Spatial organization, target-search and non-specific interactions in the nucleus
Making sense out of single-molecule microscopy data

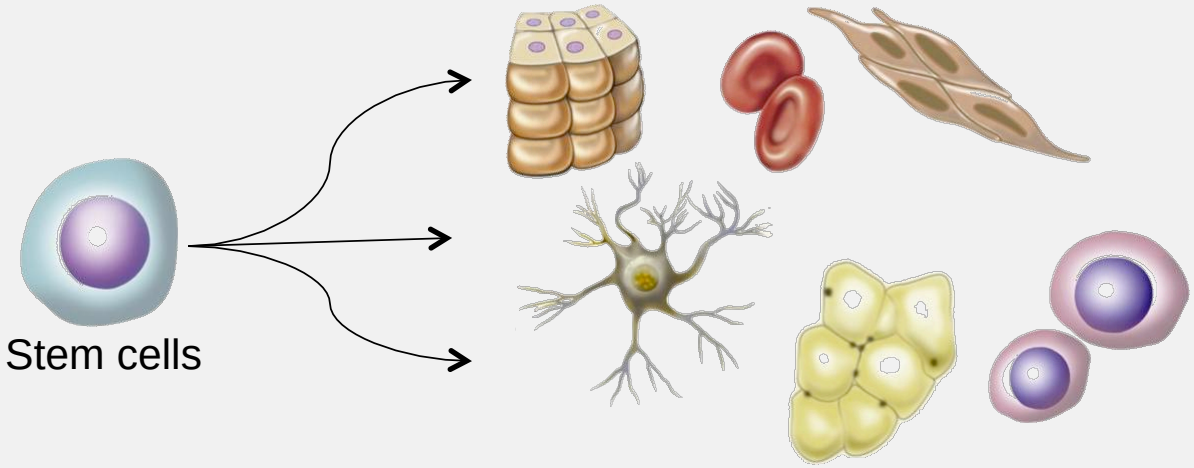
Ignacio Izeddin

Institut Langevin – Ondes et Images
ESPCI ParisTech, PSL, CNRS UMR 7588

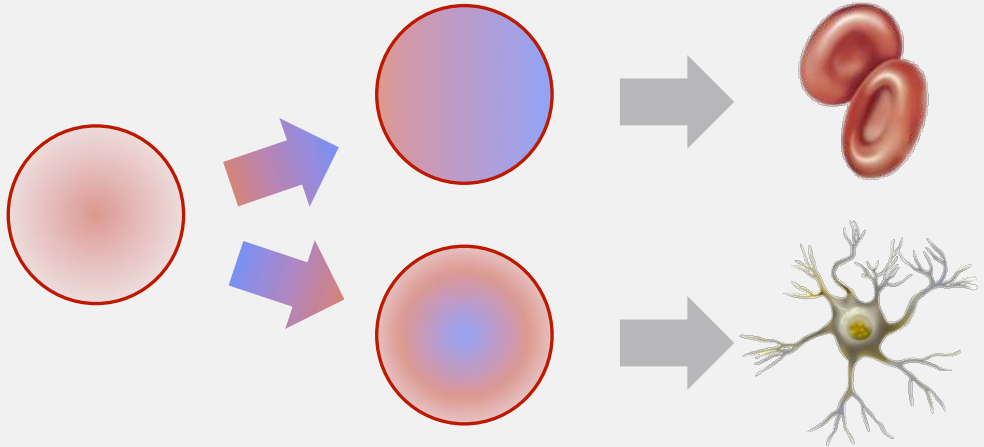
Rencontres scientifiques des Grands Causses, 29-09-2015
Modélisation physique de l'organisation nucléaire et de ses pathologies

Gene expression

Gene expression drives organism development and cellular adaptation to the environment

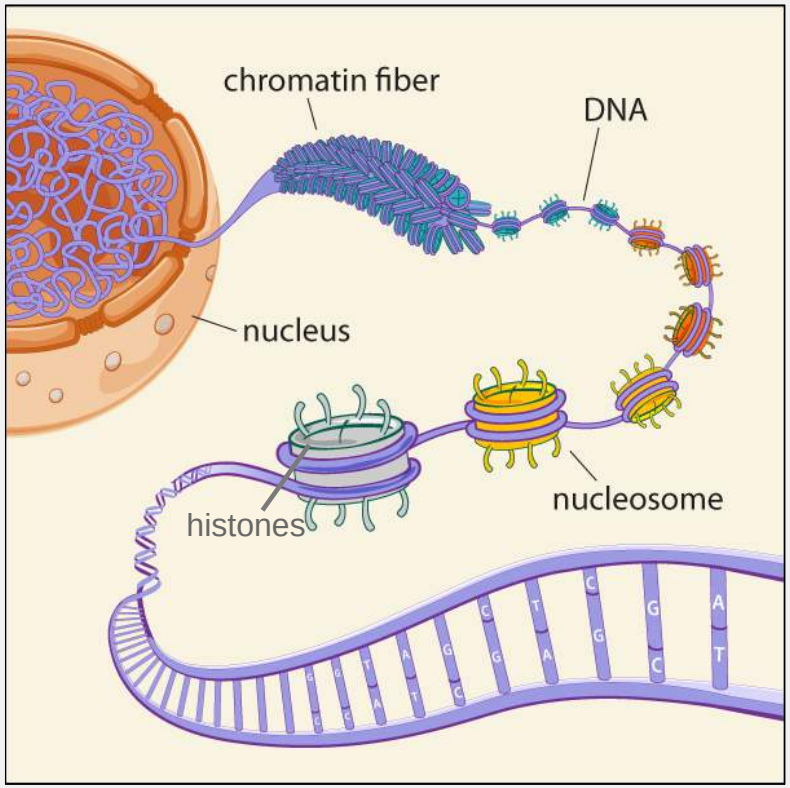


Transcription regulation

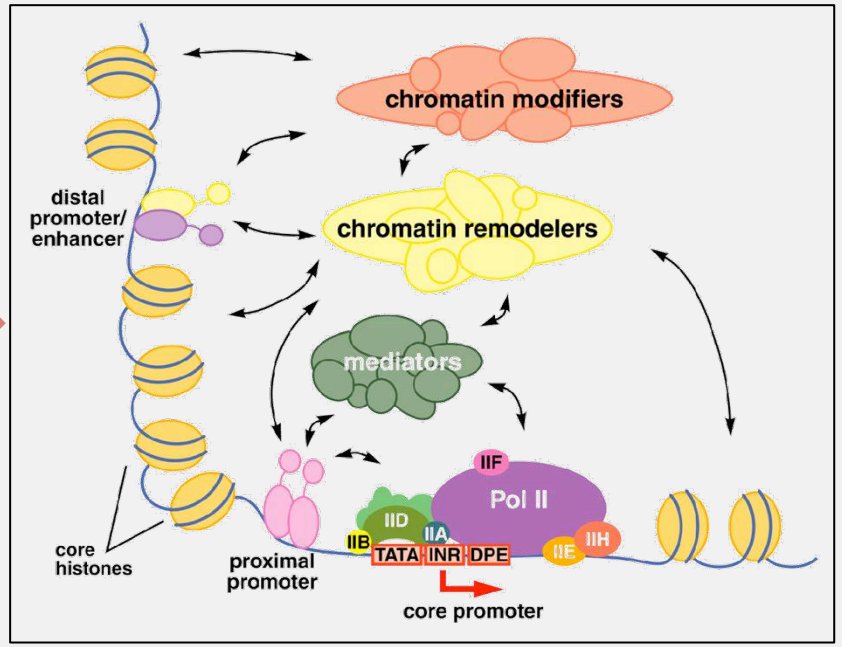


Gene expression regulation

Nuclear architecture

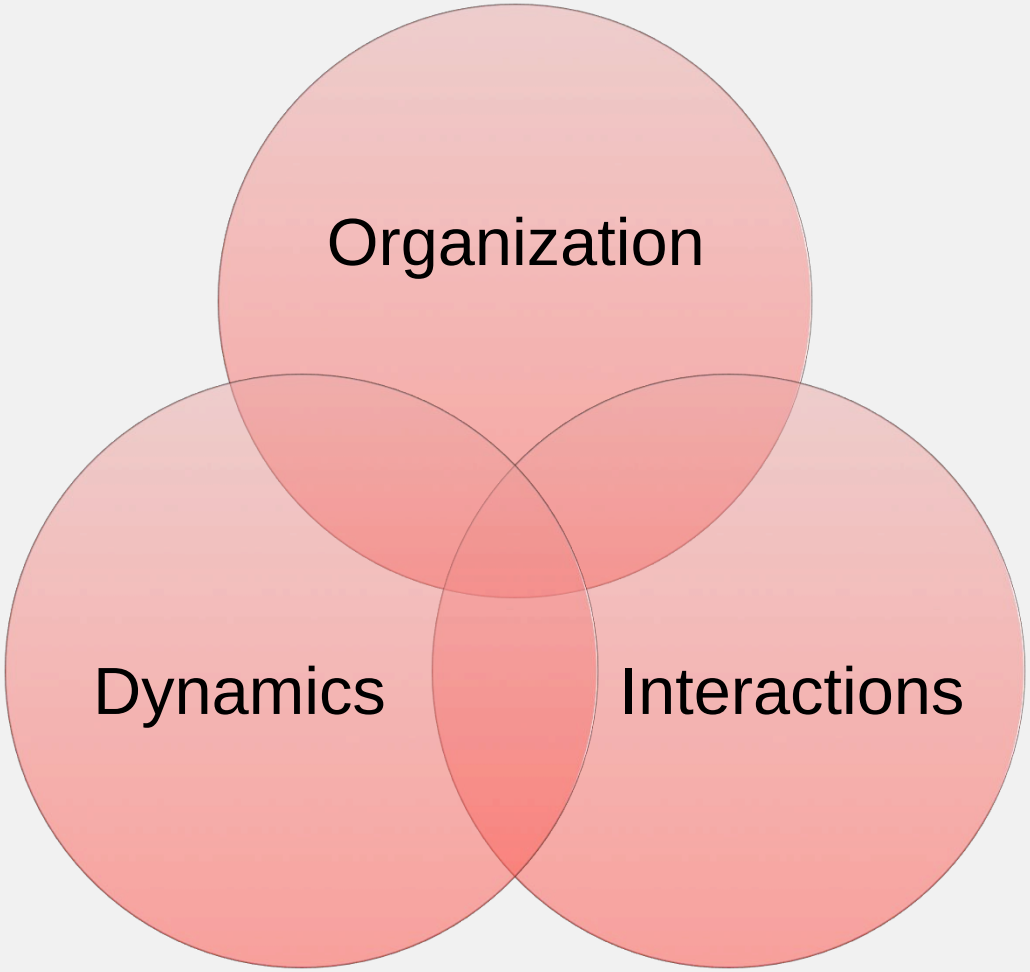


Gene expression control machinery



Resolving spatial and temporal relationships

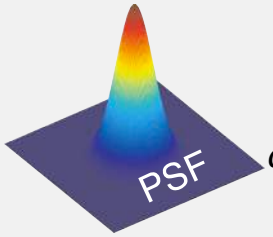
Gene expression regulation



Resolving spatial and temporal relationships

Single-molecule (live) cell imaging tools
Quantitative approach to transcription regulation

Single-molecule detection

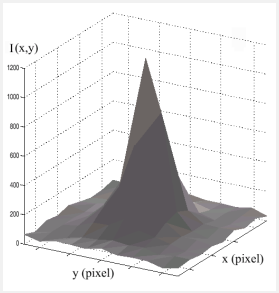


Point Spread Function

Diffraction of light

Abbe's equation

$$d = \frac{\lambda}{2n\sin(\alpha)} \approx 300nm$$

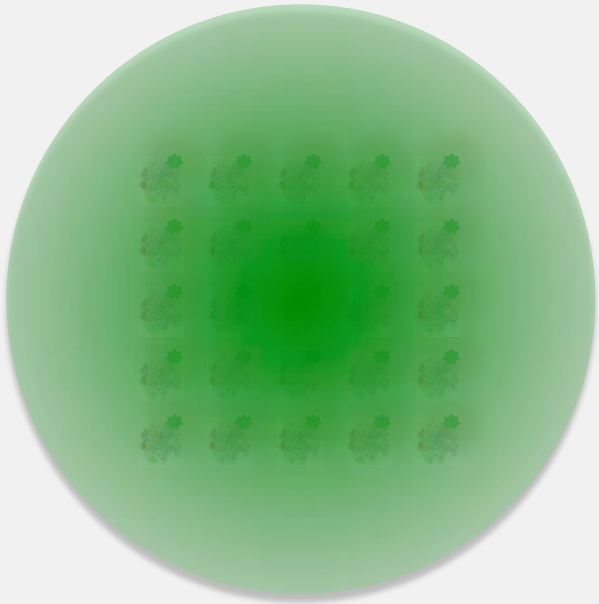


Experimental SM detection

Localization precision

limited by the number of photons

$$\sigma \propto \frac{1}{\sqrt{N}} \approx 10nm$$

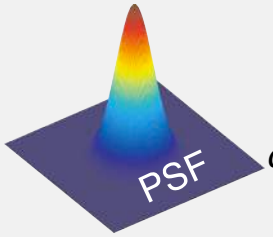


Classical fluorescence microscopy resolution is limited by diffraction

Localization-based super-resolution (PALM / STORM)

Nobel prize for the development of super-resolved fluorescence microscopy

Single-molecule detection

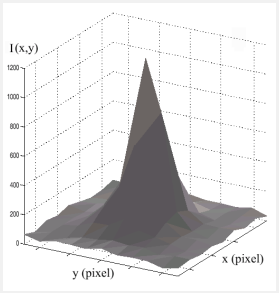


Point Spread Function

Diffraction of light

Abbe's equation

$$d = \frac{\lambda}{2n\sin(\alpha)} \approx 300nm$$



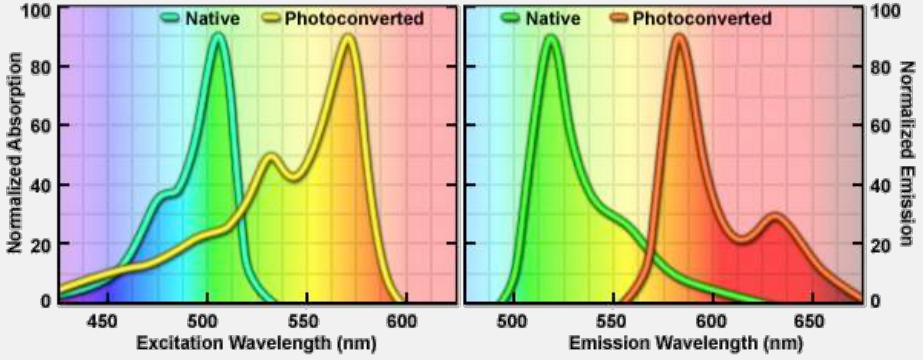
Experimental SM detection

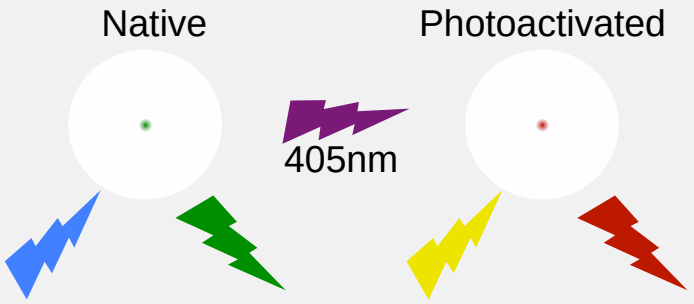
Localization precision

limited by the number of photons

$$\sigma \propto \frac{1}{\sqrt{N}} \approx 10nm$$

Photoactivatable fluorophores



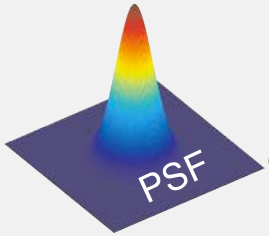


Temporal isolation of single molecules

Localization-based super-resolution (PALM / STORM)

Nobel prize for the development of super-resolved fluorescence microscopy

Single-molecule detection

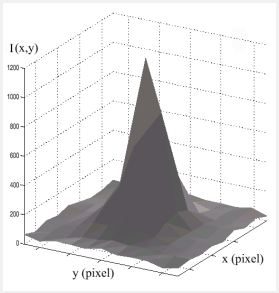


Point Spread Function

Diffraction of light

Abbe's equation

$$d = \frac{\lambda}{2n\sin(\alpha)} \approx 300nm$$



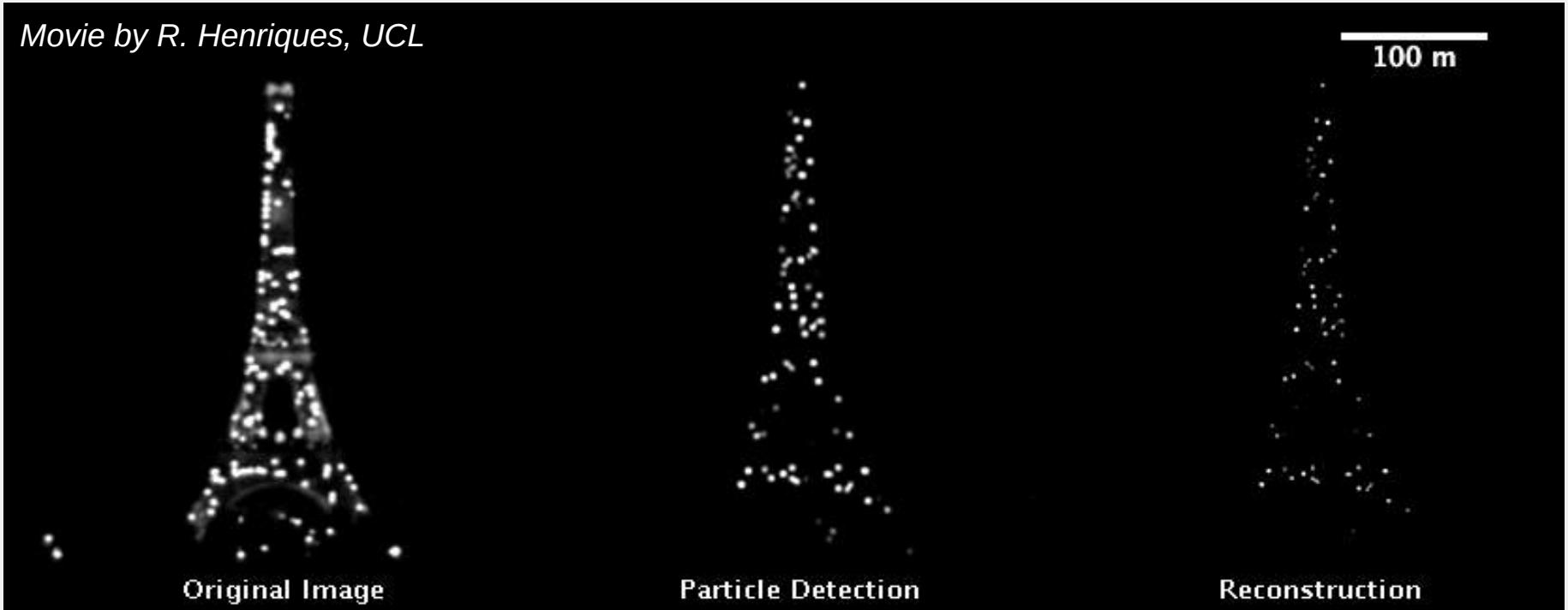
Experimental SM detection

Localization precision

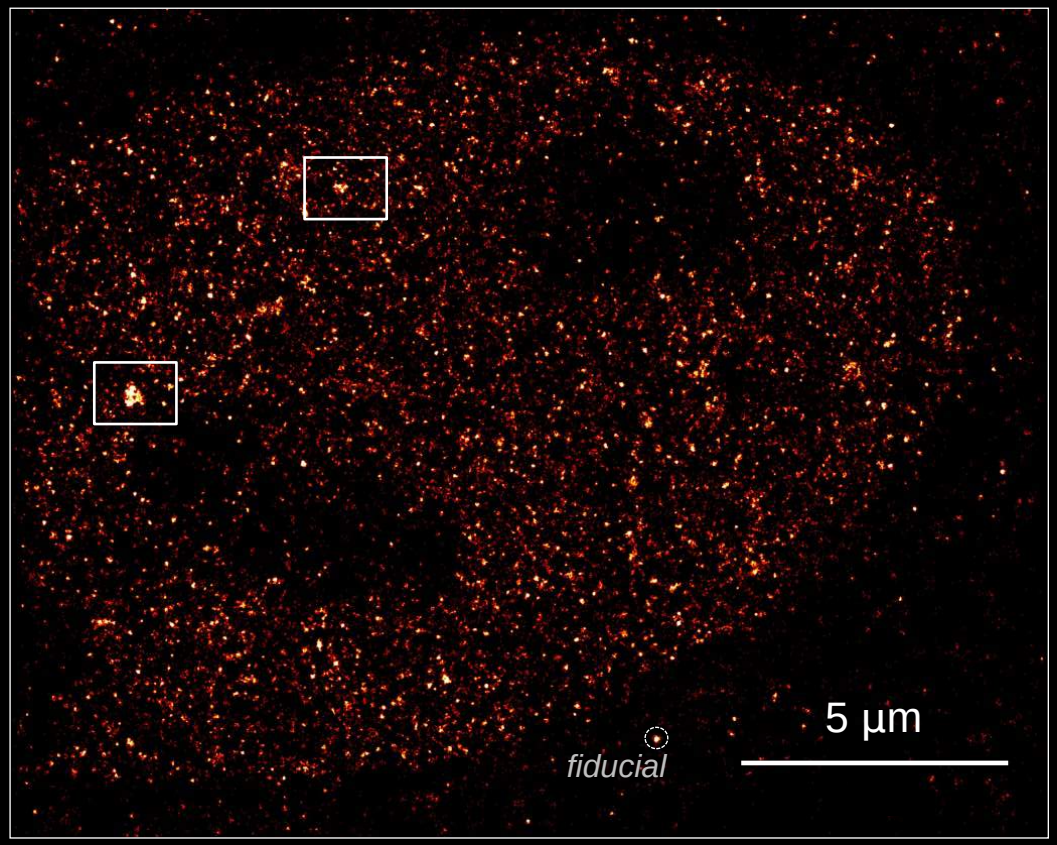
limited by the number of photons

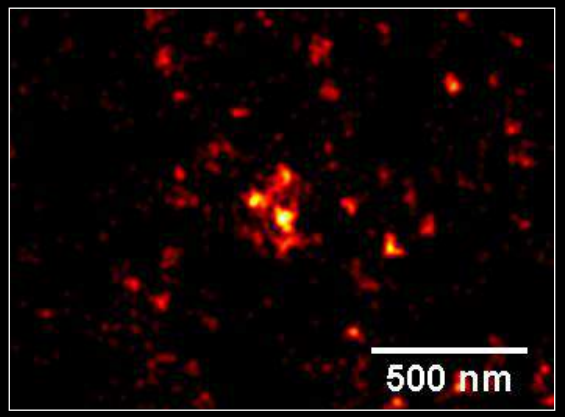
$$\sigma \propto \frac{1}{\sqrt{N}} \approx 10nm$$

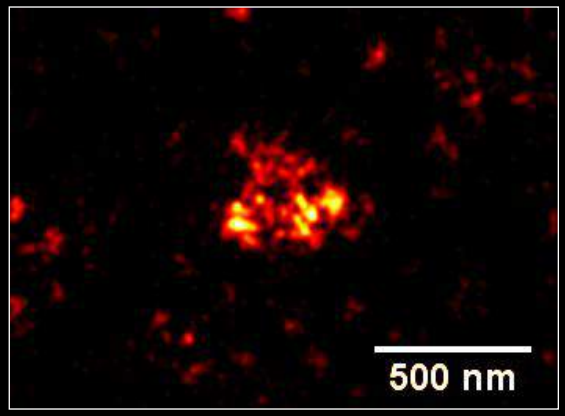
Movie by R. Henriques, UCL



High resolution mapping of RNA Pol II



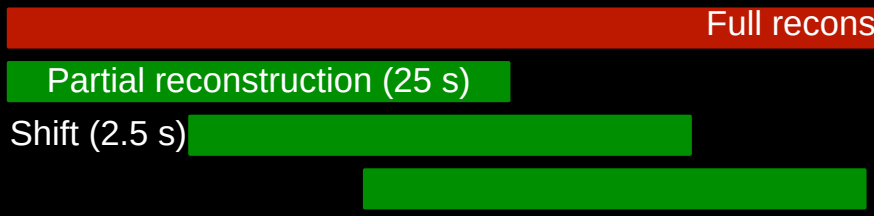


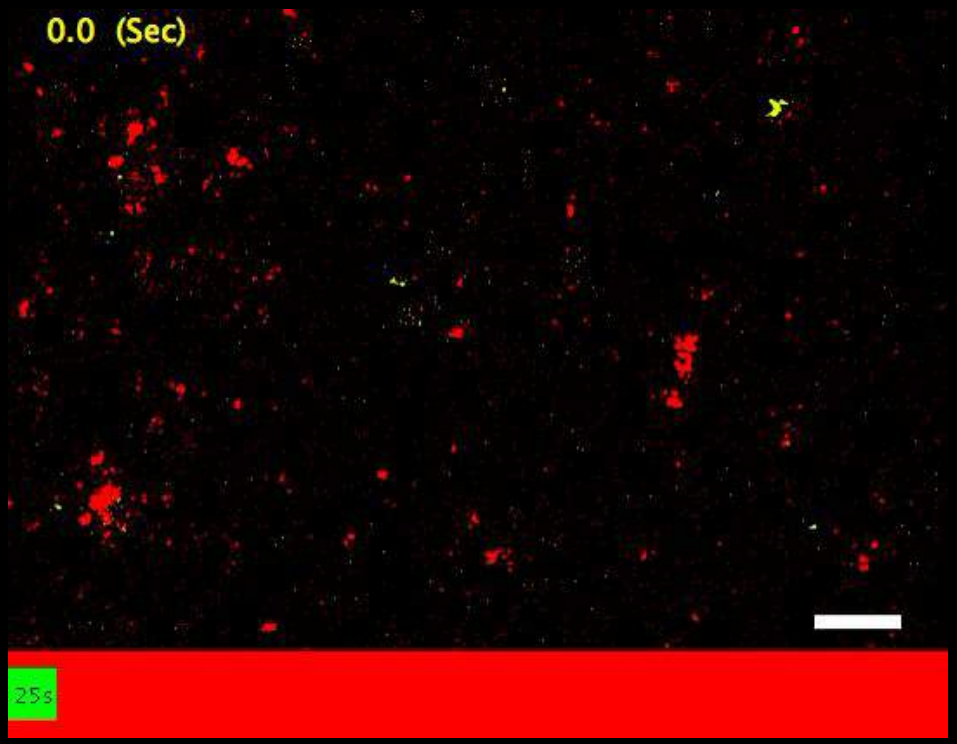


How dynamic are RNA Pol II clusters?

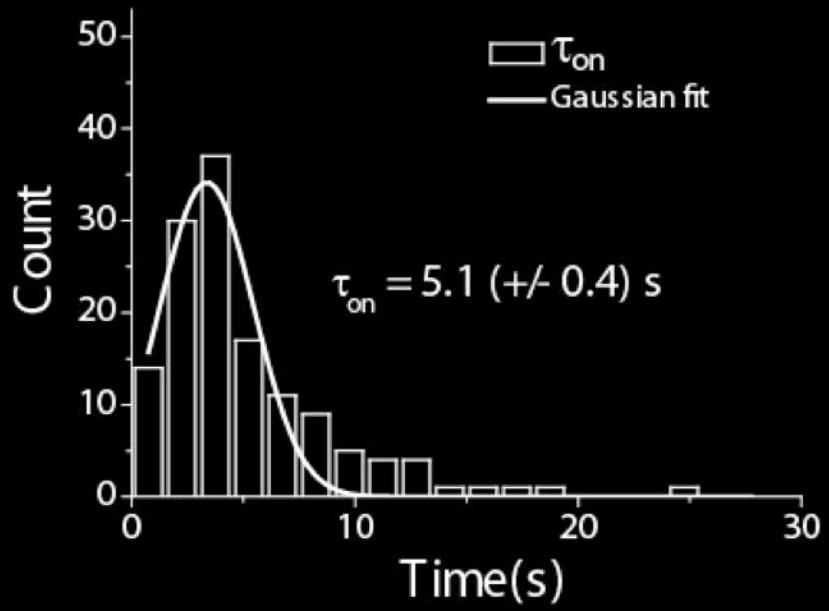
Cisse et al, Science (2013)

High resolution mapping of RNA Pol II in live cells





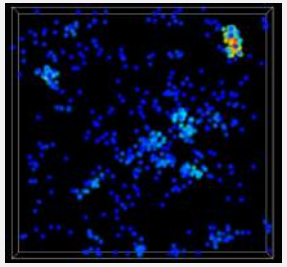
Average cluster lifetime ~ 5s

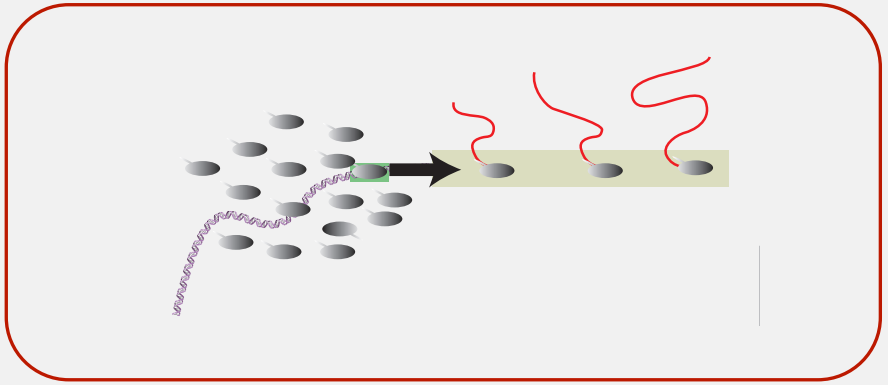


How dynamic are RNA Pol II clusters?

Cisse et al, Science (2013)

High resolution mapping of RNA Pol II in live cells



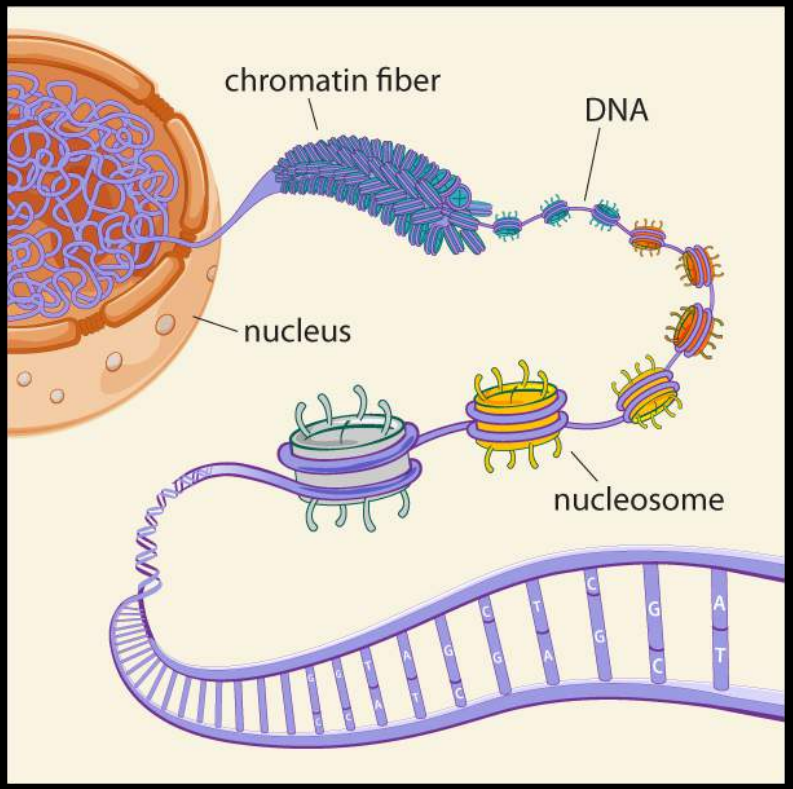


- RNA Pol II distribution in the nucleus
  - is not homogenous
  - is highly dynamic

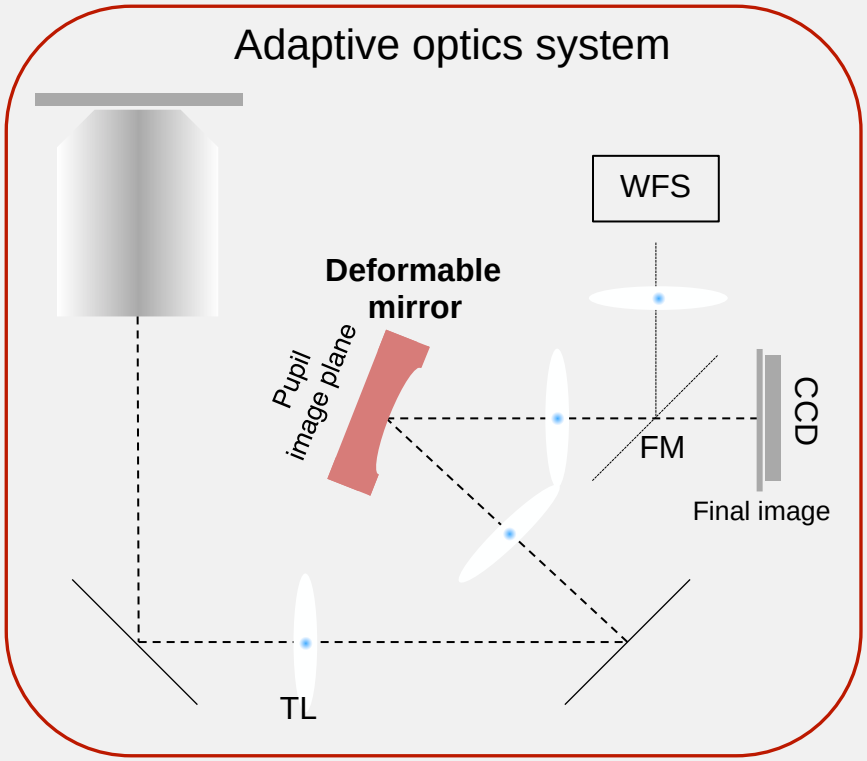
- Clusters are transient (~5 s)
  - Lifetime not compatible with elongation time (minutes)!

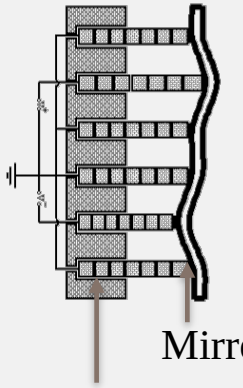
Cisse et al, Science (2013)

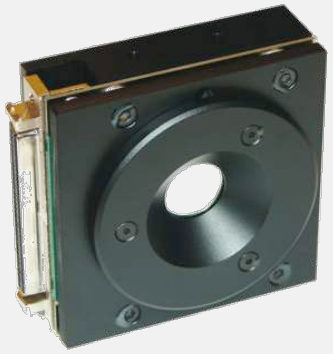
How complex is the available space in the nucleus?



Adaptive optics for single-molecule microscopy
Higher budget of photons and access to 3D data

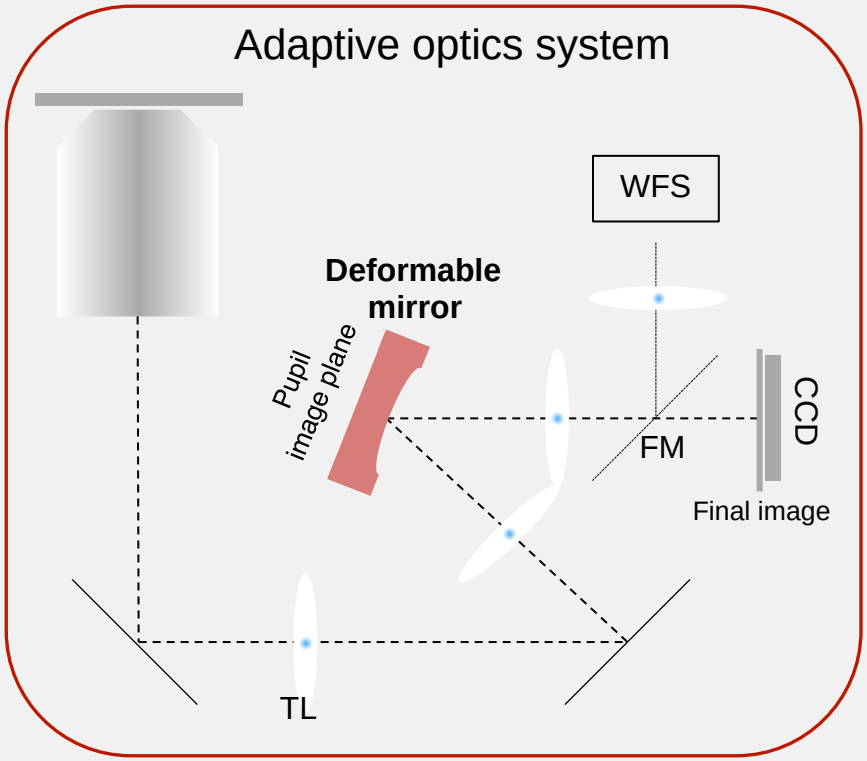


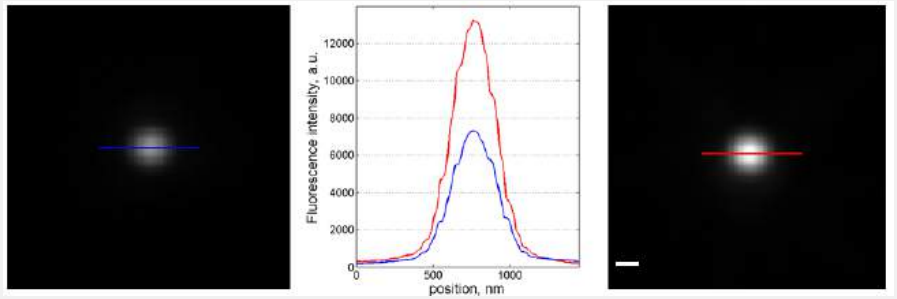


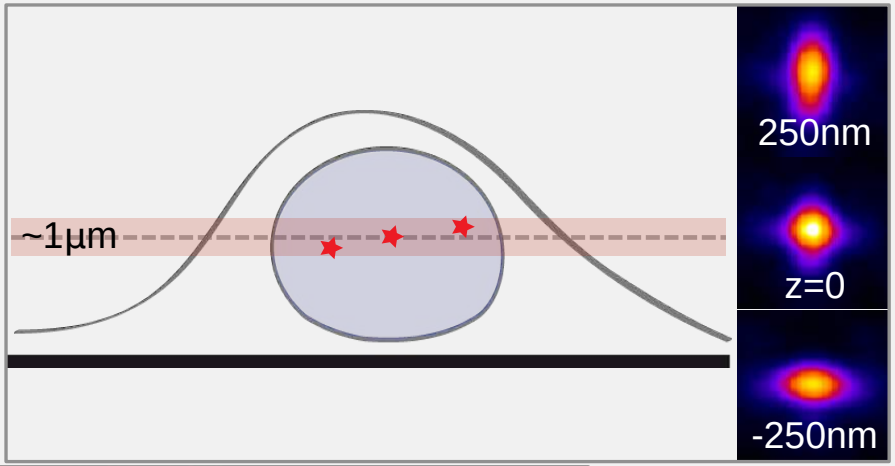


Imagine Optic™

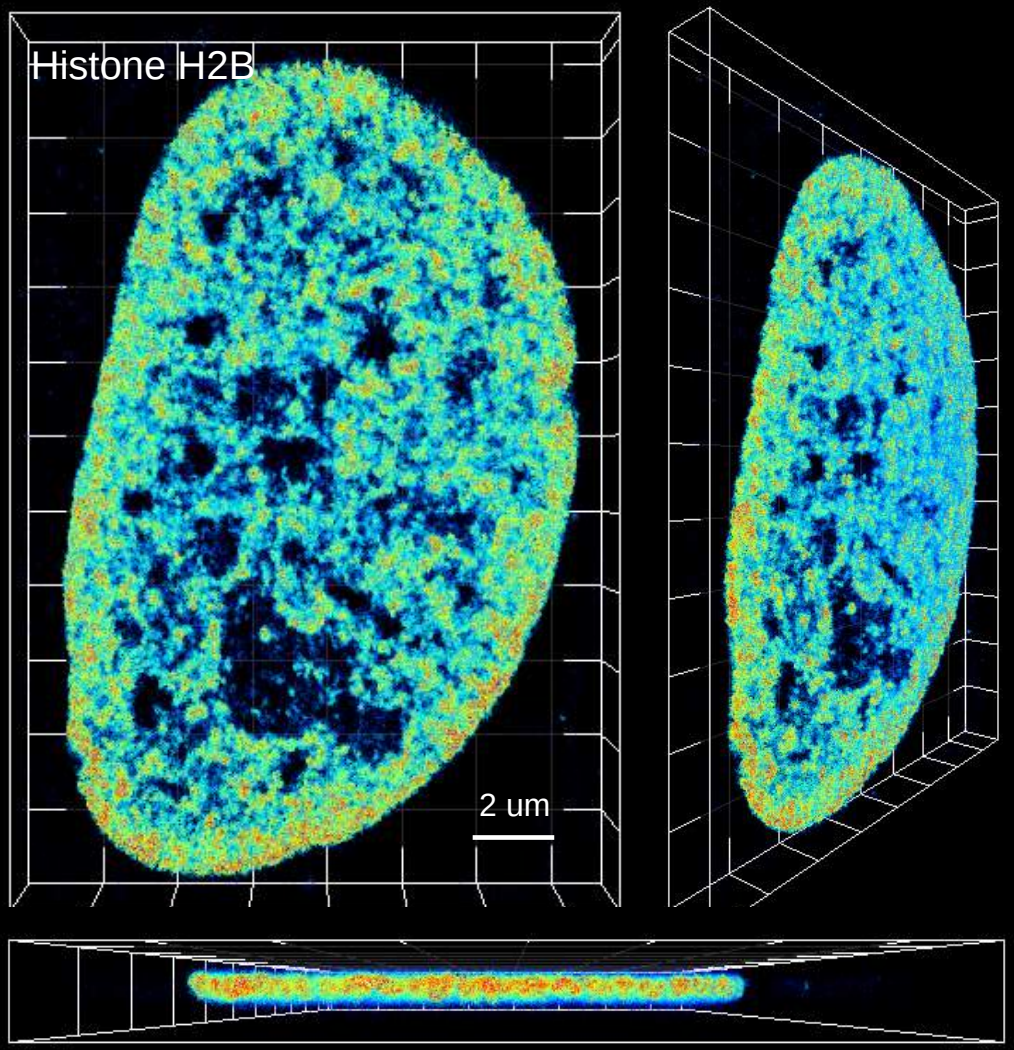
Adaptive optics for single-molecule microscopy
Higher budget of photons and access to 3D data



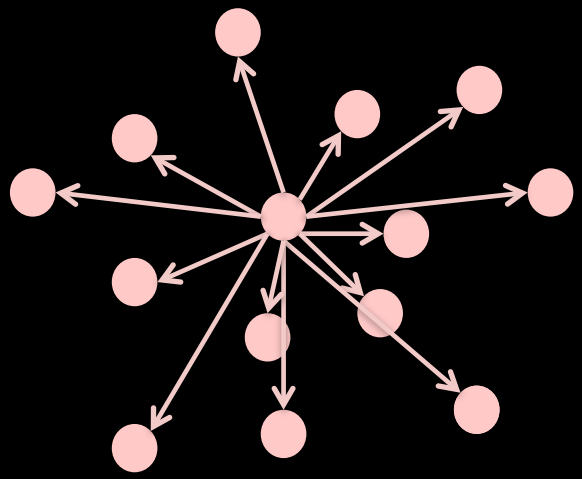




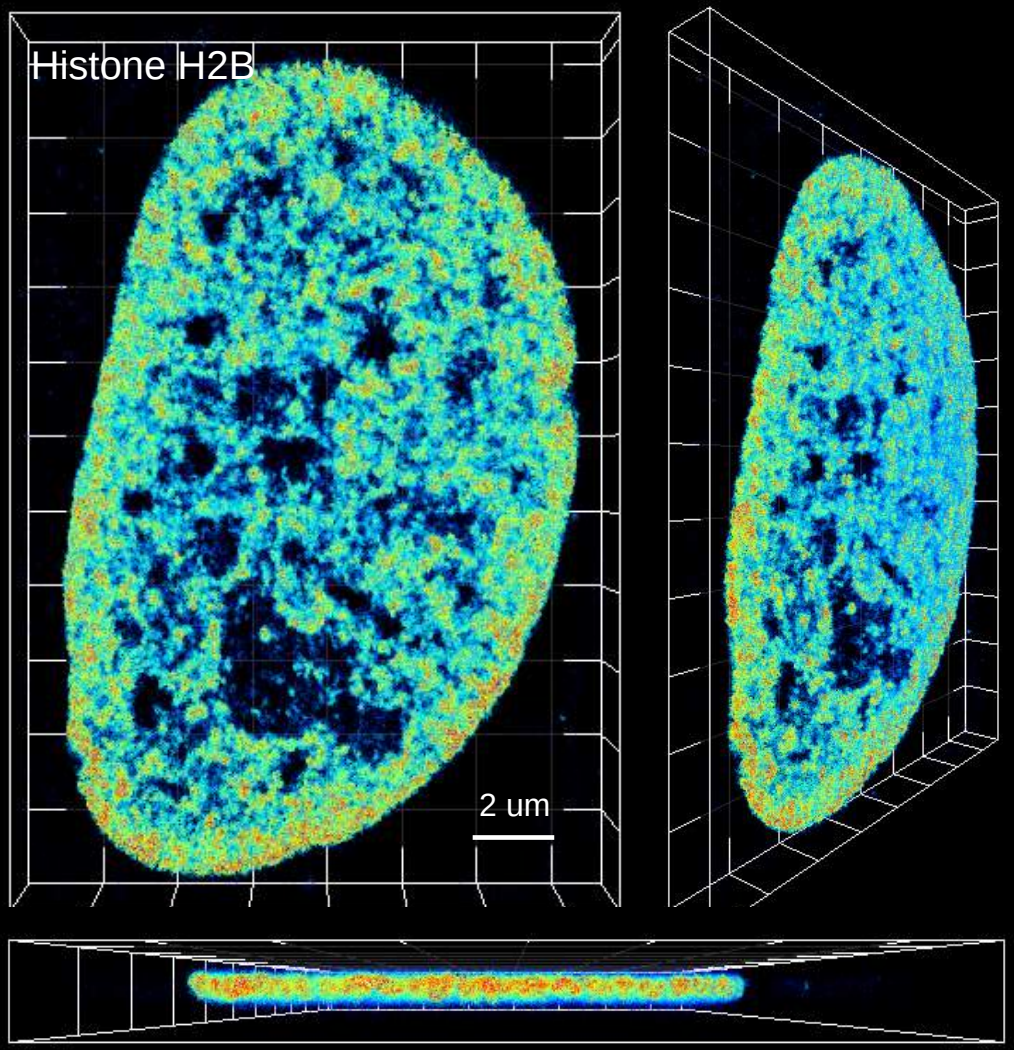
How complex is the available space in the nucleus?



K Ripley distribution function

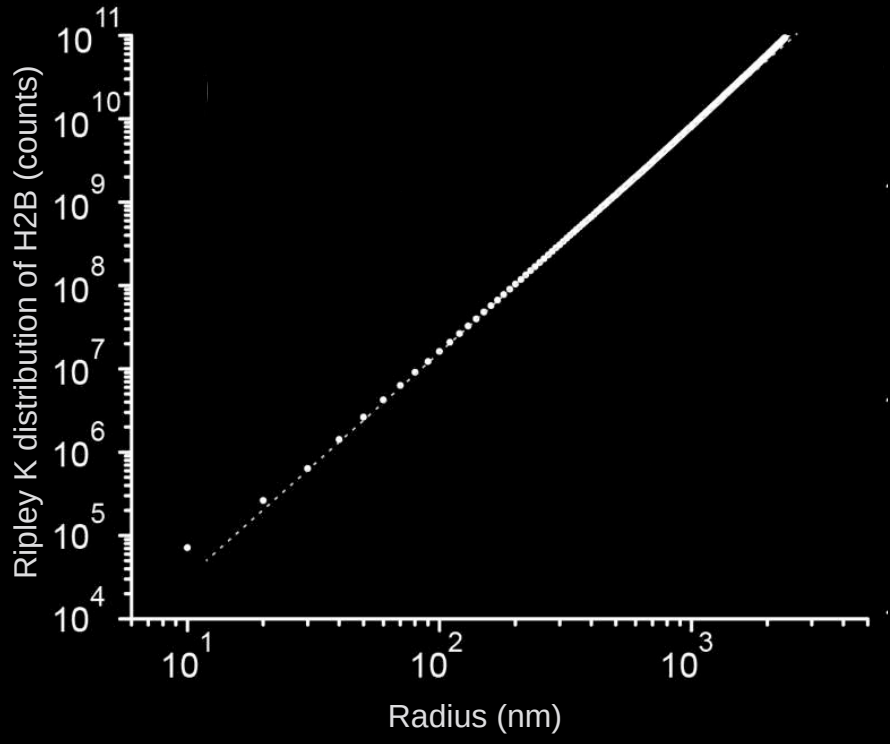


How complex is the available space in the nucleus?



$K \propto K_{random}^{0.889}$

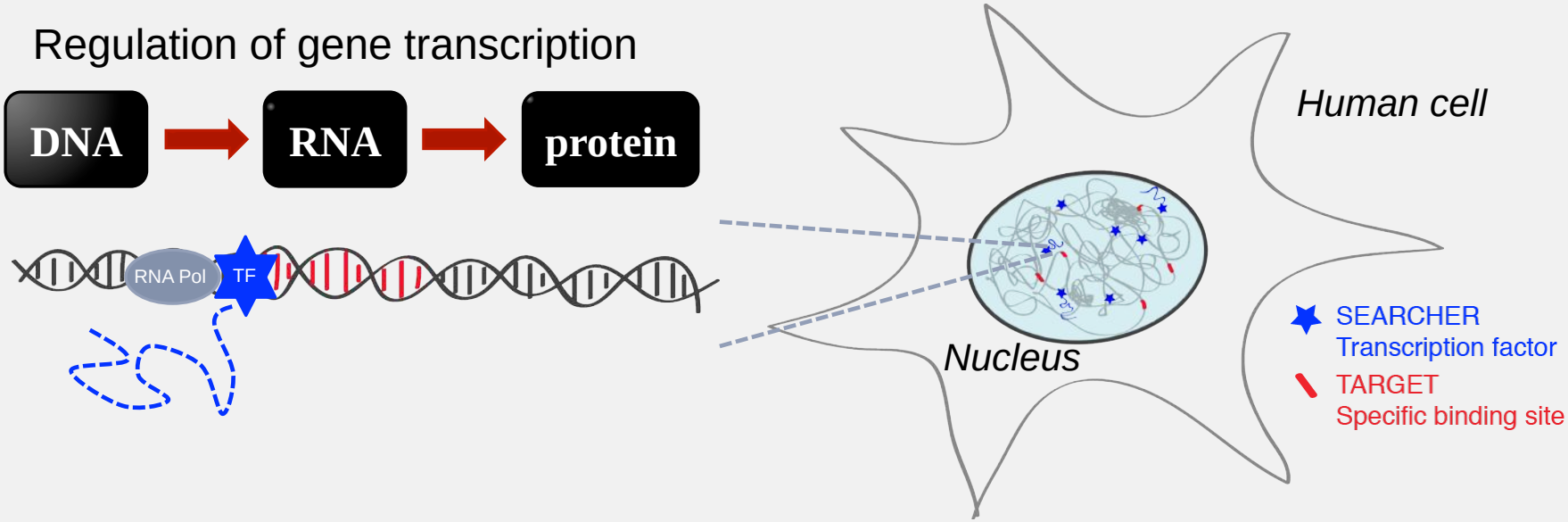
Fractal dimension Df = 2.7



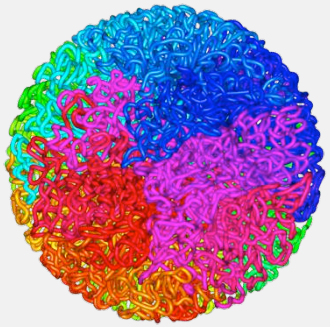
Chromatin has a fractal spatial distribution (30nm – 3µm) with fractal dimension of 2.7

Récamier et al., Nucleus (2014)

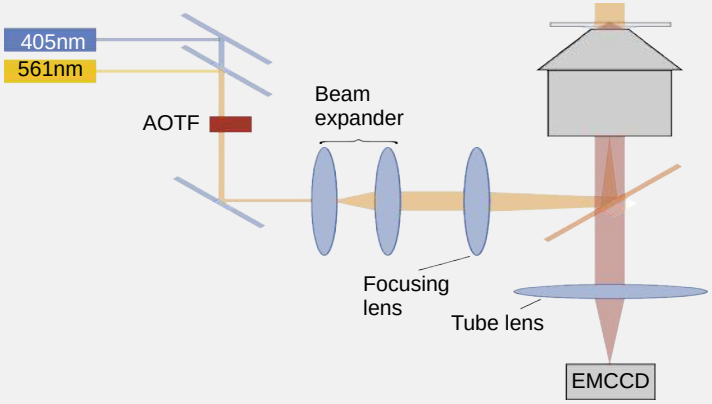
Regulation of gene transcription, a problem of target search

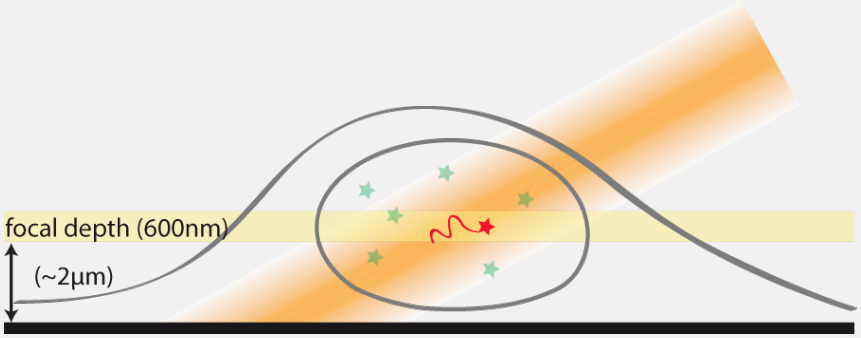


Proteins diffuse in the nucleoplasm
What is the nature of their space exploration?

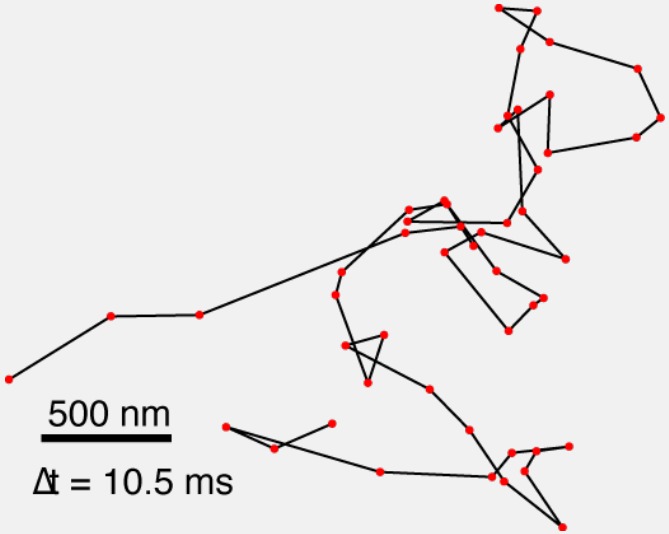


Single-molecule tracking in the nucleus of live cells





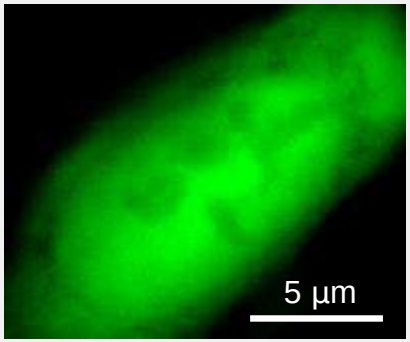
- Transcription factors fused to photoactivatable fluorescent protein Dendra2
- Photoactivation allows us to isolate single molecules within a dense populations



Different nuclear proteins with a wide range of diffusive behaviors

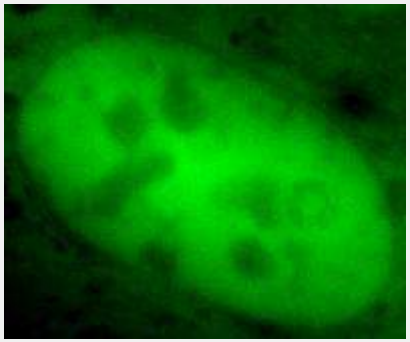
Human osteosarcoma cells (U2OS)

Dendra2

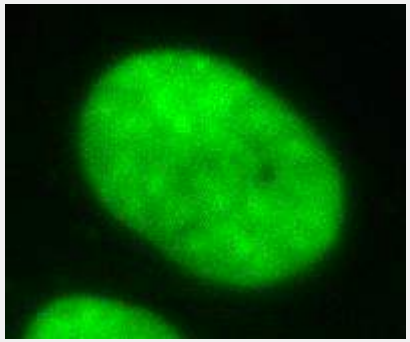


Free (inert) fluorophore

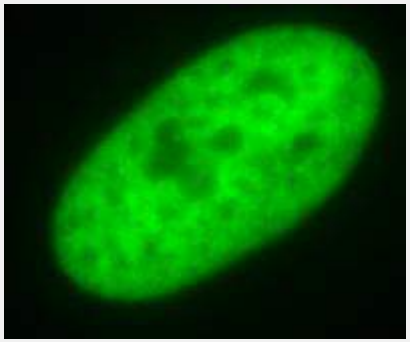
c-Myc



P-TEFb



Histone H2B

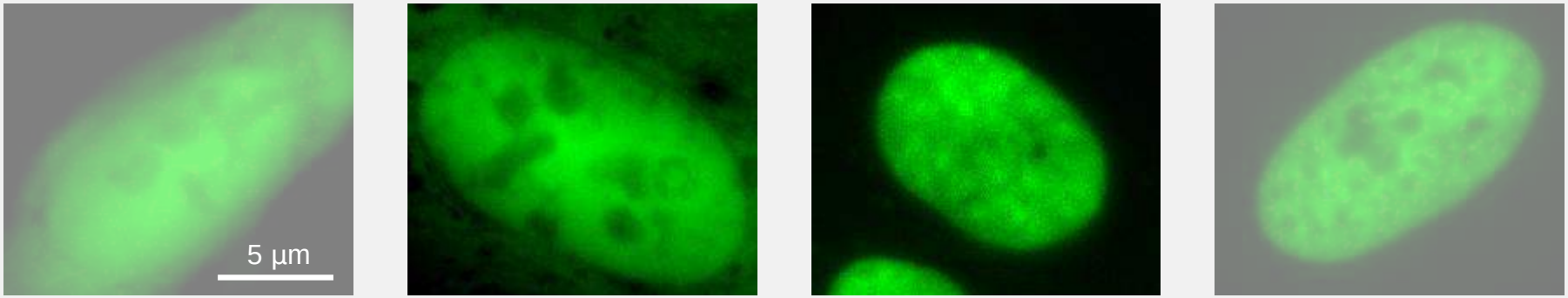


“Bound” protein

Different nuclear proteins with a wide range of diffusive behaviors

Human osteosarcoma cells (U2OS)

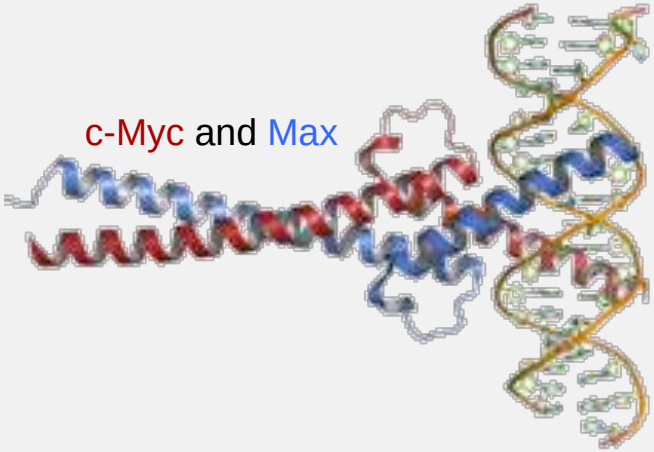
Dendra2 | c-Myc | P-TEFb | Histone H2B



Free (inert) fluorophore

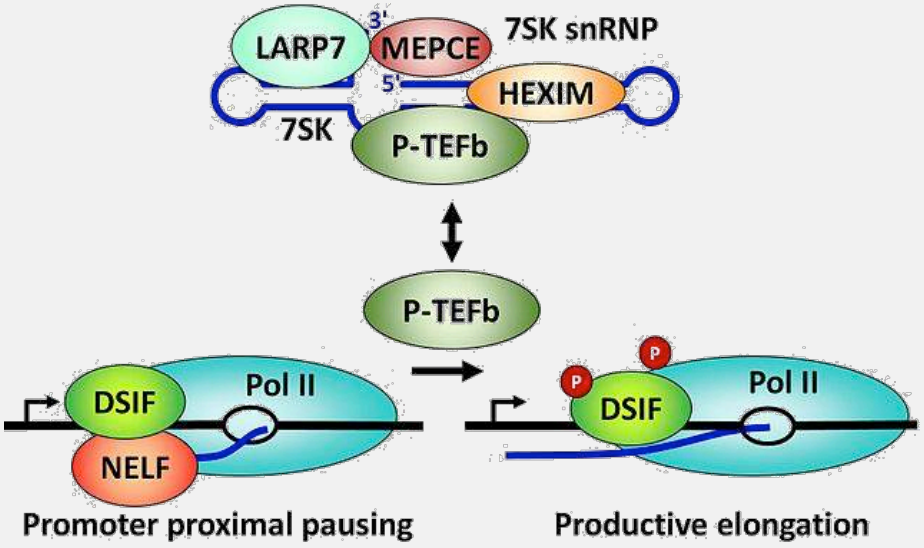
"Bound" protein

c-Myc is a proto-oncogenic TF



c-Myc upregulates transcription of nearly all genes
(Lin et al., Cell, 2012; Nie et al., Cell, 2012)

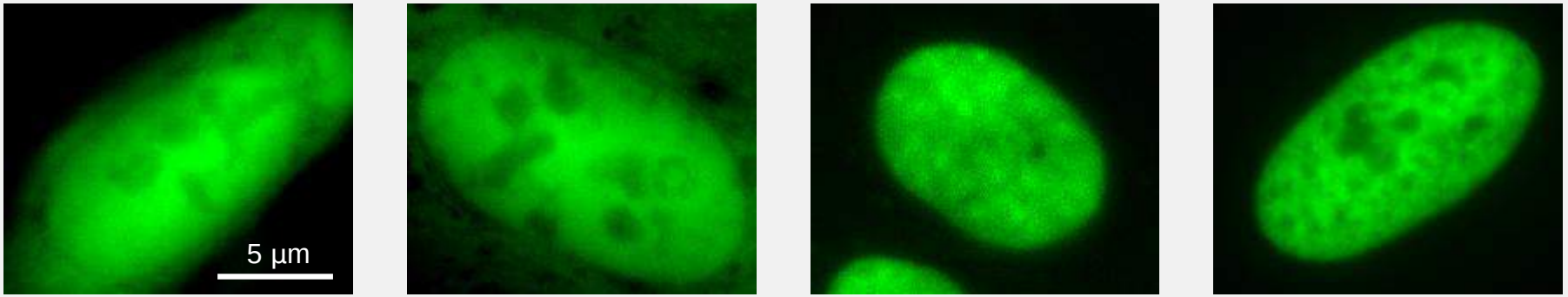
Positive transcription elongation factor P-TEFb



Different nuclear proteins with a wide range of diffusive behaviors

Human osteosarcoma cells (U2OS)

Dendra2	c-Myc	P-TEFb	Histone H2B



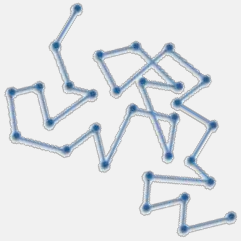
Free (inert) fluorophore

“Bound” protein

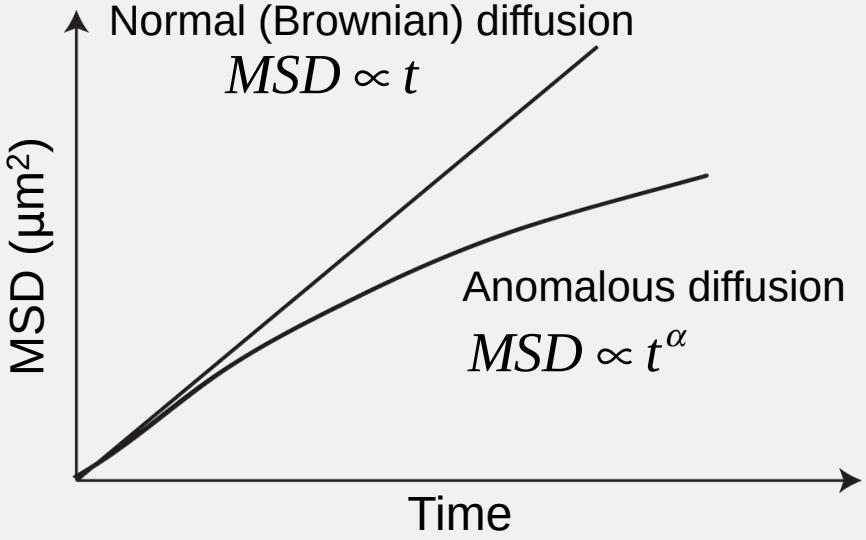
500 nm

How does a protein explore the available space?

The mean square displacement (MSD)
is a measure of the amount of explored space

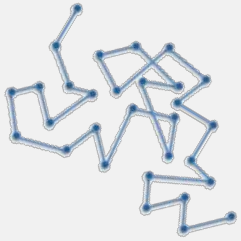


$$MSD = \langle (x(t) - x_0)^2 \rangle$$

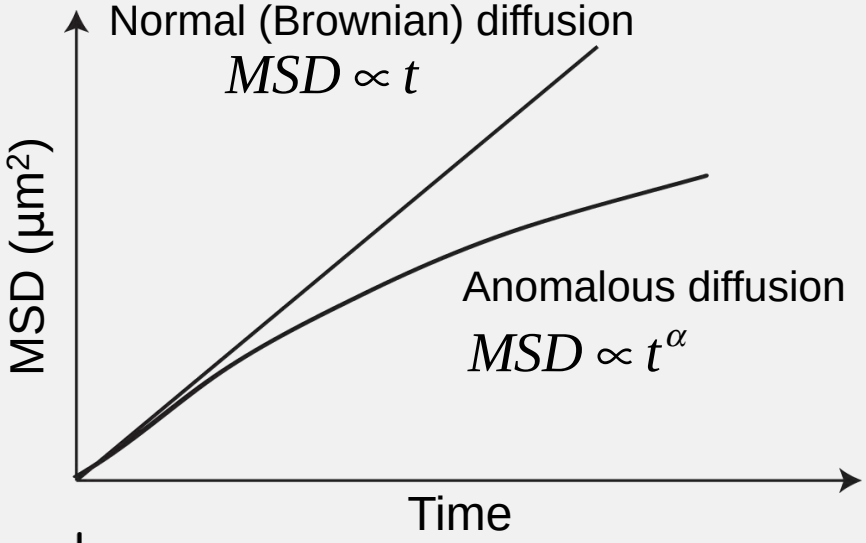


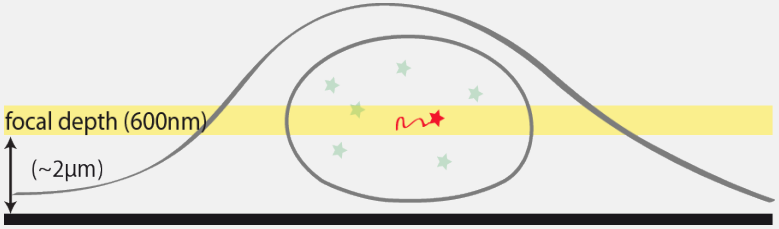
How does a protein explore the available space?

The mean square displacement (MSD)
is a measure of the amount of explored space

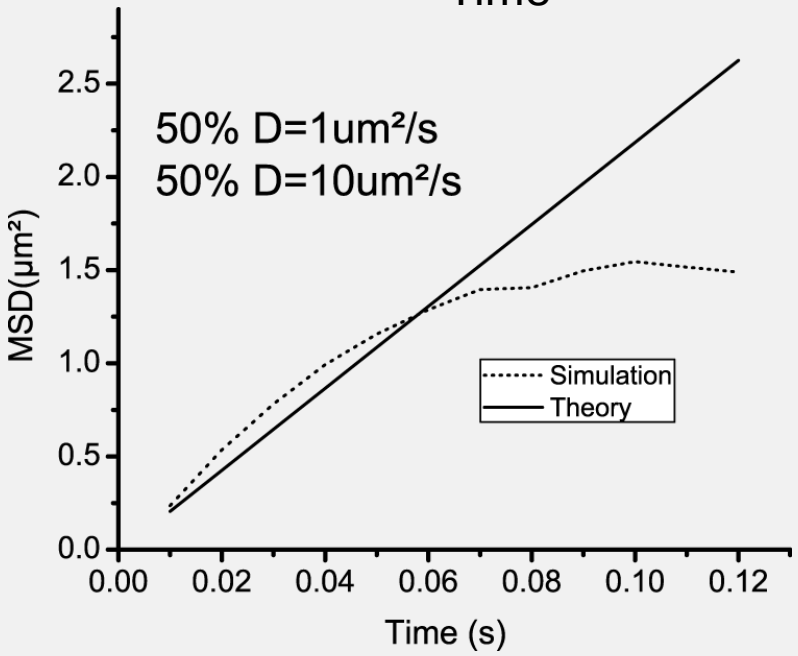


$$MSD = \langle (x(t) - x_0)^2 \rangle$$

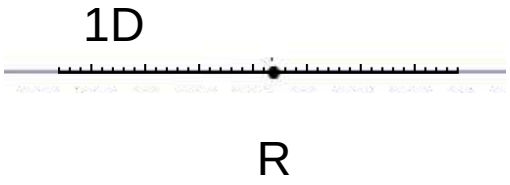


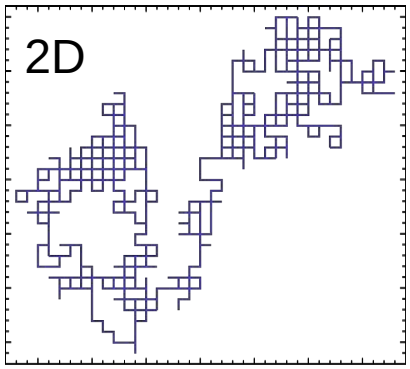


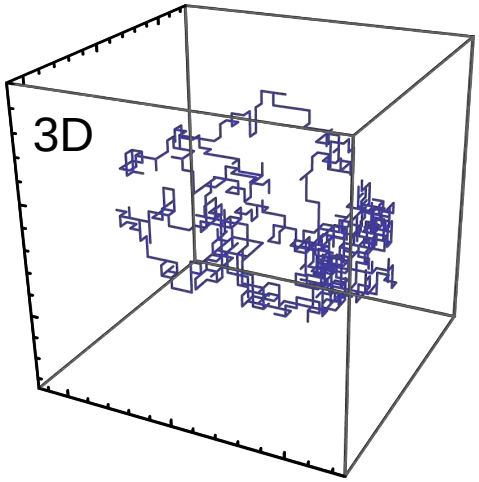
Several diffusive populations
population exclusion effect



Geometry of random exploration and space sampling







Df = 1
Dw = 2 (Brownian)

Df = 2
Dw = 2 (Brownian)

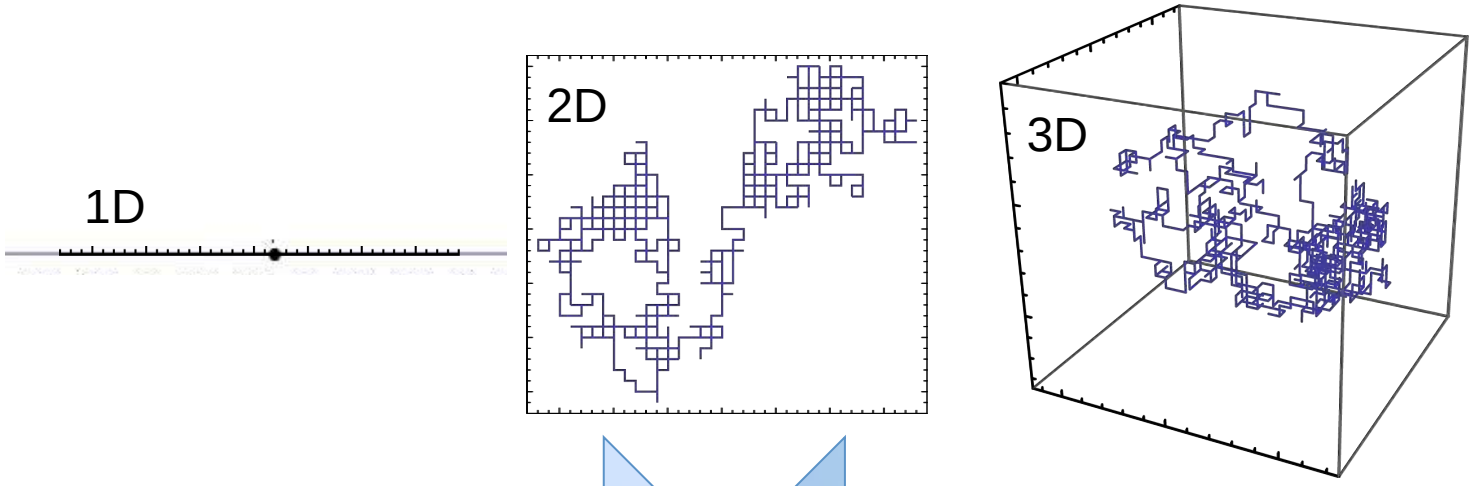
Df = 3
Dw = 2 (Brownian)

- Accessible sites $Na \propto R^{Df}$, where Df is the fractal dimension
- Visited sites $Nv \propto R^{Dw}$, where Dw is the dimension of the walk
  Dw can be defined by the scaling law of the MSD $\propto t^{2/Dw}$

Non-compact exploration: Dw < Df
Compact exploration: Dw > Df

P.G. De Gennes, J. Chem. Phys. (1982)

Geometry of random exploration and space sampling



Spatial Oversampling

Compact exploration

Target Search **depends on the initial distance** to the target (in a closed volume)

Spatial Undersampling

Non-compact exploration

Target Search is **independent on the initial distance** to the target (in closed volume)

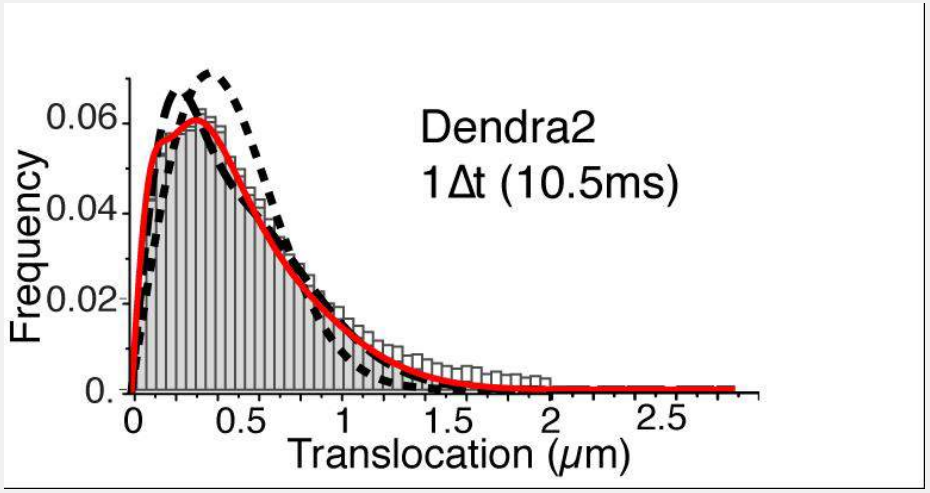
Bénichou et al, Nat. Chem 2010; Condamin et al, Nature 2007

Dendra2 (“free” protein) and c-Myc undergo normal diffusion

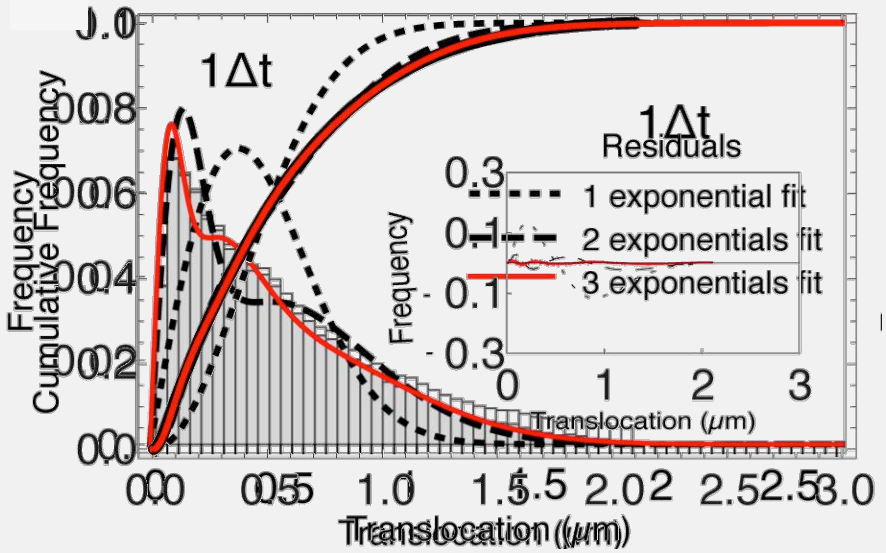
Dendra2 and c-Myc, heterogeneous normal diffusion

	Dendra2	c-Myc
D_1 (immobile)	4%	9.5%
D_2	2.6μm^2/s, 24%	0.5μm^2/s, 20%
D_3	13μm^2/s, 72%	13.5μm^2/s, 70%

Fitting the (cumulative) distribution of translocations



Dendra2



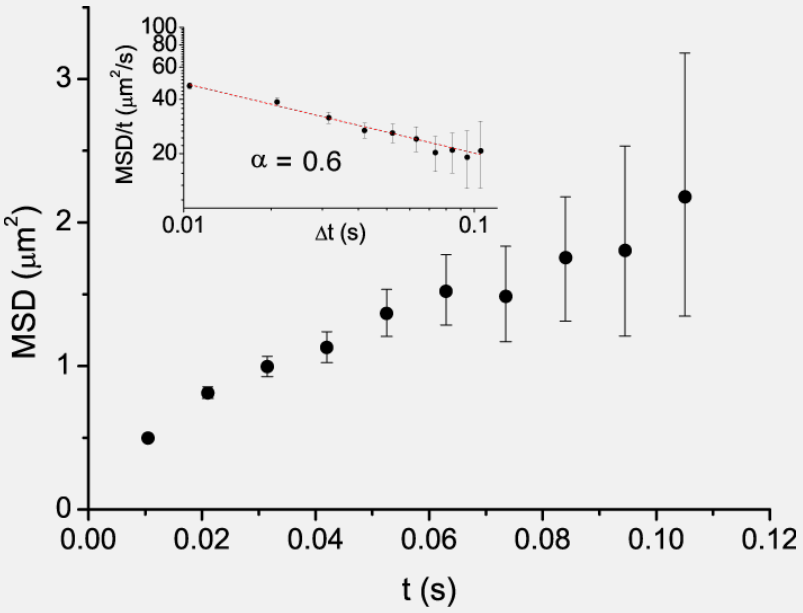
c-Myc

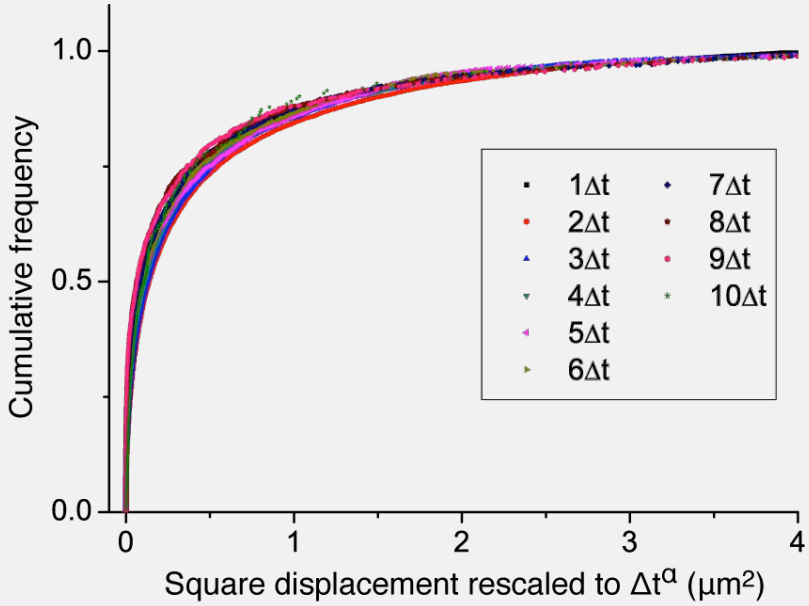
c-Myc undergoes normal diffusion; P-TEFb, subdiffusion

Dendra2 and c-Myc, heterogeneous normal diffusion

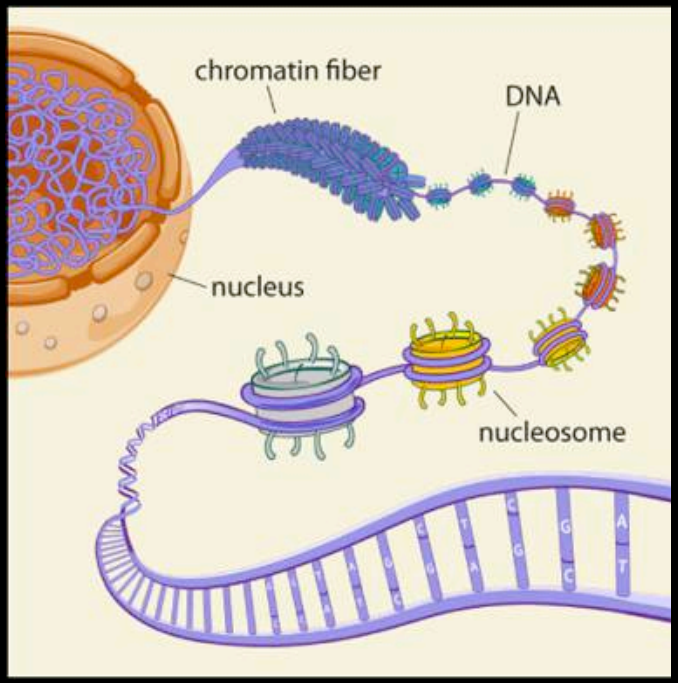
	Dendra2	c-Myc
D_1 (immobile)	4%	9.5%
D_2	2.6μm^2/s, 24%	0.5μm^2/s, 20%
D_3	13μm^2/s, 72%	13.5μm^2/s, 70%

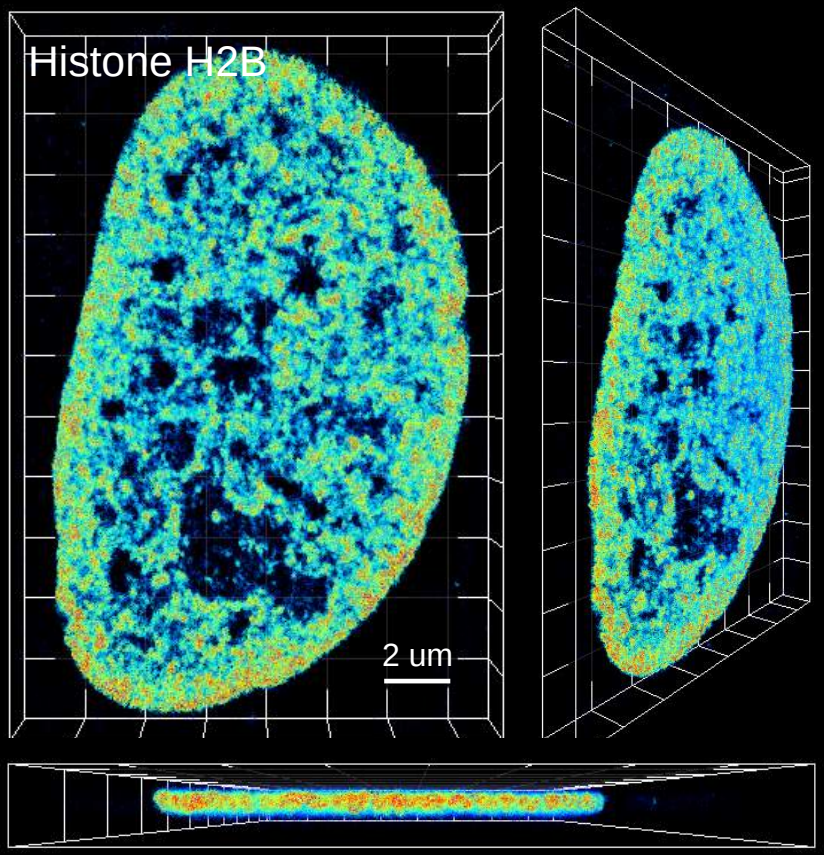
P-TEFb, anomalous diffusion





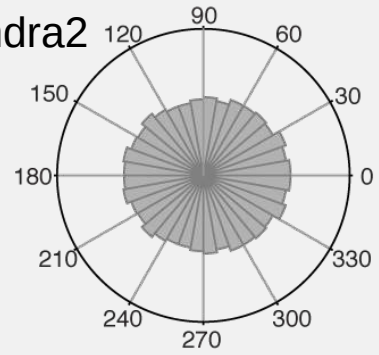
The nucleus is a complex environment



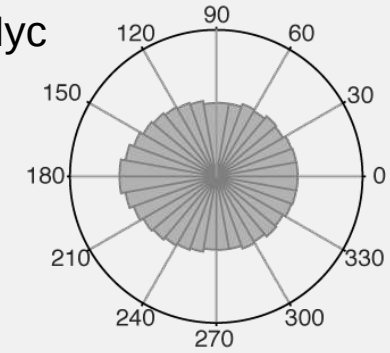


Angular θ distribution to understand the geometry of exploration

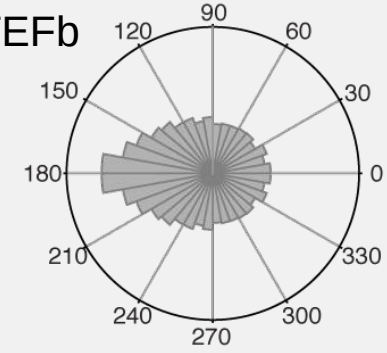
Dendra2

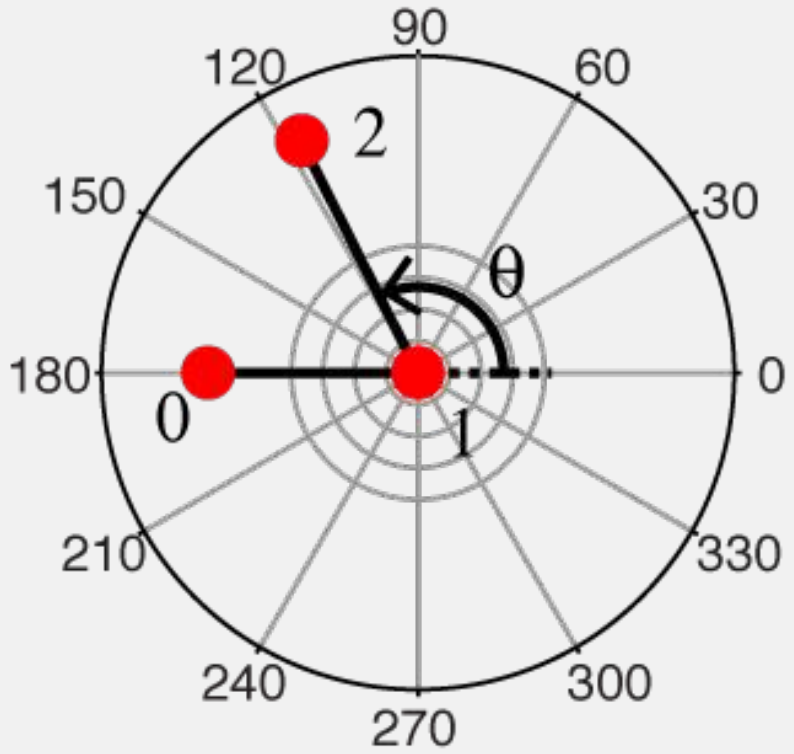


c-Myc



P-TEFb

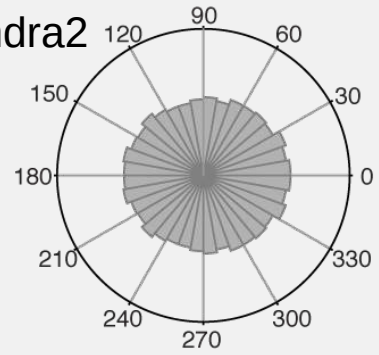


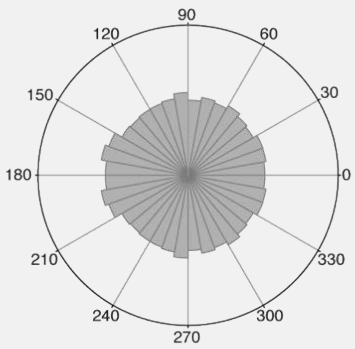


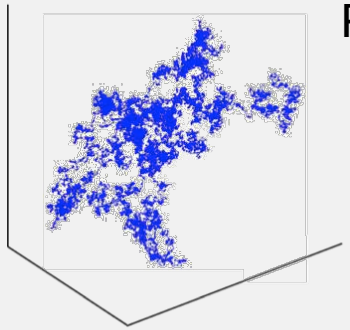
Izeddin et al, eLife (2014)

Angular θ distribution to understand the geometry of exploration

Dendra2

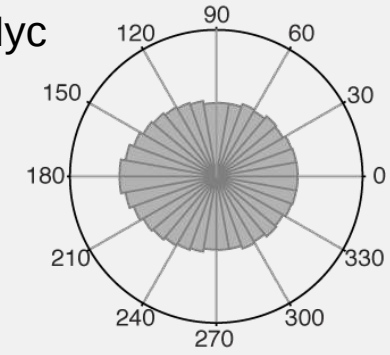


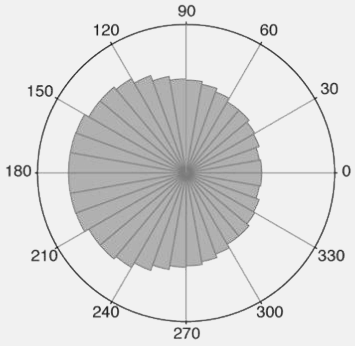


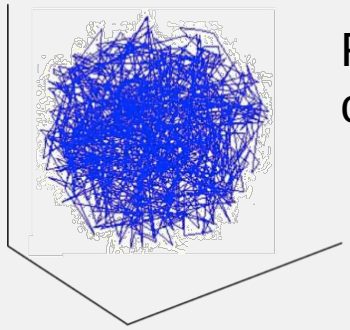


Free diffusion

c-Myc

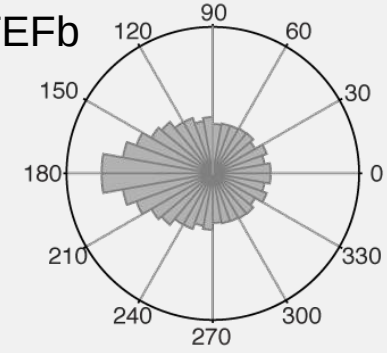


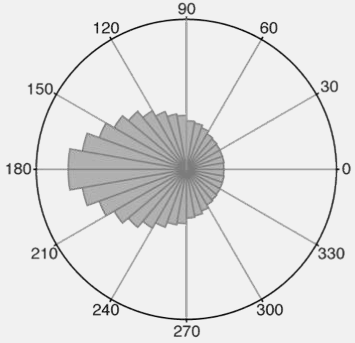


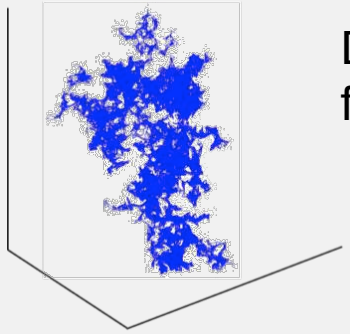


Free diffusion in confined volume

P-TEFb







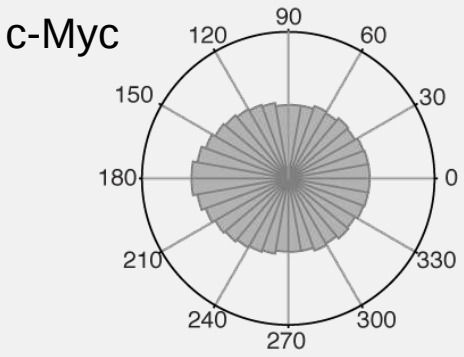
Diffusion on fractal structure

Experiments

Simulations

Izeddin et al, eLife (2014)

Nuclear exploration of c-Myc



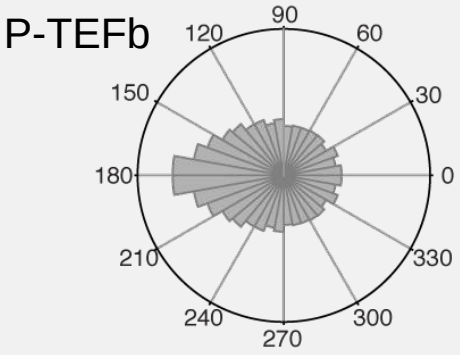
Dynamics (MSD) + Exploration geometry
Experiments + Simulations

- C-Myc is a non-compact explorer
- C-Myc explores the 3D nucleoplasm without constraints

Every target of c-Myc can be visited with the same probability, as in a homogenous medium

Izeddin et al, eLife (2014)

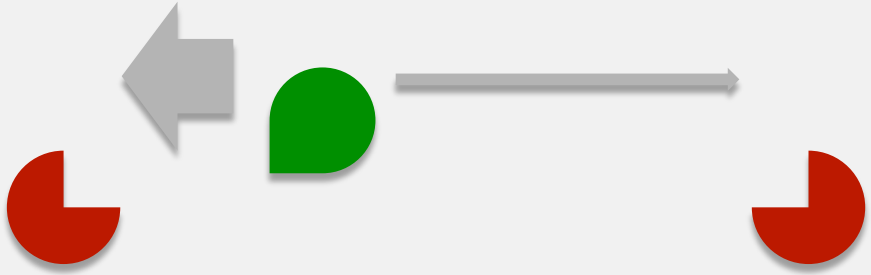
Nuclear exploration of P-TEFb



Dynamics (MSD) + Exploration geometry
Experiments + Simulations

- P-TEFb is a compact explorer
- C-Myc explores the 3D nucleoplasm without constraints

Interacting partners closer to P-TEFb have a larger probability of interaction



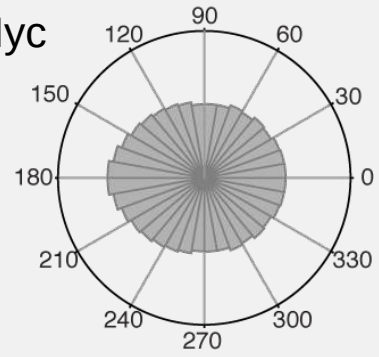
Du to oversampling, P-TEFb visits iteratively its partners

Izeddin et al, eLife (2014)

Compact vs non-compact exploration and the kinetics of the target search process

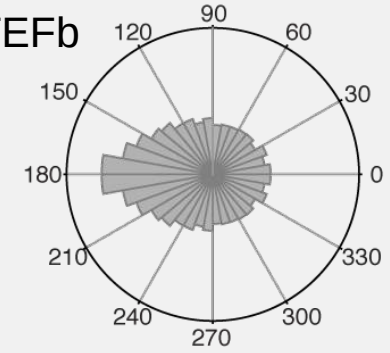
Non-compact

c-Myc

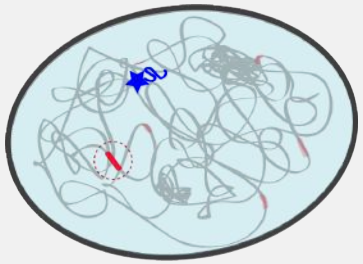


Compact

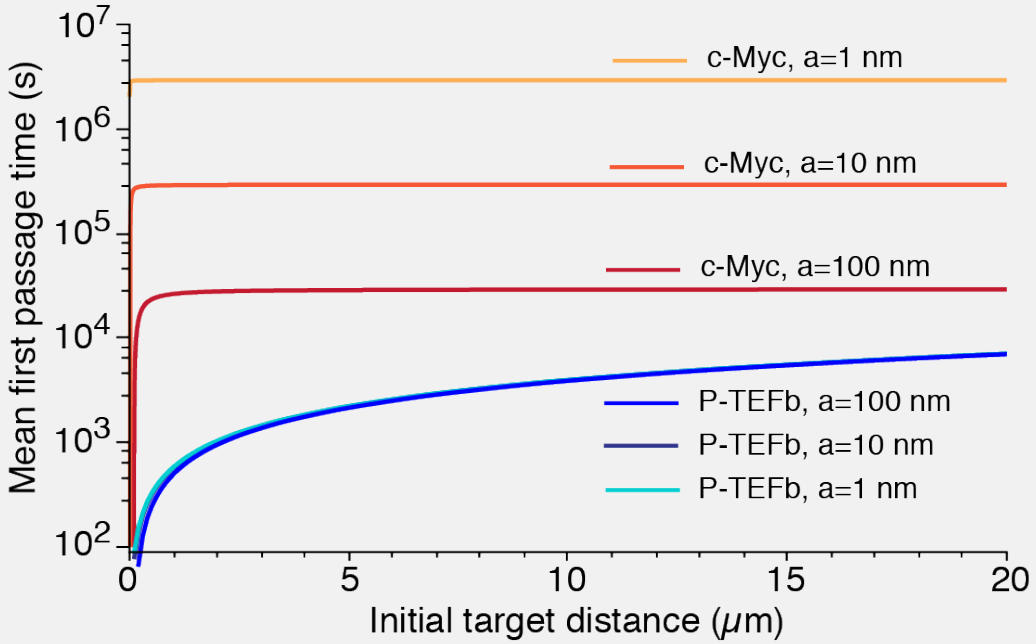
P-TEFb



Search time as a function of distance and target size

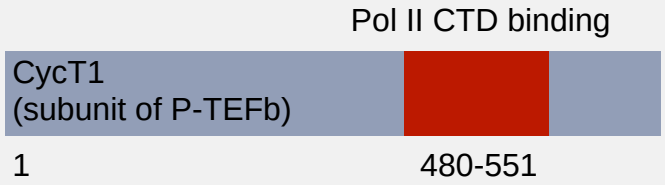


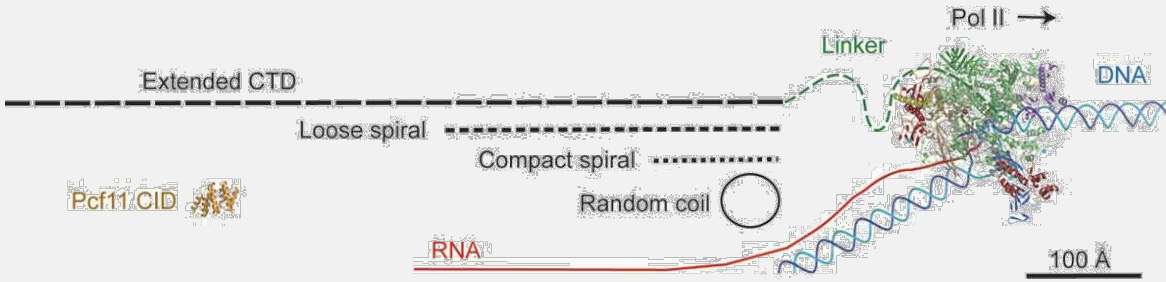
SEARCHER
Transcription factor
TARGET
Specific binding site
Target size (a)



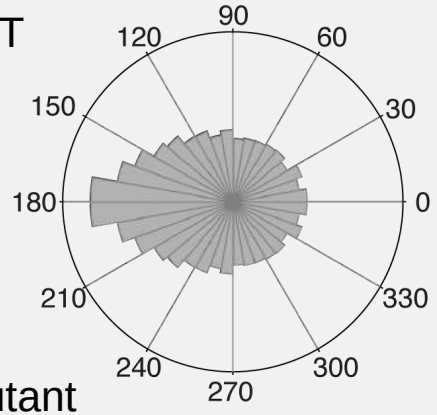
Izeddin et al, eLife (2014)

Binding to the CTD of the RNA Pol II guides P-TEFb exploration

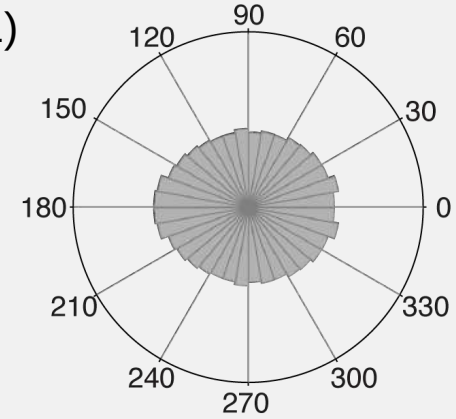


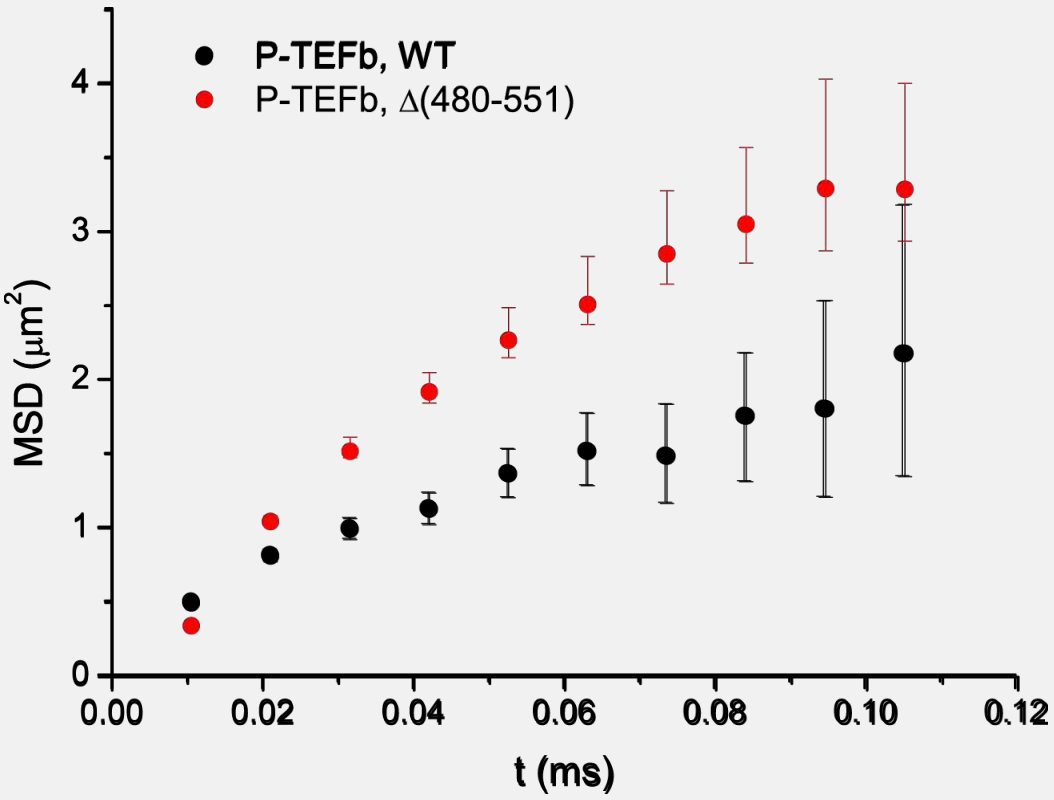


P-TEFb, WT



P-TEFb mutant
Δ(480-551)





Nuclear architecture, exploration and regulation of transcription regulation kinetics

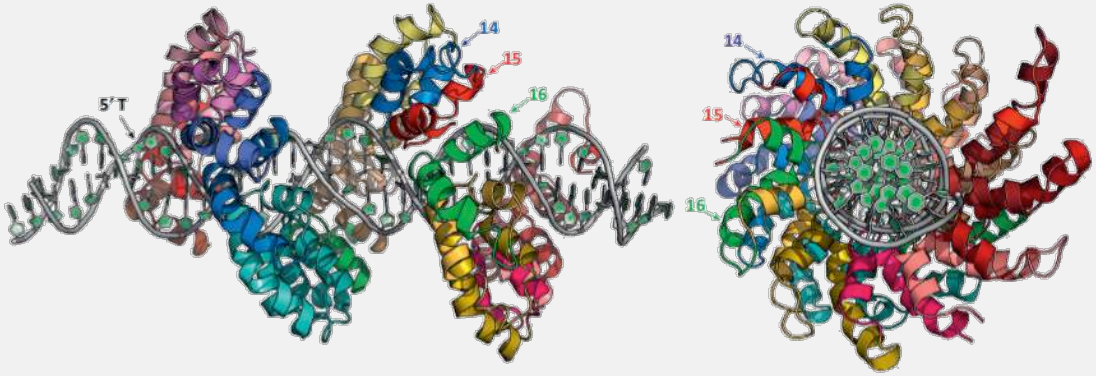
- Nuclear components are organized in the nucleus in a non-random manner
- Nuclear components have a “complex” dynamics

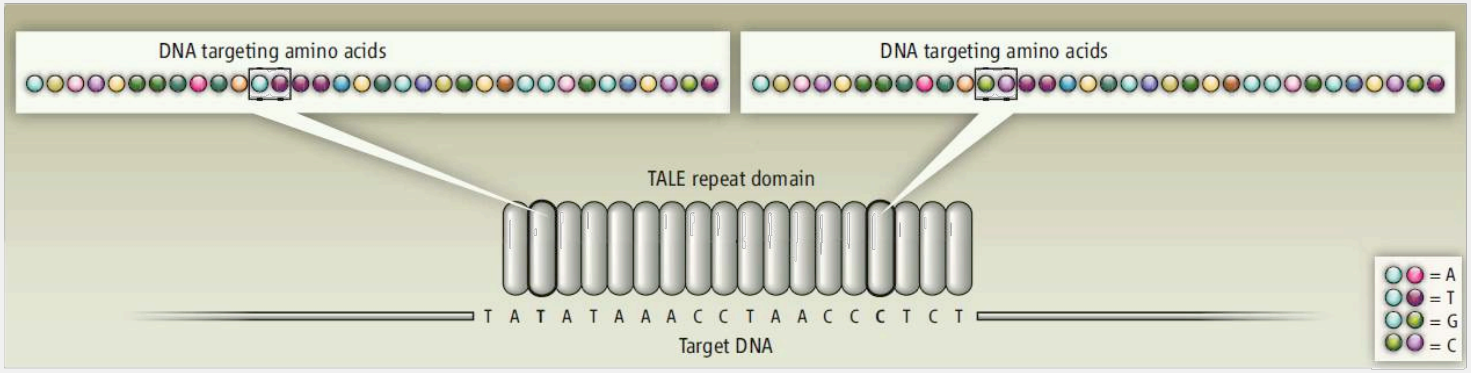
- Exploration strategy is protein-dependent
- Interactions with molecular partners define the explorations mode

Is protein mobility in the nucleus dominated by non-specific DNA-protein / protein-protein interactions?

Non-specific DNA-protein interactions
TALE proteins and LacI/LacO in mammalian cells

Thanks to **TALE** proteins (Transcription Activator-Like Effectors)
it is now possible to bind and modify virtually any site in the genome.

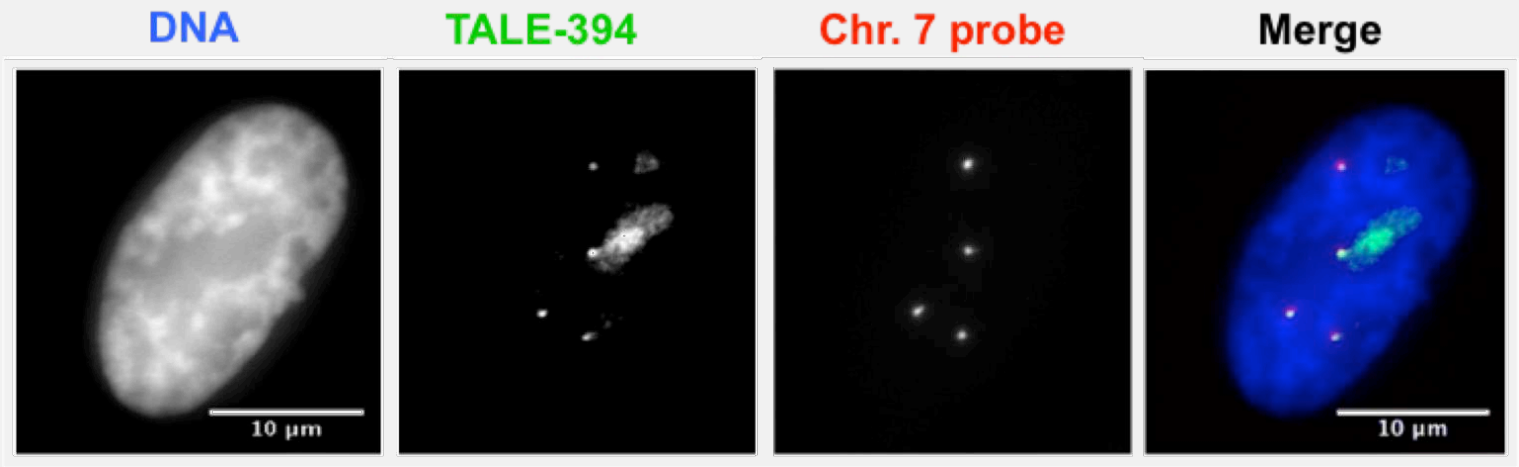




One-to-one code. Each TALE repeat targets a specific DNA base: A, T, C, or G. The repeats are the same series of 34 animo acids (color-coded) except at positions 12 and 13, which differ depending on the base being targeted.

Non-specific DNA-protein interactions
TALE proteins and LacI/LacO in mammalian cells

Our TALEs are designed to target the human alpha-satellite repeats from the centromeric region of chromosome 7 (Judith Lopes, MNHN)



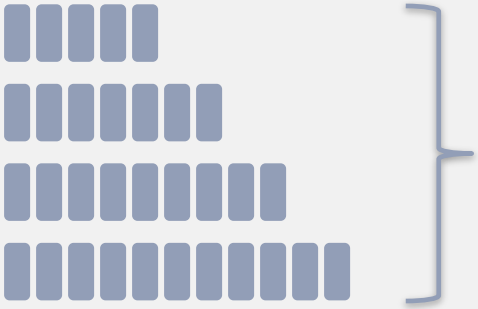
U2OS transfected by the TALE-394
DNA stained with Hoechst, TALE-394-GFP,
FISH on the alpha-satellite repeats of chr. 7

In collaboration with Judith Lopes (MNHN)

TALEs with recognition motifs of varying length

Set of TALEs engineered to target the same alpha-satellite repeated sequences, with recognition motifs of 10, 14, 18, 22 and 26 base pairs

TALE-GFP-Halo + TMR

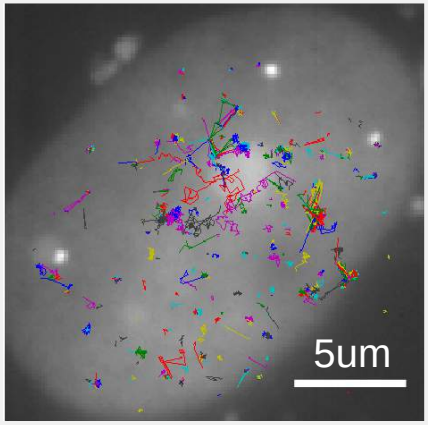


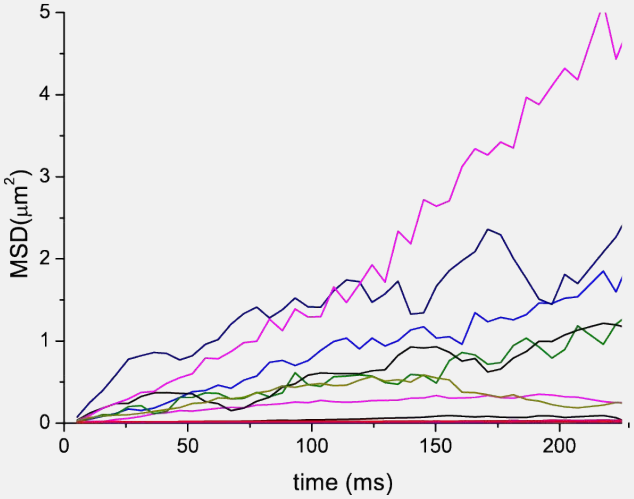
DNA recognition motif

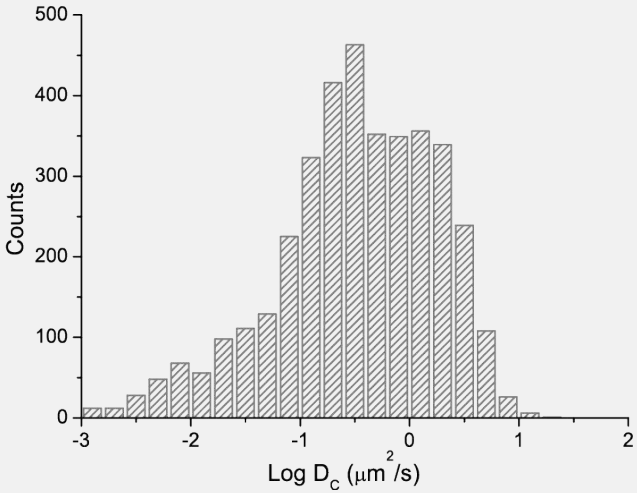
We can “tune” the strength of non-specific interactions
To what extent the diffusion properties are affected?

Dynamics of non-specific DNA-binding proteins

Dynamics of **TALEs** (18bp) in the nucleoplasm of U2Os cells

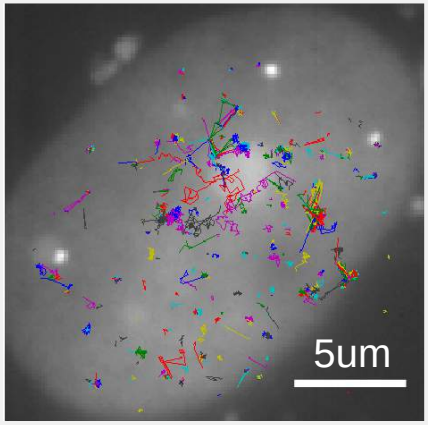


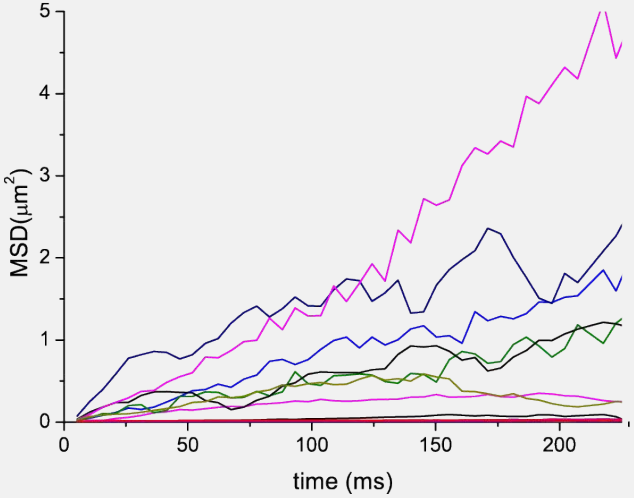


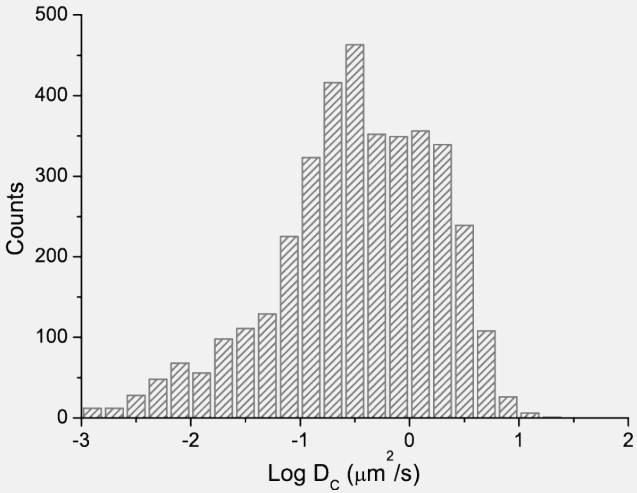


Dynamics of non-specific DNA-binding proteins

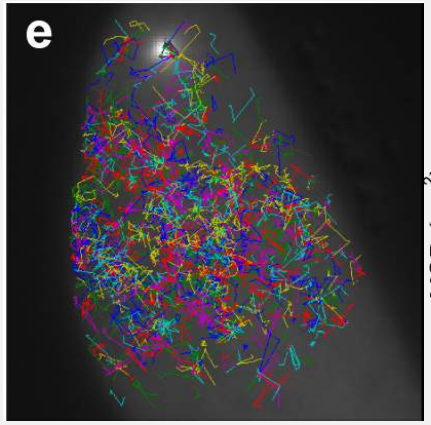
Dynamics of **TALEs** (18bp) in the nucleoplasm of U2Os cells

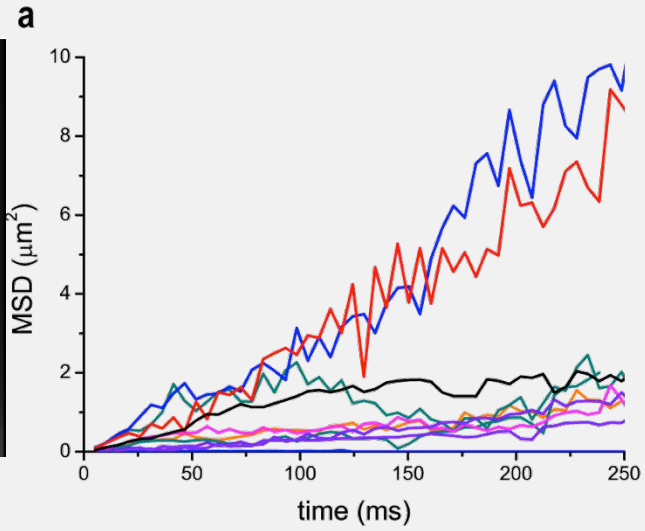


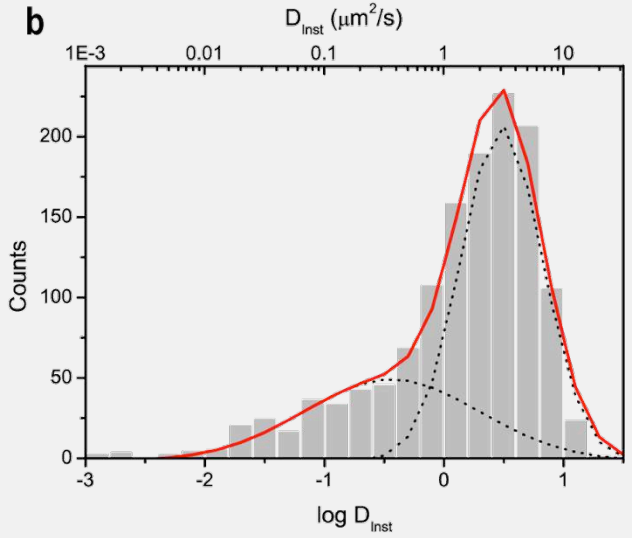




Dynamics of **LacI** in the nucleoplasm of U2Os cells

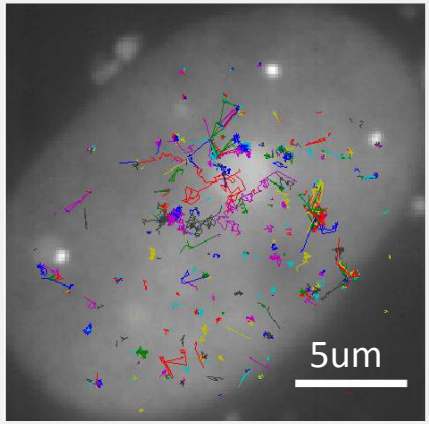


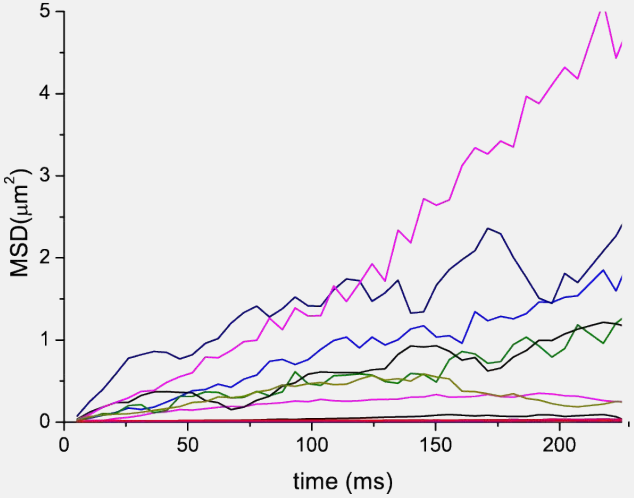


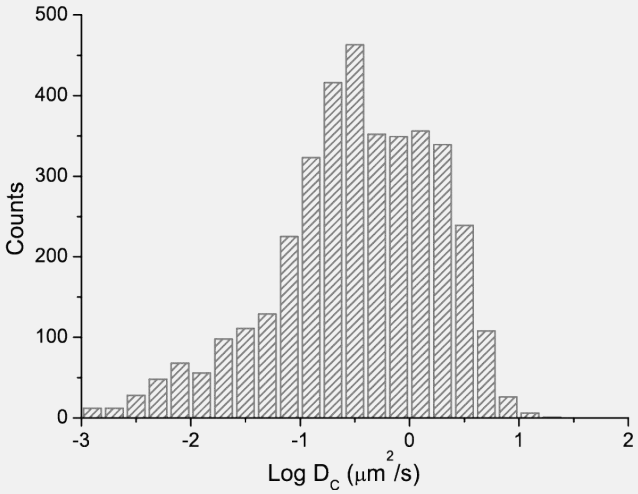


Residence time of non-specific DNA interactions

Diffusivity

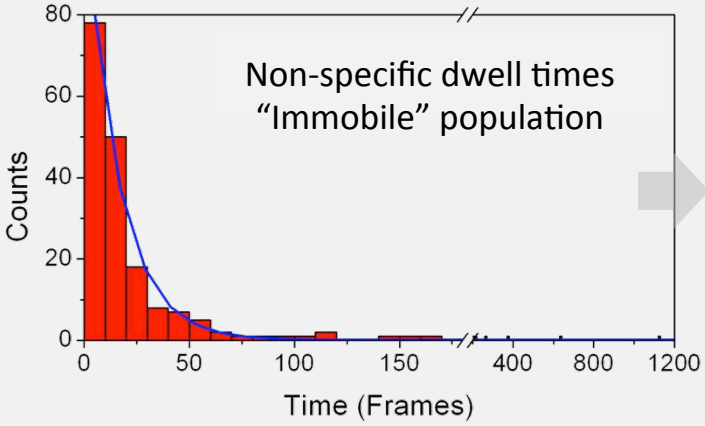


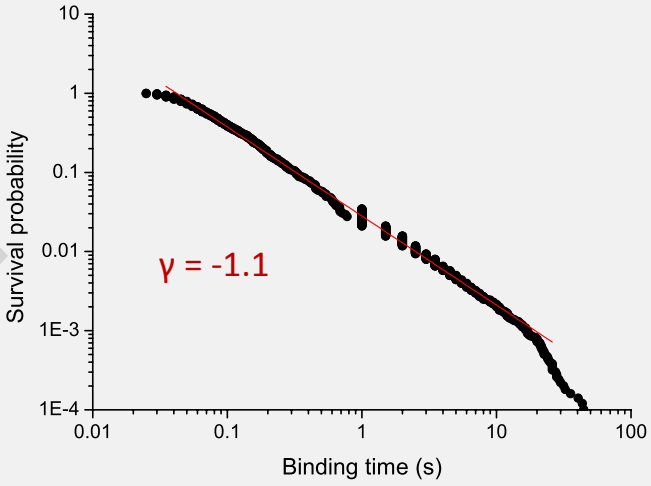




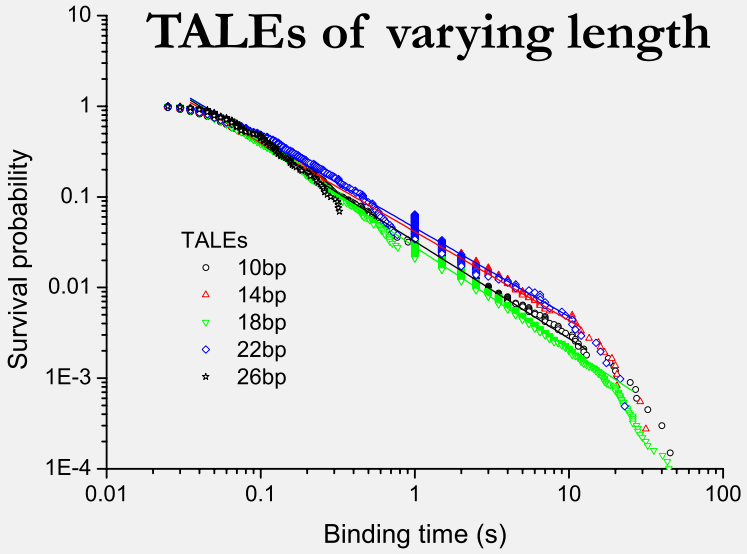
Choose immobile traces
(bound proteins)

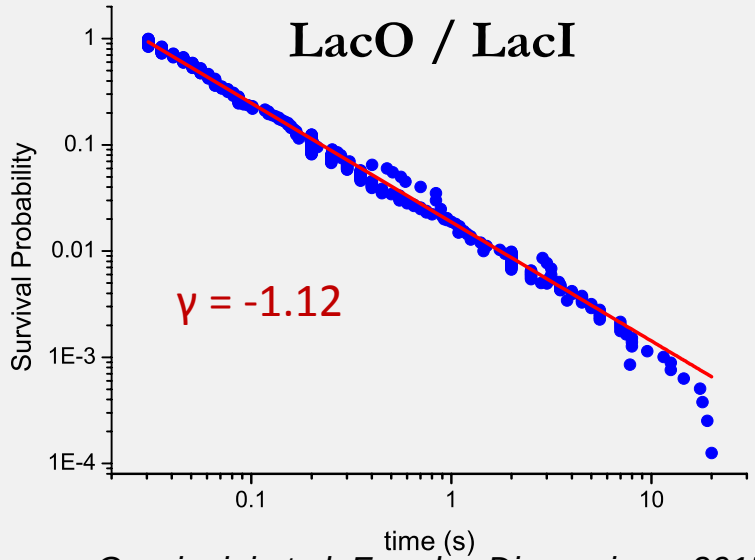
Non-specific binding



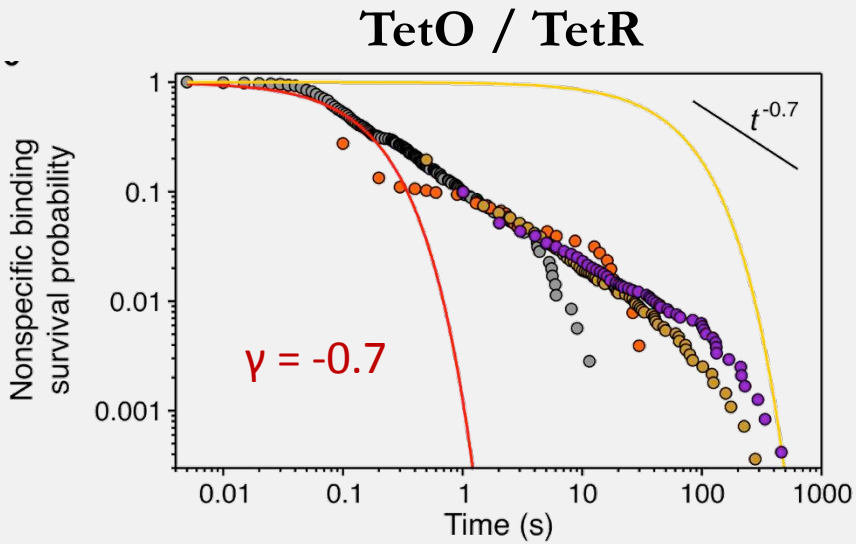


Power-law distribution of binding times (survival probability) for several protein-DNA interacting systems





Caccianini et al, Faraday Discussions. 2015

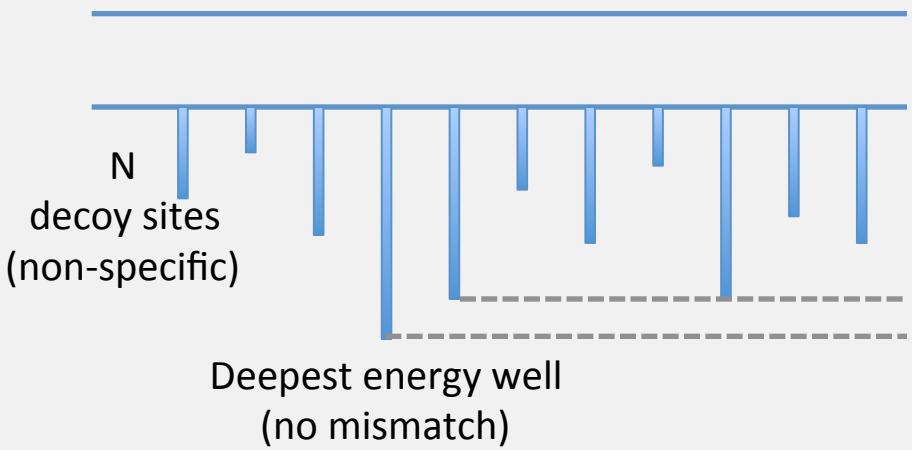


Normanno et al, Nat Comm. 2015

	Slope γ
TetR	-0.7
LacI	-1.12
TALE	-1.0 [-1.07, -1.0, -1.12,-1.0, -1.1]

What is the origin of the power-law distribution of binding times?

Simple energetic model for protein-DNA interaction



Energy of the freely diffusing protein

Energy of decoy interaction, Ens

For each mismatch, an energy penalty e

N = 1e8 decoy sites
e < 1, energy penalty per mismatch (constant or Poisson distribution)
L, length of the sequence
M, maximum number of mismatches
m, number of mismatches

Ens = e (L-(M+1))

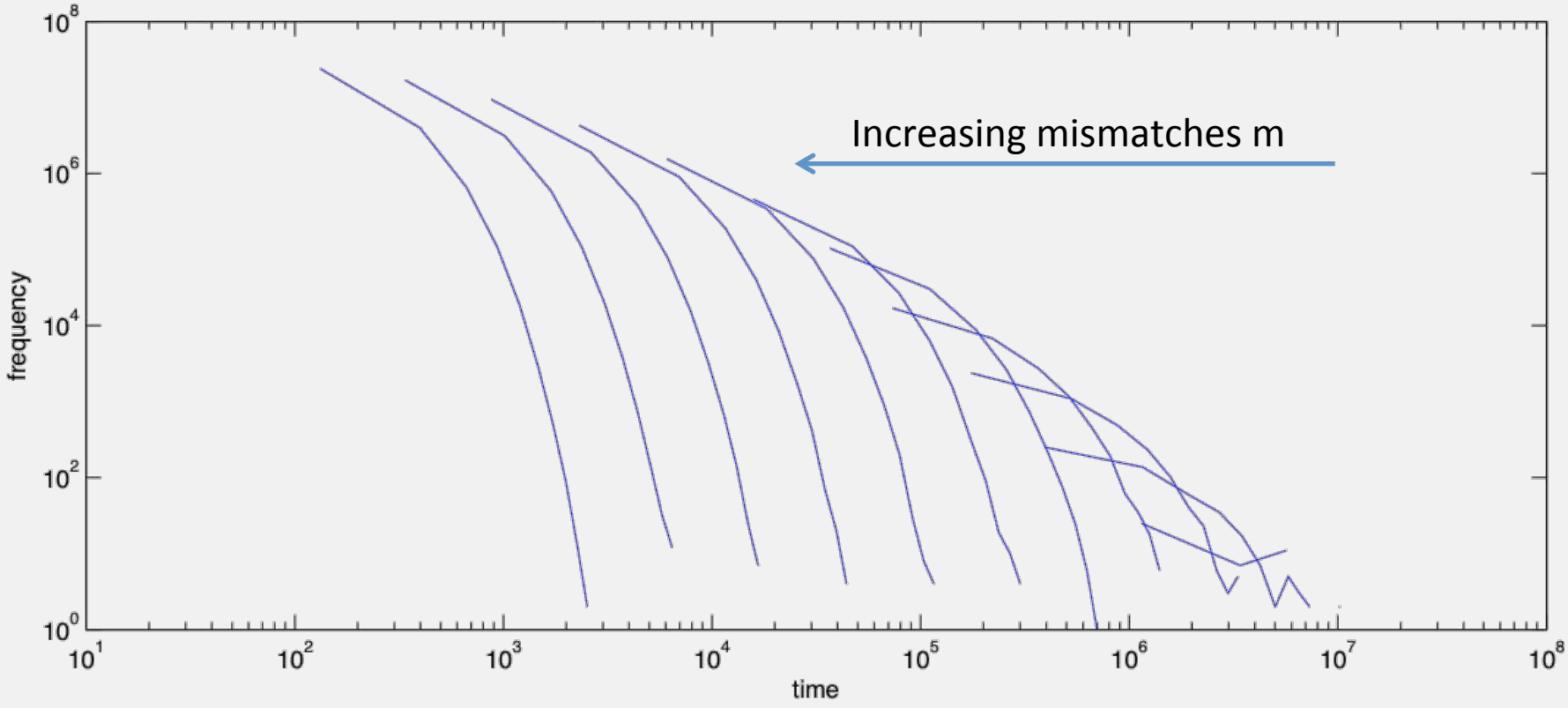
m = 0 … M

E = e(L-m)

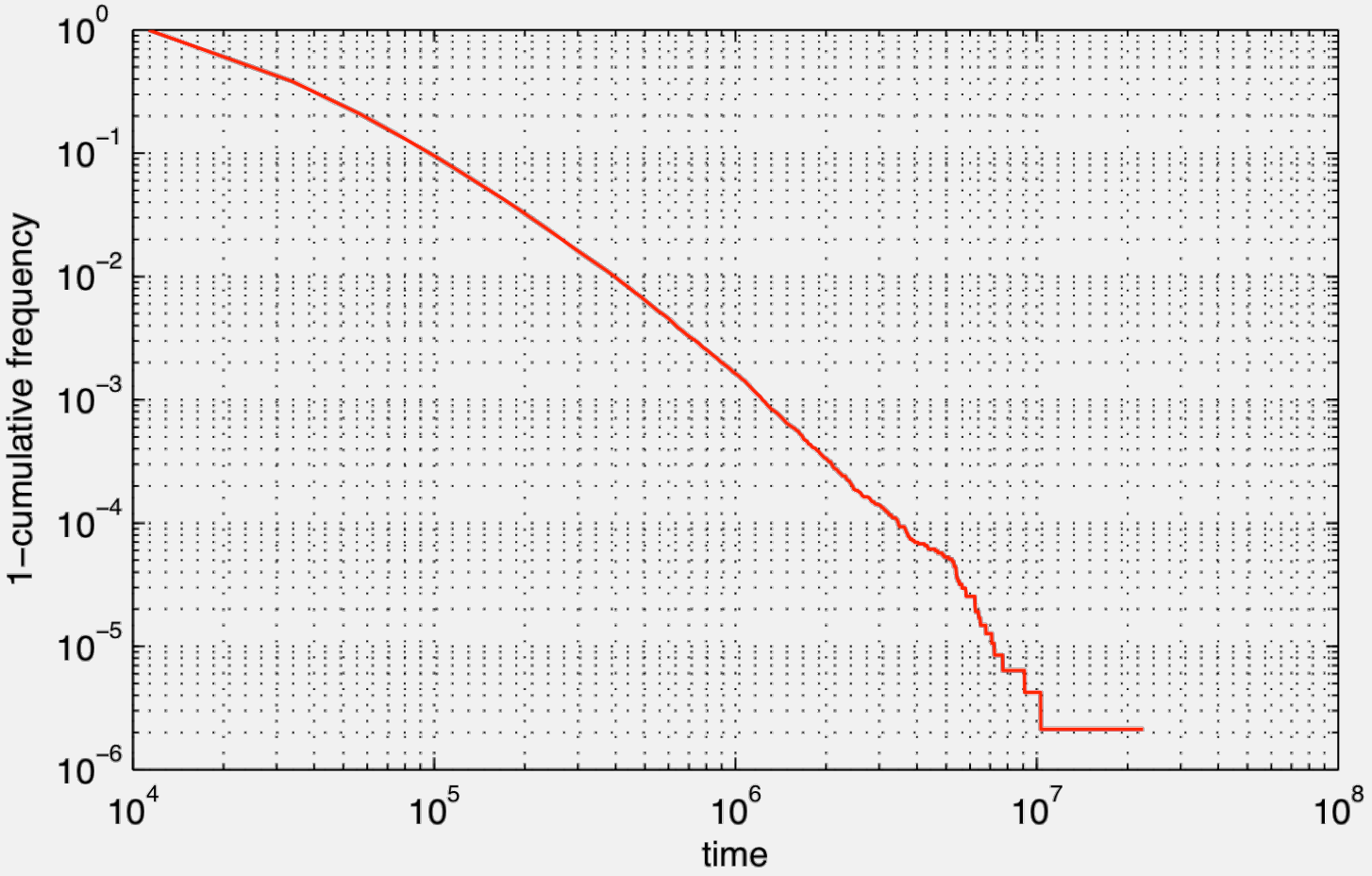
P = C(L,m) (1/4)^(L-m) (3/4)^m

tau = tau0 exp(-E)

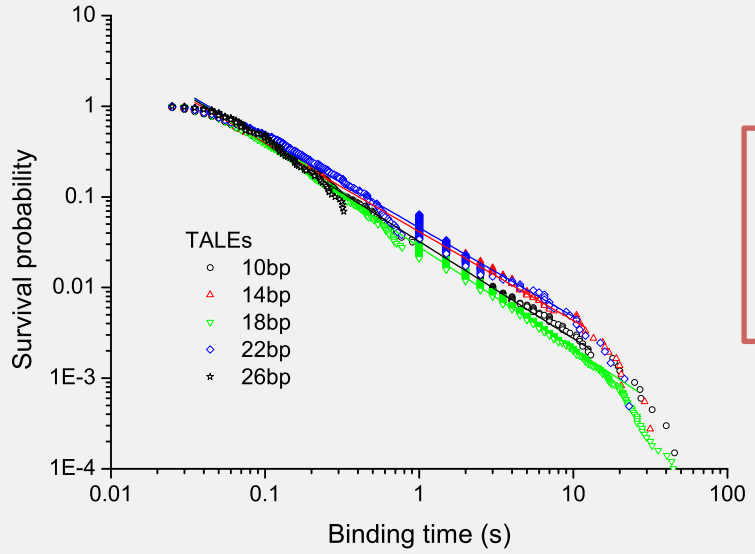
For each mismatch (m = 0 to M) we obtain an exponential distribution of residence times



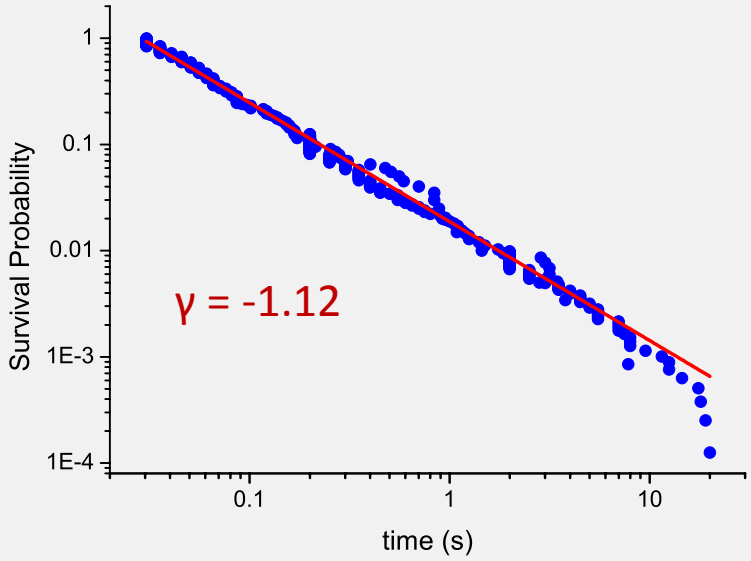
The envelope of which results in a power-law-like distribution of binding times



Comparison with experimental data
TALE proteins

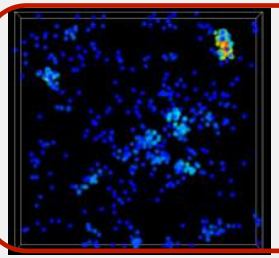


For TALEs ranging from L=10 to L=26, to obtain γ ≈ -1.1, the energy penalty ranges from e ≈ -2 kT to e ≈ -3.3 kT

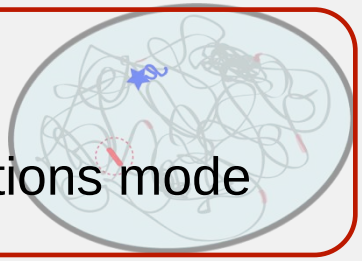


For L = 20bp (LacI), to obtain γ ≈ -1.1, the energy penalty is e ≈ -2.5 kT

Nuclear architecture, exploration and regulation of transcription regulation kinetics



- Nuclear components are organized in the nucleus in a non-random manner
- Nuclear components have a “complex” dynamics



- Exploration strategy is protein-dependent
- Interactions with molecular partners define the explorations mode

- Distribution of non-specific binding times scales as a power law for several DNA-binding proteins
- Possible cause: large variability of non-specific sequences encountered by the diffusing proteins

Acknowledgements

Functional Imaging of Transcription, UC Berkeley

- Xavier Darzacq
- Ibrahim Cissé (MIT), Vincent Récamier, Lana Bosanac

Imagerie et Contrôle Optique de l'Organisation Cellulaire, Physico-Chimie Curie, Institut Curie

- Maxime Dahan (Institut Curie)
- Laura Caccianini

Collaborators:

Muséum National d'Histoire Naturelle, CNRS UMR 7196 / INSERM U1154

- Judith Lopes

Centre de recherche en cancérologie de Marseille

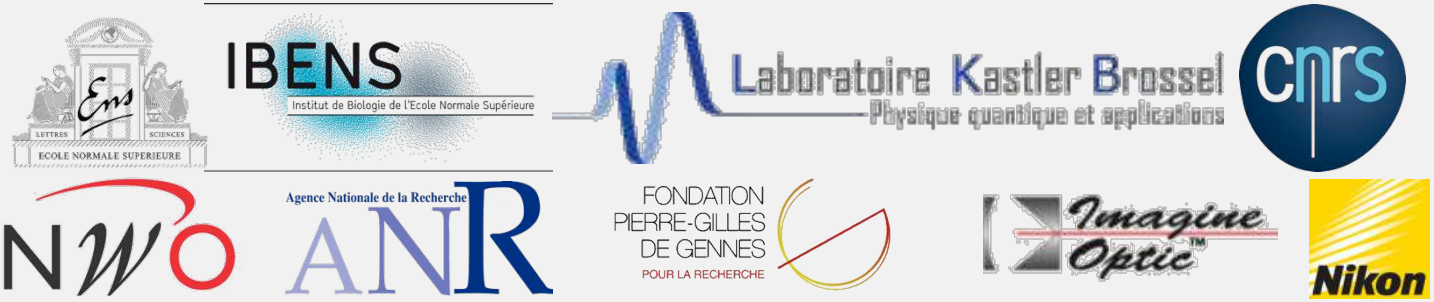
- Davide Normanno

LPTMC, CNRS UMR7600, UPMC, Jussieu

- Raphaël Voituriez and Olivier Bénichou

Massachusetts Institute of Technology (MIT), Cambridge, USA

- Leonid Mirny



Blank

3D diffusion and the non-specific binding times reported for different TFs in several cellular lines / organisms

Protein	Cell	τ_{1D}	τ_{3D}	Ref.
LacI	U2OS	182 ms	2.1 s	This study
LacI	*E. coli*	< 5 ms	< 1 ms	Elf *et al.*
TetR	U2OS	2 s	6 s	Normanno *et al.*
p53	H1299	1.7 s	1.8 s	Mazza *et al.*
GR	MCF / U2OS	1.5 s	–	Gebhardt *et al.*
Sox2	Mouse ES	0.8 s	3.3 s	Chen *et al.*