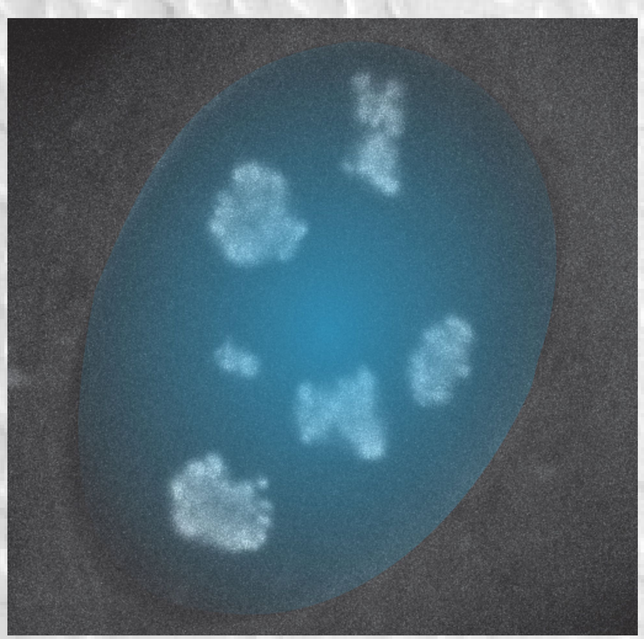
Protein mobility and interactions in the cell nucleus

—

Diffusion and single particle tracking



dkfz.

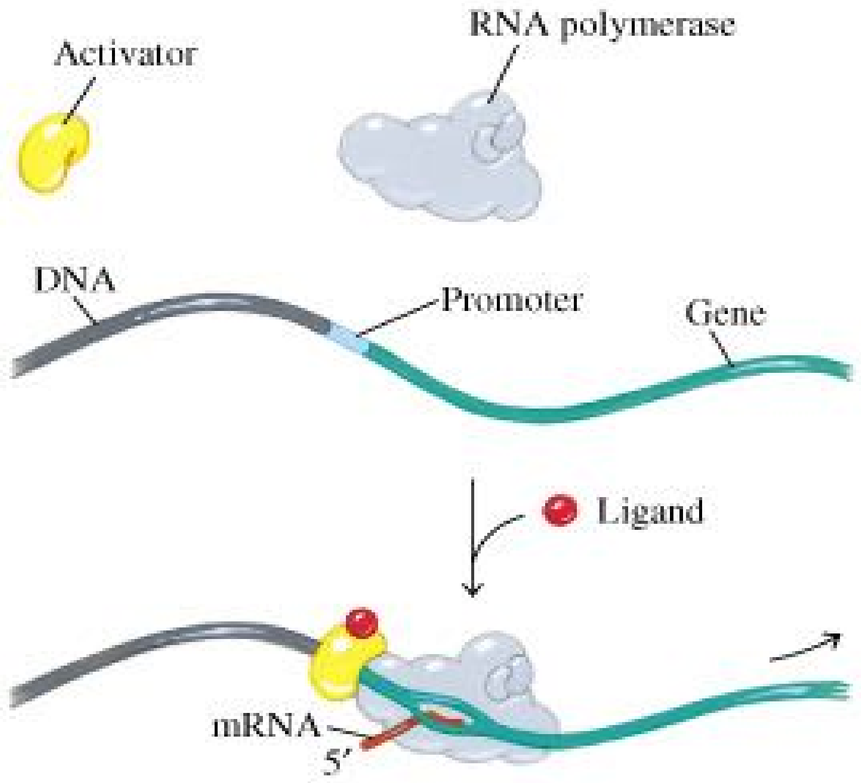
Research for a Life without Cancer

Division of Chromatin Networks
DKFZ & Bioquant, Heidelberg

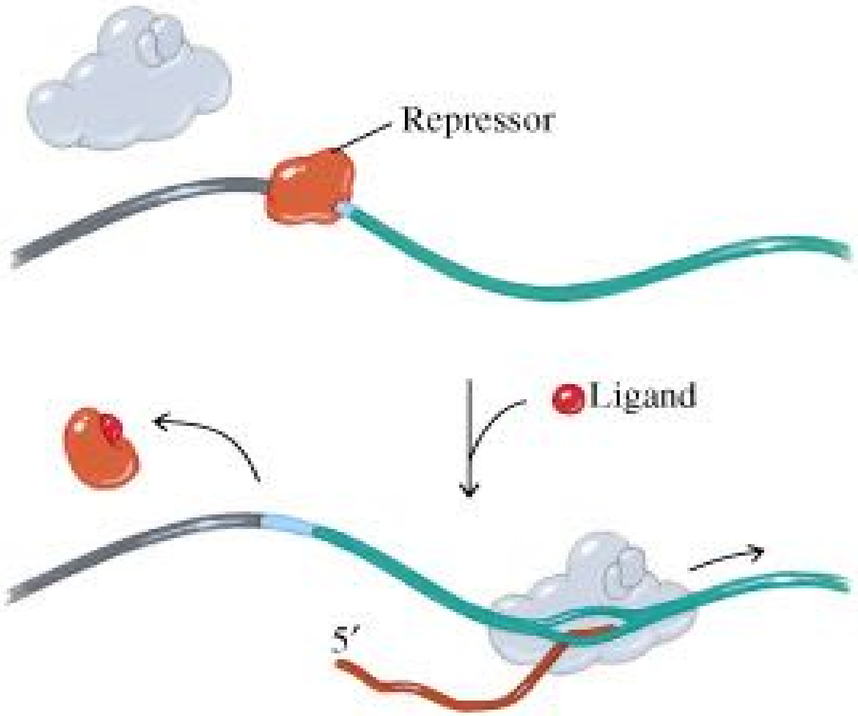
BioQuant
MODEL base of LIFE

Transcription regulation in bacteria

Activator



Repressor



Regulator concentration -> promoter site occupancy -> transcription level

The basic description of protein binding to DNA

$$AB \underset{k_{on}}{\overset{k_{off}}{\rightleftharpoons}} A+B$$

k_{off} in s^{-1} is the reaction rate constant for dissociation

k_{on} in $M^{-1}\ s^{-1}$ is the reaction rate constant for binding

$$\frac{k_{off}}{k_{on}} = K_d$$

relation to the equilibrium dissociation constant

$$\frac{1}{k_{off}} = \tau$$

life time of the complex

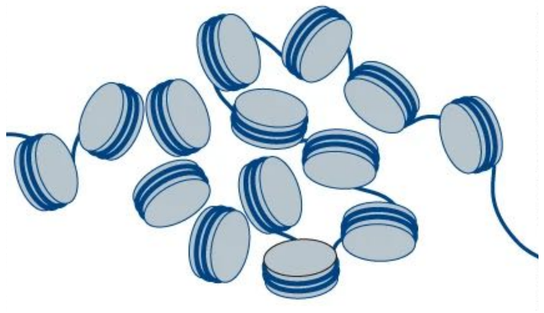
$$\frac{d[AB]}{dt} = k_{on} \cdot [A] \cdot [B] - k_{off} \cdot [AB]$$

rate equation for complex formation,

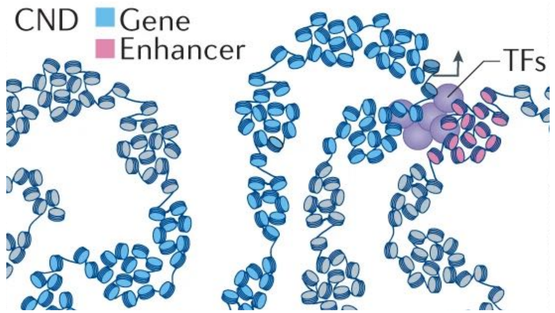
can be solved but it is already difficult

k_{on} cannot be higher than 10^8 - 10^9 $M^{-1}\ s^{-1}$ for a diffusion controlled reaction

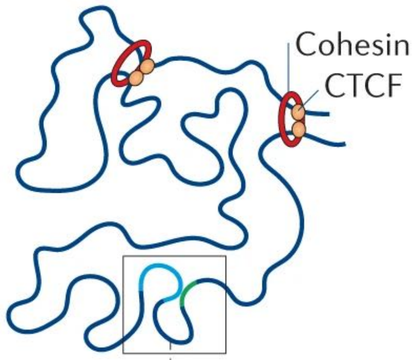
Chromatin organization provides an additional regulatory layer in eukaryotes



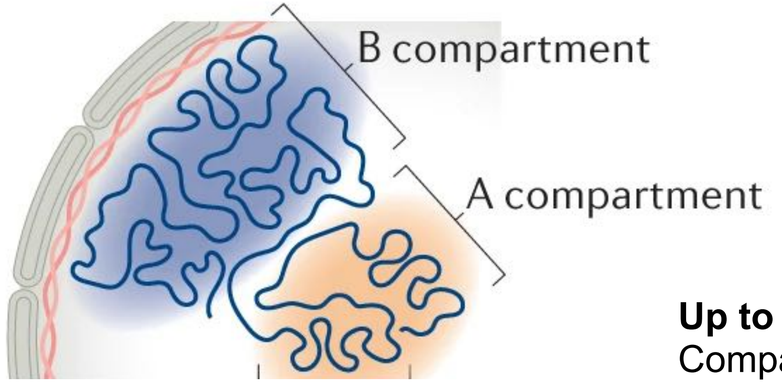
1-2 kb: Nucleosome clutches



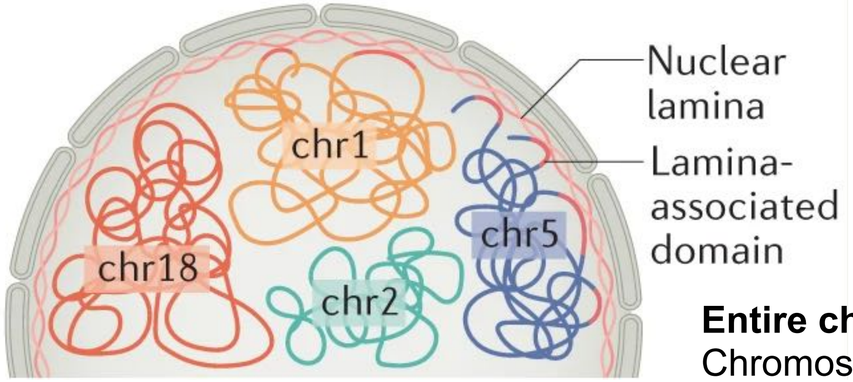
10-100 kb: Chromatin domains and functional loops (E-P contacts)



100 kb to a few Mb: Chromatin loops and topologically associating domains (TADs)



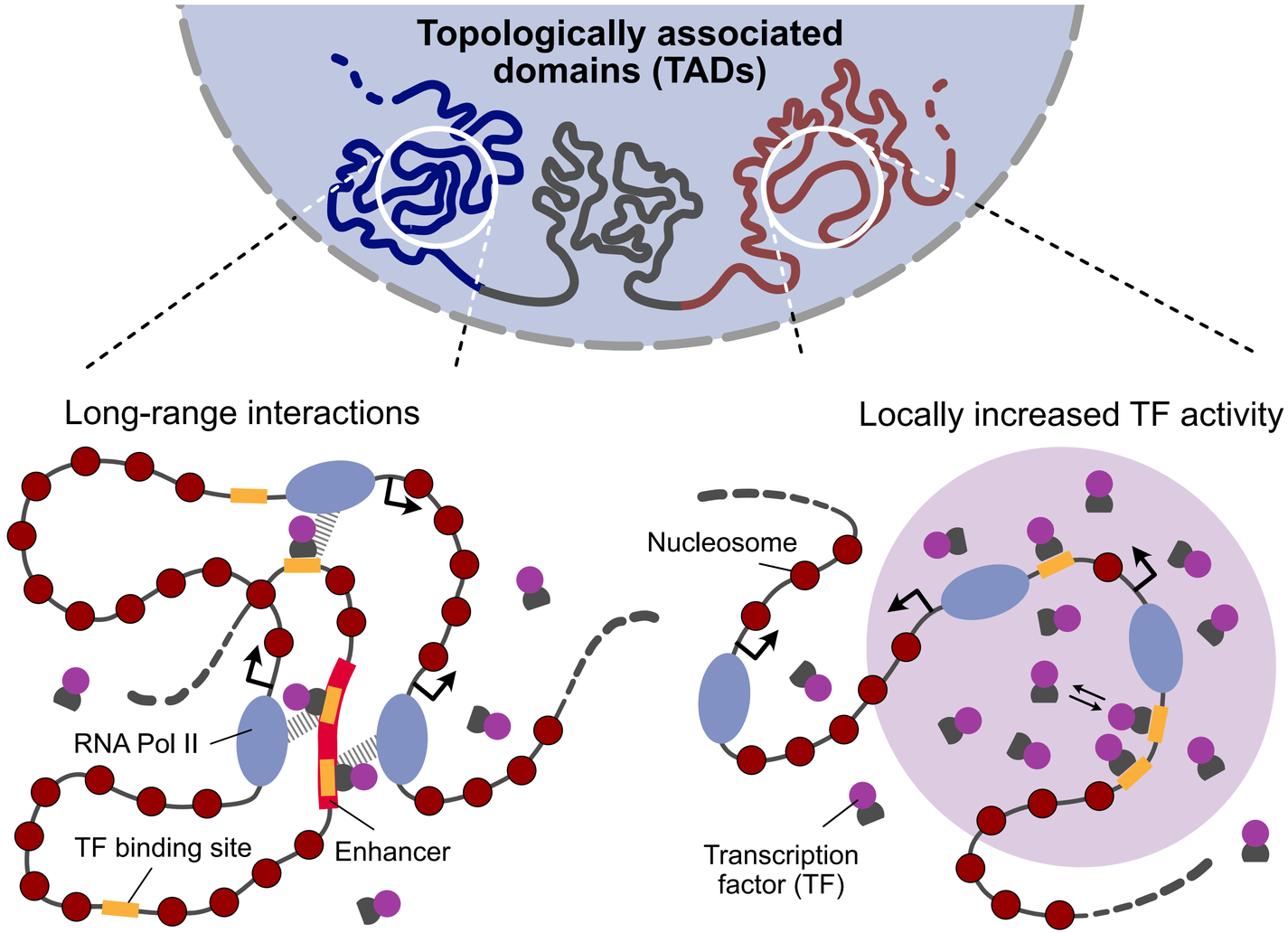
Up to 100s of Mb: Compartments and hubs



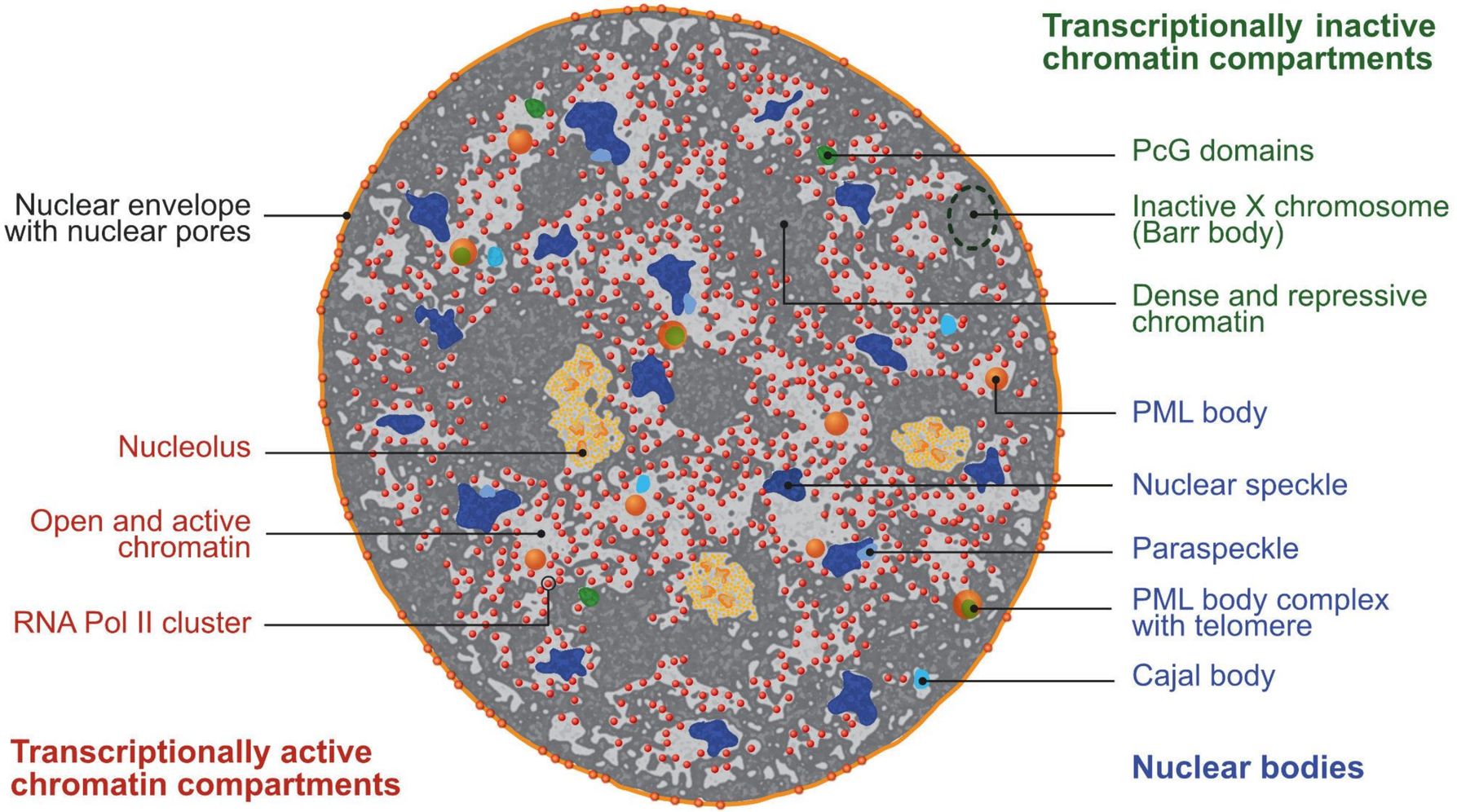
Entire chromosomes: Chromosome territories

Adapted from Fitz-James & Cavalli 2022 *Nat Rev Genet*

Genome organization can regulate gene expression



The mammalian nucleus organizes genome functions in subcompartments



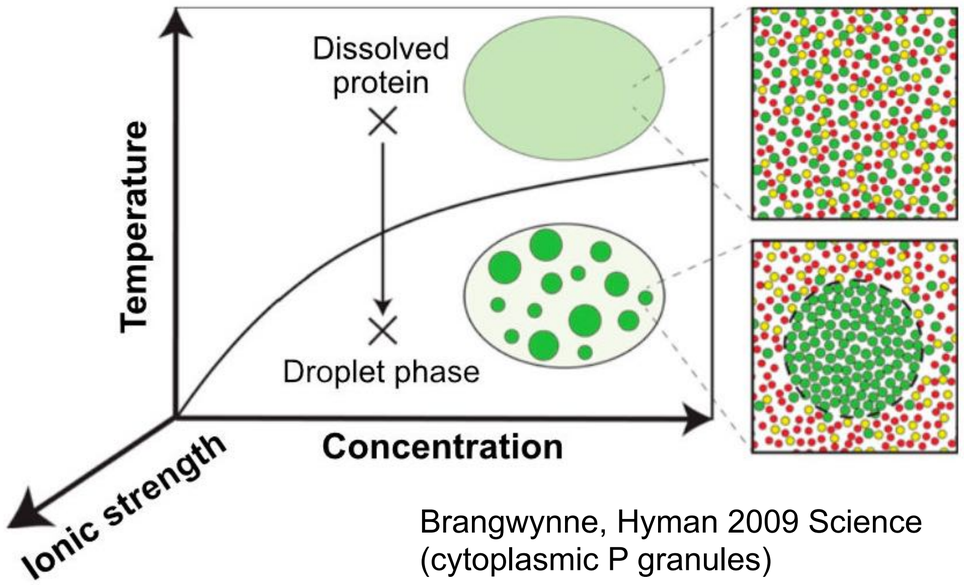
Caudron-Herger & Rippe 2012 *Curr Opin Genet Dev*

On the 10 μm length scale of the nucleus proteins mix fast by diffusion

- Thermal energy at 25 ºC: $k \cdot T = 4 \cdot 10^{-21}$ J
- GFP translocation in 1 sec (D = 30 µm^2 s^{-1}): 13 µm
- Gravitational energy *PEG* GFP (27 kDa, 10 µm): $4 \cdot 10^{-27}$ J or 0.000001 kT
- Human chr 1: 0.25 Gb DNA, 1.2 mio nucleosomes, mass: 520 GDa or 0.87 pg, *PEG* (10 µm): ~20 kT

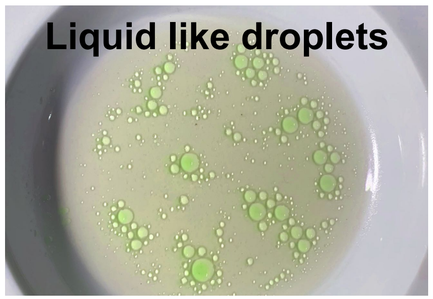
… and gravitation is irrelevant for proteins

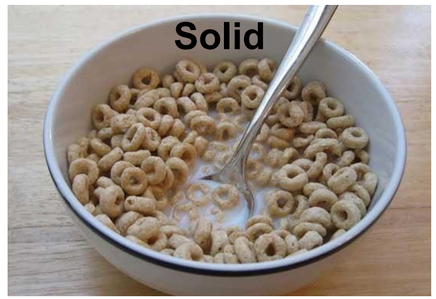
Liquid-liquid phase separation (LLPS) to form cellular subcompartments



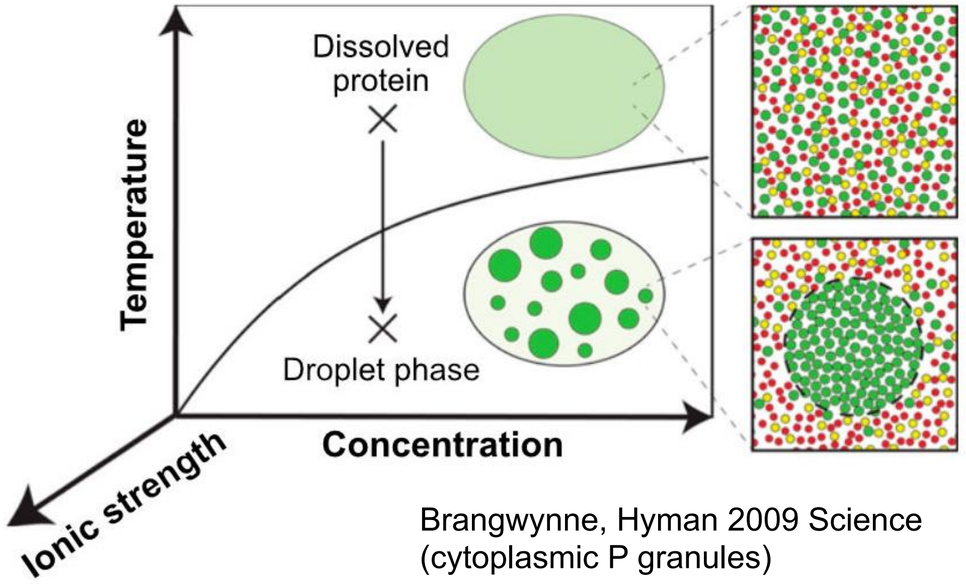
Brangwynne, Hyman 2009 Science
(cytoplasmic P granules)

or phase transitions to other states?





Liquid-liquid phase separation (LLPS) to form chromatin subcompartments



Brangwynne, Hyman 2009 Science
(cytoplasmic P granules)

Nucleolus
Brangwynne 2011 *PNAS*
Feric 2016 *Cell*
Caragine 2019 *eLife*
Frottin 2019 *Science*
Riback 2020 *Nature*

(Peri)centromeres/ heterochromatin
Larson 2017 *Nature*
Strom 2017 *Nature*
Cerase 2019 *NSMB*
Wang 2019 *Mol Cell*
Trivedi 2019 *Nat Cell Biol*
Huo 2020 *Mol Cell*

Telomeres
Shin 2018 *Cell*
Min 2019 *Genes Dev*
Jack 2022 *Dev Cell*

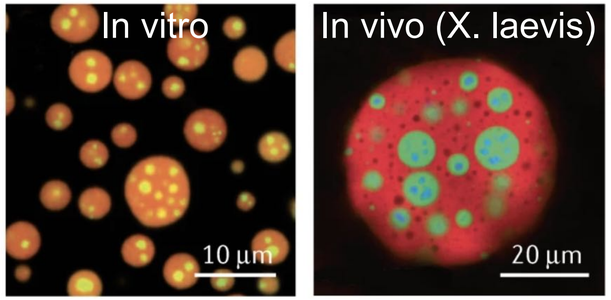
DNA repair sites
Kilic 2019 *EMBO J*
Pessina 2019 *Nat Cell Biol*

"Transcriptional condensates"
Hnisz 2017 *Cell*
Sabari 2018 *Science*
Boija 2018 *Cell*
Boehning 2018 *NSMB*

Cho 2018 *Science*
Lu 2018 *Nature*
Chong 2018 *Science*
Shrinivas 2019 *Mol Cell*
Zamudio 2019 *Mol Cell*
Klein 2020 *Science*
Wei 2020 *Nat Cell Biol*
Lu 2020 *Nat Cell Biol*
Liu 2020 *Nat Cell Biol*
Henninger 2020 *Cell*
Ma 2021 *Mol Cell*

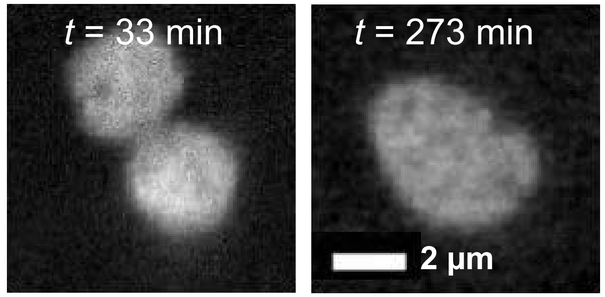
Chromatin
Gibson 2019 *Cell*
Gallego 2020 *Nature*

Droplets of nucleolar proteins



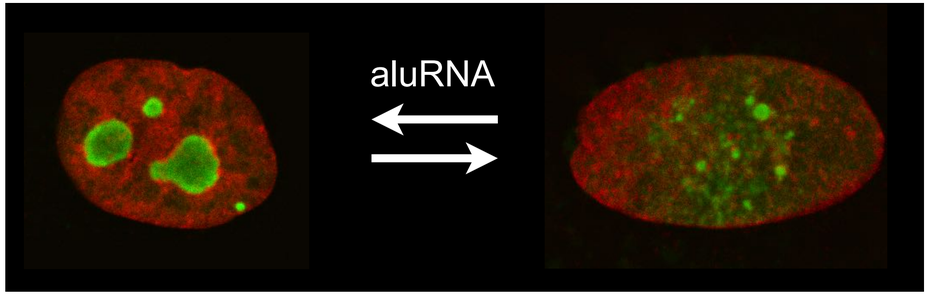
Lafontaine 2020 *Nat Rev Mol Bio*

Nucleoli fusion



Caragine 2019 *eLife*

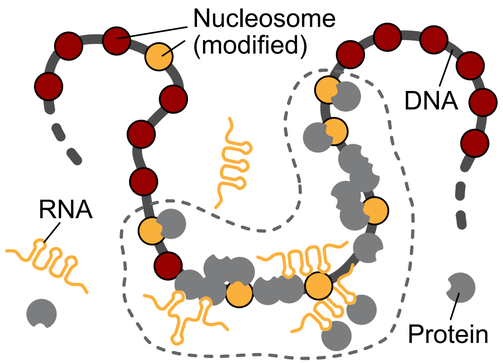
RNA dependent dispersion of nucleoli



Caudron-Herger 2015 *EMBO J;* 2016 *Nucleus*

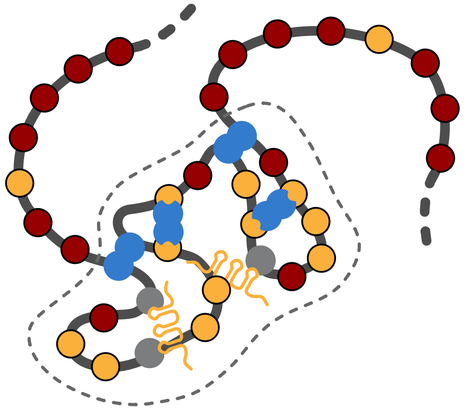
Different mechanisms to form chromatin subcompartments

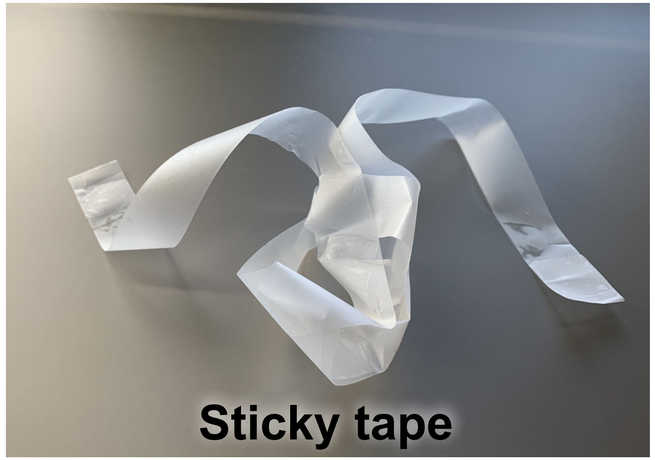
Binding site cluster



Original image by Sean McGrath

Chromatin bridging interactions





Liquid-liquid phase separation

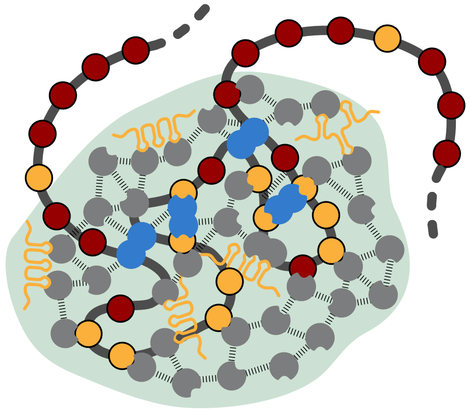
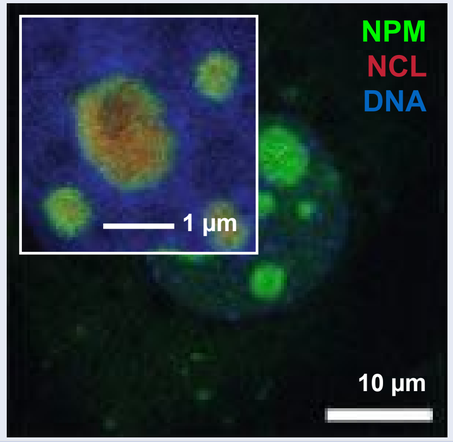
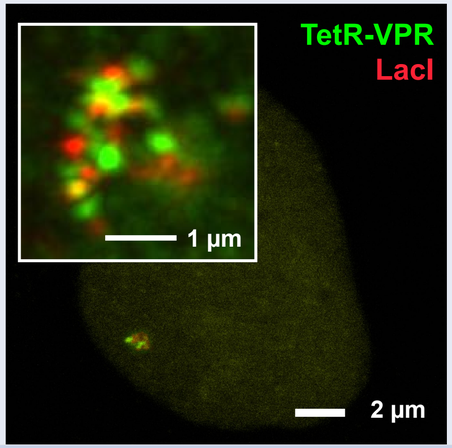
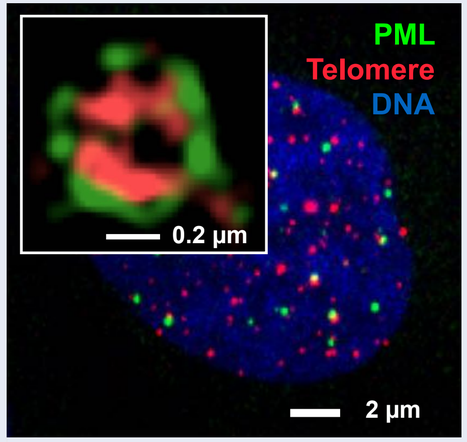
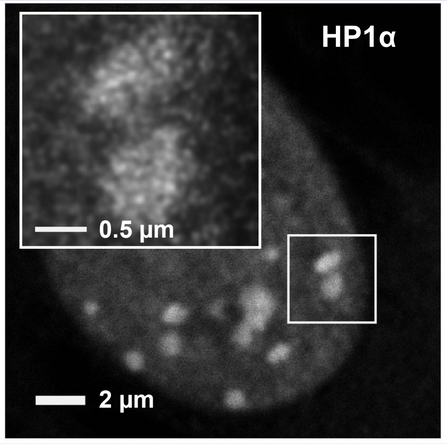


Image by Mathew Spolin

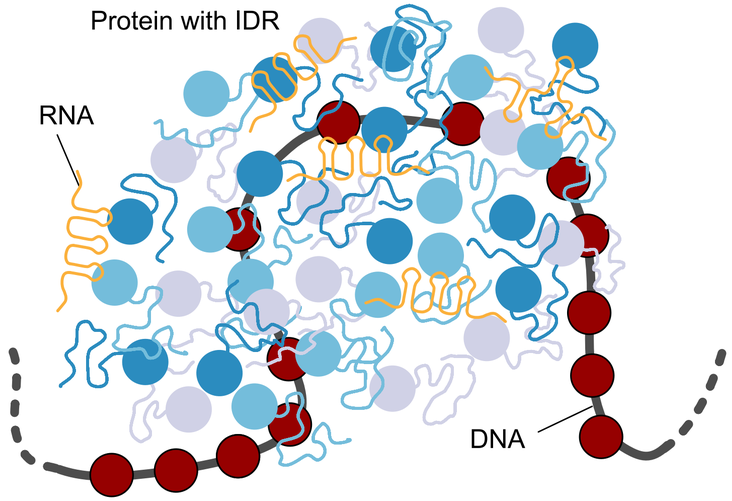
Erdel & Rippe 2018 *Biophys J;* Frank & Rippe 2020 *J Mol Biol;* Rippe 2022 *Cold Spring Harb Perspect Biol*

Mesoscale chromatin subcompartments have diverse properties

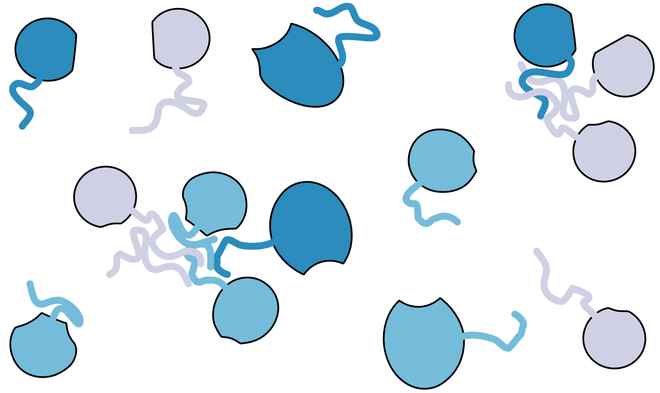
	Nucleolus (NPM, NCL, FBL)	Pol II transcription factories/reporter array	PML-telomere complexes (PML, SP100)	Mouse chromocenters (HP1α, MeCP2)
				
Structure	Tripartite	Homogeneous/granular	Patchy spherical shell	Granular (HP1, DNA)
Exchange with nucleoplasm	Seconds/minutes	Seconds/minutes	Seconds/minutes	Seconds/minutes
Internal mixing (“liquid”)	Yes	?	?	No
Coalescence/fusion	Yes	?	Yes	?
Local viscosity	Increased	?	?	Average
Accessibiliy	Chemical properties	?	?	Size
Protein/DNA & RNA/DNA ratio	High / Very high	High	?	Low
Our references	Caurdron-Herger 2015 *EMBO J*, Caurdron-Herger 2016 *Nucleus*	Caurdon-Herger 2015 *NAR*, Trojanowski 2020 *Mol Cell*	Lang 2010 *J Cell Sci*, Chung 2011 *J Cell Sci*	Müller-Ott 2014 *Mol Syst Biol* Erdel 2020 *Mol Cell*

What is the role of intrinsically disordered protein regions (IDRs)?

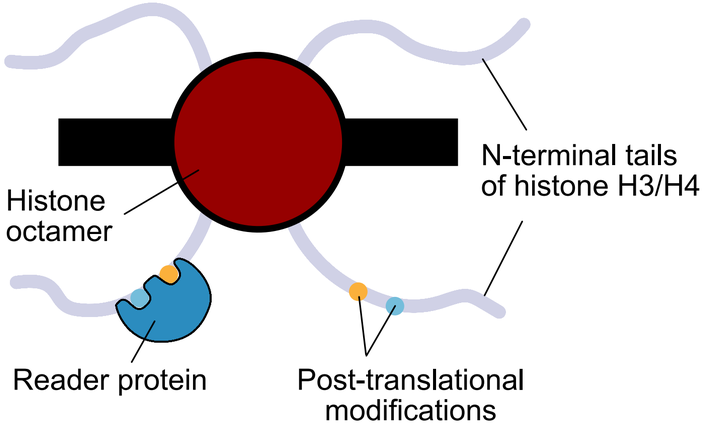
Driving LLPS?



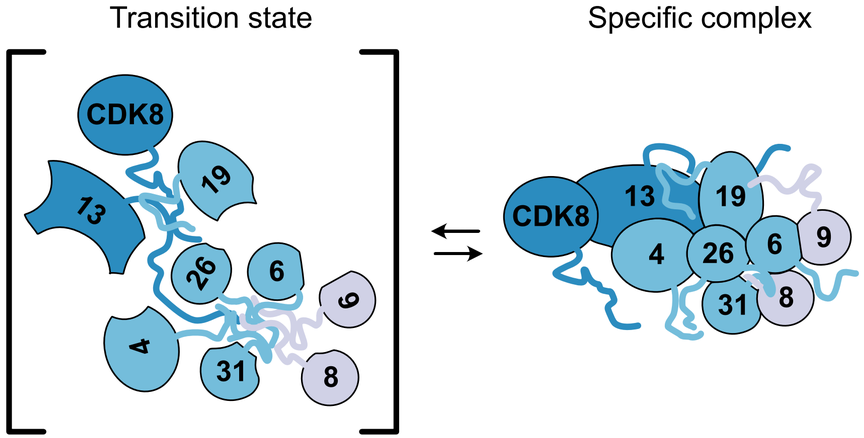
Establishing transient interactions below C_{sat}?



Making specific interactions? (the “histone code”)



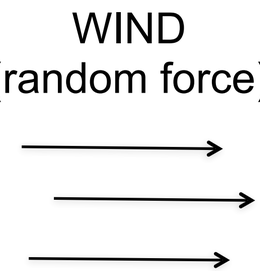
Increasing assembly kinetics of multi-subunit complexes?



Macroscopic vs microscopic world - mass vs friction

Macroscopic world: immobile = large mass

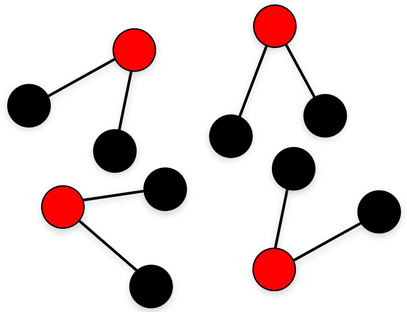
WIND
(random force)



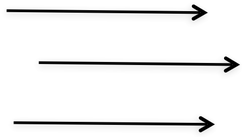
paper

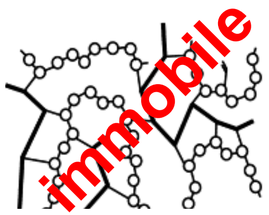
dumbbell

Microscopic world:
immobile = large friction = small diffusion coefficient

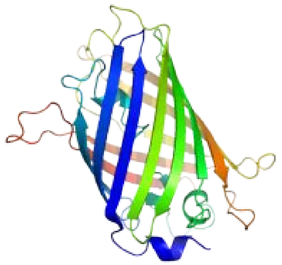


COLLISIONS
(random force)



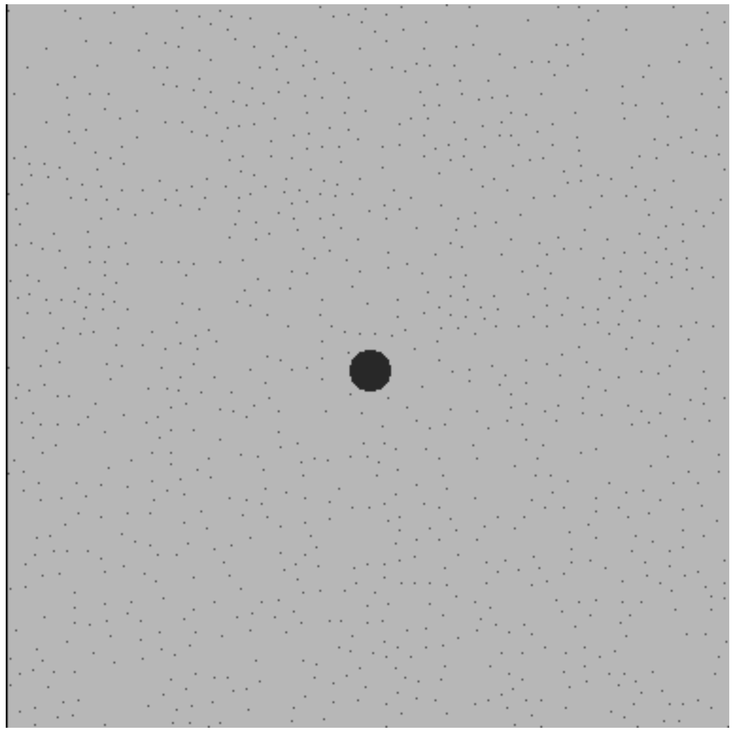


dextran



GFP protein

Movements of a macromolecule in solution by diffusion



- The macromolecule collides with water molecules and moves in a “random walk”.
- The length of the free path is much smaller than the diameter of the particle.
- The average distance from the starting point is proportional to the square root of time.

$k_B T$ is the energy available for spontaneous reactions

$$P_i \propto g_i \cdot \exp\left(\frac{-E_i}{k_B T}\right)$$

The Boltzmann equation yields the probability P_i to find a molecule with energy E_i

- g_i: number of different states with energy E_i
- k_B: Boltzmann constant
- T: Temperature

probability to find a particle with an energy

- of $k_B T$ or larger: 0.37 => processes that requires an energy of $k_B T$ occur spontaneously
- of 10 $k_B T$ *or larger*: 0.00005 => these processes will not occur spontaneously

at 298 K (25 ºC) $k_B T = 4.1 \cdot 10^{-21}$ J or $\boldsymbol{k_B T = 4.1}$ **pN·nm**

$k_B T$ refers to a single molecule
for 1 mol of particles one has to use $k_B T$ x $6.022 \cdot 10^{23} = RT$

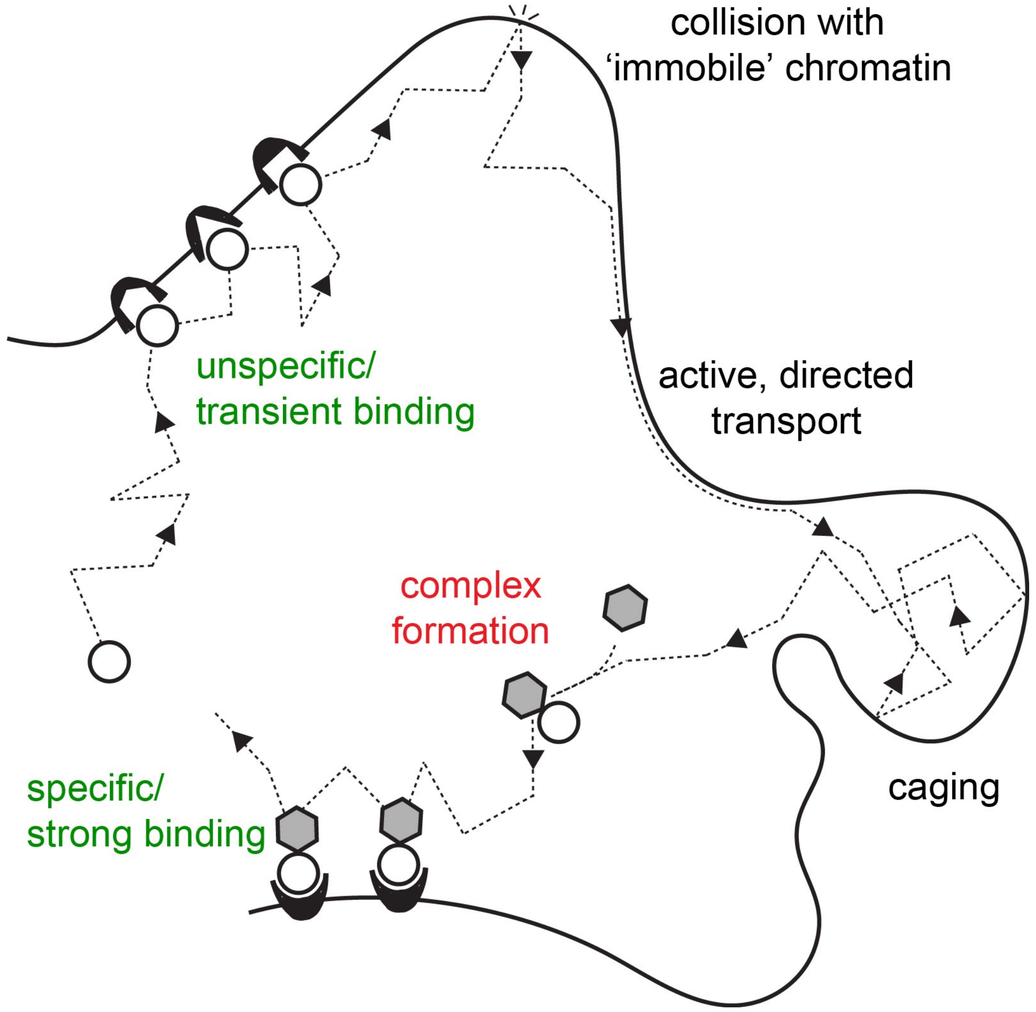
at 25 ºC with $R = 8.3$ J · mol^{-1} · K^{-1} => $\boldsymbol{RT}$ **= 2.5 kJ/mol or 0.6 kcal/mol**

Hydrolysis of ATP: 20-30 $k_B T$/ATP, 12 - 18 kcal/mol or 50 - 70 kJ/mol (physiological conditions)

Questions we want to answer

- How does the nucleus and the genome self-organize into membraneless subcompartments?
- How is structure related to function?
- How do proteins like transcription factors find their target site?
- Which role do phase separation mechanisms play?
- How can we quantitatively describe these processes?
- …

Different microscopic phenomena influence protein mobility and interactions in living cells

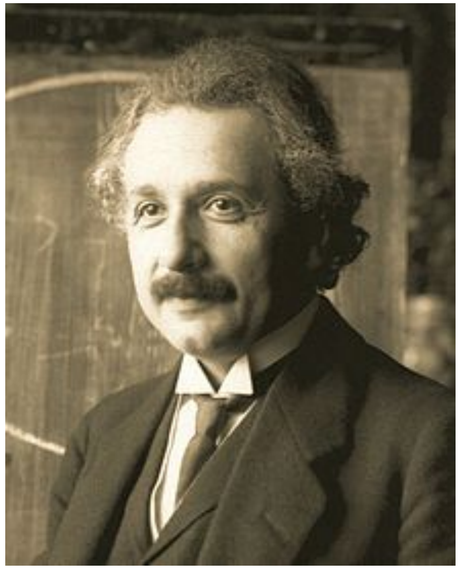


Erdel F & Müller-Ott K et al., *Chrom Res* (2010)

Diffusion

- How can we measure it?
- How can we theoretically describe it?
- What can we learn from studying it?

Robert Brown,
botanist, 1773-1858



Albert Einstein,
physicist, 1879-1955

Marian Smoluchowski,
physicist, 1872-1917

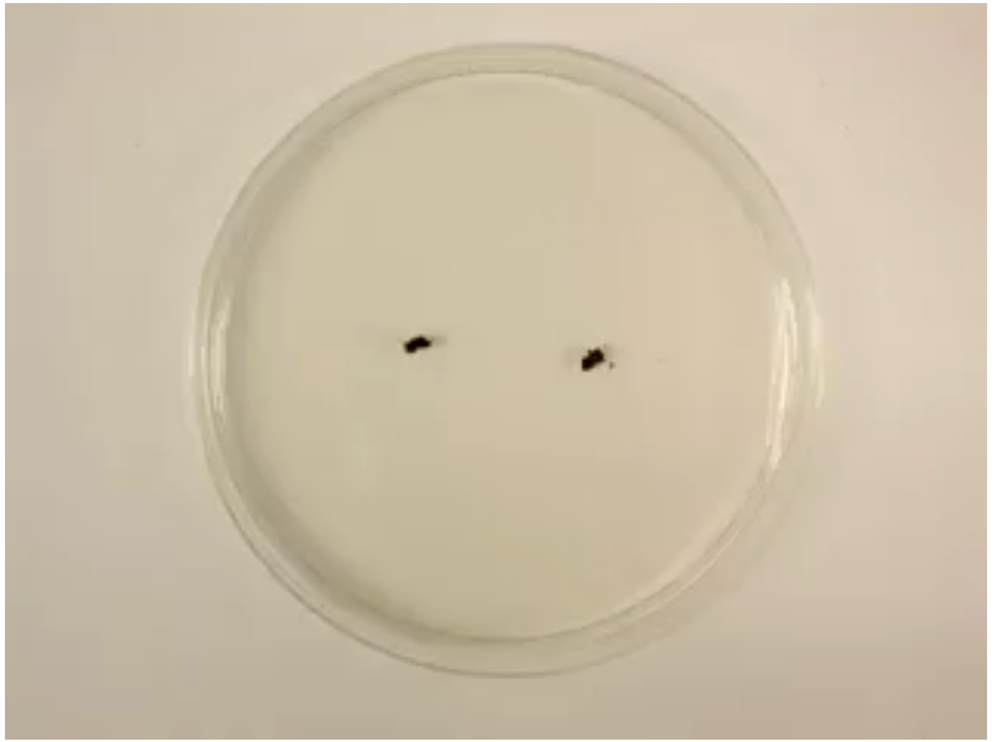
Phenomenological definition

Diffusion

=

Passive transport of particles
due to thermal energy ($k_B\ T$)

Macroscopic visualization

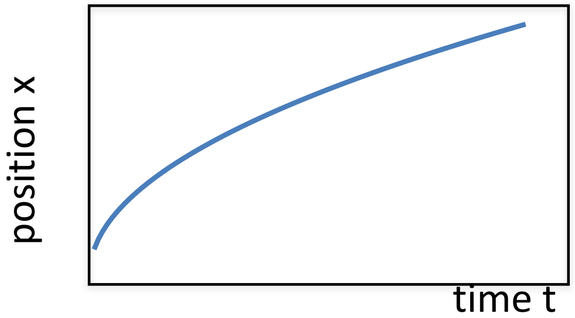


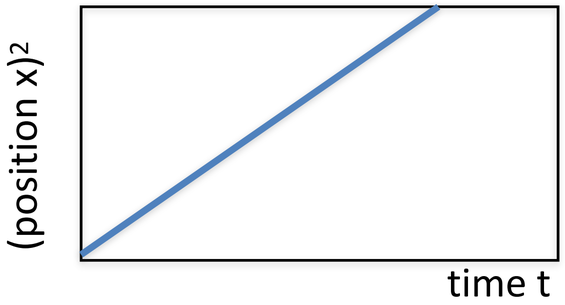
Pipette colored liquid in an agar plate and watch it over time

The diffusion coefficient

- The diffusion coefficient describes how far a particle can travel within a given time *t*:

$$x = \sqrt{6Dt}$$





- For a sphere it is given by $D = \frac{kT}{6\pi\eta r}$

→ A particle is fast at high temperature *T*, small radius *r* and low viscosity η of the liquid

Typical diffusion coefficients

- Proteins: 5-60 $\mu m^2 s^{-1}$
- mRNA: 0.04 $\mu m^2 s^{-1}$
- Telomere: 0.002 $\mu m^2 s^{-1}$

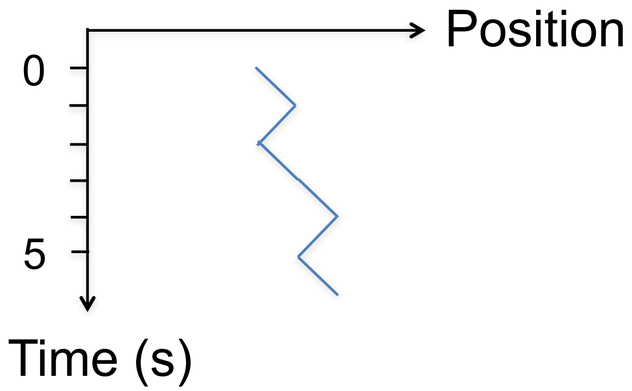
→ Proteins/RNAs need < 2 seconds to diffuse through the nucleus

→ Nuclear bodies and chromosome loci are less mobile

→ Binding to chromatin/NBs can slow them down

Microscopic world: Random Walk

- Consider a particle (or person) that moves every second randomly one step left or right

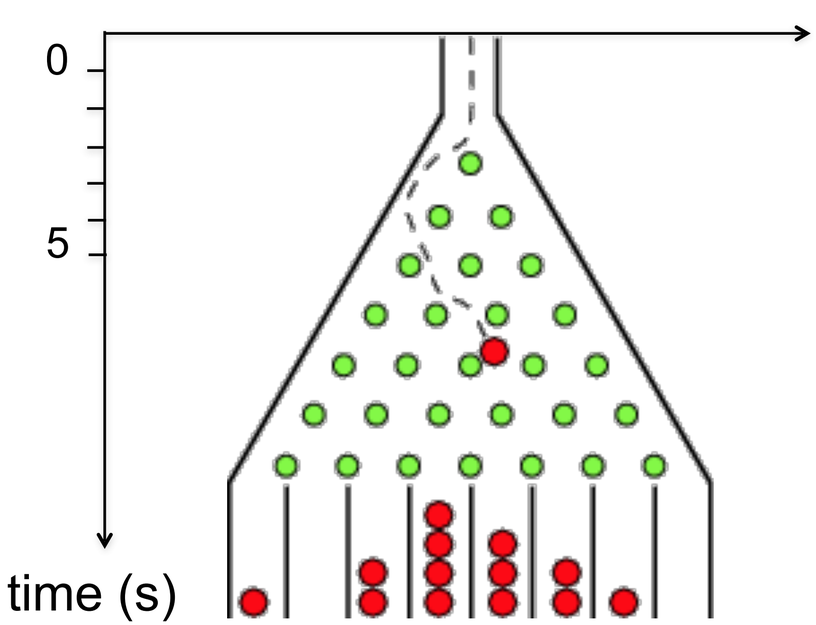


- Where is its most probable position?
- What is the most distant point it visited?

The Galton board

The “Random walk” experiment:

- Take spheres
- Throw them in a Galton board
- Look where they ended up



Spheres in red

Obstacles (nails) in green

Step width = distance between nails

A random walk in 1-dimension

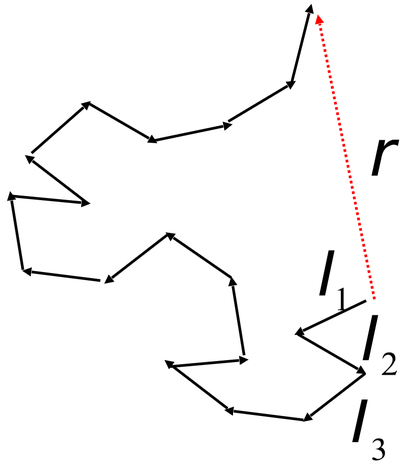
Random walk features

- The average value of *x* is zero (no net translocation/ directionality)
- The longer the time the broader the final distribution
- Width = Mean squared displacement (MSD) that equals $\langle r^2 \rangle \propto t$
- Particles explore positions between $x = 0 \ldots s \cdot t$ (with *s* being the step size), but efficiently positions up to $\sqrt{s \cdot t}$

$$\langle r^2 \rangle = 6Dt$$

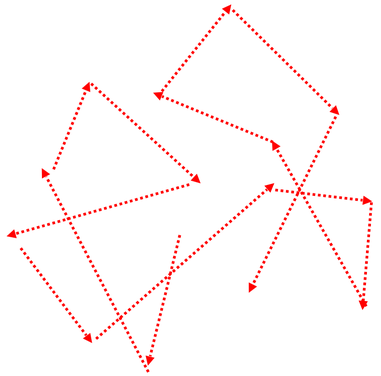
$$r = \sqrt{6Dt}$$

Diffusion in solution is a 3D random walk



$$r = \sum_{i=1}^{n} l_i$$

The displacement of the molecule after *n* steps is given by the vector *r*



$$\langle r \rangle = 0$$

The average of all vectors *r*, <*r*>, is zero

The average distance the molecule has travelled after a certain number of steps or time is larger than zero

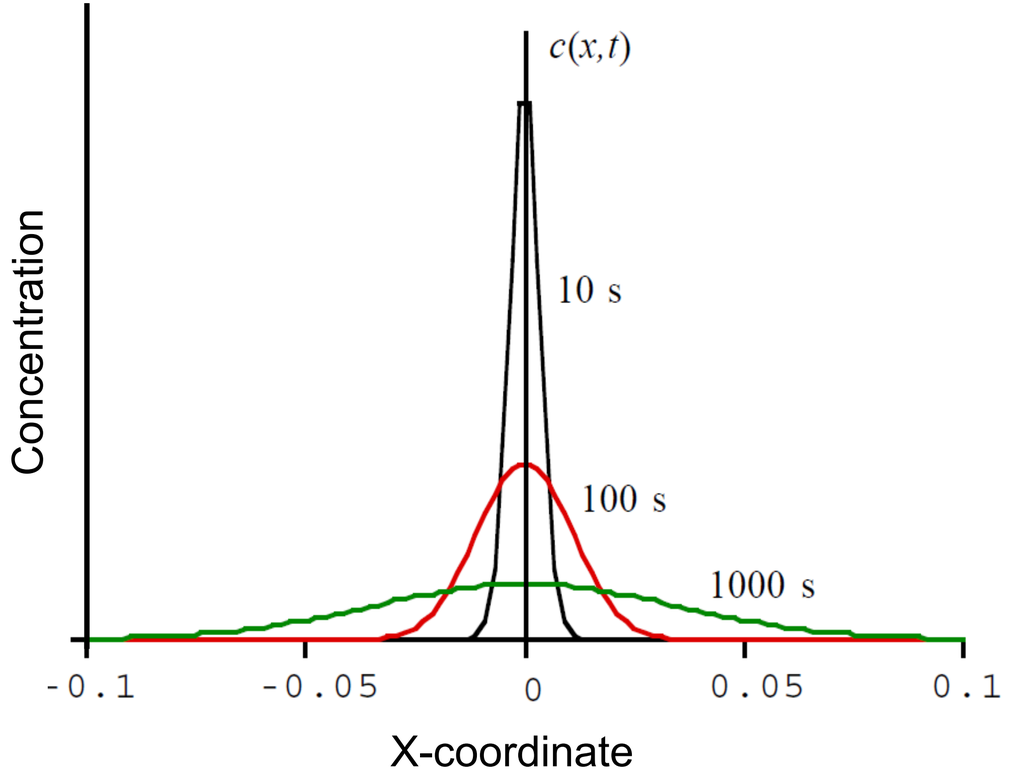
$$\langle r^2 \rangle = n \cdot l^2$$

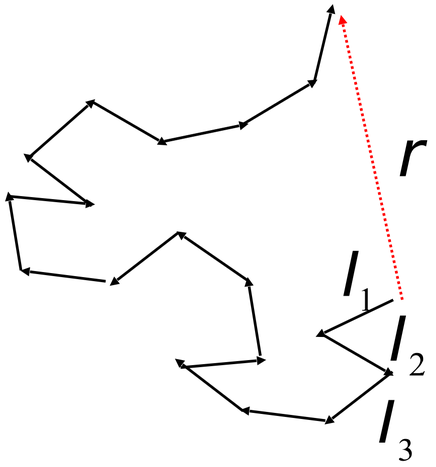
$$\sqrt{\langle r^2 \rangle} > 0$$

Particle transport by diffusion follows a random walk in time

Gaussian distribution

$$c(x,t) = \frac{1}{\sqrt{4Dt}} e^{-\frac{x^2}{4Dt}}$$





After *n* steps the molecule has moved a distance given by the vector *r*

mean squared displacement in three dimensions:

$$\left\langle r^2 \right\rangle_{x,y,z} = 6\,Dt$$

Displacement of proteins due to diffusion after a certain time

one dimension: $\left\langle d^2 \right\rangle_x = 2 \cdot D \cdot t$

two dimensions: $\left\langle d^2 \right\rangle_{x,y} = 4 \cdot D \cdot t$

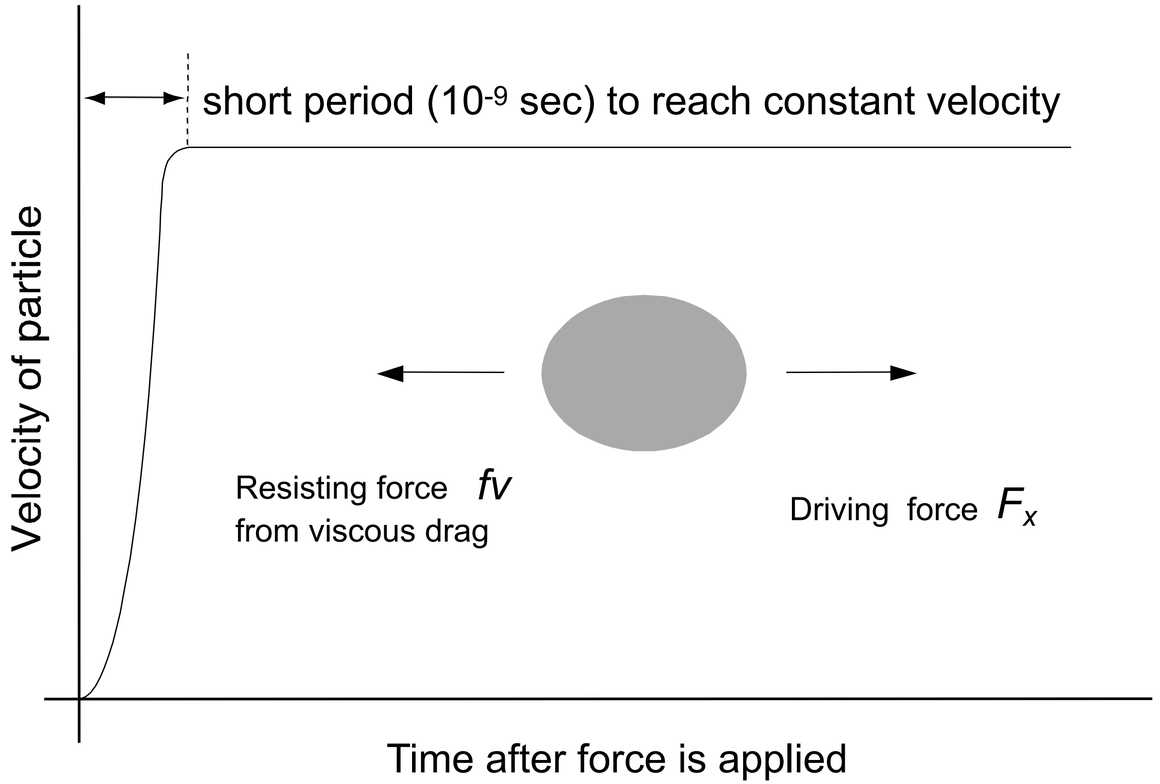
three dimensions: $\left\langle d^2 \right\rangle_{x,y,z} = 6 \cdot D \cdot t$

D (in $cm^2 \cdot s^{-1}$) is the diffusion coefficient
(reference state: pure water at 20 ˚C)

Lysozyme and other proteins (10 - 100 kDa):
$D = 10^{-6}\ cm^2 \cdot s^{-1} = 100\ \mu m^2 \cdot s^{-1}$

RNA polymerase II complex (2000 kDa): $D \approx 10^{-7}\ cm^2 \cdot s^{-1}$

Directed movement in solution induced by an external force (gravitation, centrifugation, electric field)



$$-f \cdot v + F_x = 0, \text{ or } F_x = fv$$

$\Rightarrow$ if we measure the velocity of motion produced by a known force we can determine the friction coefficient and diffusion coefficient

Parameters that describe the hydrodynamic properties of macromolecules in solution

Stokes-Einstein relation

$$D = \frac{k_B T}{f} = \frac{k_B T}{6\pi\eta r}$$

- Diffusion coefficient *D*
- Frictional coefficient *f*
- Hydrodynamic radius *r*: radius of a sphere that would have the same f or D as particle
- $k_B T$ (Boltzmann constant timers temperature): $4 \cdot 10^{-21}$ J at 25 C

- Sedimentation coefficient *s*

$$s = \frac{dr/dt}{\omega^2 r} = \frac{M \cdot (1 - \bar{v}\rho)}{N_A f_t}$$

- Partial specific volume *v* bar ($\bar{v}$)
  protein: 0.73 ml g^{-1}, DNA: 0.55 ml g^{-1}

$$\bar{v} = \frac{\partial v}{\partial m}$$

- Mass *M* of the molecules
- Density ρ of buffer, 0.9982 g ml^{-1} for water at 20 ˚C
- viscosity η of buffer, 1.002 mPa second for water at 20 ˚C

Smoluchowski limit

- For chemical reactions both reactants have to collide, which they do during their random walk
- Thus, the maximum reaction rate is determined by the diffusion coefficients:

A + B → C

$$k_{\max} = (D_A + D_B)(r_A + r_B)N_0 / 1000$$

The Smoluchowski diffusion limit of a bimolecular reaction

$$k_{\text{encounter}} = 4\pi \left(D_A + D_B\right)\left(r_A + r_B\right) N_0$$

N_0 ← Avogadro's number

$k_{\text{encounter}}$ in $M^{-1}\ s^{-1}$

$D_A + D_B$: diffusion coefficients

$r_A + r_B$: reaction radii

$$A + B \xrightarrow{k_{\text{encounter}}} AB$$

$$r_A = r_B \text{ and } D_A = D_B = \frac{k_B T}{6\pi\eta r}$$ for two identical particles

$$k_{on} = \frac{8}{3} \cdot \frac{k_B T}{\eta} \cdot N_0$$

independent of radius r

$$k_{on} = 6.4 \cdot 10^9\ \mathrm{M^{-1}\,s^{-1}}$$

$k_B T / \eta = 4 \cdot 10^{21}\ \text{J} / 1 \cdot 10^{-3}\ \text{Pa s}$

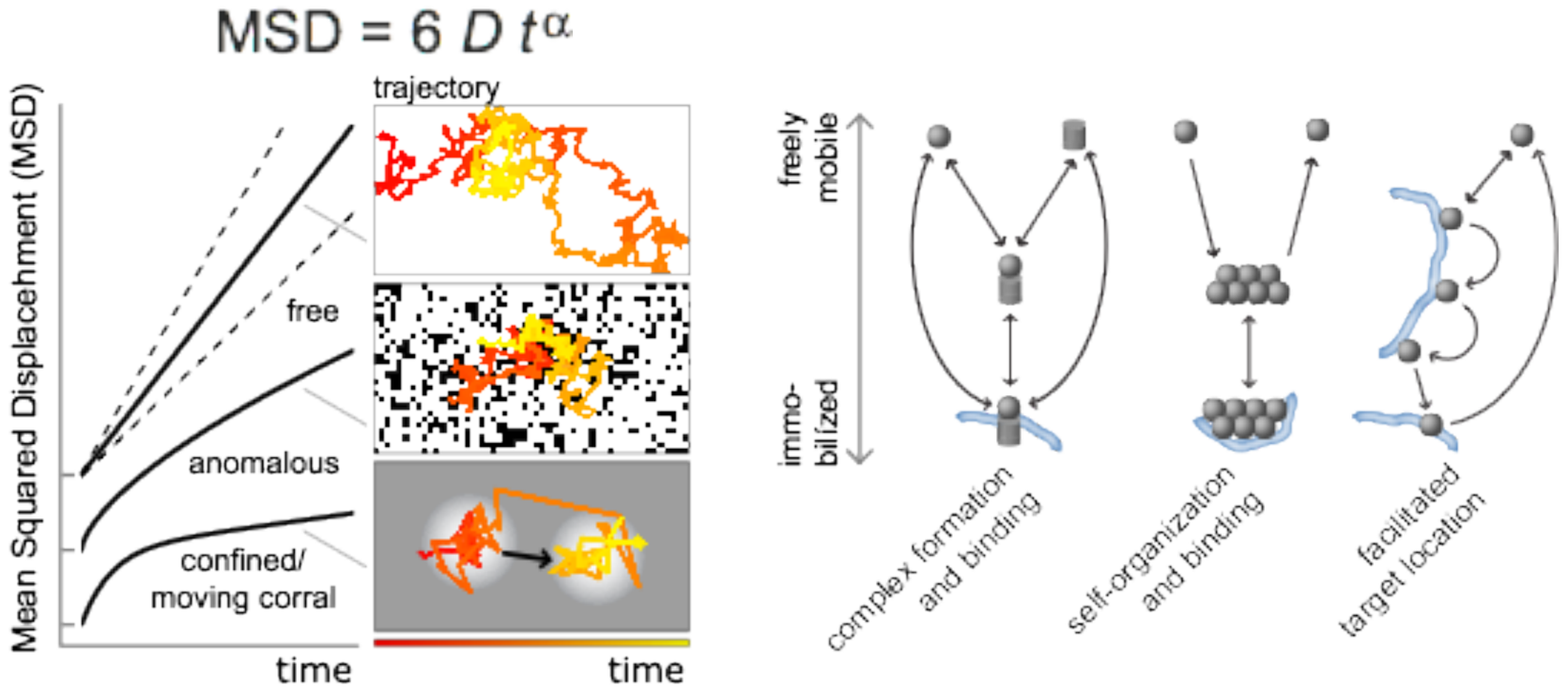
$= 4 \cdot 10^{-18}\ \text{m}^3\ \text{s}^{-1} = 4 \cdot 10^{-15}\ \text{liter s}^{-1}$

$N_o = 6.022 \cdot 10^{23}\ \text{mol}^{-1}$

Summary I

- Thermal energy is very important in cellular systems (= 10 μm scale)
- All macromolecules diffuse quickly but can bind very strongly to immobile structures like chromatin or membranes
- Friction (determined by shape) instead of mass determines particle mobility in the cell
- Without mediating transport in a particular direction diffusion allows for exploring the environment
- Diffusion limits the maximum speed of reaction to the "Smoluchowski limit" of around 10^9 M-1 cm-1 where every collision leads to a reaction product.

Mean squared displacement (MSD) and protein mobility



Dependence of diffusion coefficient *D* and molecular mass *M*

protein: $D \propto M^{-\frac{1}{3}}$

double mass *M* => 0.8 fold lower *D*

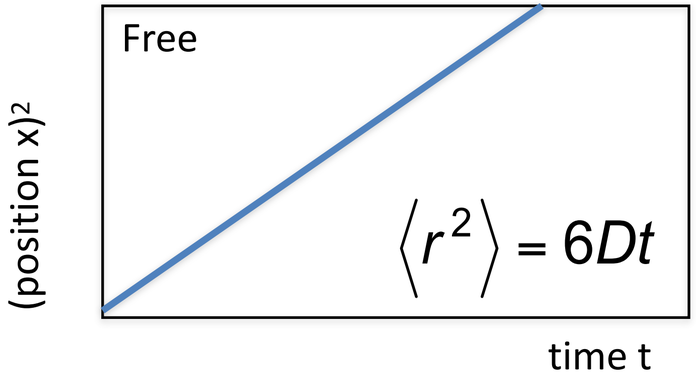
DNA: $D \propto M^{-\frac{1}{2}}$

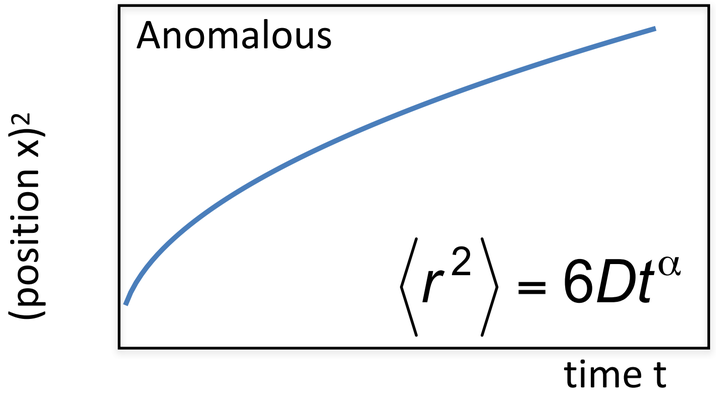
double mass *M* => 0.7 fold lower *D*

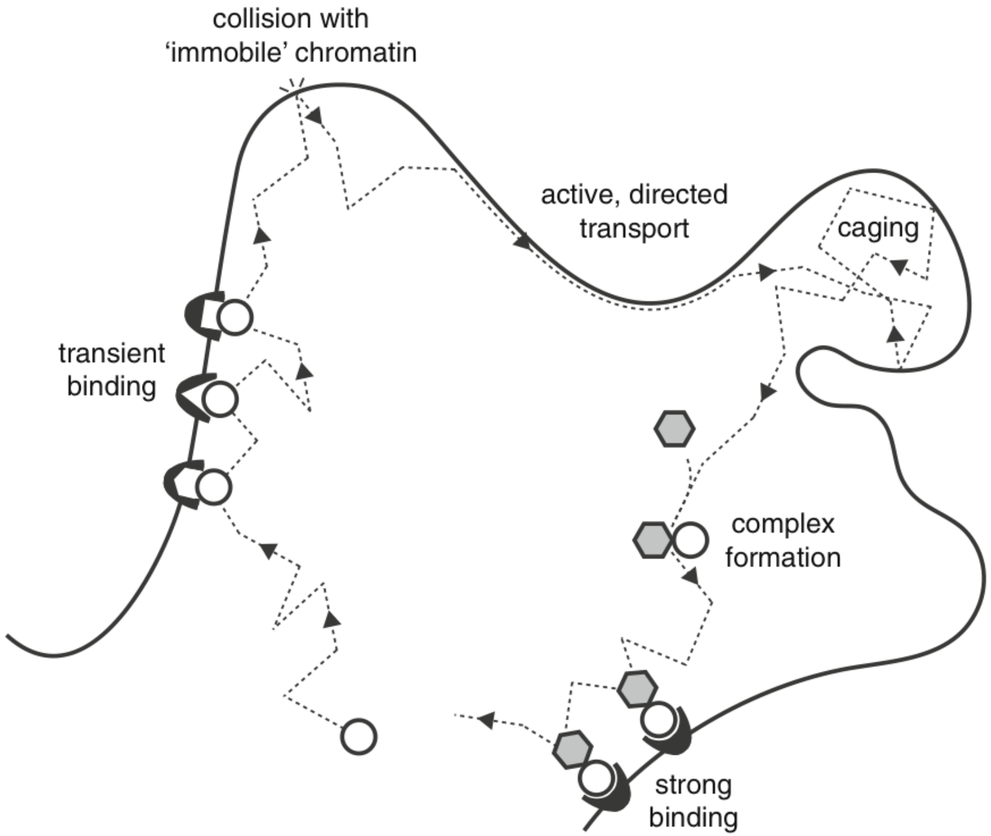
Wachsmuth, M., Caudron-Herger, M. and Rippe, K. (2008). *Biochim. Biophys. Acta* **1783**, 2061-2079.

Anomalous diffusion

Free diffusion vs anomalous/obstructed diffusion

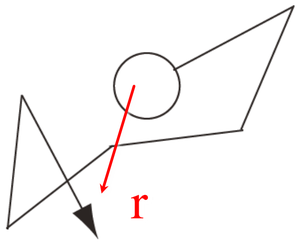






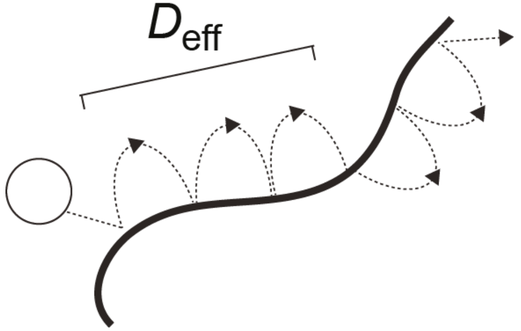
Determining diffusion coefficient D, kinetic binding rates k_{on} and k_{off}, and the apparent equilibirium constant K_{eq}^*

diffusion without binding,
α = 1 for free diffusion



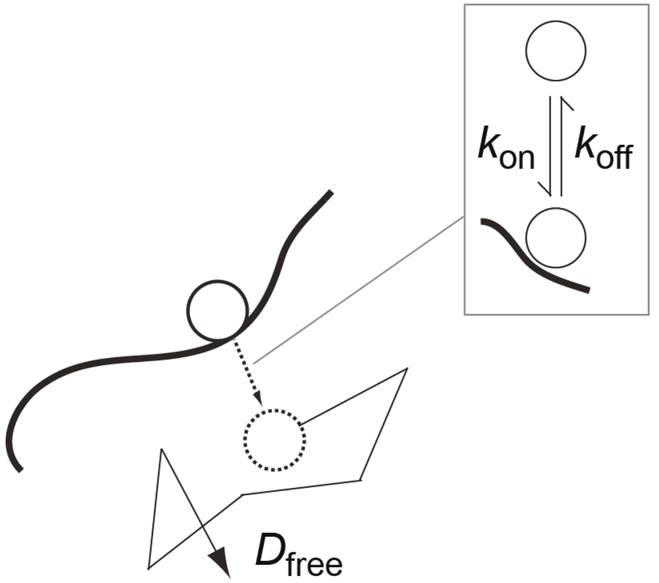
$$\left\langle r^2 \right\rangle = 6Dt^{\alpha}$$

transient chromatin binding



$$D_{eff} = \frac{D}{1 + K^*_{eq}}$$

strong chromatin binding



$$K^*_{eq} = \frac{k^*_{on}}{k_{off}} = \frac{k_{on} \cdot [S]_{eq}}{k_{off}}$$

Erdel, Müller-Ott, Baum, Wachsmuth & Rippe (2011) *Chromosome Res* **19**, 99–115.

Finding home …

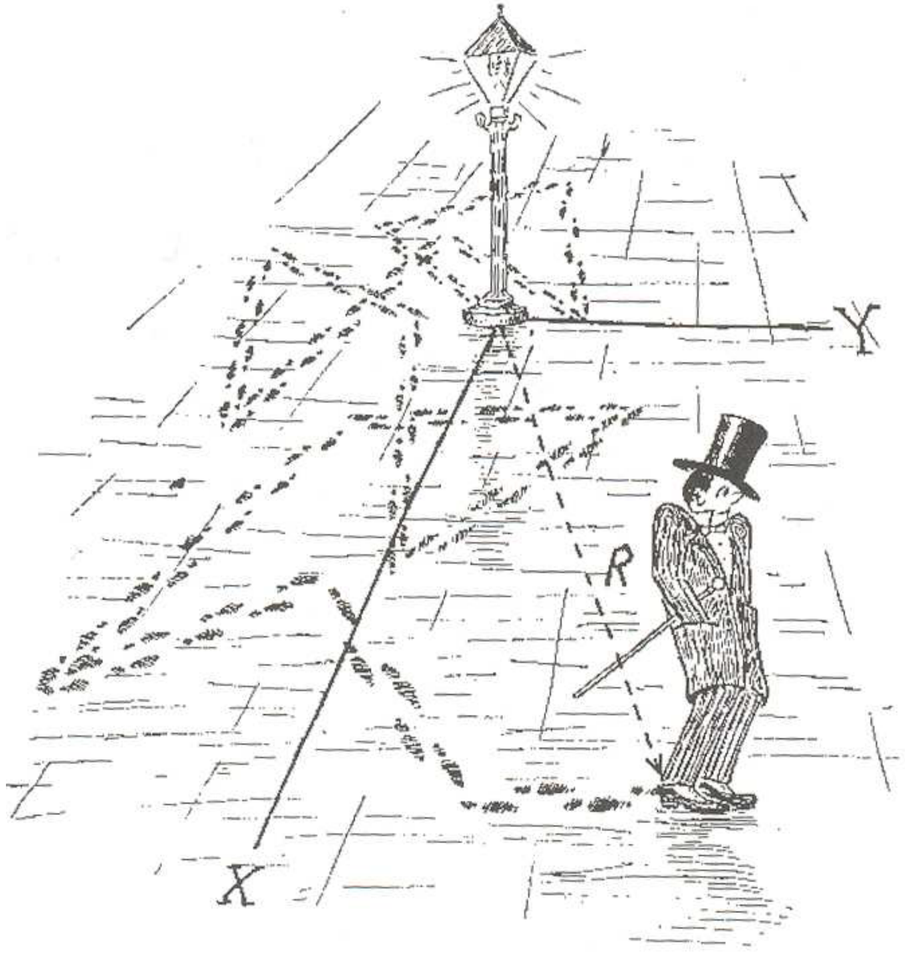
Drunkard:

“Will I ever, ever get home again?”

George Pólya (1921):

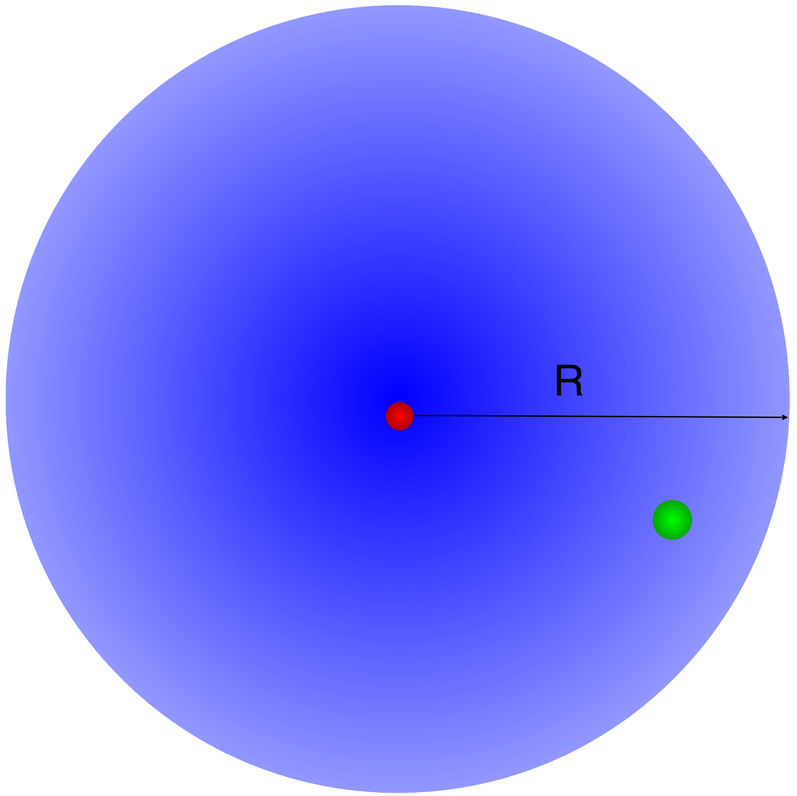
“You cant’t miss; just keep going

and stay out of 3D!”



G. Gamow, One, Two, Three…Infinity,
The Viking Press, New-York, 1955

The mean diffusion time to reach a small target of radius r in the middle of a cell of radius R with R >> r



one dimension: $$\tau_1 = \frac{R^2}{3D_1}$$

two dimensions: $$\tau_2 = \frac{R^2}{3D_2} ln\left(\frac{R}{r}\right)$$

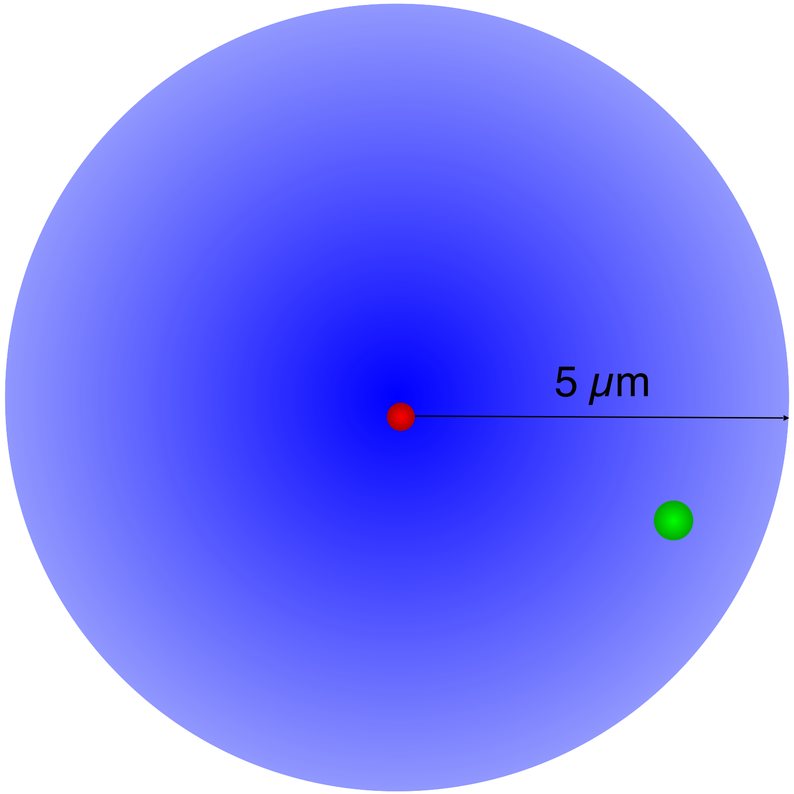
three dimensions: $$\tau_3 = \frac{R^2}{3D_3} \cdot \frac{R}{r}$$

Protein with diffusion constant D

target with radius r

Berg, O.G., and von Hippel, P.H. (1985). Diffusion-controlled macromolecular interactions. Annu Rev Biophys Biophys Chem *14*, 131-160.

How long would it take on an average for RNA polymerase II to find its target promoter?



one dimension: $\tau \approx 4$ sec

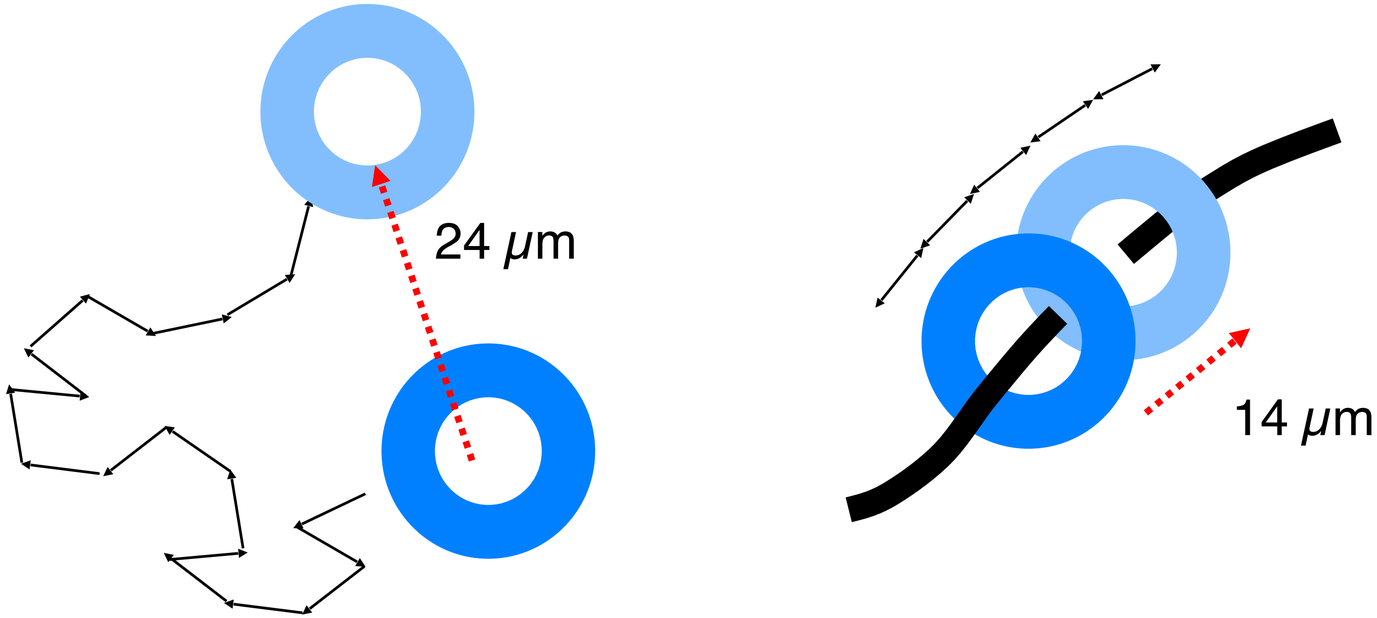
two dimensions: $\tau \approx 40$ sec

three dimensions: $\tau \approx 2000$ sec

Pol II complex ($M \approx 2\,000$ kDa)
$D_{eff} \approx 2 \cdot 10^{-8}$ cm$^2 \cdot$s^{-1}

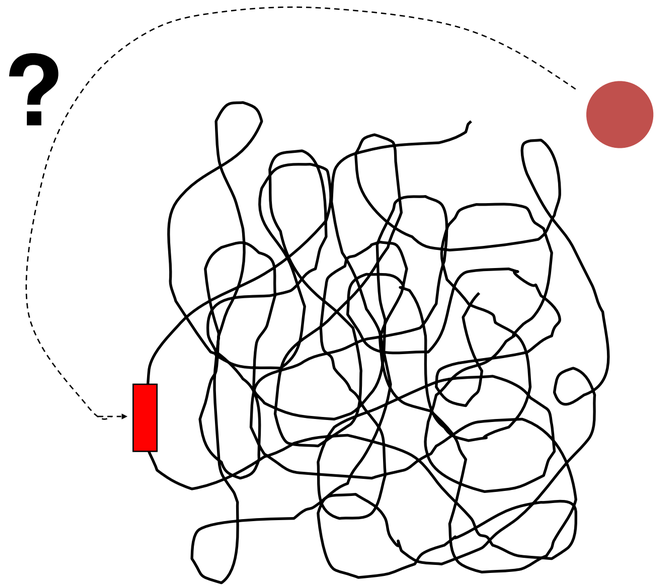
target promoter ($r = 5$ nm)

For a protein with D = 10^{-6} $cm^2 \cdot s^{-1}$ the displacement after 1 sec would be 24 µm in 3 dimensions and 14 µm in 1 dimension



Facilitated diffusion of proteins on DNA

the problem: How will a protein find its binding site on a long DNA?



Model for facilitated diffusion of lac repressor

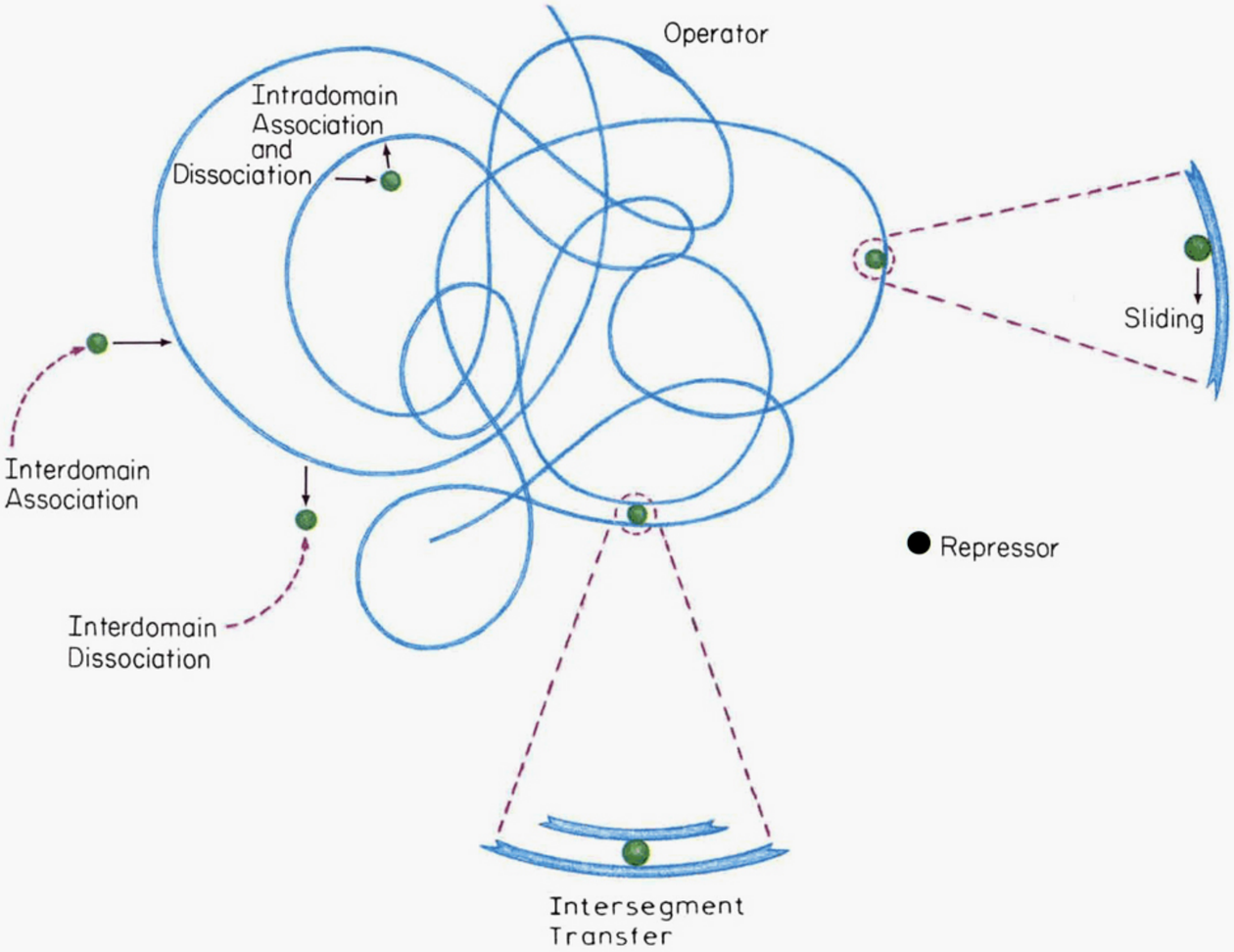
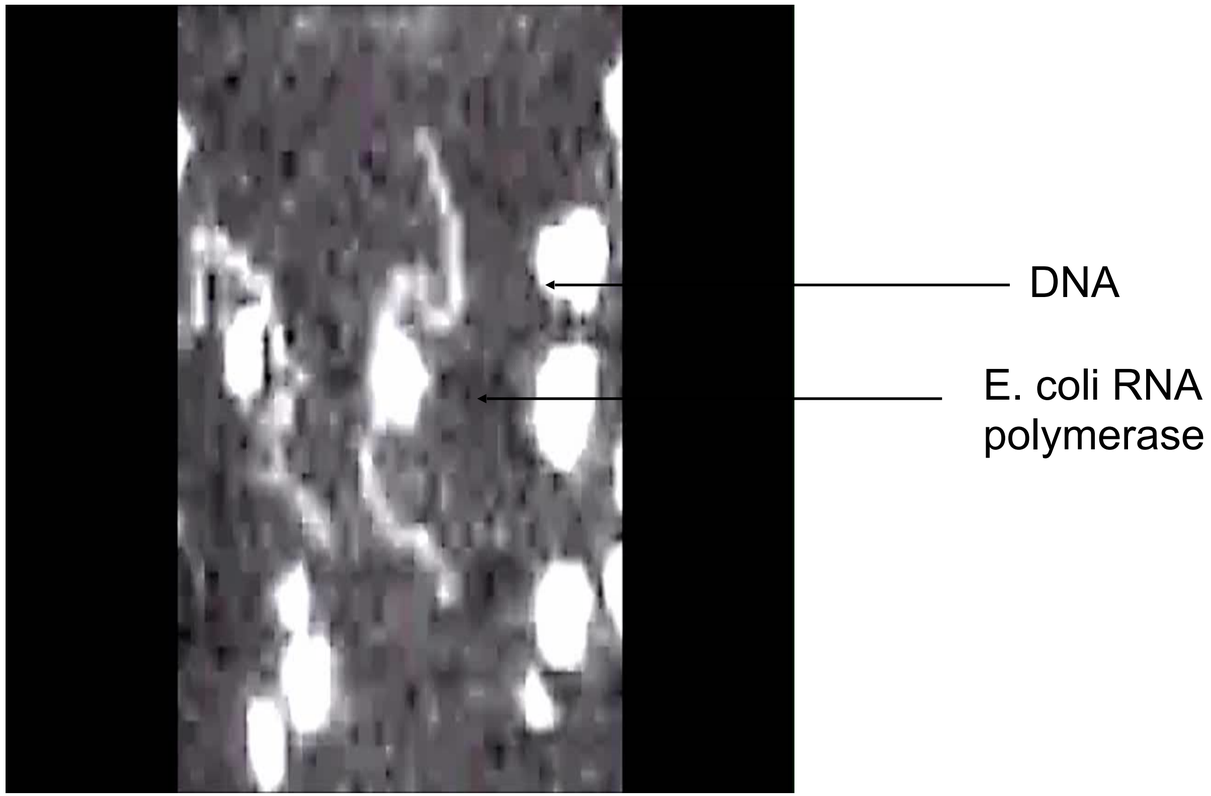


FIG. 1. **Schematic view of *lac* repressor interacting with a large operator-containing DNA molecule in dilute solution.** (The DNA molecules are well separated into "domains" under these conditions.) The (*upper*) expanded view shows repressor bound to a segment of non-operator DNA, on which it can either "slide" or engage in intradomain dissociation-association processes in seeking its specific (operator) target site. The (*lower*) expanded view shows a repressor molecule double bound to two DNA segments; this corresponds to the intermediate state in the intersegment transfer process.

RNA polymerase finds its promoter by “sliding” along the DNA as visualized by microscopy



Guthold, M. et al. (1999). Direct observation of one-dimensional diffusion and transcription by Escherichia coli RNA polymerase. Biophys J 77, 2284-2294.

Summary II

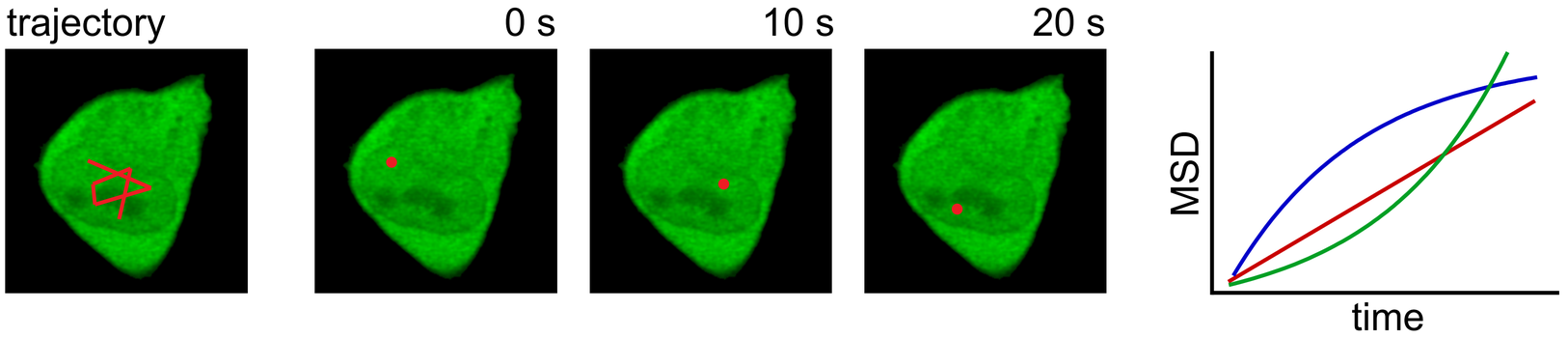
- Free diffusion is fast (seconds) on the length scale of the cell (μm) for translocations but finding a target site in 3D can be slow (minutes to hours)
- Target search can become much faster if the space that has to be searched by the protein is reduced. This can be accomplished by restricting the accessible space or by reducing the dimensionality of the search process from 3D to 1D or 2D.
- Proteins like *lac* repressor or RNA polymerase can "slide" by 1D diffusion along DNA, which could speed up finding their specific bindings sites.

Single Particle tracking (SPT): nuclear bodies, chromatin loci, proteins, RNA

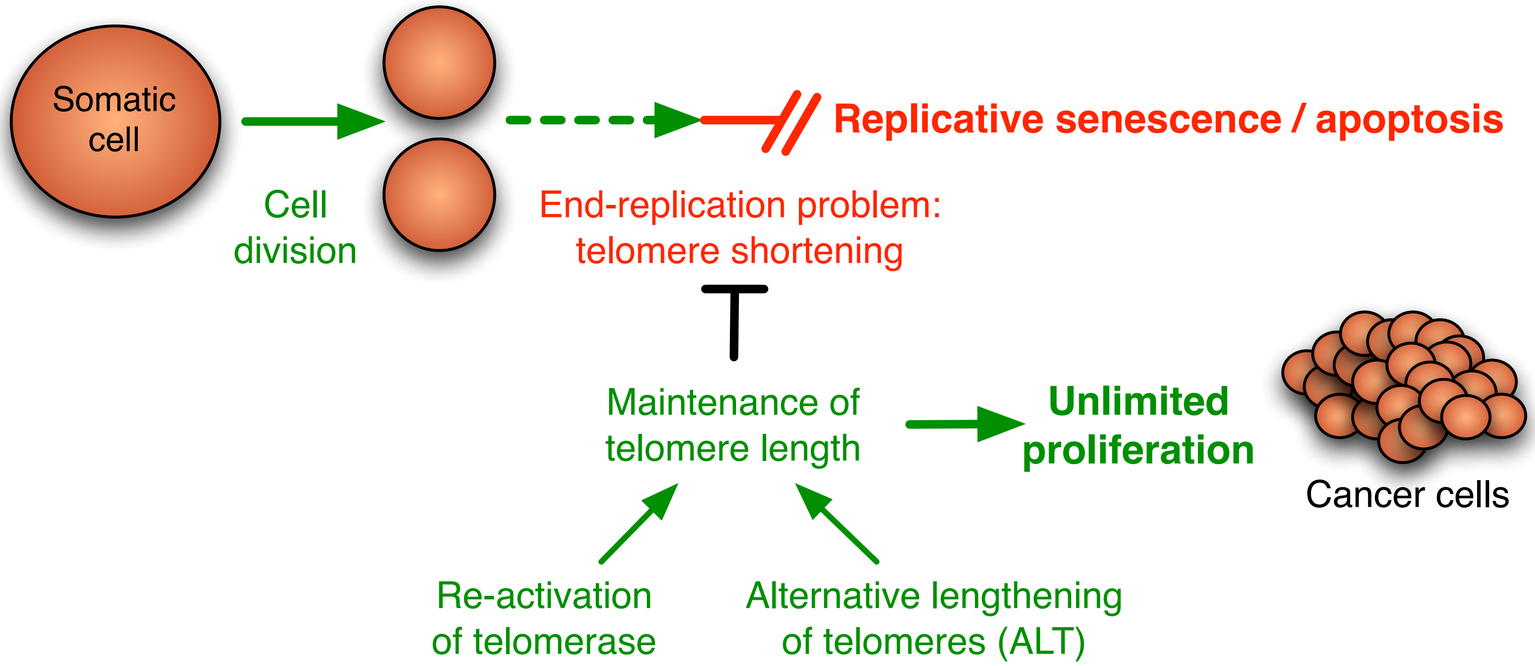
Easiest approach to measure mobility: Directly watch single particles over time but you need

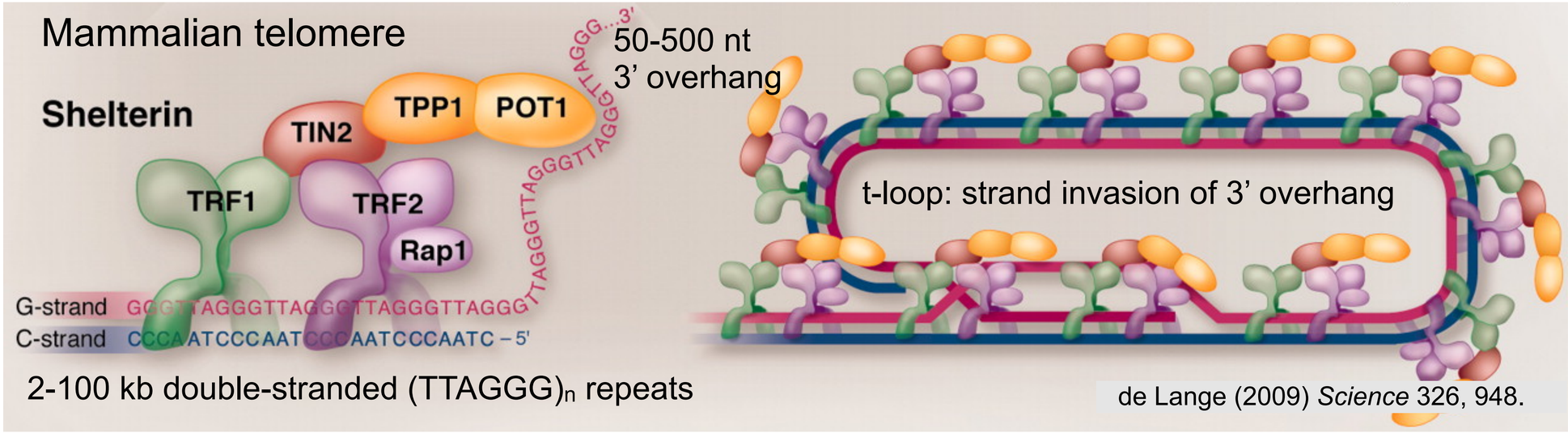
- Low concentration → Individual particles
- Bright particles → detect single particles (molecules)
- Stable fluorescence → sufficient
- High spatial resolution (~20 nm) → resolve trajectory in space
- High temporal resolution (~50 ms) → resolve trajectory in time

SPT

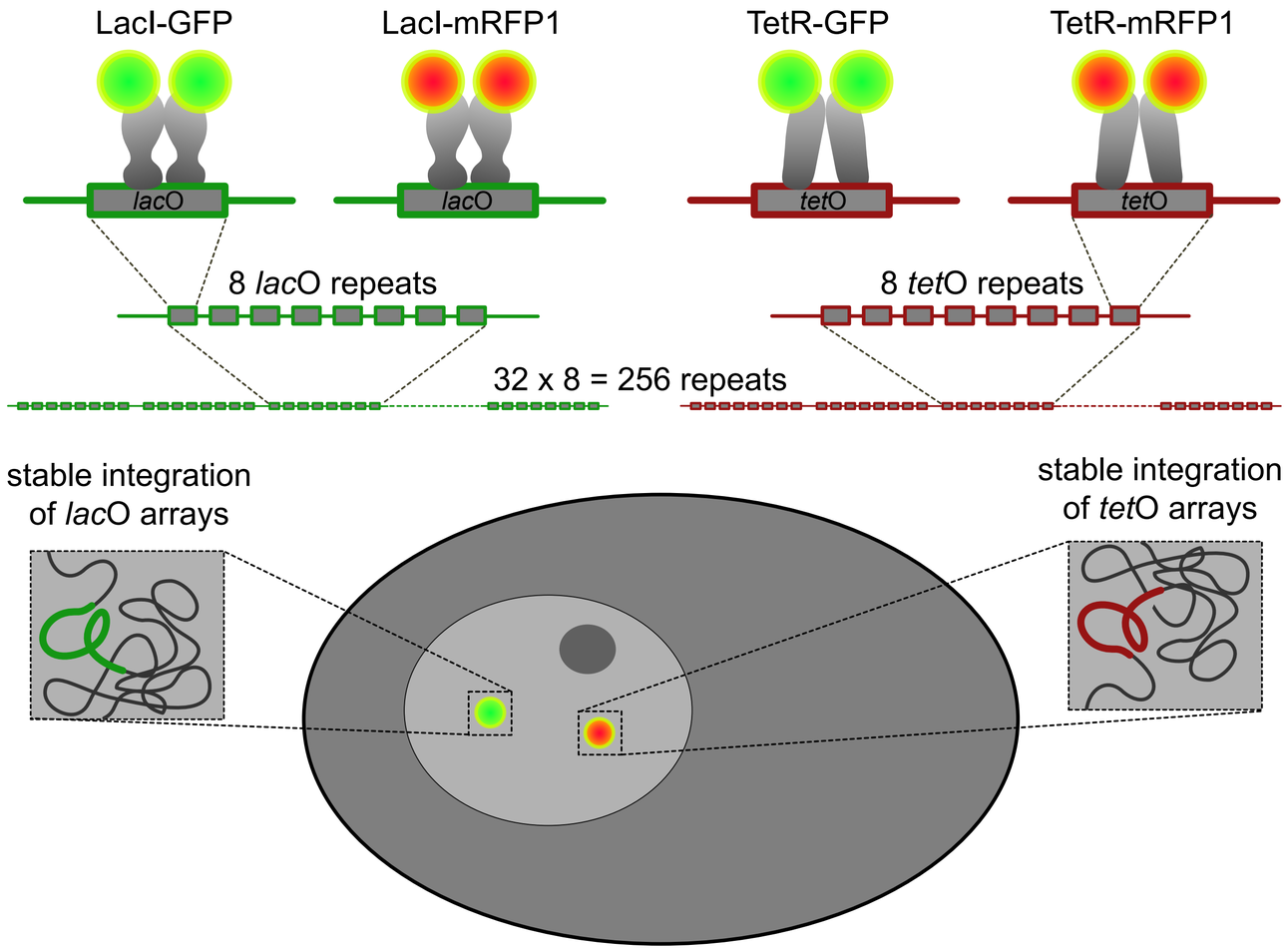


Tracing telomeres in living cells



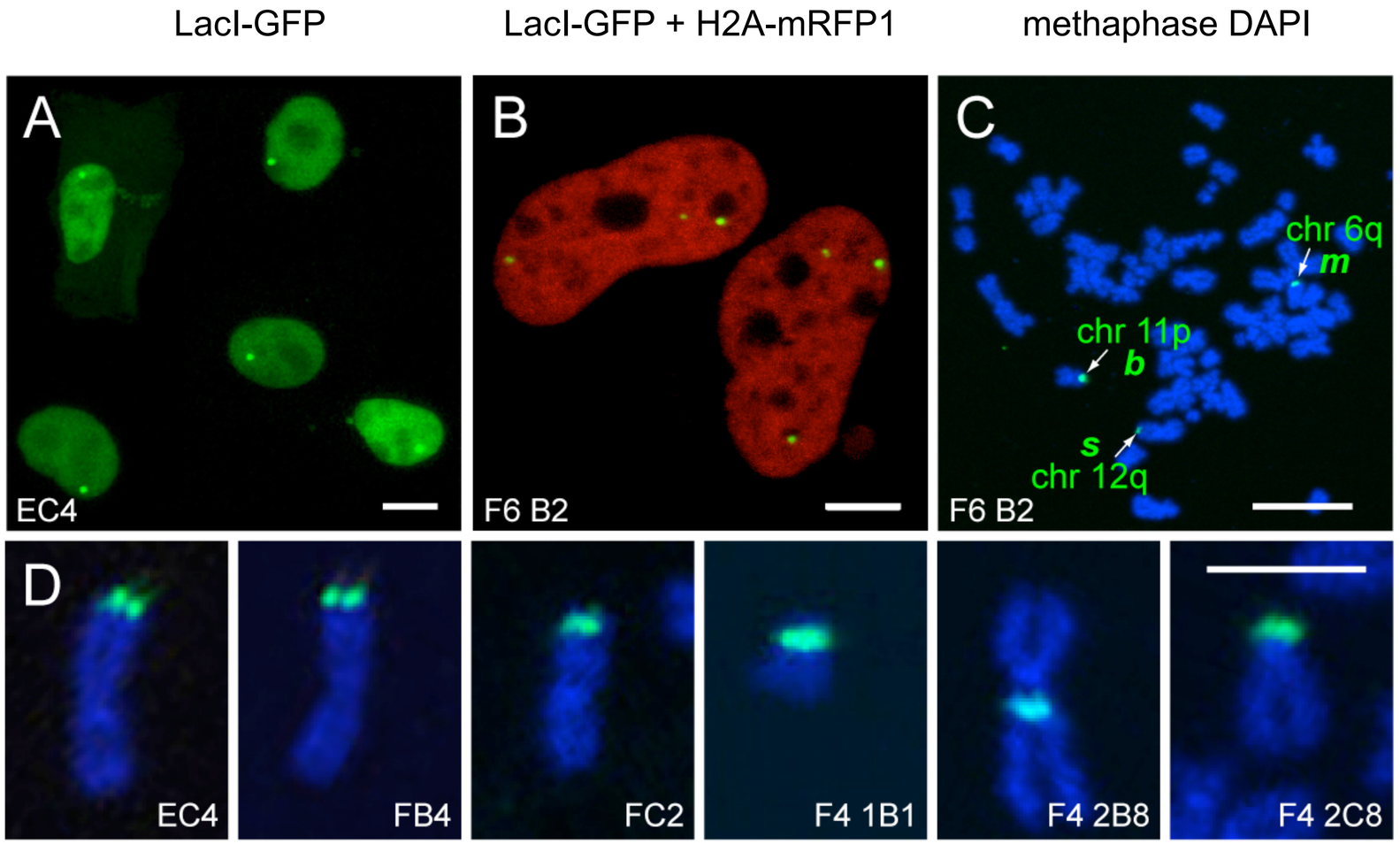


Tracing specific telomeres in living cells to study their dynamics



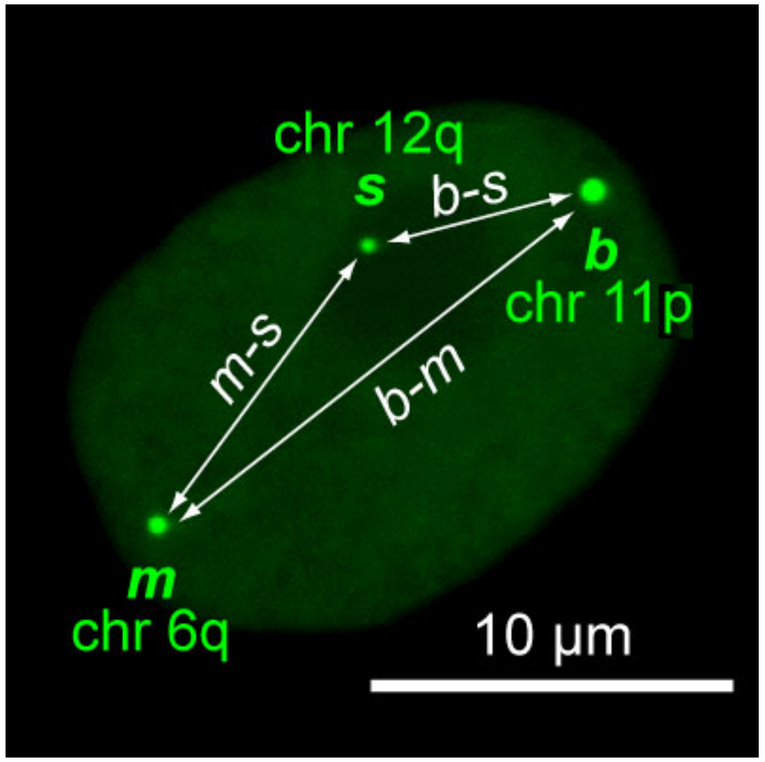
Jegou et al. (2009). *Mol. Biol. Cell* **20**, 2070.

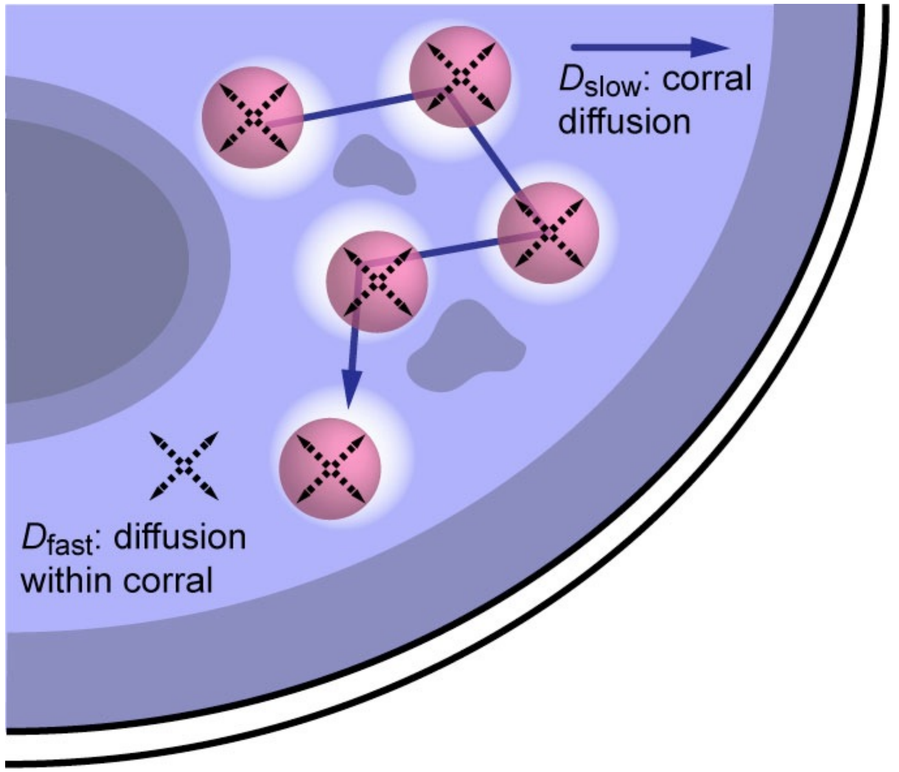
In vivo labeling of telomeres in human osteosarcoma U2OS cells with LacI-GFP via integrated *lac*O repeats



Metaphase FISH reveals preferred *lac*O integration into telomeres

The telomere mobility is derived from distance changes between two loci according to a “moving corral” model

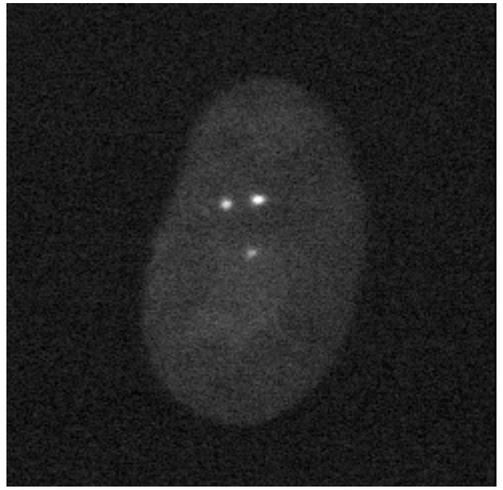




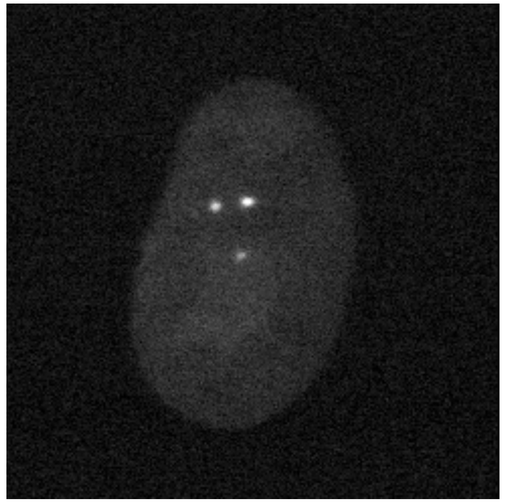
$$MSD = \left\langle r_c^{\,2} \right\rangle \cdot \left(1 + \frac{2n\, D_{slow}\, \Delta t}{\left\langle r_c^{\,2} \right\rangle} \right) \cdot \left[1 - \exp\left(- \frac{2n\, D_{fast}\, \Delta t}{\left\langle r_c^{\,2} \right\rangle} \right) \right]$$

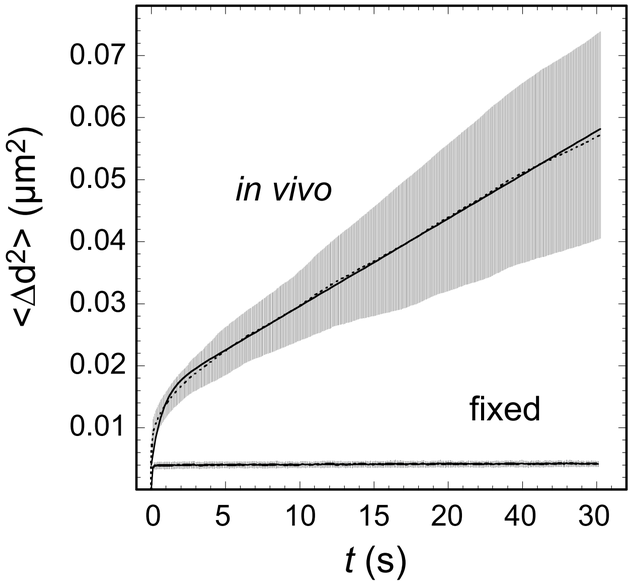
Mobility of telomeres at the second and the minute time scale

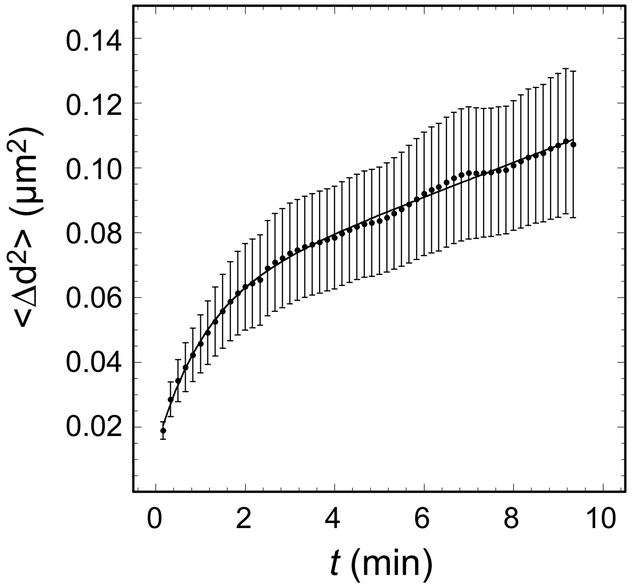
second scale
“real time” (Δt = 70 msec)



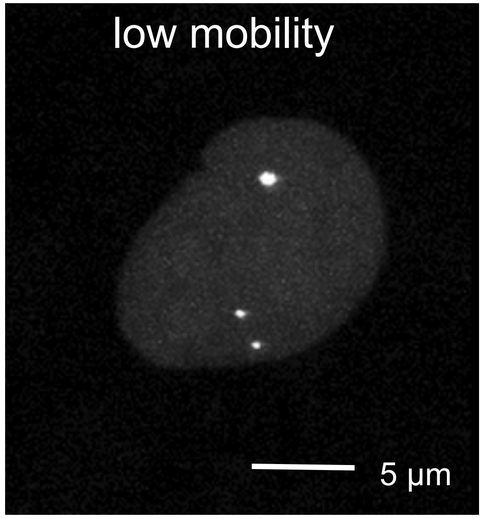
minute scale
25x higher speed

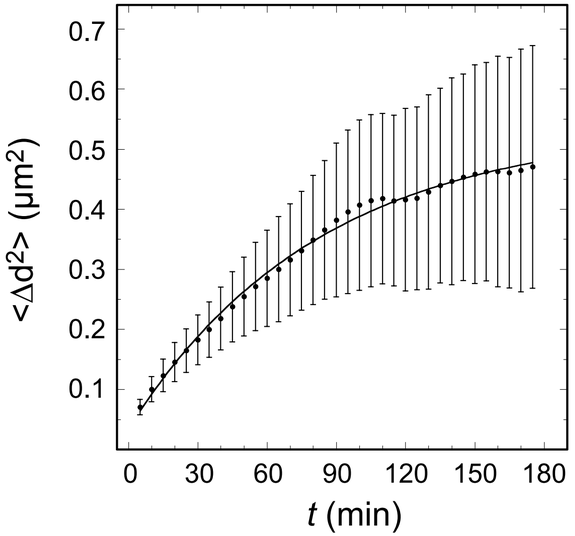


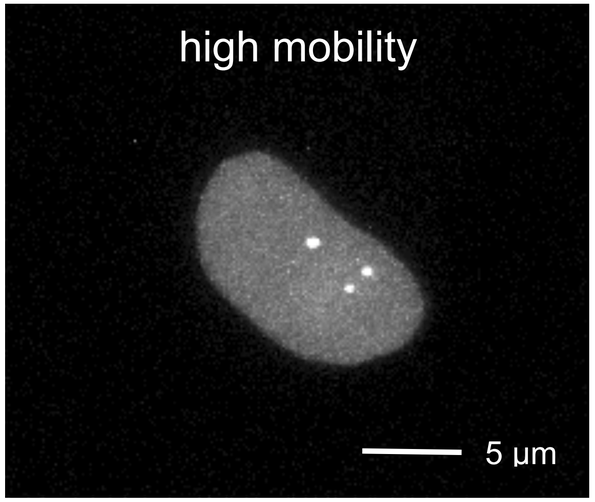


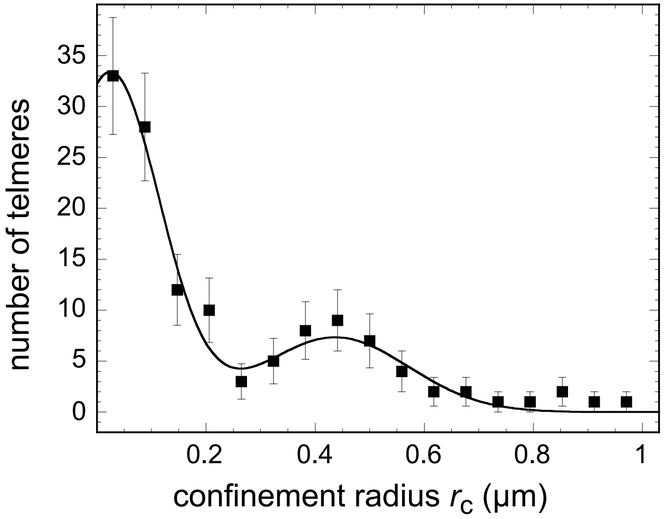


Mobility measurements over ~3 h (looped, 700x higher speed) show a state of extended mobility for some telomeres

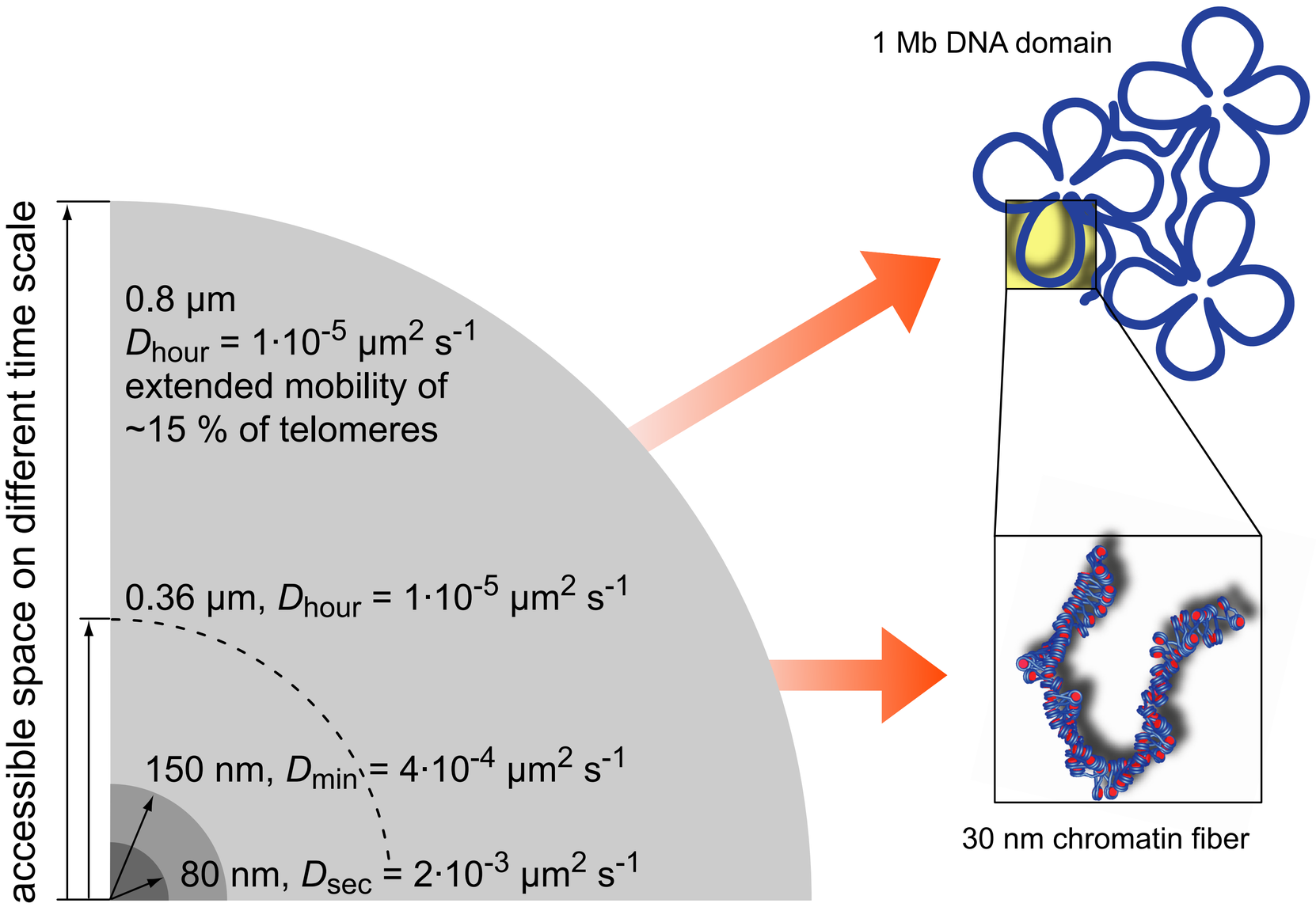




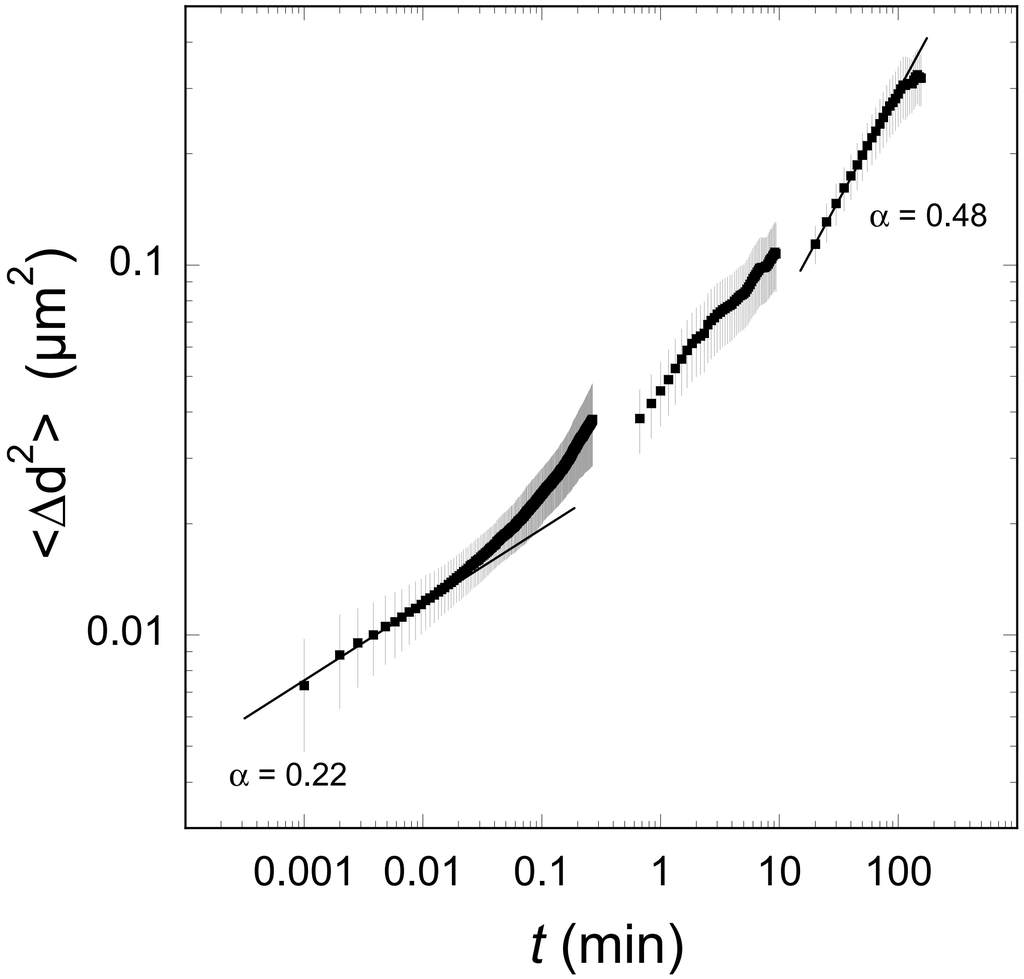




Telomere mobility over different time scales

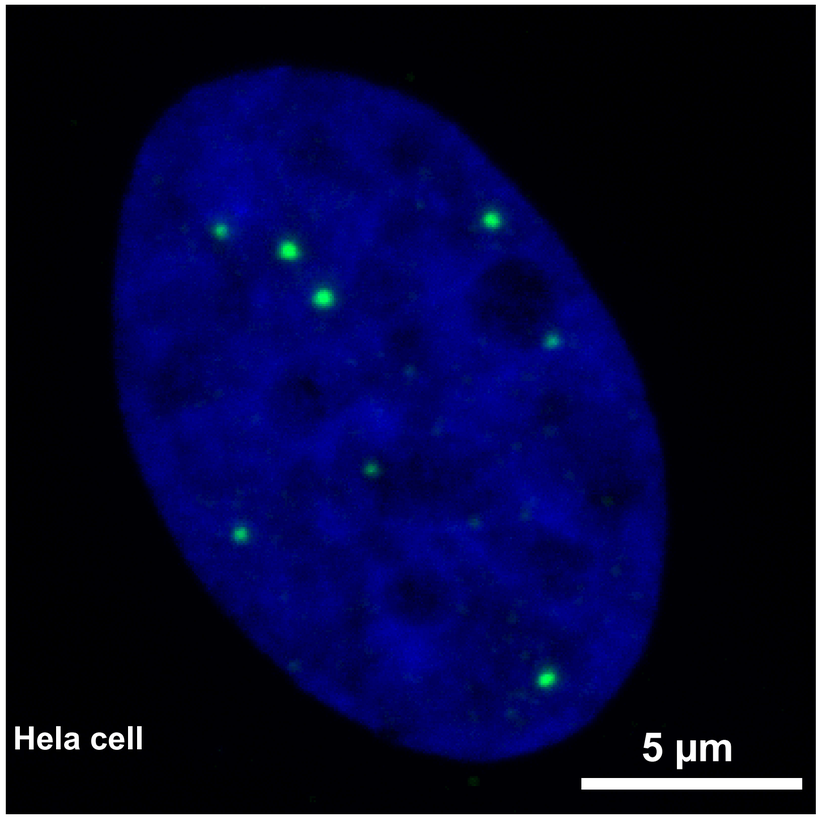


Telomere mobility is that of a polymer in a crowded environment according to the “reptation model”

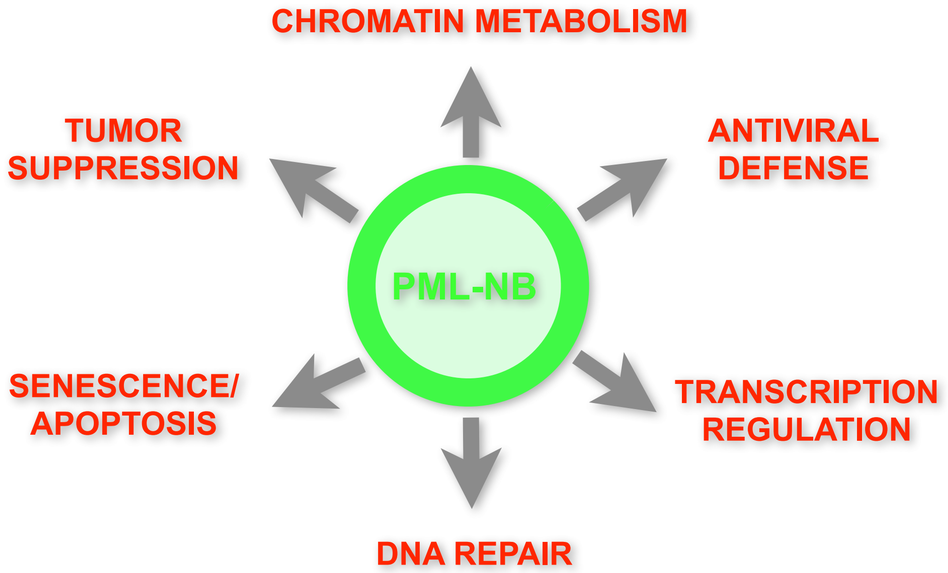


$$\mathrm{MSD} = \langle d^2(\Delta t) \rangle = 2n\Gamma\Delta t^{\alpha}$$

The promyelocytic leukemia (PML) nuclear body

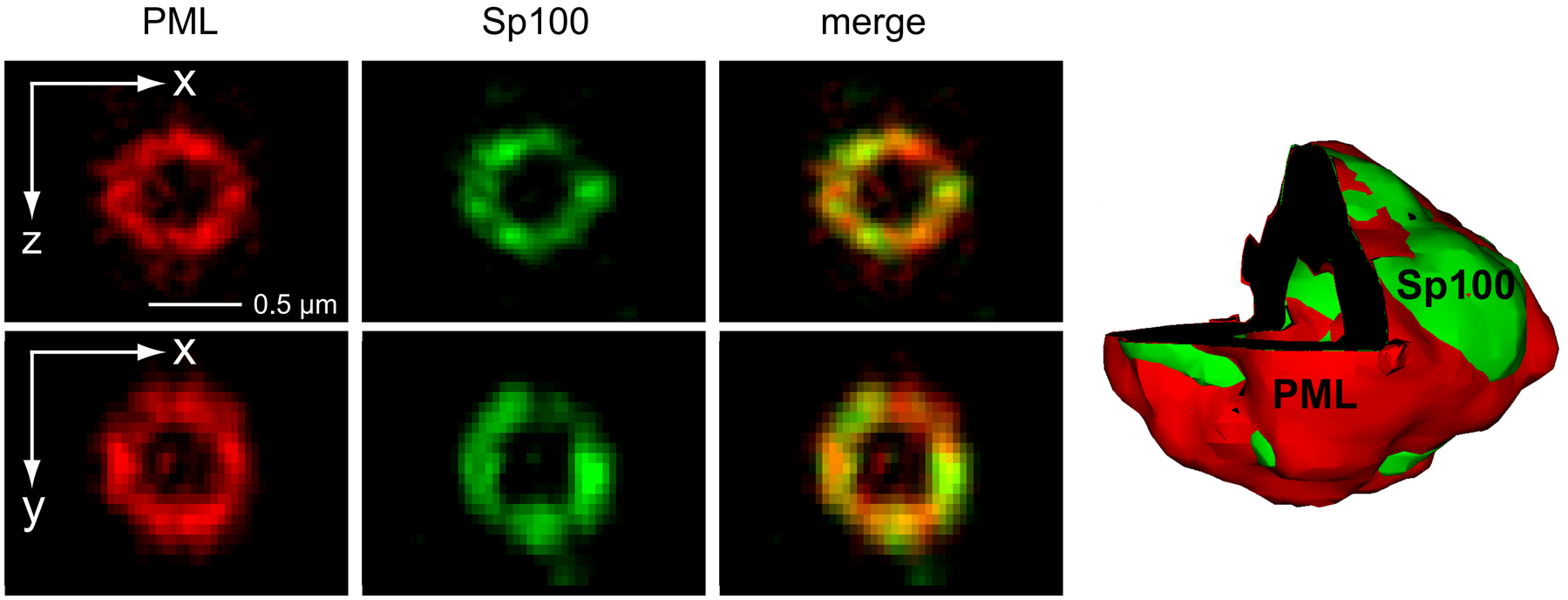


blue: DAPI; green: anti PML immunostaining

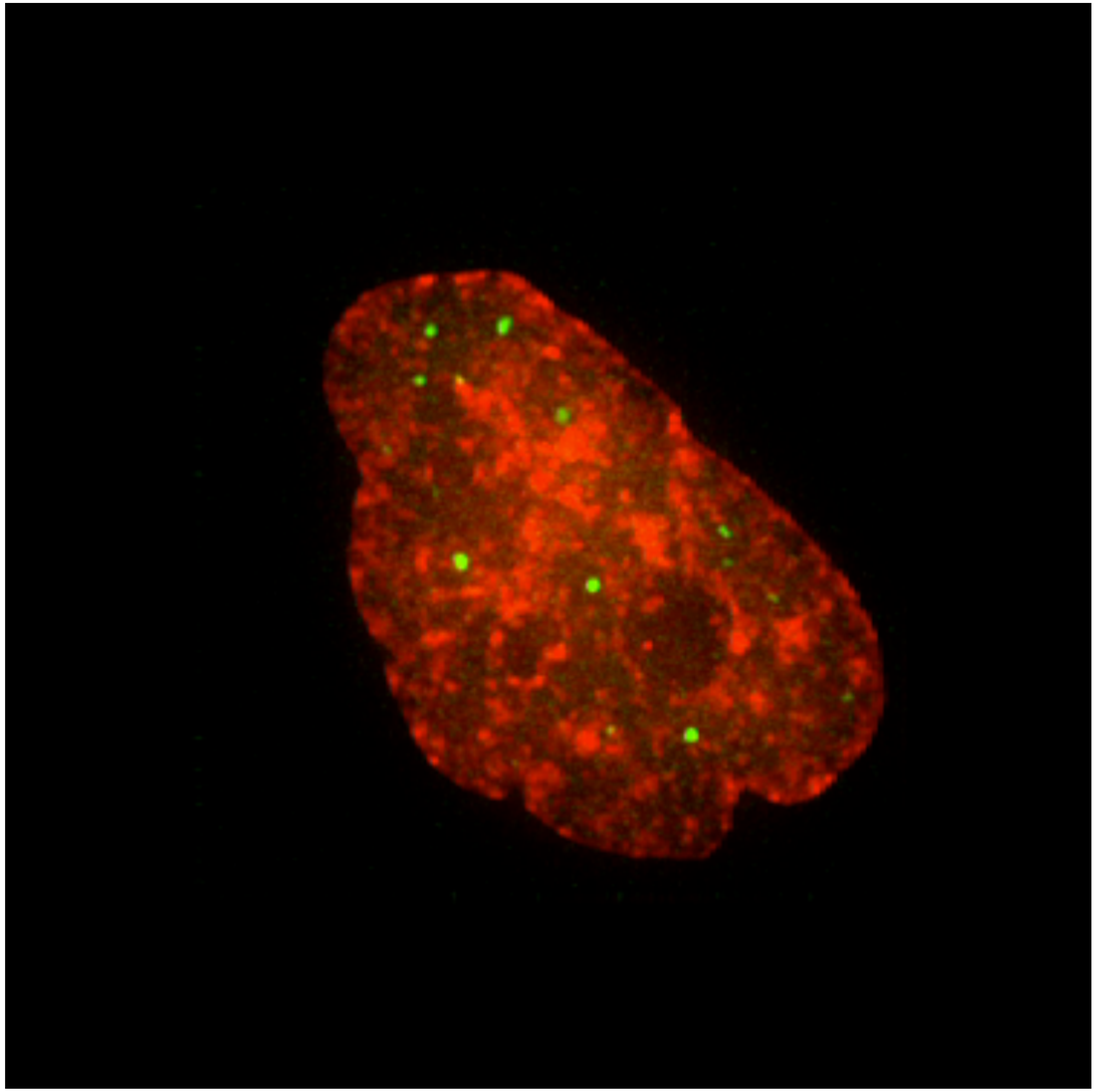


Görisch, Wachsmuth, Ittrich, Bacher, Rippe & Lichter (2004). Nuclear body movement is determined by chromatin accessibility and dynamics. *Proc Natl Acad Sci USA* **101**, 13221–13226 (2004).

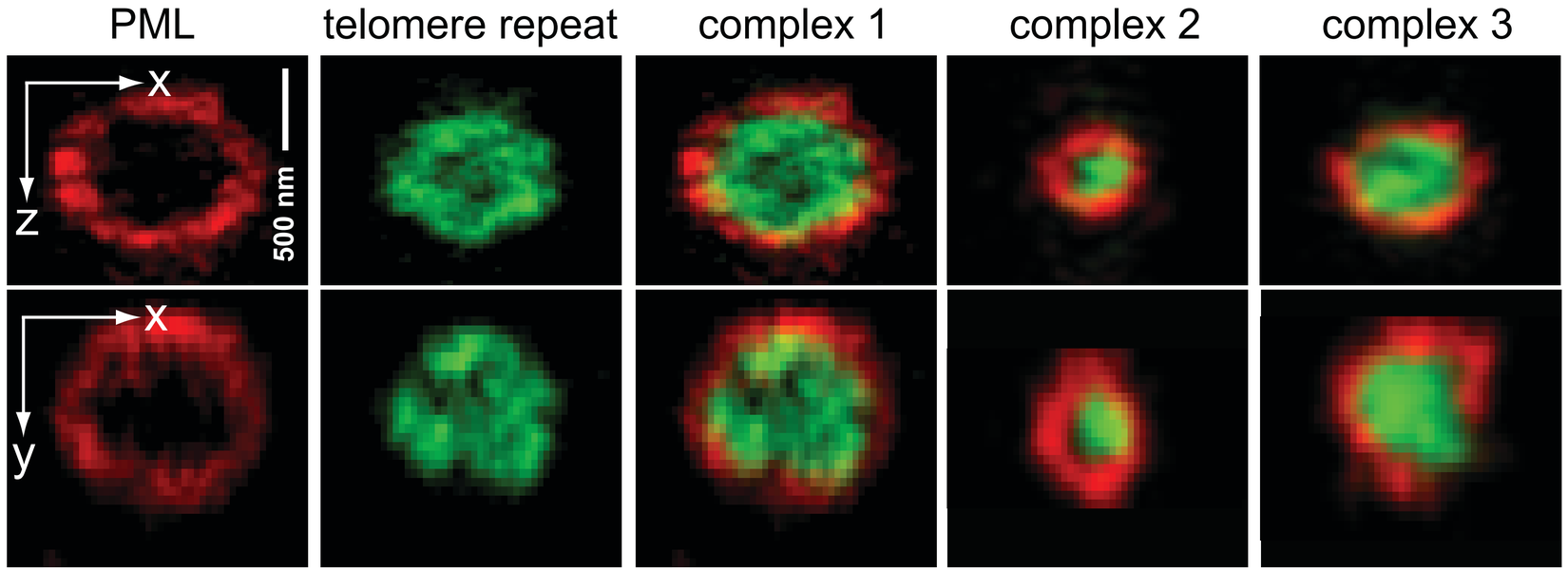
PML and Sp100 proteins form distinct patches in the spherical shell of the PML nuclear body

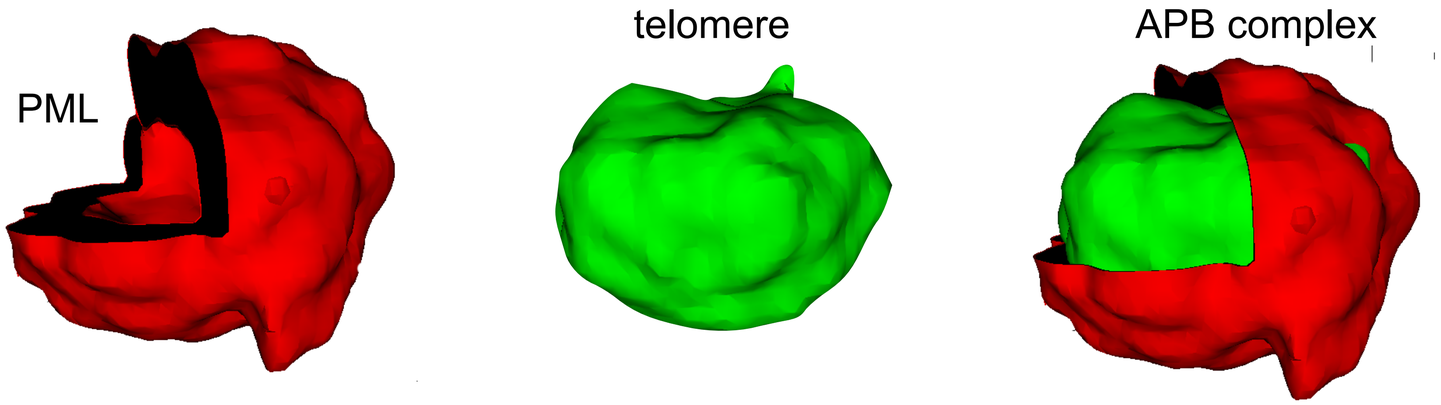


Movements of PML bodies (green) in the nucleus



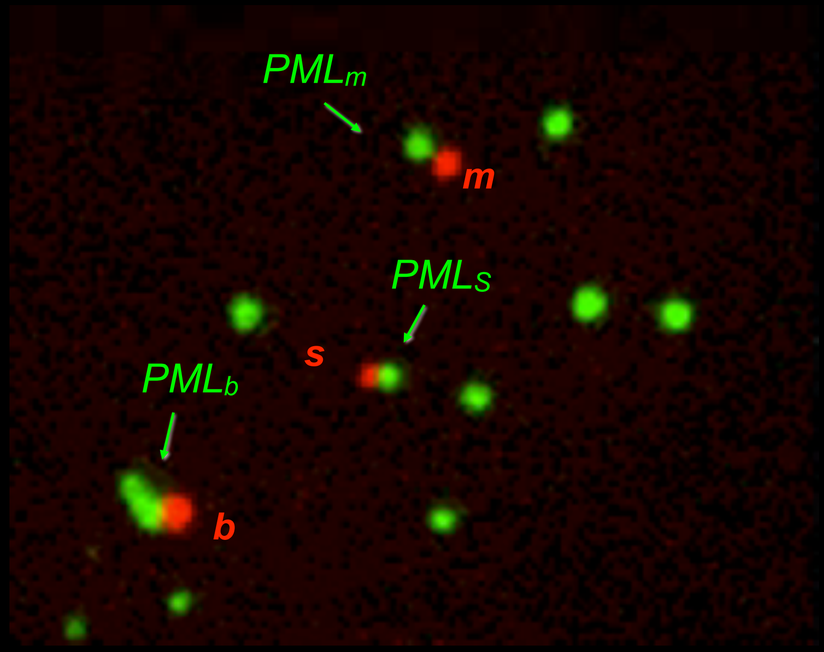
High resolution fluorescence microcopy images of PML nulcear bodies at telomeres in human cells





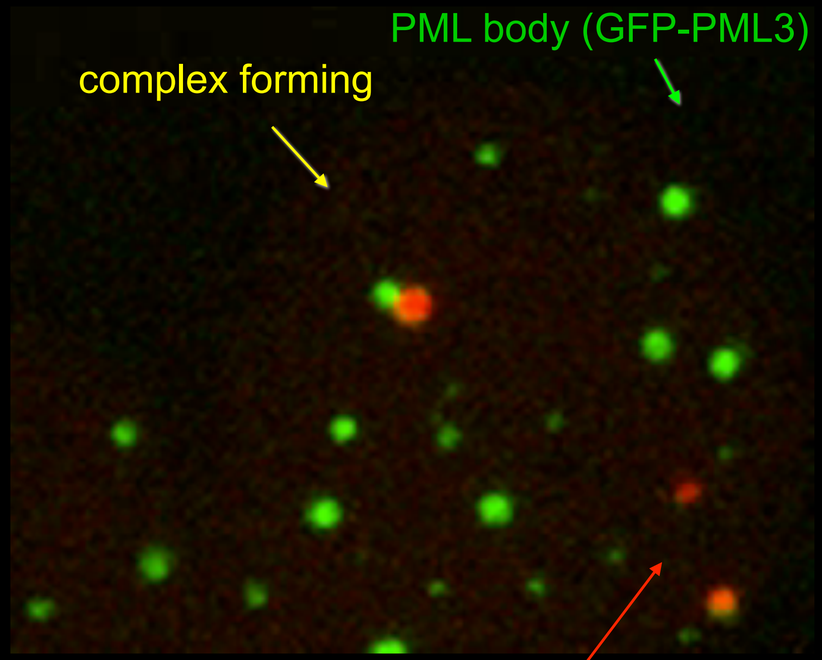
Lang J Cell Sci 2010

Mobility of PML bodies (green) close to telomeres (red)



70x higher speed

Complex formation between a PML body (green) and a telomere (red)



50x higher speed

Summary III

- Measuring the mean-squared displacement over time by tracking single particles describes their mobility.
- The observed mobility represents the intrinsic random translocations by diffusion but also the environment (obstacles, spatial confinement, binding interactions).
- Resolving the different contributions to particle mobility needs appropriate theoretical models (anomalous diffusion, confined diffusion, reaction-diffusion analysis)
- Acquiring a sufficient number of trajectories for single particle tracking (especially for single molecules) is technically challenging and thus ensemble methods are an alternative.