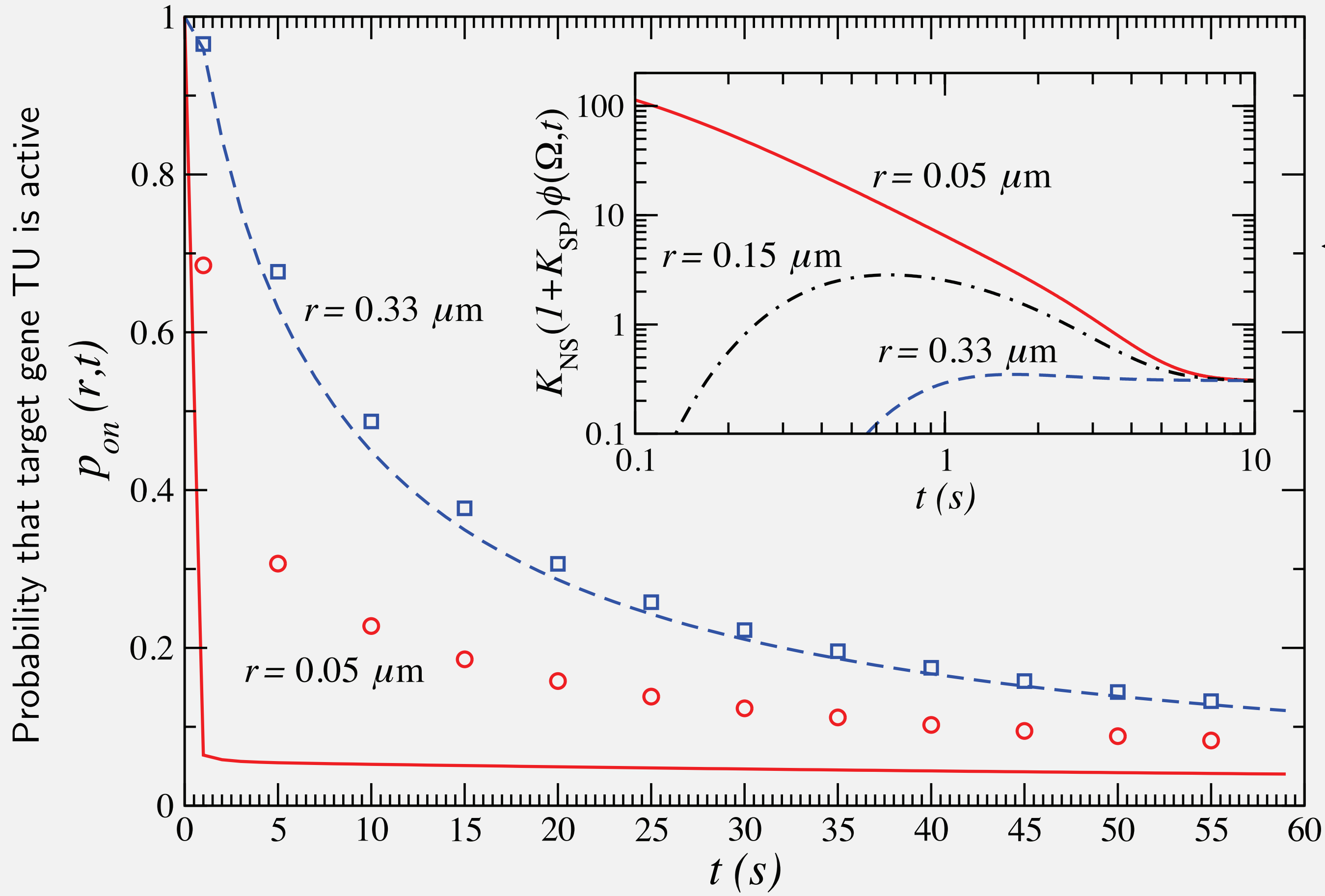
Kolesov et al, PNAS (2007)

Transient intracellular signalling is geometry-controlled



O Pulkkinen & RM, PRL (2013)

Result 1: transient response to repression



← TF concentration @ TU

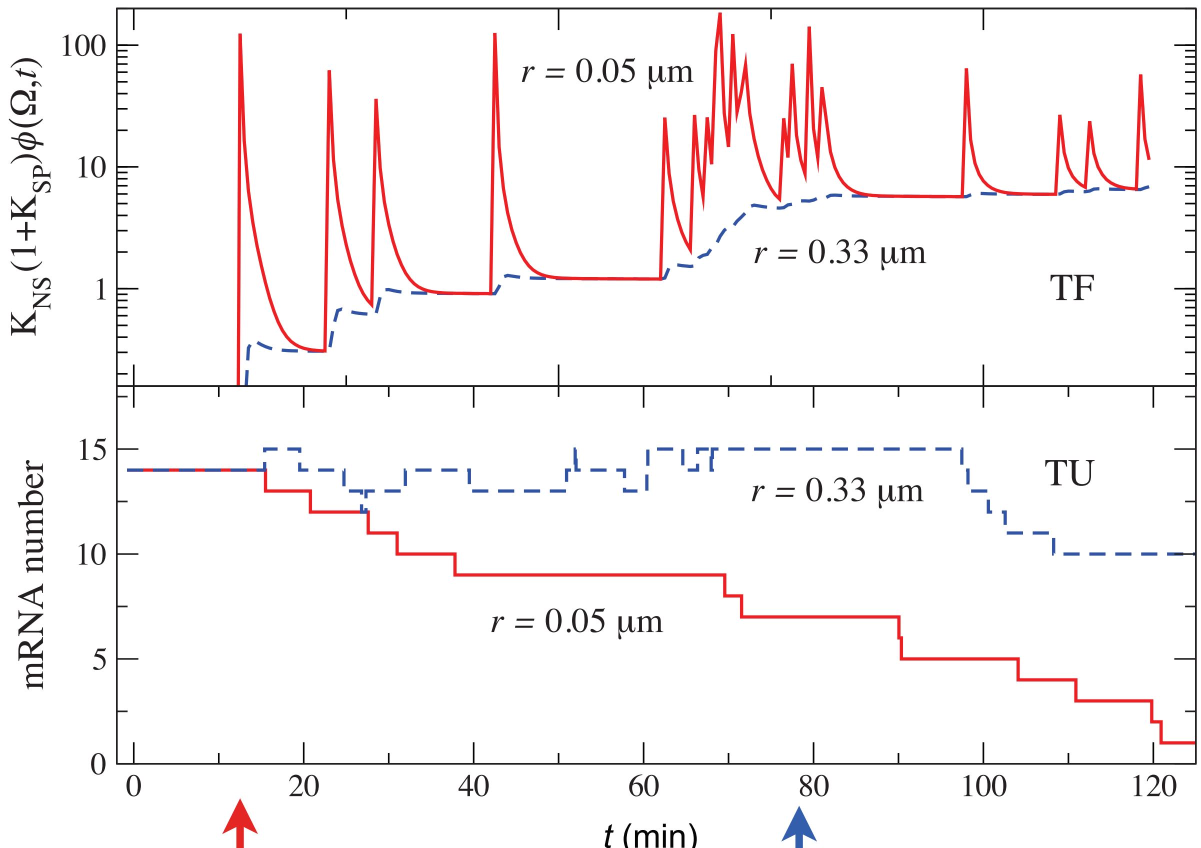
Mean field approximation (full & dashed lines):

$$p_{on}(r,t) = \left\langle \frac{1 + K_{\rm NS}\rho_{\rm TF}(r,t)}{1 + \tilde{K}\rho_{\rm TF}(r,t)} \right\rangle \approx \frac{1 + K_{\rm NS}\langle\rho_{\rm TF}(r,t)\rangle}{1 + \tilde{K}\langle\rho_{\rm TF}(r,t)\rangle}$$

O Pulkkinen & RM, PRL (2013)

Result 2: time dependence of gene response: bursts!

Sample paths from MC simulations:



First repressor binding at r = 0.05 μm

First repressor binding at r = 0.33 μm

TF density at the target operator

Target expression

O Pulkkinen & RM, PRL (2013)

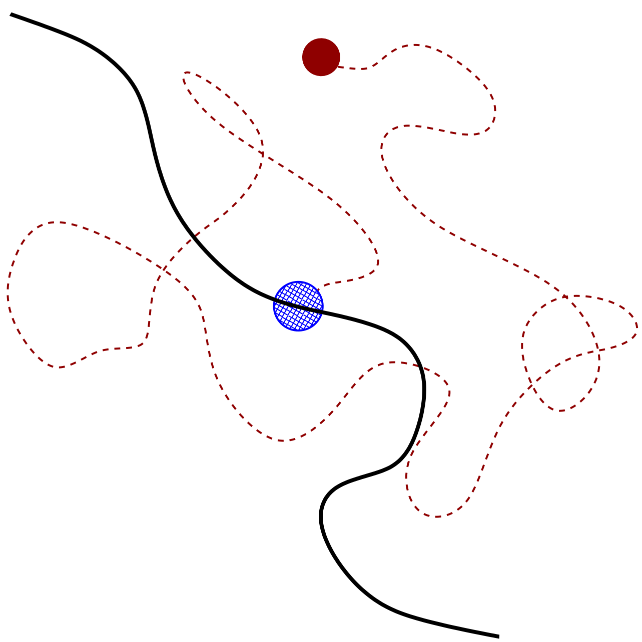
Numerical analysis confirms relevance of proximity effect

P Kar, AV Cherstvy & RM, PCCP (2018)

Brownian impromptu: from mere means to distributions

Search rate for particle with diffusivity D_{3d} to find an immobile target of radius a (assuming immediate binding):

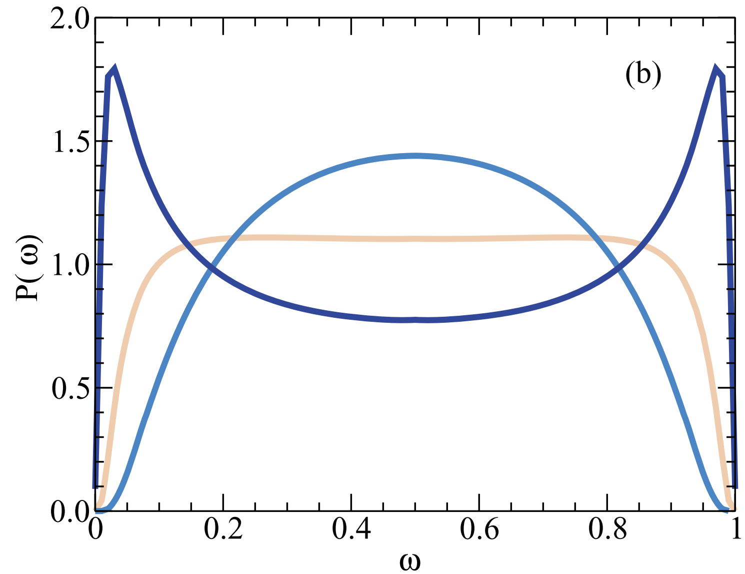
$$k_{\rm on} = 4\pi D_{3d} a$$



Uniformity index for two independent first-passage times τ_1, τ_2:

$$\omega = \frac{\tau_1}{\tau_1 + \tau_2}$$

$\curvearrowright$ $\omega = 1/2$ means good reproducibility



M v Smoluchowski, Physikal Zeitschr (1916)

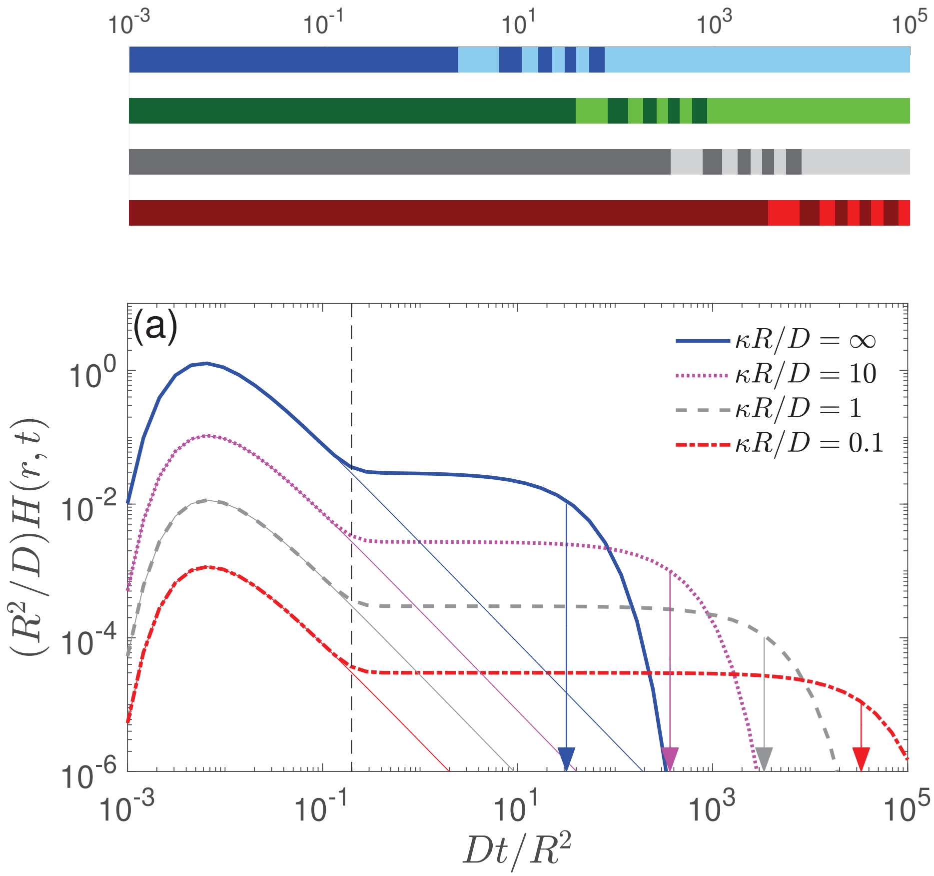
T Mattos, C Mejía-Monasterio, RM & G Oshanin, PRE (2012)

Beyond mere means: first-passage time distributions

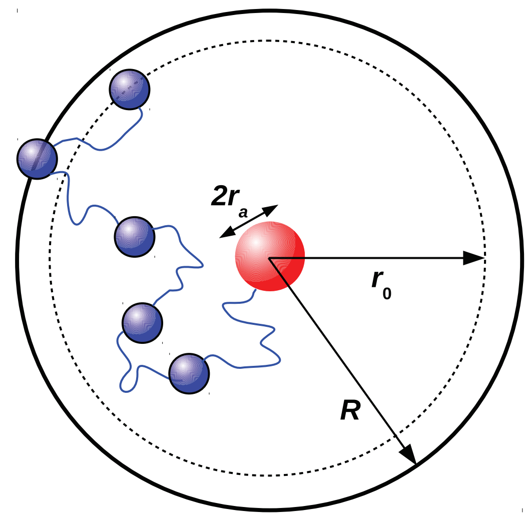
Even in μm-sized bacteria cells biochemical reactions are distance-dependent (geometry-controlled) [Kolesov et al, PNAS (2007); O Pulkkinen & RM, PRL (2013)]

Geometry- versus reaction-control in finite-reactivity scenario

Full first passage time density:



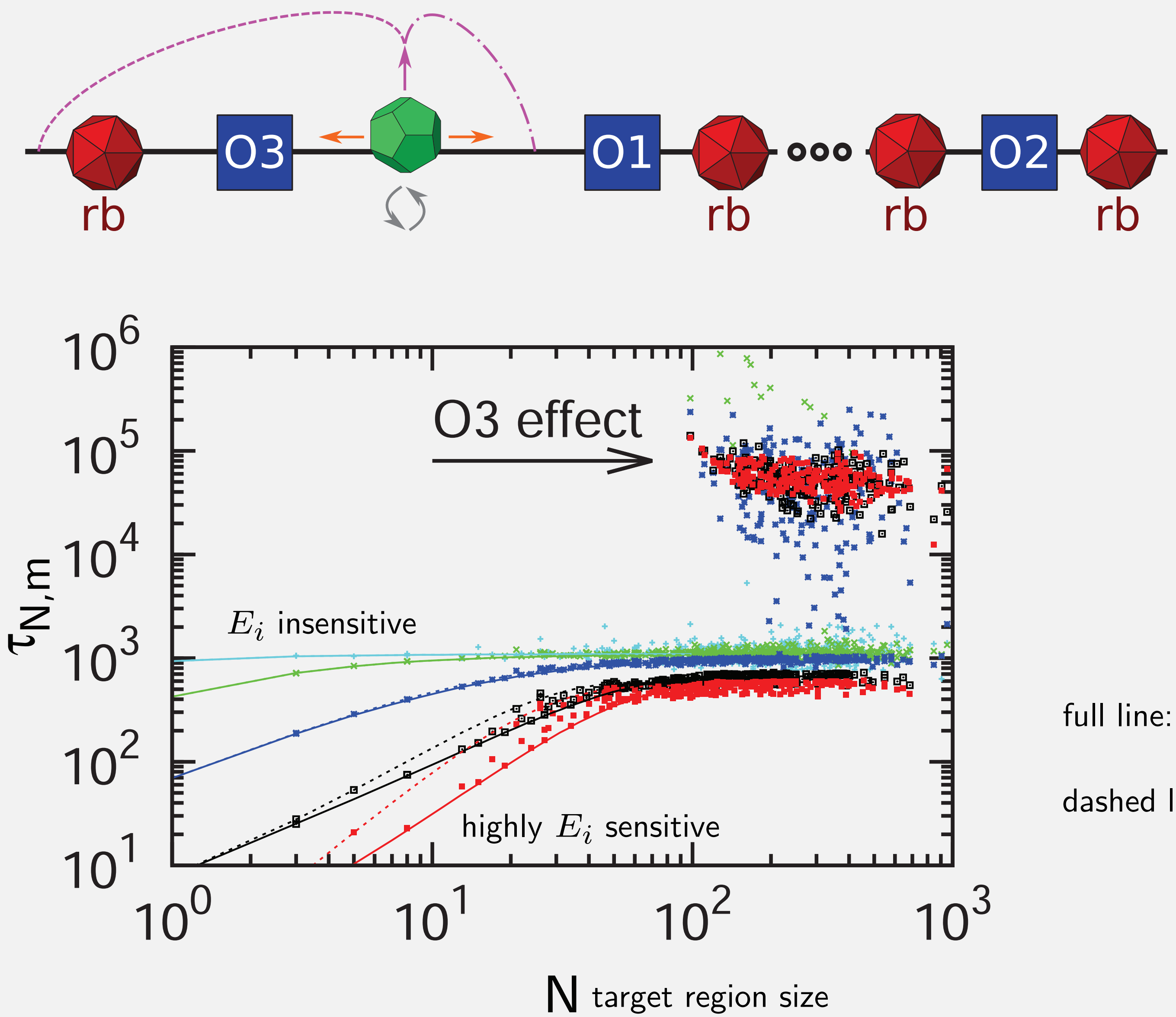
Direct vs indirect trajectories:



$$\langle t\rangle=\frac{(r_0-r_a)(2R^3-r_0r_a[r_0+r_a])}{6Dr_0r_a}+\frac{R^3-r_a^3}{3\kappa r_a}$$

D Grebenkov, RM & G Oshanin, Comm Chem (2019), PCCP (2018); A Godec & RM, PRX (2016), Sci Rep (2016)

Sequence (binding energy) effects on target search time

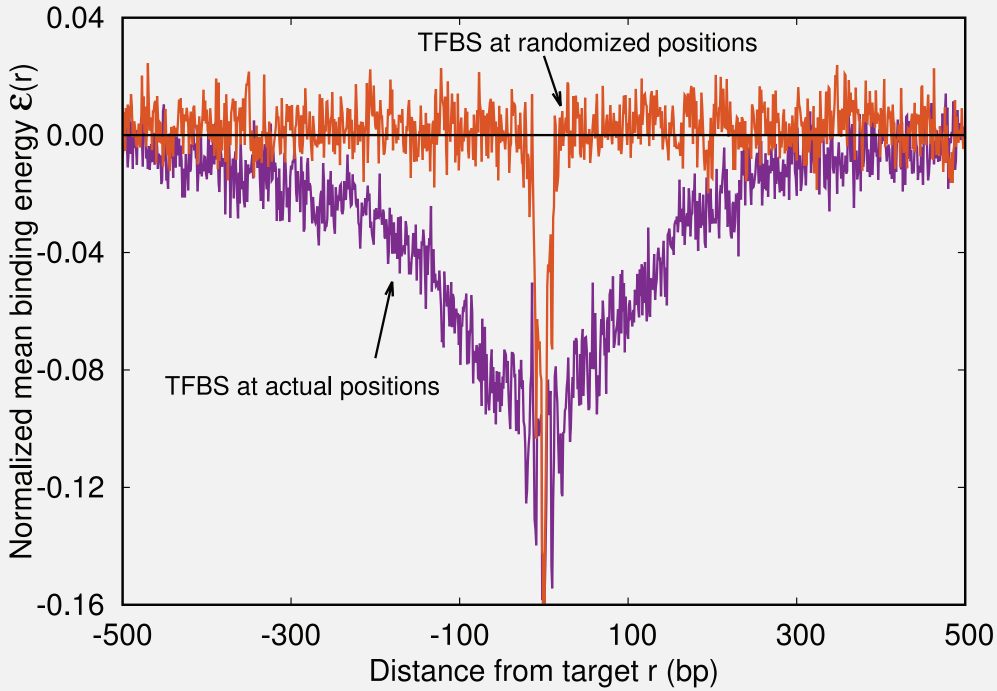


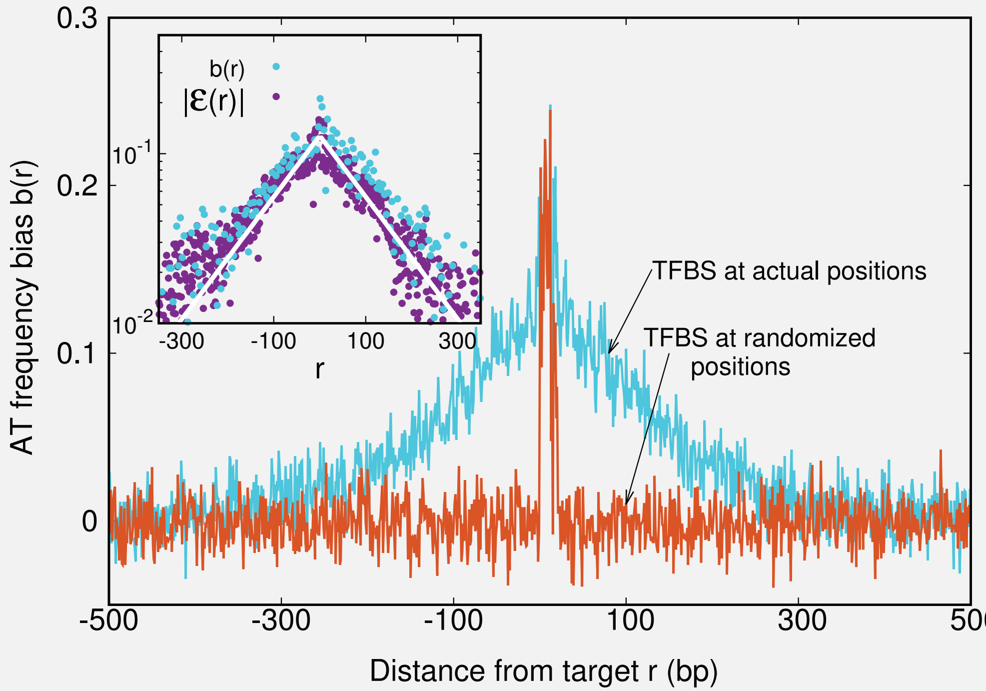
full line: centred target

dashed line: target @ boundary

M Bauer, E. S. Rasmussen, M. A. Lomholt, & RM, Sci Rep (2015)

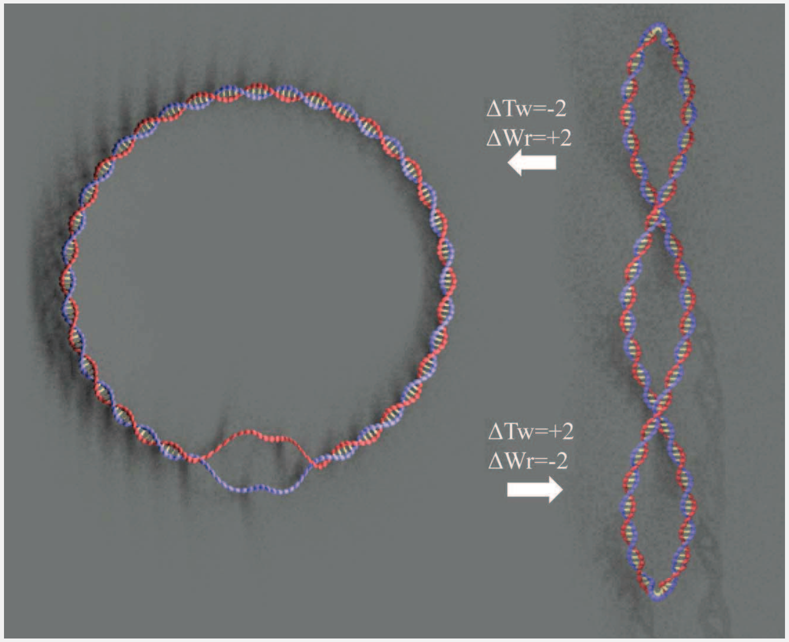
Energetic funnel facilitated diffusion

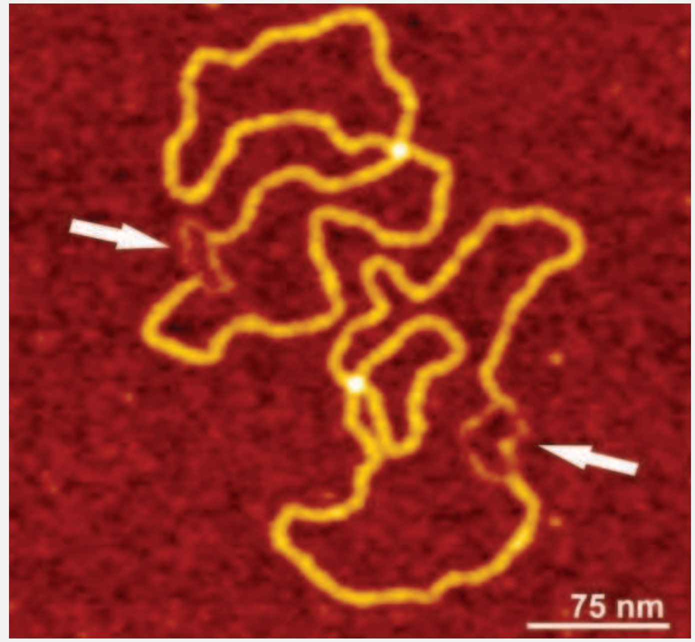




M Cencini & S Pigolotti, NAR (2017)

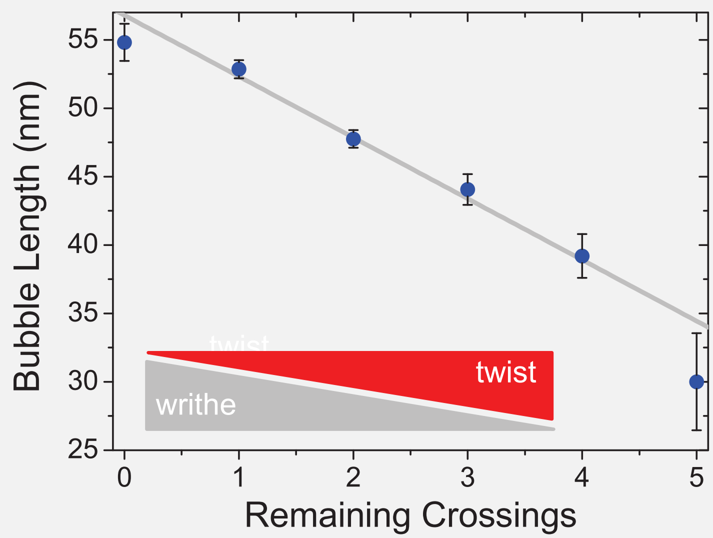
Weak regions at gene starts promote DNA denaturation





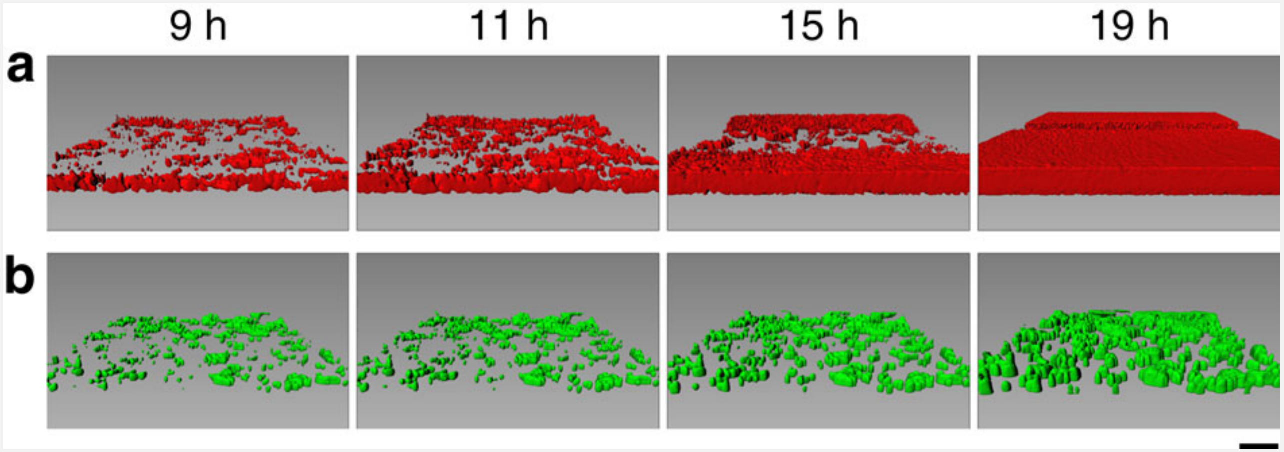
DNA superhelical density:

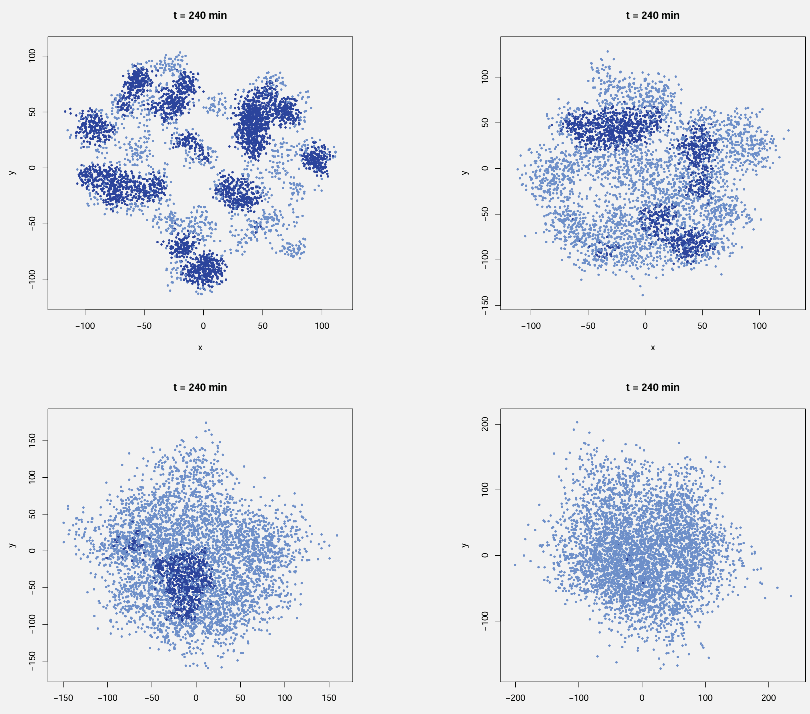
$$\sigma = \frac{\mathrm{Lk} - \mathrm{Lk}_0}{\mathrm{Lk}_0} \approx -0.06$$

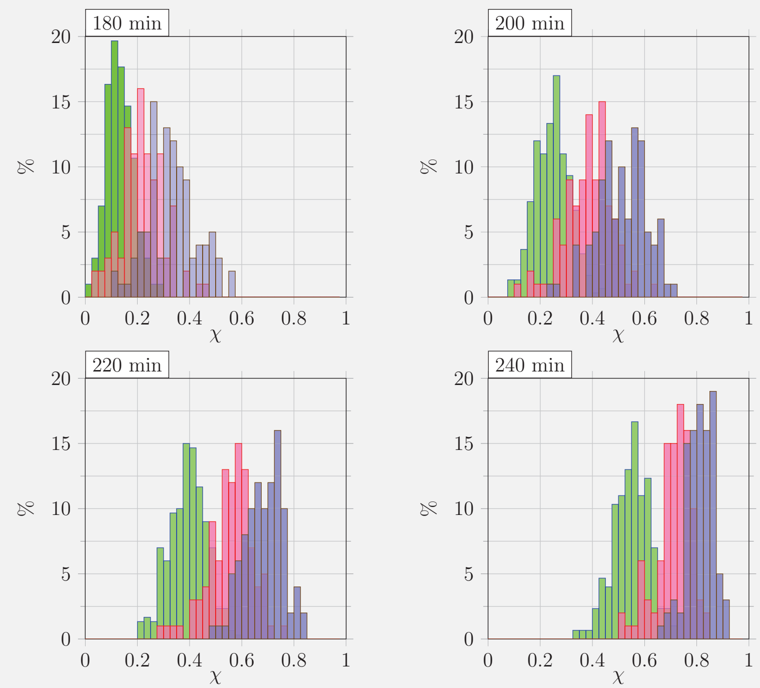


J-H Jeon, J Adamczik, G Dietler & RM, PRL (2010); J Adamczik, J-H Jeon, RM & G Dietler, Soft Matter (2012)

Quorum sensing dynamics in heterogeneous populations







SH Hong et al, Nat Comm **3**, 613 (2011); O Kindler, O Pulkkinen, AG Cherstvy & RM, Sci Rep (2019)

Acknowledgements

Eli Barkai (Bar-Ilan U Ramat Gan)
Carsten Beta (U Potsdam)
Andrey Cherstvy, Aleksei Chechkin (U Potsdam)
Aljaz Godec (MPIBC Göttingen)
Denis Grebenkov (École Polytechnique)
Jae-Hyung Jeon (POSTECH Pohang)
Michael Lomholt (Syddansk U Odense)
Marcin Magdziarz (Politechnika Wrocławska)
Lene Oddershede (NBI Københavns U)
Gleb Oshanin (Sorbonne)
Gianni Pagnini (BCAM Bilbao)
Christine Selhuber-Unkel (U Kiel)
Flavio Seno (Universitá di Padova)
Igor Sokolov (Humboldt U Berlin)
Ilpo Vattulainen (Helsingin Yliopisto)
Matthias Weiss (U Bayreuth)
Agnieszka Wyłomańska (Politechnika Wrocławska)

Universität Potsdam

TAMPERE UNIVERSITY OF TECHNOLOGY

SUOMEN AKATEMIA

FNP Fundacja na rzecz Nauki Polskiej

DFG Deutsche Forschungsgemeinschaft

FIDIPRO Finland Distinguished Professor Programme