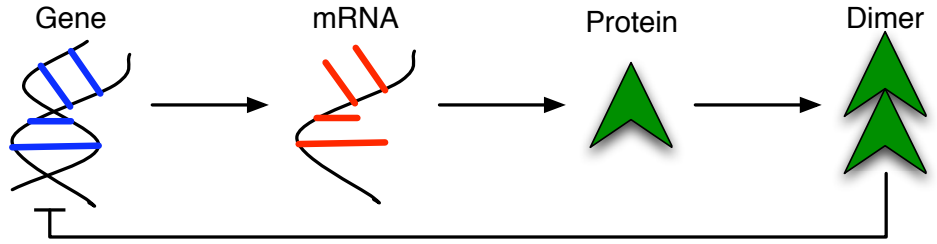
# Influence of Nuclear Substructure on Gene Regulation and Expression

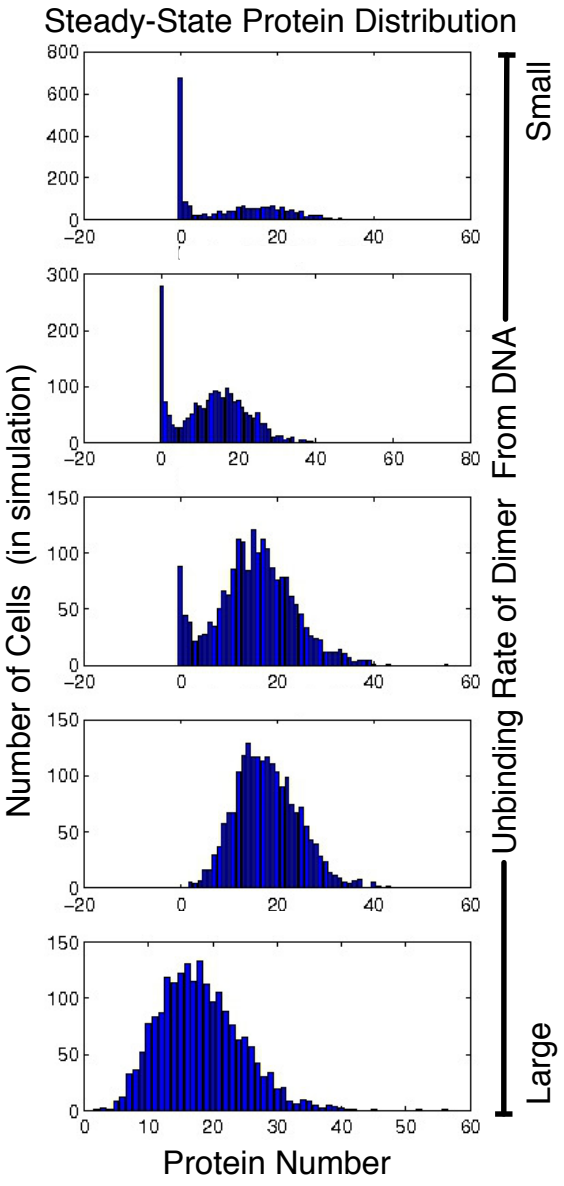
Samuel Isaacson

in collaboration with:
Carolyn Larabell (UCSF/LBL),
David McQueen (NYU),
Charles Peskin (NYU)

# Why is stochasticity in biochemical reactions important?

- Present in many cellular processes. Evolution.
- Arises from discreteness of chemical species populations.
- For example, gene expression.
  - Seen experimentally and theoretically.
  - See Arkin, Collins, Elowitz...
- Can serve useful biological purpose. See, for example, competence in *Bacillus subtilis*.



# What does the inside of a eukaryotic cell look like?

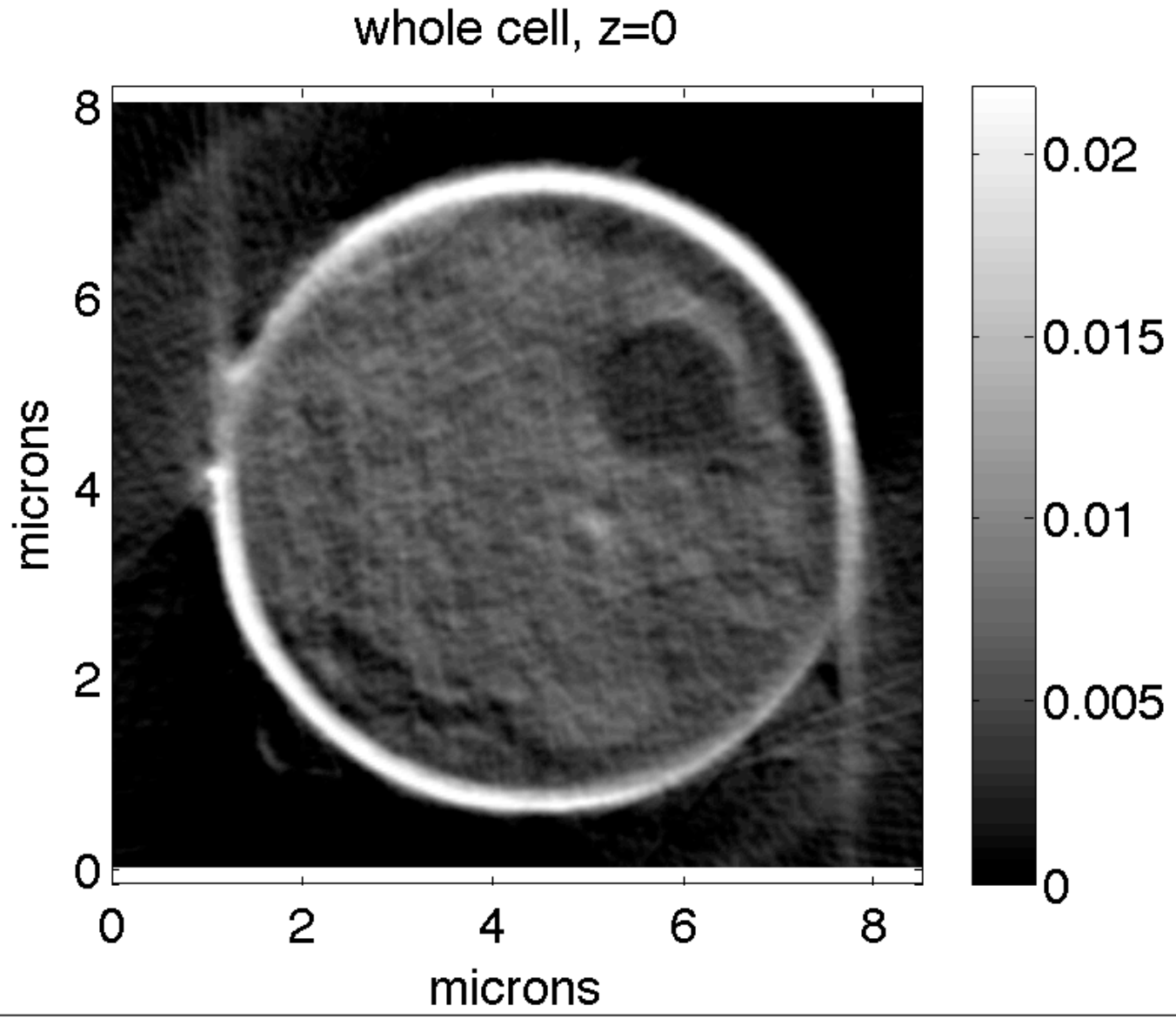
- This is an X-ray CT image of mouse olfactory epithelial cell.
- In this example a mouse cell is imaged inside a glass capillary.
- Pixel intensity is proportional to density of material in pixel.

Mouse Olfactory Epithelial Cell

# What does the inside of a eukaryotic cell look like?

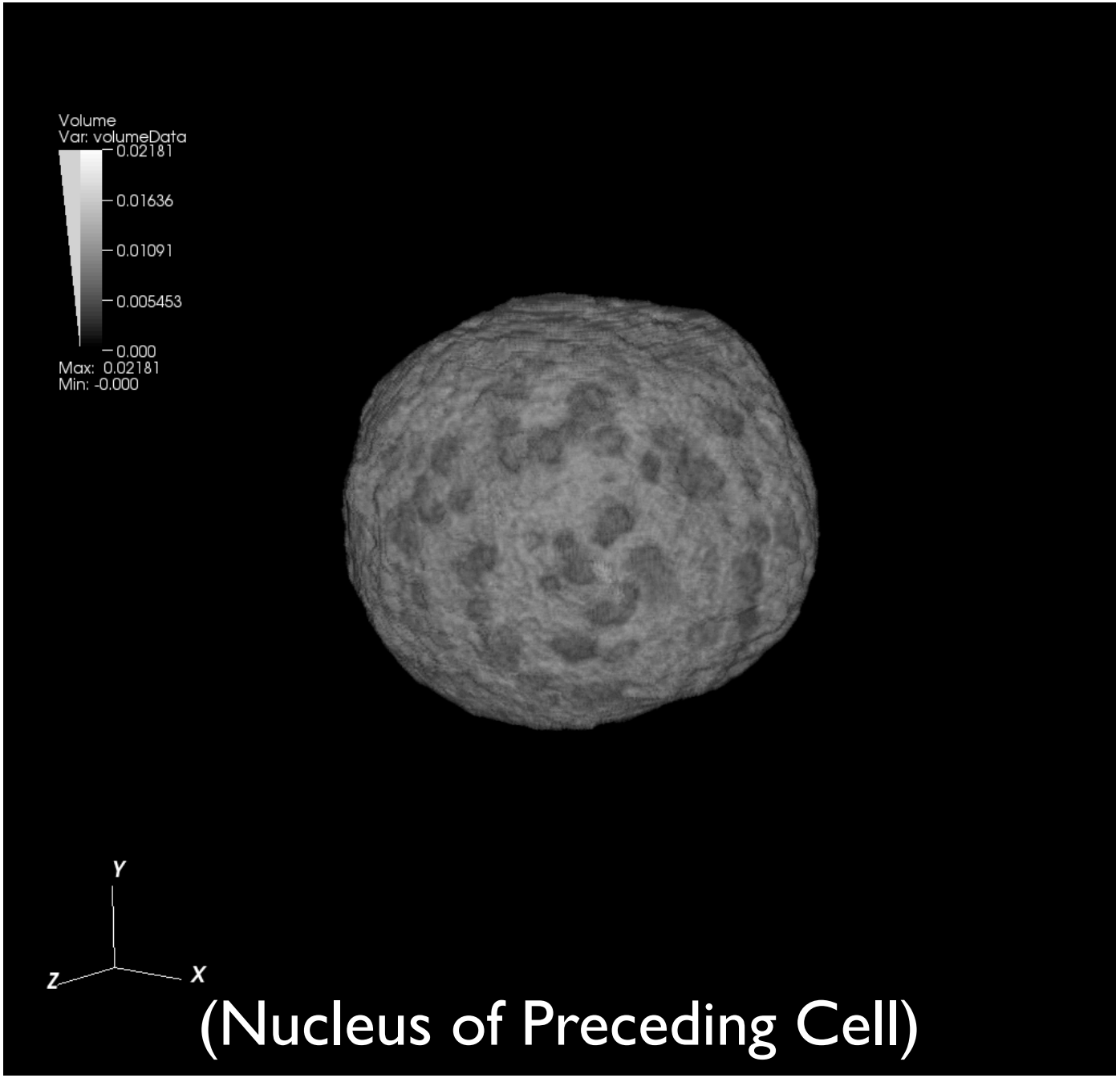
- This is an X-ray CT image of mouse olfactory epithelial cell.
- In this example a mouse cell is imaged inside a glass capillary.
- Pixel intensity is proportional to density of material in pixel.

Mouse Olfactory Epithelial Cell



# What does a nucleus look like in these reconstructions?

# What does a nucleus look like in these reconstructions?



# How does subcellular structure influence the dynamics of biochemical processes within cells?

- We are exploring this question by studying the effect of three-dimensional subcellular structure on gene regulation and expression.

# How does subcellular structure influence the dynamics of biochemical processes within cells?

- We are exploring this question by studying the effect of three-dimensional subcellular structure on gene regulation and expression.
- The following processes are influenced by the spatial organization of cellular substructure:
  - Gene Regulation: proteins searching for a DNA binding site within the nucleus.
  - mRNA export: mRNPs searching for nuclear pores and then translocating out of nucleus.
  - Protein import: movement of cytosolic proteins to nuclear pores and translocation into the nucleus.

# How does subcellular structure influence the dynamics of biochemical processes within cells?

- We are exploring this question by studying the effect of three-dimensional subcellular structure on gene regulation and expression.
- The following processes are influenced by the spatial organization of cellular substructure:
  - Gene Regulation: proteins searching for a DNA binding site within the nucleus.
  - mRNA export: mRNPs searching for nuclear pores and then translocating out of nucleus.
  - Protein import: movement of cytosolic proteins to nuclear pores and translocation into the nucleus.

**What is the influence of volume exclusion due to chromatin on the time required for regulatory proteins to find DNA binding sites?**

# What effect does explicitly modeling the three-dimensional spatial extent of chromatin have on gene regulation?

- Could chromatin help to funnel regulatory proteins towards binding sites thereby decreasing the time to find specific binding sites vs. models that do not incorporate the spatial chromatin organization?
- Could molecules instead get trapped in the chromatin matrix and meander far from binding sites?
- What differences arise when studying movement of nuclear proteins with volume exclusion due to chromatin vs. treating the nucleus as an empty volume?
- Do the observed dynamics in the 3D stochastic reaction-diffusion model for eukaryotic gene expression and regulation of Isaacson and Peskin (SISC 2006) change when chromatin and a realistic nuclear membrane surface are incorporated?

# Outline:

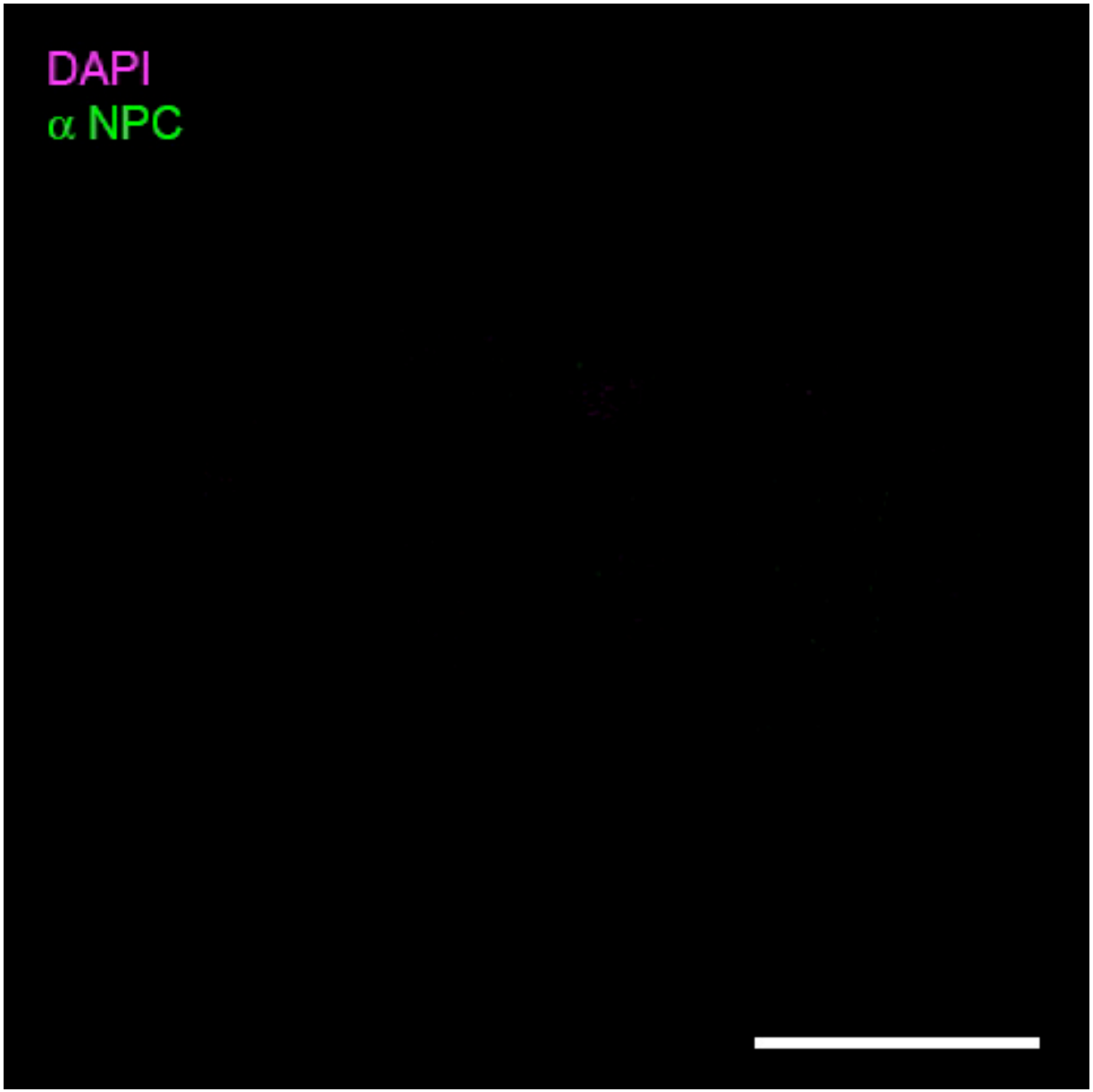
- How can we construct a mathematical model of the time to find a binding site using high-resolution DAPI fluorescence data?
- What happens if we study this question with more quantitative X-ray CT data?
- What properties of nuclear substructure might contribute to our observed results?

# Mouse myoblast cell nucleus

(Structured Illumination Microscopy of Schermelleh et al. Science 2008)

# Mouse myoblast cell nucleus

(Structured Illumination Microscopy of Schermelleh et al. Science 2008)



# How to model volume exclusion due to chromatin?

- We assume that regions of higher DAPI fluorescence intensity are more repulsive (harder to diffuse into).
- Volume exclusion due to chromatin is modeled as a repulsive potential, $\phi_i$, experienced by diffusing regulatory proteins.
- We have considered several functional relationships between the potential $\phi_i$ and the DAPI fluorescence intensity, $I_i$.
- For the remainder of the talk we choose a linear relationship,

$$\phi_i = \phi_{\max} I_i,$$

We subsequently call $\phi_{\max}$ the *volume exclusivity*.

# How do we model the search process of a protein for a DNA binding site?

For now, we consider a spatially-continuous model:

- We assume the protein diffuses in the repulsive potential, $\phi(\boldsymbol{x})$.
- Motion of the protein within the nucleus is then governed by a Fokker-Planck partial differential equation.
- The binding site is modeled as a sphere of radius $r_{\text{bind}}$ about the the point, $\boldsymbol{x}_{\text{bind}}$.

- Let $D$ be the diffusion constant of the protein.
- Denote by $\rho(\boldsymbol{x}, t)$ the probability density the protein has not yet found the binding site, and is located at $\boldsymbol{x}$, at time $t$.

# How do we model the search process of a protein for a DNA binding site? (2)

- Within the nucleus the Fokker-Planck equation describing the protein's motion is:

$$\frac{\partial \rho}{\partial t}(\boldsymbol{x}, t) = D\nabla \cdot \left[\nabla \rho(\boldsymbol{x}, t) + \frac{1}{k_{\mathsf{B}} T}\rho(\boldsymbol{x}, t)\nabla \phi(\boldsymbol{x}, t)\right]$$

- An absorbing boundary condition is used to model binding:

$$\rho(\boldsymbol{x}, t) = 0, \quad |\boldsymbol{x} - \boldsymbol{x}_{\mathsf{bind}}| = r_{\mathsf{bind}}$$

- At the nuclear membrane a no-flux boundary condition prevents the protein from diffusing out of the nucleus:

$$-D\left[\nabla \rho(\boldsymbol{x}, t) + \frac{1}{k_{\mathsf{B}} T}\rho(\boldsymbol{x}, t)\nabla \phi(\boldsymbol{x}, t)\right] \cdot \boldsymbol{\eta}(\boldsymbol{x}) = 0, \quad \forall \boldsymbol{x} \in \text{membrane},$$

where $\boldsymbol{\eta}(\boldsymbol{x})$ denotes the normal to the nuclear membrane at $\boldsymbol{x}$.

# How do we model the search process of a protein for a DNA binding site? (2)

- Within the nucleus the Fokker-Planck equation describing the protein's motion is:

$$\frac{\partial \rho}{\partial t}(\boldsymbol{x}, t) = D\nabla \cdot \left[\nabla \rho(\boldsymbol{x}, t) + \frac{1}{k_{\mathsf{B}} T}\rho(\boldsymbol{x}, t)\nabla \phi(\boldsymbol{x}, t)\right]$$

- An absorbing boundary condition is used to model binding:

$$\rho(\boldsymbol{x}, t) = 0, \quad |\boldsymbol{x} - \boldsymbol{x}_{\mathsf{bind}}| = r_{\mathsf{bind}}$$

- At the nuclear membrane a no-flux boundary condition prevents the protein from diffusing out of the nucleus:

$$-D\left[\nabla \rho(\boldsymbol{x}, t) + \frac{1}{k_{\mathsf{B}} T}\rho(\boldsymbol{x}, t)\nabla \phi(\boldsymbol{x}, t)\right] \cdot \boldsymbol{\eta}(\boldsymbol{x}) = 0, \quad \forall \boldsymbol{x} \in \text{membrane},$$

  where $\boldsymbol{\eta}(\boldsymbol{x})$ denotes the normal to the nuclear membrane at $\boldsymbol{x}$.

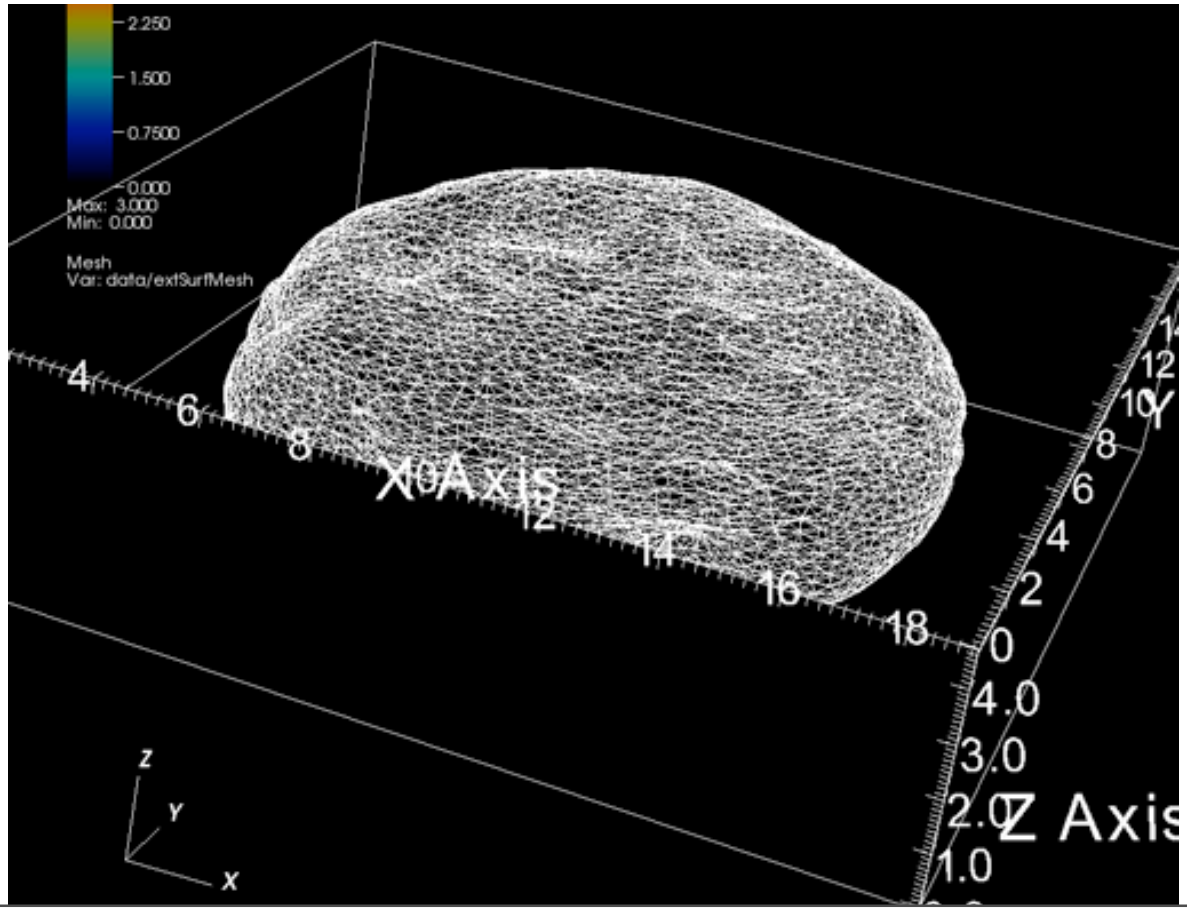
- **However, since the underlying imaging data and potential are defined on a lattice, we use a discrete approximation to this equation.**

# What is the spatially discrete model?

- We use a discretization of the Fokker-Planck equation that has the form of a reaction-diffusion master equation (RDME).
- To account for the nuclear membrane we use an extension of the embedded boundary method of Isaacson et al., SISC (2006).
- Discretization takes into account the intersection of the nuclear membrane with the natural mesh given by the imaging voxels.

# What is the spatially discrete model?

- We use a discretization of the Fokker-Planck equation that has the form of a reaction-diffusion master equation (RDME).
- To account for the nuclear membrane we use an extension of the embedded boundary method of Isaacson et al., SISC (2006).
- Discretization takes into account the intersection of the nuclear membrane with the natural mesh given by the imaging voxels.



# How do we discretize the Fokker-Planck equation?

Consider the one-dimensional Fokker-Planck equation

$$\frac{\partial \rho}{\partial t}(x,t) + \frac{\partial F}{\partial x}(x,t) = 0,$$

where the flux, $F(x,t)$, is given by

$$F(x,t) = -D\left(\frac{\partial \rho}{\partial x}(x,t) + \frac{\rho(x,t)}{k_{\mathrm{B}}T}\frac{\partial \phi}{\partial x}(\boldsymbol{x})\right).$$

# How do we discretize the Fokker-Planck equation?

Consider the one-dimensional Fokker-Planck equation

$$\frac{\partial \rho}{\partial t}(x,t) + \frac{\partial F}{\partial x}(x,t) = 0,$$

where the flux, $F(x,t)$, is given by

$$F(x,t) = -D\left(\frac{\partial \rho}{\partial x}(x,t) + \frac{\rho(x,t)}{k_{\mathrm{B}}T}\frac{\partial \phi}{\partial x}(\boldsymbol{x})\right).$$

Let

- $p_i(t) \approx \rho(ih,t)$, the probability density the regulatory protein is at location $ih$ at time $t$.
- $F_i(t) \approx F(ih,t)$ and $\phi_i \approx \phi(ih)$.

# How do we discretize the Fokker-Planck equation?

Consider the one-dimensional Fokker-Planck equation

$$\frac{\partial \rho}{\partial t}(x,t) + \frac{\partial F}{\partial x}(x,t) = 0,$$

where the flux, $F(x,t)$, is given by

$$F(x,t) = -D\left(\frac{\partial \rho}{\partial x}(x,t) + \frac{\rho(x,t)}{k_{\mathrm{B}}T}\frac{\partial \phi}{\partial x}(\boldsymbol{x})\right).$$

Let

- $p_i(t) \approx \rho(ih,t)$, the probability density the regulatory protein is at location $ih$ at time $t$.
- $F_i(t) \approx F(ih,t)$ and $\phi_i \approx \phi(ih)$.

We look for a discretization of the form

$$\frac{dp_i}{dt}(t) + \frac{1}{h}\left(F_{i+1/2} - F_{i-1/2}\right) = 0.$$

# How do we discretize the Fokker-Planck equation? (2)

We also assume that

$$F_{i+1/2} = \alpha(\phi_{i+1} - \phi_i)\, p_i - \beta(\phi_{i+1} - \phi_i)\, p_{i+1}.$$

where $\alpha$ and $\beta$ are functions to be determined. Notice

$$\begin{aligned} F_{i+1/2} = \quad & [\alpha(\phi_{i+1} - \phi_i) - \beta(\phi_{i+1} - \phi_i)] \left( \frac{p_{i+1} + p_i}{2} \right) \\ & - [\alpha(\phi_{i+1} - \phi_i) + \beta(\phi_{i+1} - \phi_i)] \left( \frac{p_{i+1} - p_i}{2} \right). \end{aligned}$$

We *choose* the second term to approximate the diffusive component of the flux, so that the standard three-point discrete Laplacian is recovered:

$$\alpha(\phi_{i+1} - \phi_i) + \beta(\phi_{i+1} - \phi_i) = \frac{2D}{h}.$$

# How we can use detailed balance to determine $\alpha$ and $\beta$?

Following Wang et al. (JTB 2003), in thermodynamic equilibrium we expect the probability density the molecule is at position $x$ to be proportional to the Boltzmann distribution.

$$\rho^{\mathsf{eq}}(x) \propto e^{-\phi(x)/k_{\mathsf{B}}T}.$$

We therefore require that

$$p_{i+1}^{\mathrm{eq}} = p_i^{\mathrm{eq}} e^{(\phi_i - \phi_{i+1})/k_{\mathsf{B}}T},$$

where $p_i^{\mathrm{eq}} = \lim_{t\to\infty} p_i(t)$. Moreover, at thermodynamic equilibrium detailed balance requires that

$$F_{i+1/2} = 0.$$

Combining these last two equations we obtain a second equation for $\alpha$ and $\beta$.

# What are the discretization weights?

We now have two equations for $\alpha$ and $\beta$. Solving them gives

$$\alpha\left(\phi_{i+1}-\phi_i\right)=\frac{2D}{h}\frac{1}{e^{(\phi_{i+1}-\phi_i)/k_{\mathsf{B}}T}+1},$$

$$\beta\left(\phi_{i+1}-\phi_i\right)=\frac{2D}{h}\frac{1}{e^{(\phi_i-\phi_{i+1})/k_{\mathsf{B}}T}+1}.$$

# What are the discretization weights?

We now have two equations for $\alpha$ and $\beta$. Solving them gives

$$\alpha\left(\phi_{i+1}-\phi_i\right)=\frac{2D}{h}\frac{1}{e^{(\phi_{i+1}-\phi_i)/k_{\mathsf{B}}T}+1},$$
$$\beta\left(\phi_{i+1}-\phi_i\right)=\frac{2D}{h}\frac{1}{e^{(\phi_i-\phi_{i+1})/k_{\mathsf{B}}T}+1}.$$

- These two functions then determine the flux, $F_{i+1/2}$.
- The overall discretization is second order in space.
- To discretize the three-dimensional Fokker-Planck equation in complex geometries
  - We extend the one-dimensional discretization to three-dimensions.
  - We use the expressions for the three-dimensional flux from this discretization in the Cartesian grid embedded boundary method of Isaacson et al. (SISC 2006).

# What is the 3D spatially discrete master equation model?

- $p(\boldsymbol{i}, t) =$ the probability the regulatory protein is in voxel $\boldsymbol{i}$ at time $t$. $(\boldsymbol{i} = (i_1, i_2, i_3))$.
- $D =$ the diffusion constant of the protein.
- $\beta =$ the bimolecular reaction rate constant, units $\mu\text{m}^3/s$.
- $\boldsymbol{i}_0$ the voxel containing the binding site.

# What is the 3D spatially discrete master equation model?

- $p(\boldsymbol{i}, t) =$ the probability the regulatory protein is in voxel $\boldsymbol{i}$ at time $t$. $(\boldsymbol{i} = (i_1, i_2, i_3))$.
- $D =$ the diffusion constant of the protein.
- $\beta =$ the bimolecular reaction rate constant, units $\mu\text{m}^3/s$.
- $\boldsymbol{i}_0$ the voxel containing the binding site.

Then

$$\frac{dp}{dt}(\boldsymbol{i}, t) = D \sum_{\boldsymbol{j}} \left[k_{\boldsymbol{i},\boldsymbol{j}} p(\boldsymbol{j}, t) - k_{\boldsymbol{j},\boldsymbol{i}} p(\boldsymbol{i}, t)\right] - \frac{\beta}{V_{\boldsymbol{i}_0}} \delta_{\boldsymbol{i},\boldsymbol{i}_0} p(\boldsymbol{i}_0, t),$$

where

$$\begin{cases} k_{\boldsymbol{i},\boldsymbol{j}} = \frac{2DA_{\boldsymbol{i},\boldsymbol{j}}}{h_d V_{\boldsymbol{j}}} \frac{1}{\exp\left((\phi_{\boldsymbol{i}} - \phi_{\boldsymbol{j}})/k_{\text{B}}T\right)+1}, & \boldsymbol{i} \text{ a neighbor of } \boldsymbol{j}, \\ 0, \quad \text{else.} \end{cases}$$

# How to solve the system of differential-difference equations?

- We could solve the system of ODEs numerically.
- For the voxel mesh determined by the image planes we get a 480x480x37 system of ODEs.
- This approach will be impractical when we have more than a few diffusing proteins.
- Also, if the binding site location is itself a random variable, then these equations contain a random coefficient.

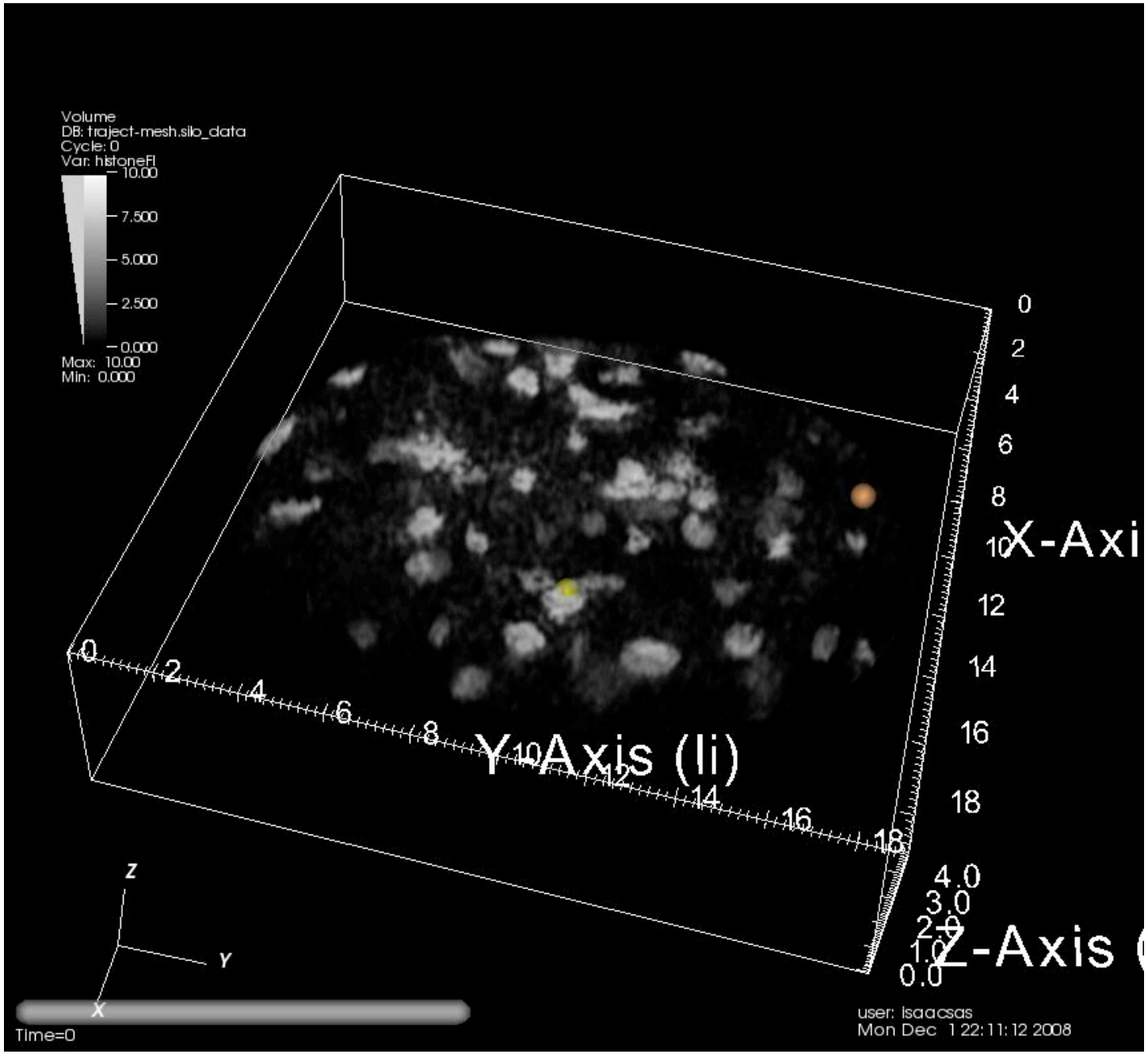
# How to solve the system of differential-difference equations?

- We could solve the system of ODEs numerically.
- For the voxel mesh determined by the image planes we get a 480x480x37 system of ODEs.
- This approach will be impractical when we have more than a few diffusing proteins.
- Also, if the binding site location is itself a random variable, then these equations contain a random coefficient.

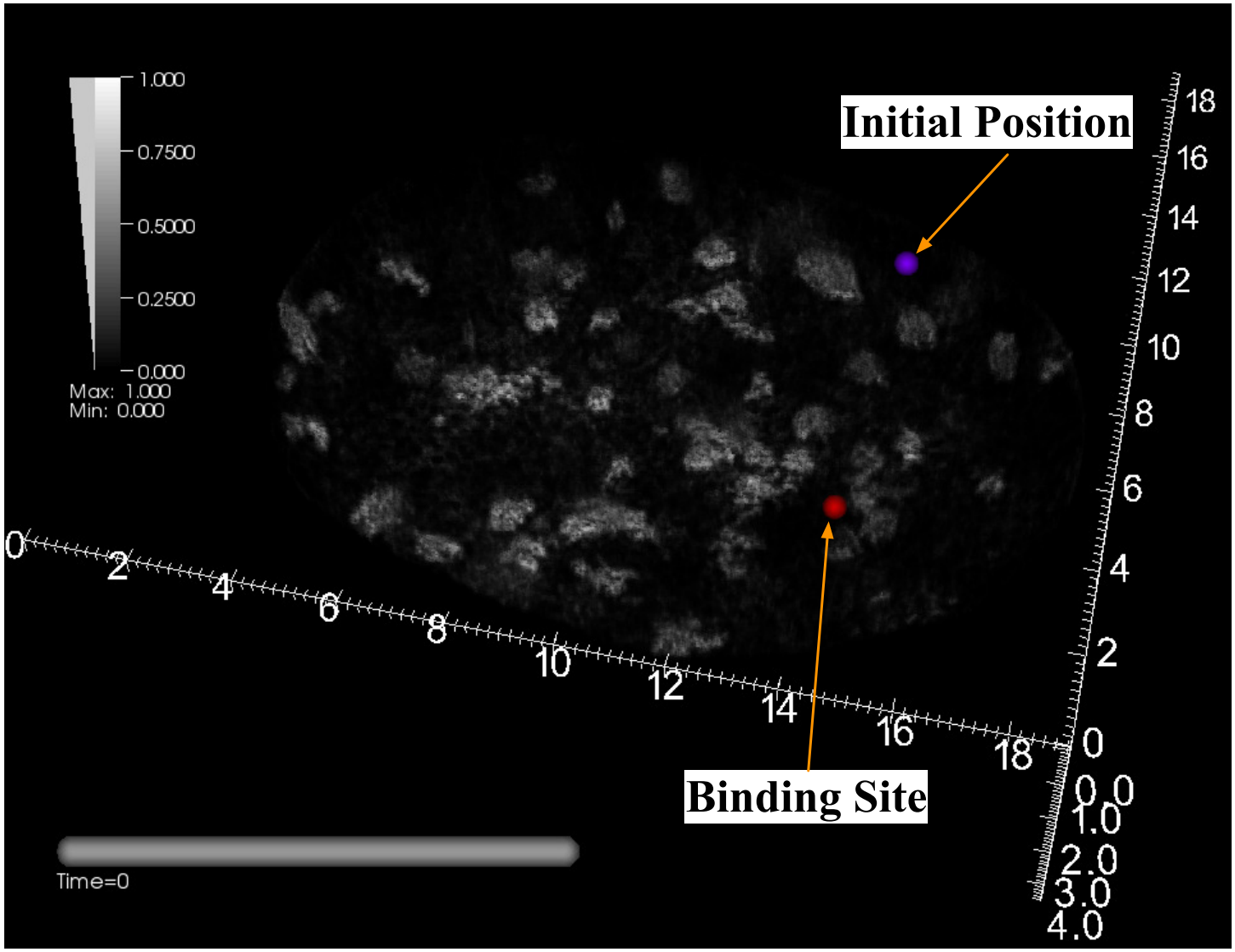
- Instead, we create realizations of the stochastic process described by the ODEs, that of a molecule undergoing a continuous time random walk on the voxel lattice in the potential, $\phi_{\boldsymbol{i}}$.
- We can exactly simulate this process by the Gillespie method.
- In the simulations, the protein hops from voxel $\boldsymbol{i}$ to $\boldsymbol{j}$ with probability per unit time $k_{\boldsymbol{j},\boldsymbol{i}}$.
- When entering a voxel, $\boldsymbol{i}_0$, containing a binding site, the protein can bind with probability per unit time $\beta/V_{\boldsymbol{i}_0}$.

What does a typical protein search process for a DNA binding site look like?
In our stochastic reaction-diffusion model we find:

What does a typical protein search process for a DNA binding site look like?
In our stochastic reaction-diffusion model we find:

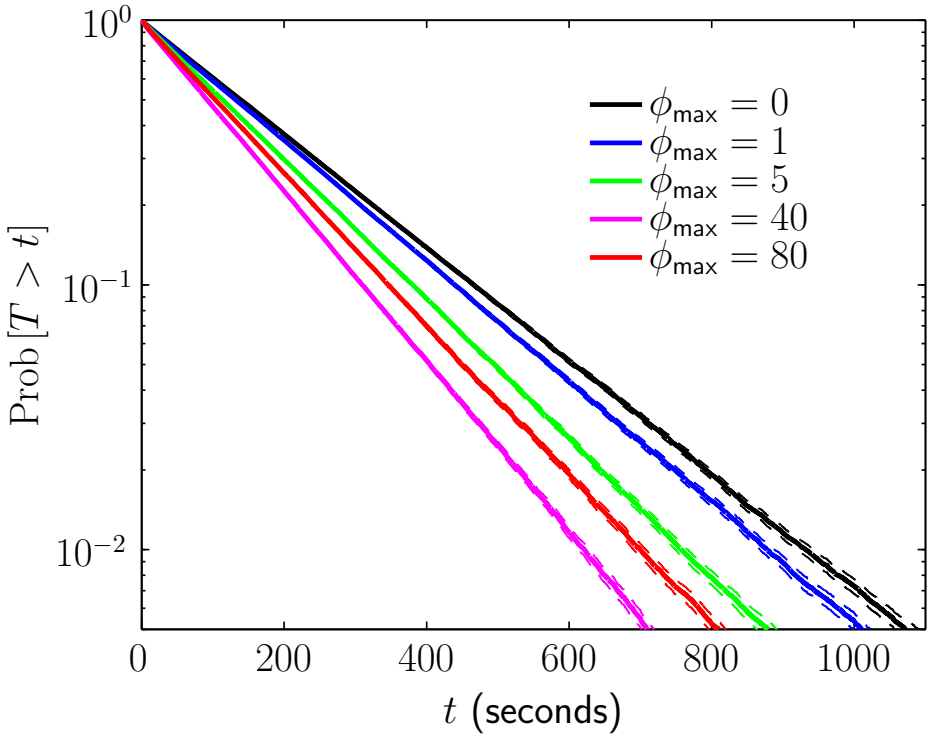


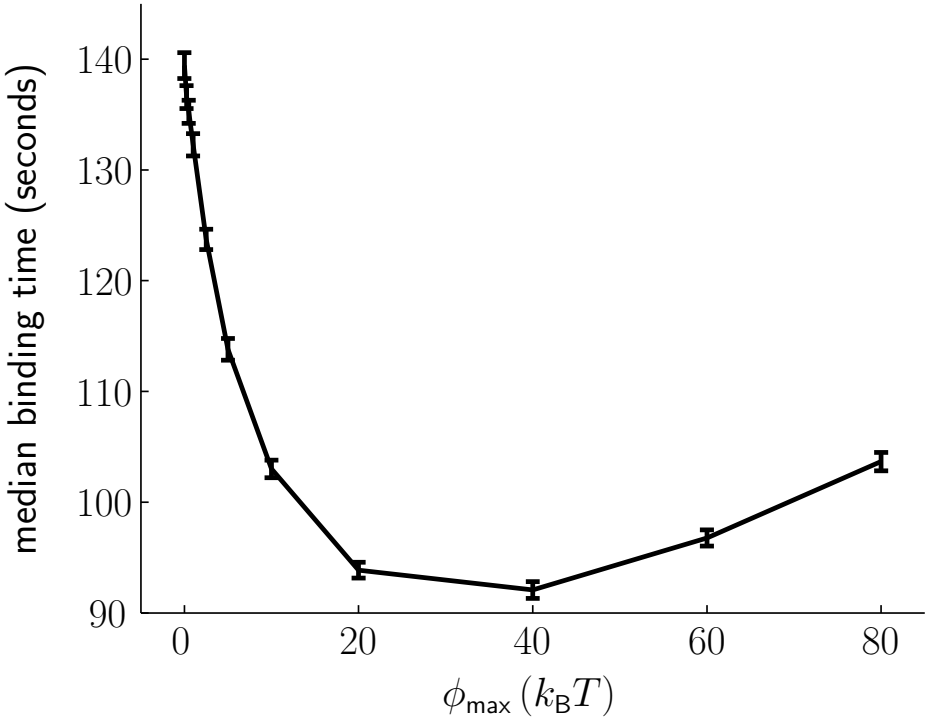
# How does volume exclusion influence the time needed to find a specific binding site?



# How does volume exclusion influence the time needed to find a specific binding site?

For a diffusion constant of 10 $\mu$m$^2$/s, and the initial and binding site positions of the previous slide:




See Isaacson, McQueen, and Peskin, PNAS (2011).

# Can we estimate the binding time distribution?

The preceding distributions appear to be well-approximated by an exponential distribution, *i.e.*

$$\text{Prob}\left[T > t\right] \approx 1 - e^{-D\lambda t}.$$

We can estimate $\lambda$ using the numerically calculated medians from our Monte Carlo simulations. For an exponential distribution

$$T_{\text{med}} = \frac{\ln(2)}{D\lambda}.$$

# Can we estimate the binding time distribution?

The preceding distributions appear to be well-approximated by an exponential distribution, *i.e.*

$$\text{Prob}\left[T > t\right] \approx 1 - e^{-D\lambda t}.$$

We can estimate $\lambda$ using the numerically calculated medians from our Monte Carlo simulations. For an exponential distribution

$$T_{\text{med}} = \frac{\ln(2)}{D\lambda}.$$

We have investigated several other ways to estimate the observed binding rate, $D\lambda$:

- Smoluchowski diffusion limited reaction rate arguments.
- Eigenvalues of the transition rate matrix for the master equation.
- Perturbation theory.

# Can we estimate the binding time distribution?(2)

Denote by $V$ the volume of the nucleus.

1. Standard diffusion limited reaction-rate theory

$$\lambda \approx \frac{4\pi r_{\mathrm{b}}}{V}.$$

2. A lattice diffusion limited reaction-rate theory gives

$$\frac{1}{\lambda} \approx \frac{16}{(2\pi)^3} \frac{V}{h_x h_y h_z} \iiint_0^{\pi/2} \frac{d\theta d\phi d\psi}{\left(\frac{\sin(\theta)}{h_x}\right)^2 + \left(\frac{\sin(\phi)}{h_y}\right)^2 + \left(\frac{\sin(\psi)}{h_z}\right)^2}.$$

3. Denote by $\boldsymbol{p}(t)$ the vector with components $\boldsymbol{p_i}(t) = p(\boldsymbol{i}, t)$. Then

$$\frac{d\boldsymbol{p}}{dt}(t) = L\boldsymbol{p}(t),$$

where $L$ denotes the transition matrix. We can estimate the binding rate by the smallest eigenvalue of $L$.

# Can we estimate the binding time distribution?(3)

For example, when $\phi_{\max} = 0$ we find the following estimates for the rate $D\lambda$:

| **Monte Carlo** | **DLR** | **Latt. DLR** | **Trans. Mat. Eval.** |
|---|---|---|---|
| .005038 $\pm$ .000039 | .0094 to .0326 | .00524 | .0050 |

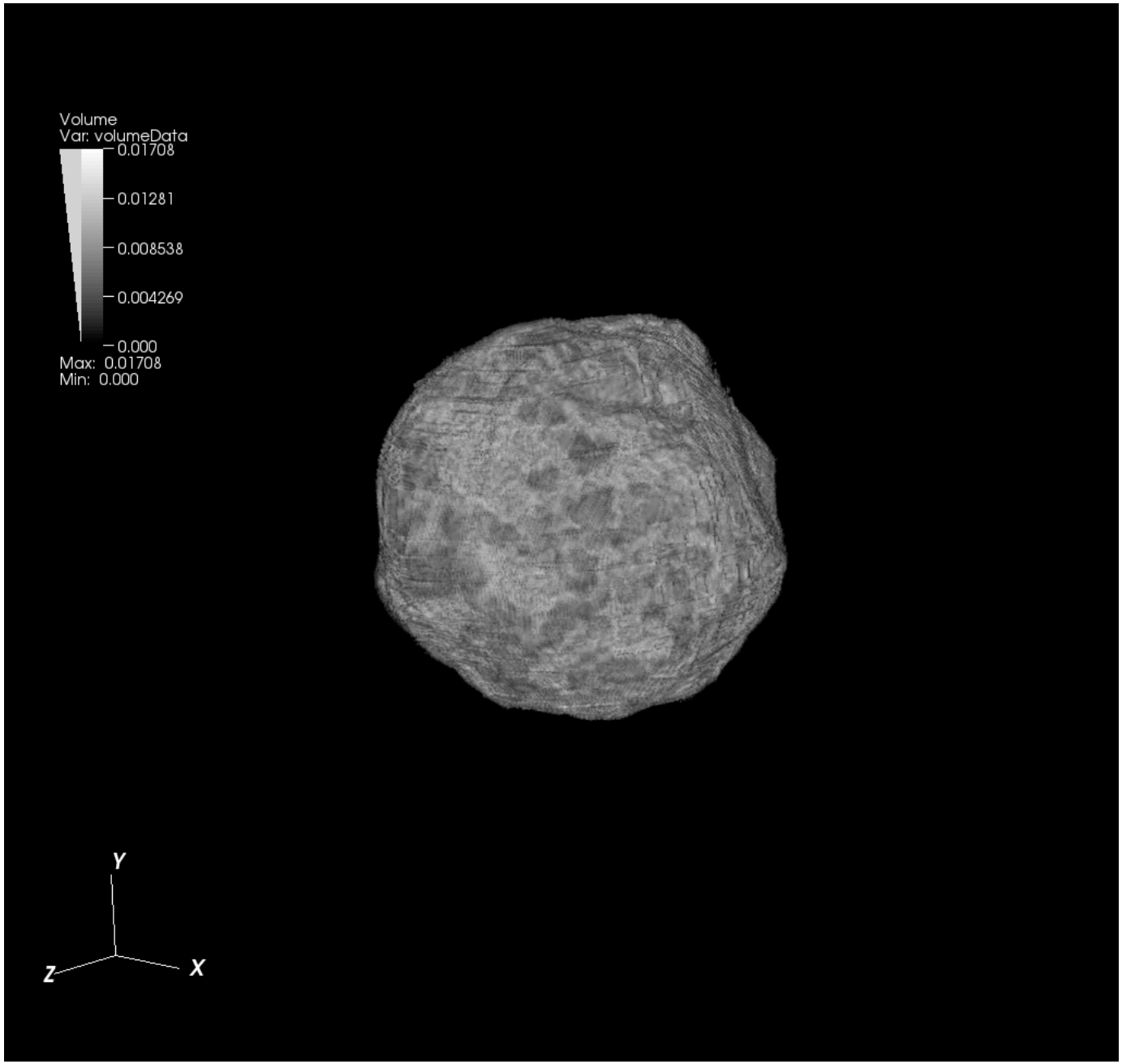
We have more recently derived estimates for the smallest eigenvalue of $L$ through perturbation theory arguments.

# How could we do better?

- This data only shows the densest regions of DNA/chromatin.
- To get a better feel for the amount of nuclear DNA we are collaborating with the Larabell Lab to use their X-ray CT imaging data.
- The new X-ray data gives linear absorption coefficients within each voxel that are *proportional* to the amount of organic material in that voxel.

# Volume Rendering of Nuclear LACs from 02-11

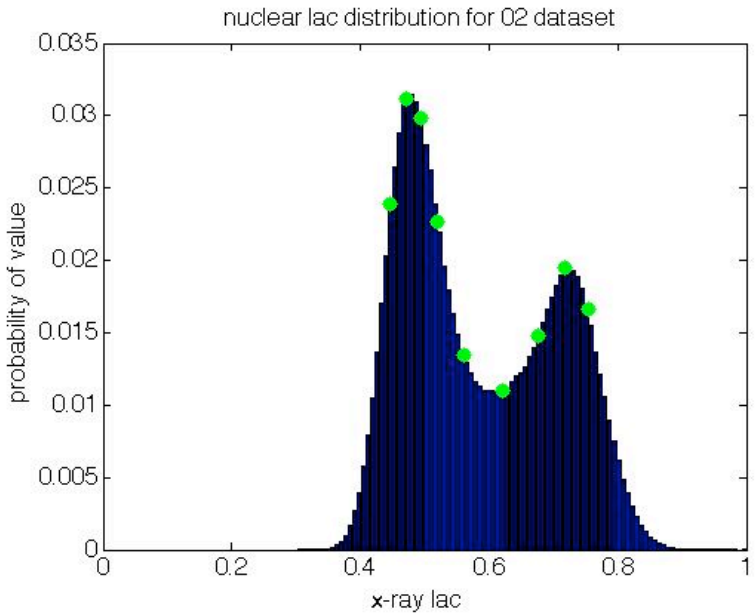
# Volume Rendering of Nuclear LACs from 02-11



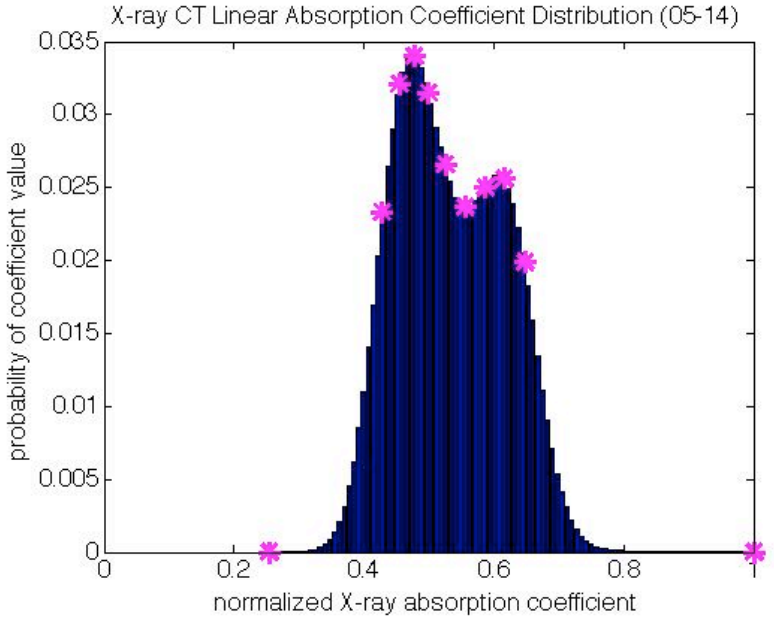
# Normalized Histogram of Normalized Nuclear LACs.

Markers denote every tenth percentile.

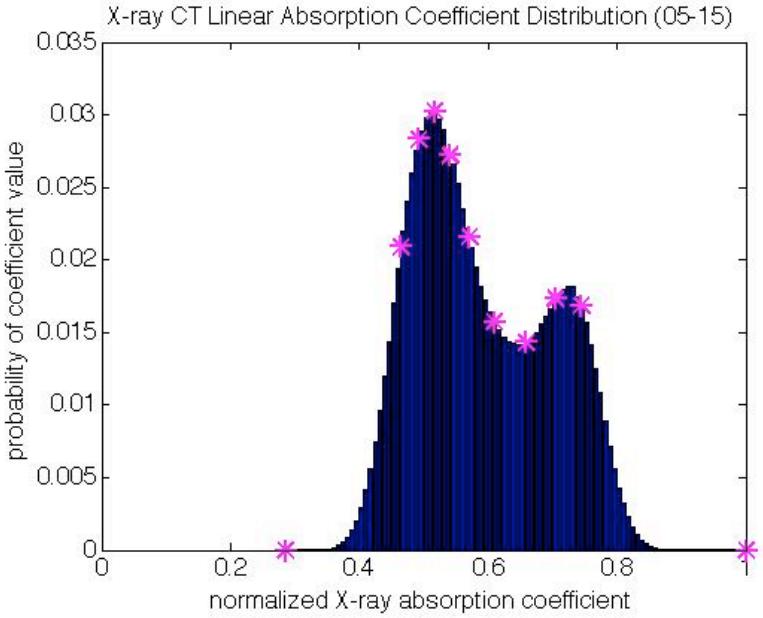
## data02-11



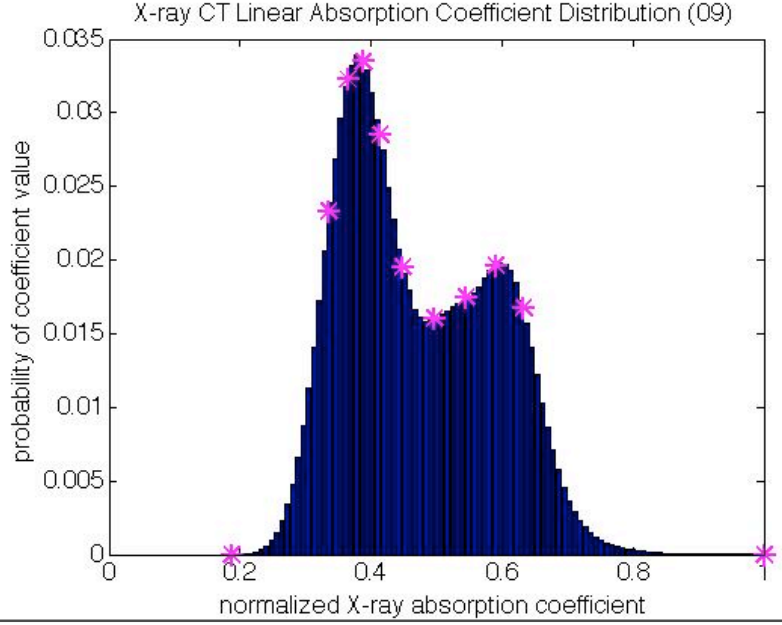

## data05-14




## data05-15




## data09-03

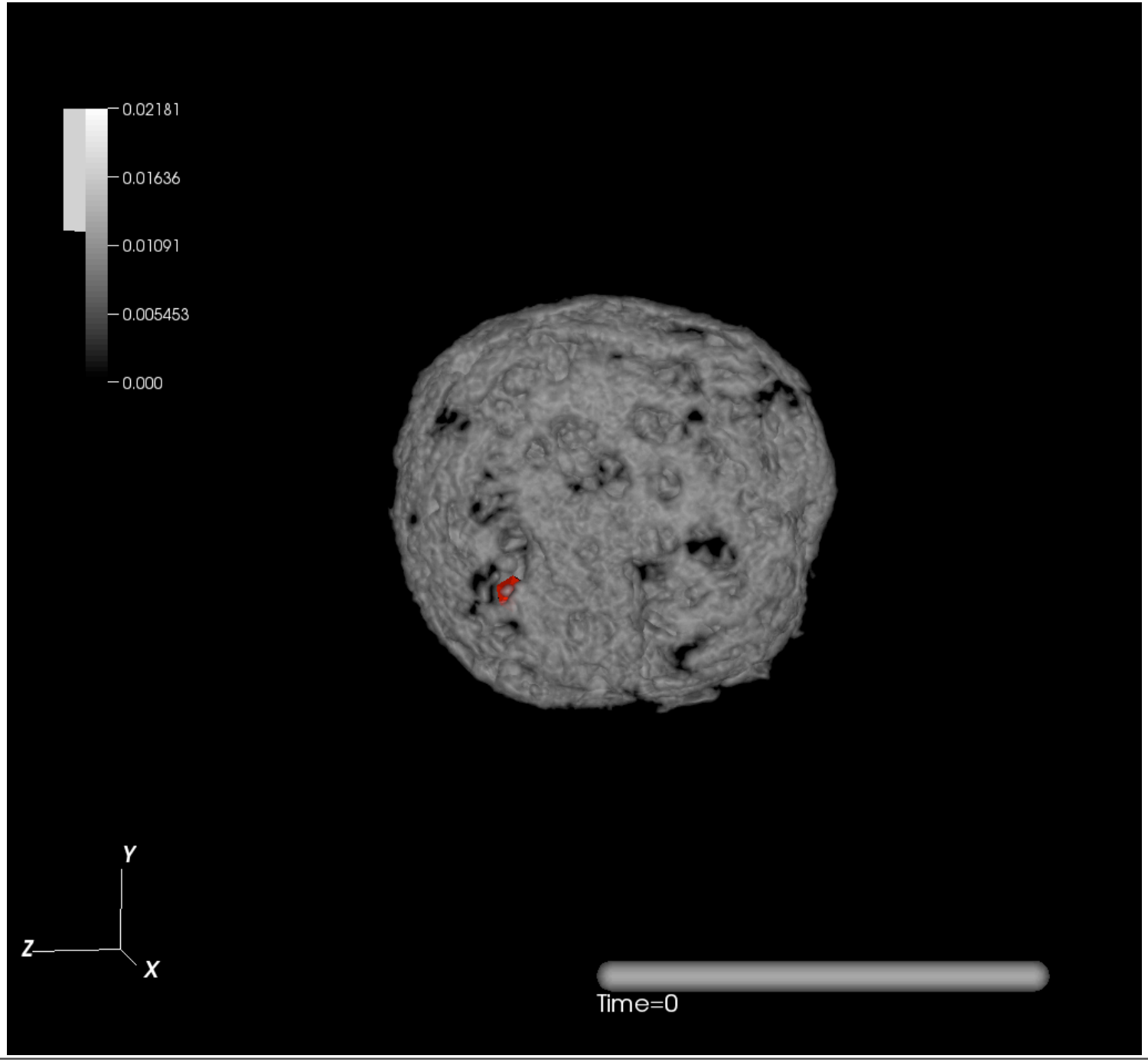


# How do we model the search for a binding site using this new data?

- Still model volume exclusion as a repulsive potential felt by the diffusing protein.
  - The potential of each voxel is proportional to the normalized measured LAC of that voxel.
  - Again, when the volume exclusivity is zero the protein simply diffuses (the chromatin is not felt by the protein).
  - As this parameter is increased the maximum strength of the potential increases, and it becomes substantially more difficult for the protein to enter regions with large LACs.
- We will subsequently use SIM to refer to the structured illumination data and X-ray for the new X-ray CT data.

# Example simulation with new X-ray CT imaging data:

# Example simulation with new X-ray CT imaging data:

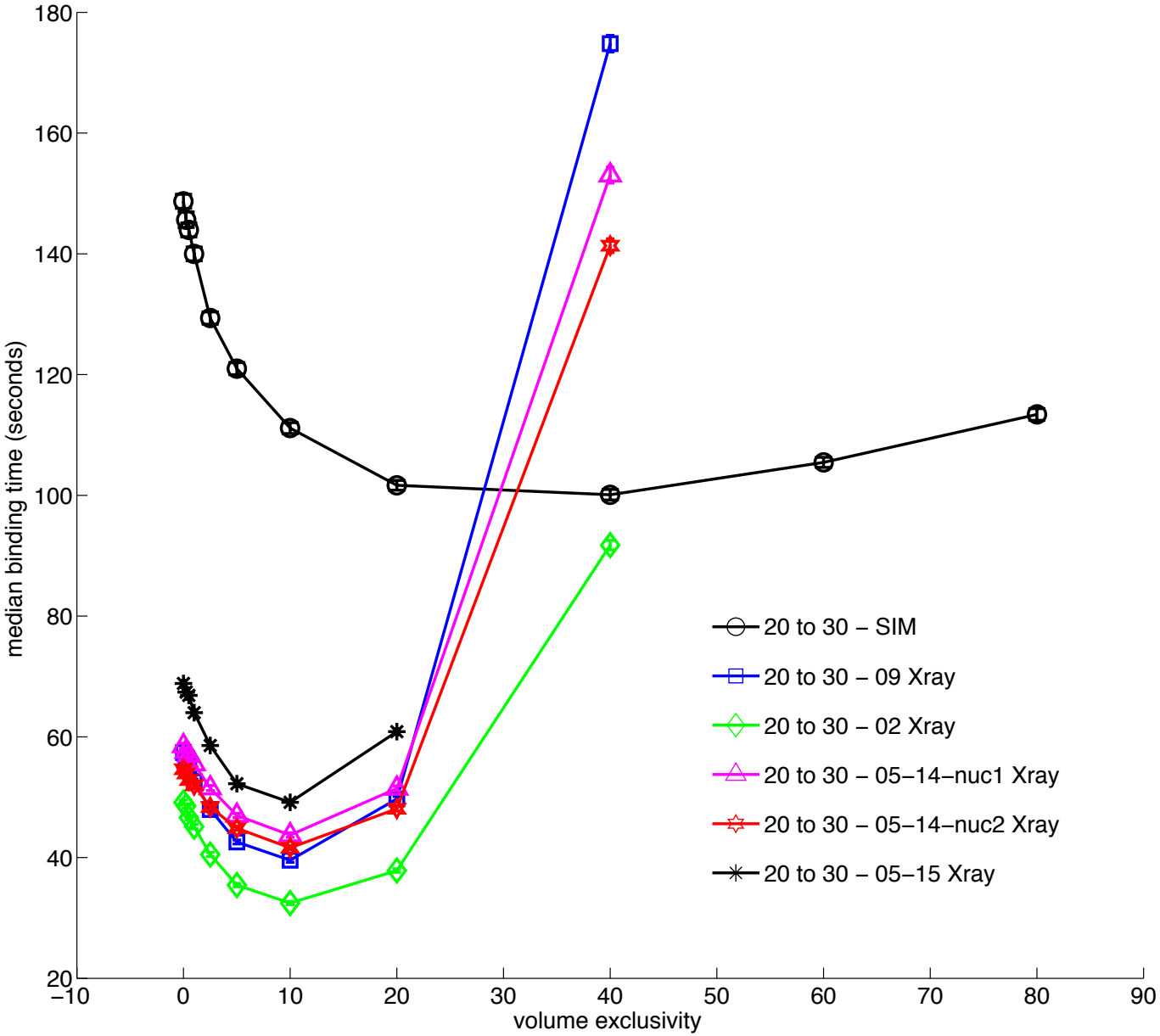


## How does binding site position influence the time needed to find a specific binding site?

- Assume the protein initiates its search from a specific pore.
  - For the SIM data, in each simulation we sample the protein's initial position from a uniform distribution of all possible pore locations.
  - For the X-ray CT data, we sample the position randomly from voxels that are on the boundary of the nucleus.
- We restrict the binding site position to subregions of the nucleus with specified intensity/LAC levels.
  - Regions of very high intensity/LAC may correspond to heterochromatin.
  - Regions of low intensity/LAC to euchromatin.

# Median Binding Times when Binding Sites are Placed in the 20th to 30th percentile of intensity values.

- Binding site is a randomly chosen voxel in the 20th to 30th percentile of intensity values for each simulation.
- Initial position is chosen randomly from all pores for the SIM data, and randomly from all voxels on the nuclear membrane in the Xray data.



## Speedup in fastest median binding time vs. no volume exclusion case.

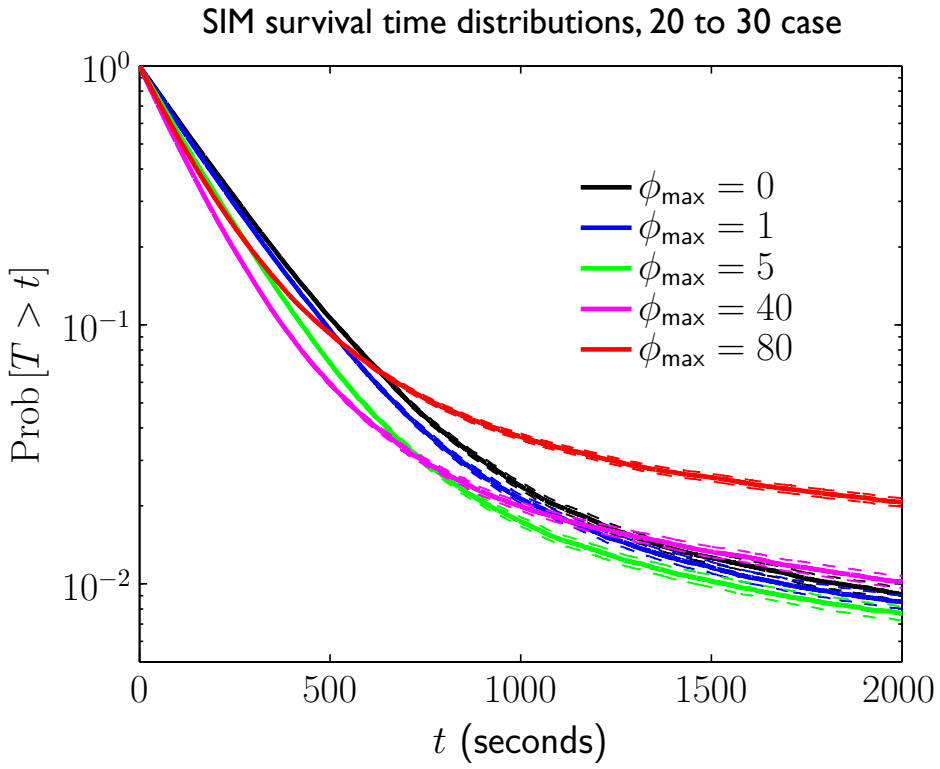
(For binding sites in the 20th to 30th percentile of nuclear voxel intensities.)

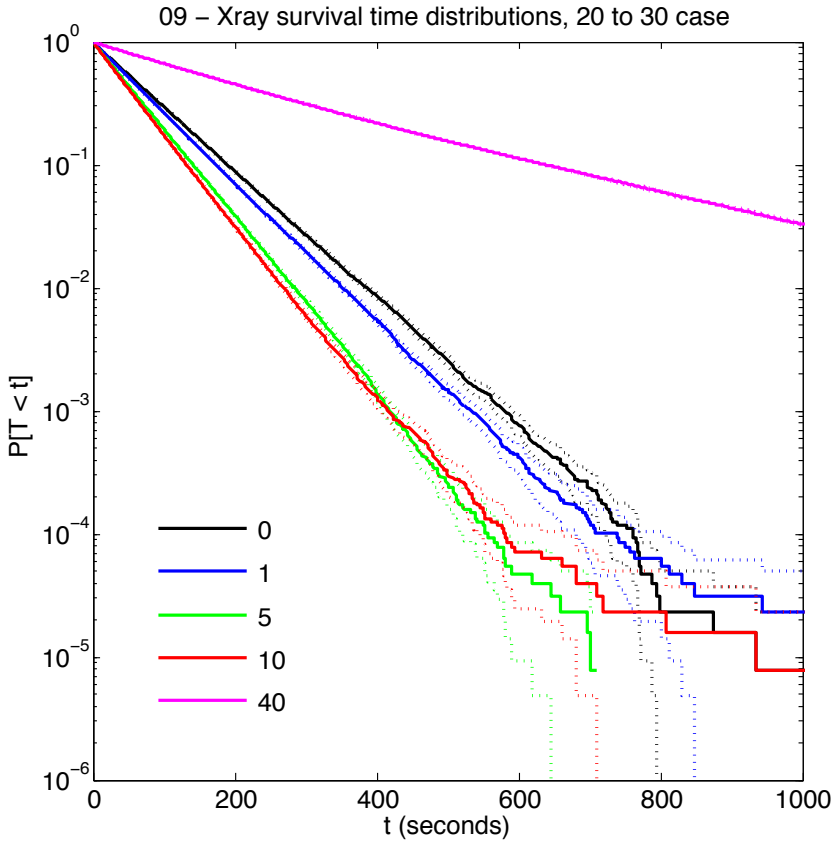
| Nuclei | Volume Exclusivity of minimum binding time (KbT) | Percent speedup vs. no volume exclusion |
|---|---|---|
| SIM data | 40 | 32.69 |
| Xray 09 | 10 | 31.09 |
| Xray 02 | 10 | 33.93 |
| Xray 05-14 | 10 | 25.22 |
| Xray 05-14 Nucleus 2 | 10 | 23.83 |
| Xray 05-15 | 10 | 28.61 |

# What do the binding time distributions look like in this case?

Here

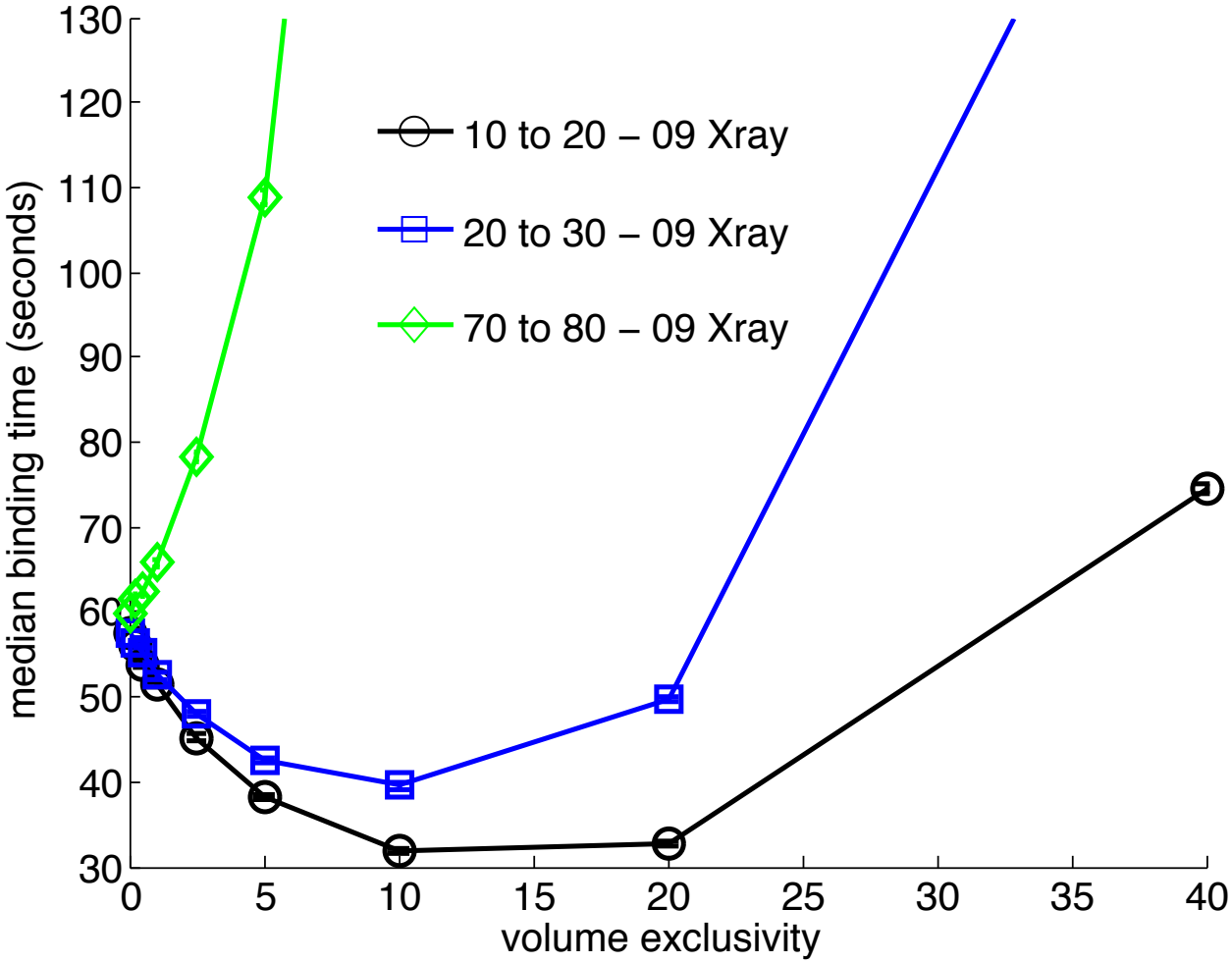
- P[T<t] = probability binding time, T, is less than t.
- Also, note that each legend gives the volume exclusivity for a given curve.



# What happens when the binding site is localized to regions of higher density?

(Legend gives percentile range where binding site was randomly placed.)



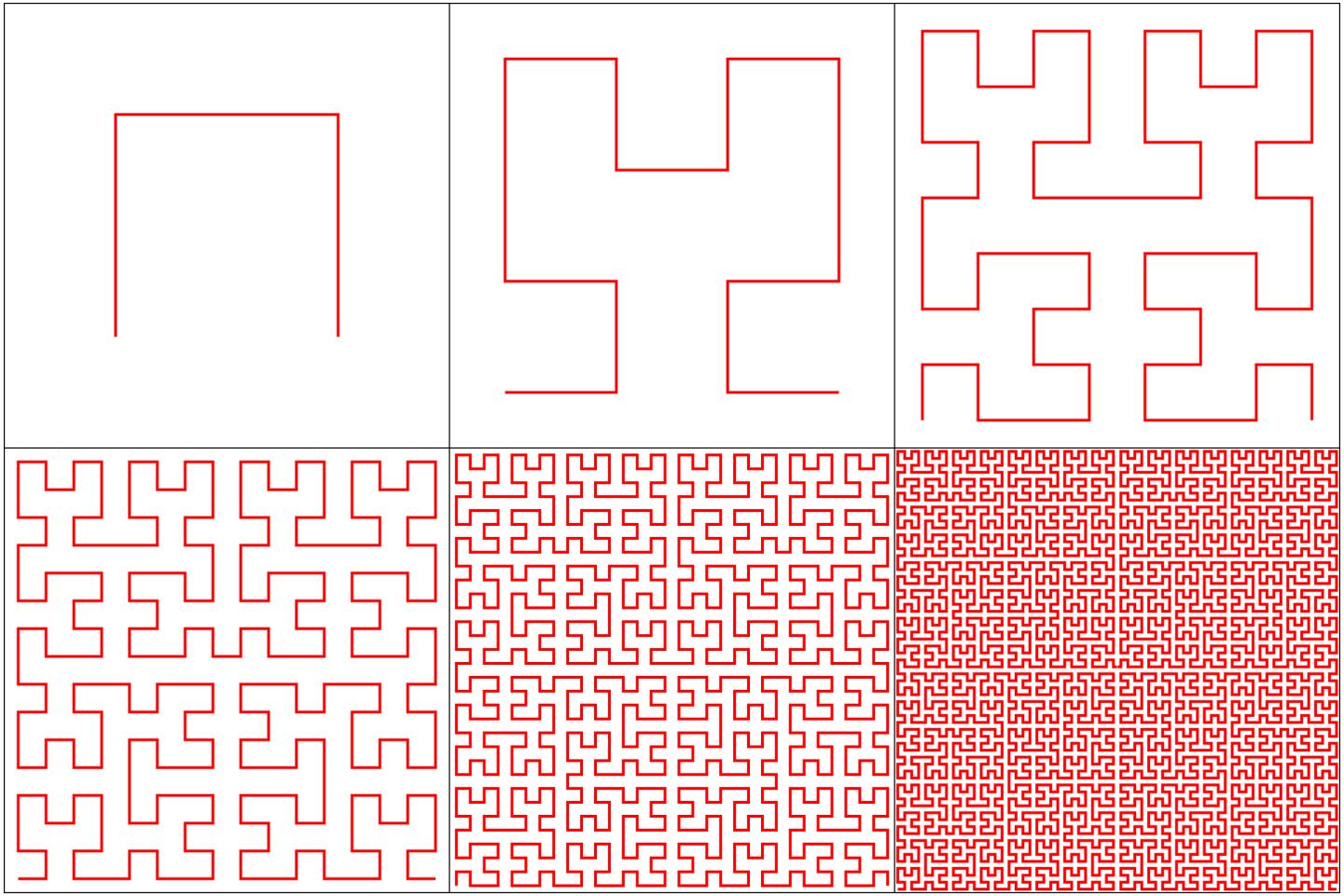

This is similar to what we see for the SIM data.

# What "fractal dimension" of euchromatin gives the most consistent binding times to our potential model?

- Based on a threshold LAC, we remove voxels above this LAC from the free space in the nucleus. (i.e. the protein is not allowed to diffuse into them).
  - We then calculate the box-counting / fractal dimension of the remaining voxels (i.e. the "free space" / euchromatin).
- We set the volume exclusivity to zero so that in the "free space" the protein simply diffuses (i.e. no longer feels the varying density of chromatin).
- Binding sites are still chosen randomly from the subset of free space voxels that are also within the 20th to 30th percentile of intensity values from ***all*** nuclear voxels.
- We compare the minimum time to find the binding site over all volume exclusivities vs. the time for a subset of nuclear voxels.

# What is the box counting dimension / fractal dimension?

- We want to understand / measure how an object fills space.
- For example, consider the Hilbert Curve, a space-filling curve:



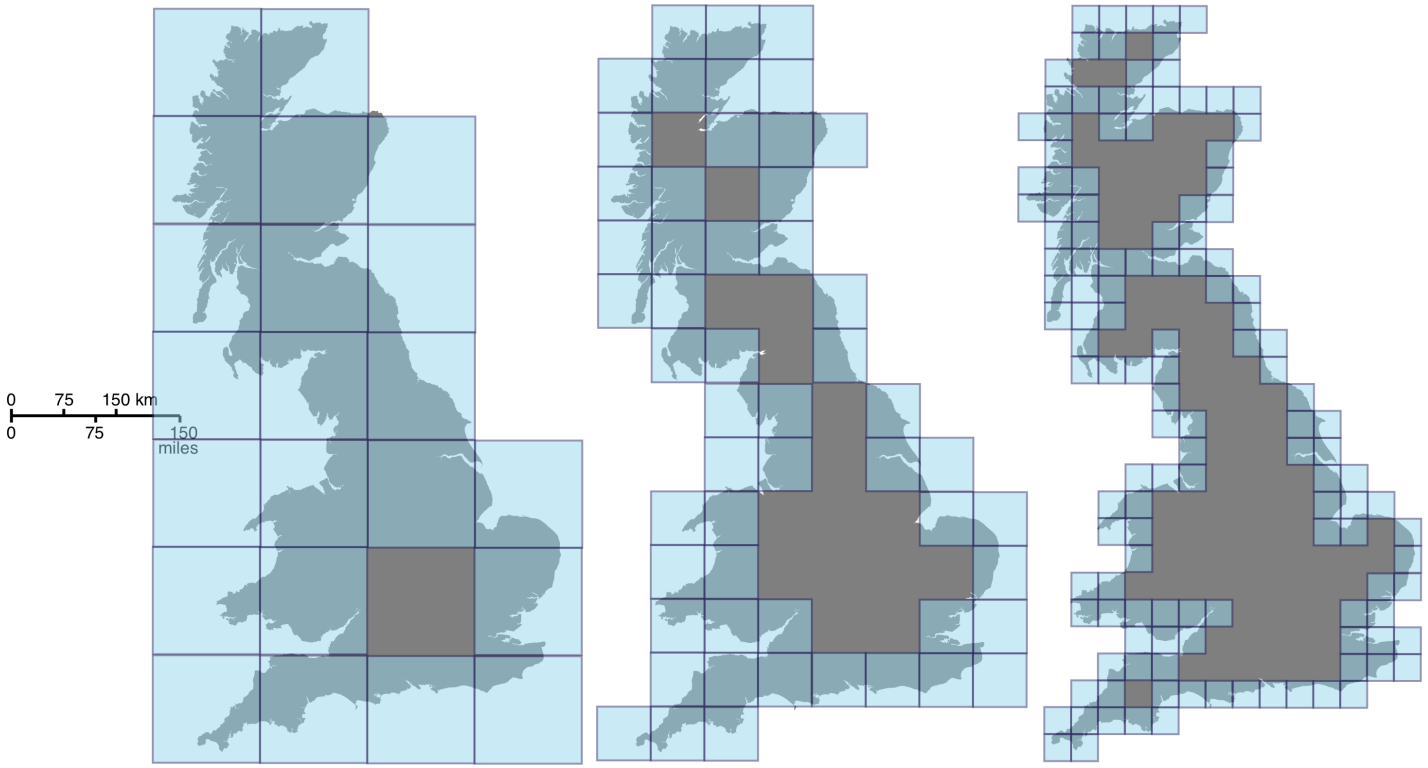
The limiting curve is one-dimensional, but touches every point in the box. It fills space like a 2D object.

(From Wikipedia)

# What is the box counting dimension / fractal dimension?

- The box counting dimension gives a measure of how an object fills the underlying space.
- It is given by covering the object with $N$ boxes of size $h$
- We then count how many, $M(h)$, of the boxes contain a piece of the object.
- It is assumed that as $h \to 0$, $M(h) \sim h^{-d}$, for some *box counting dimension*, $d$.

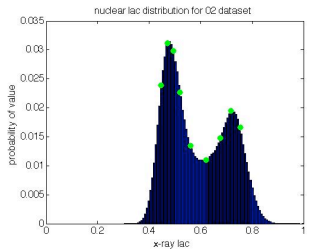
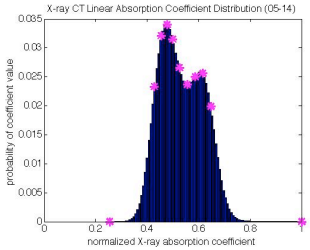
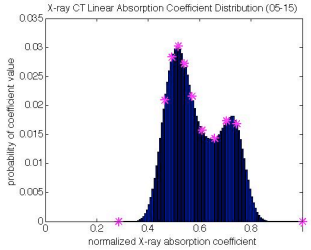
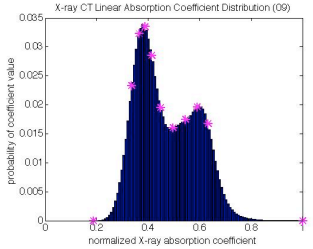
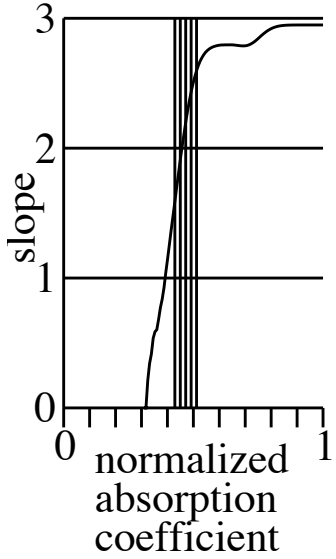
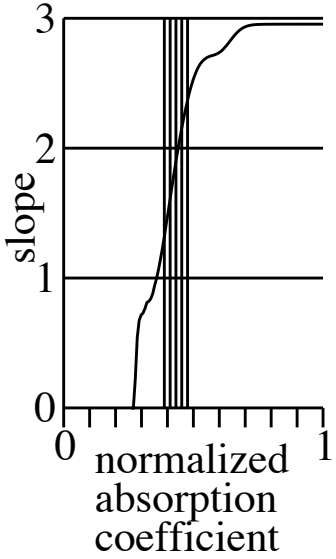
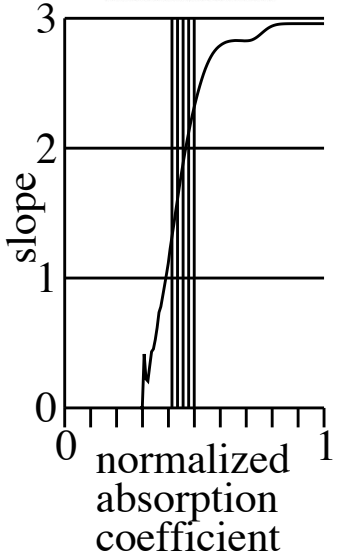
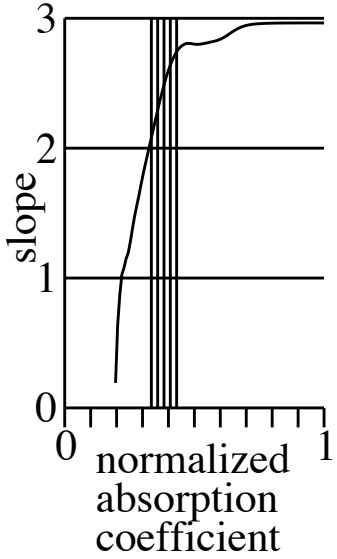
$$d = \lim_{h \to 0} \frac{\log(M(h))}{\log(1/h)}$$




From Wikipedia

# What is the Predicted Fractal Dimension of "Free Space" vs Threshold LAC Value

(slope equals the finest level box counting / fractal dimension)

| | data02-11 | data05-14 | data05-15 | data09-03 |
|---|---|---|---|---|
| Distribution of Nuclear LACs |  |  |  |  |
| Fractal Dimension vs Threshold LAC |  |  |  |  |

For all four nuclei we see that the effective box counting dimension of the voxels at or before the first "peak" in the histogram is between two and three.

# What does the remaining free space look like?

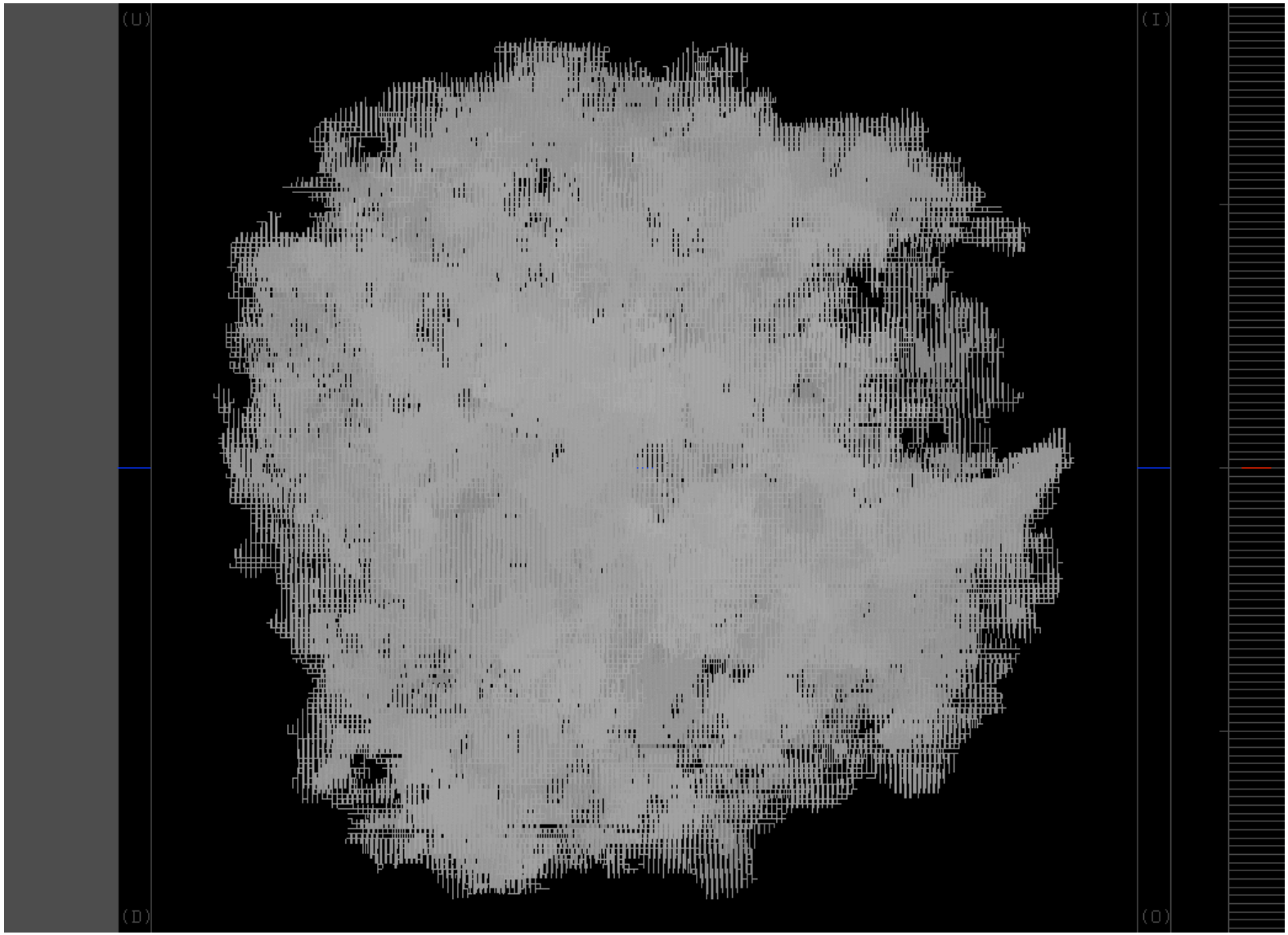
Here we threshold near the first peak of the LAC distribution (*i.e.* the euchromatin peak)

Free space:

# What does the remaining free space look like?

Here we threshold near the first peak of the LAC distribution (*i.e.* the euchromatin peak)

Free space:



# How are the threshold level, dimension of free space, and predicted binding time related?

| Thresholded Cell / Level | Percent of intensity values at or before the threshold value. | Finest Level Box Counting Dimension (*i.e.* estimated fractal dimension) | Percent difference of median times with thresholding from minimum median time without thresholding. |
|---|---|---|---|
| Cell 09, Level 24 | 27.21 | 2.49 | -6.45 |
| Cell 09, Level 27 | 37.02 | 2.64 | -21.17 |
| Cell 09, Level 30 | 45.22 | 2.74 | -22.14 |
| Cell 02, Level 27 | 28.93 | 2.43 | 2.40 |
| Cell 02, Level 30 | 37.32 | 2.60 | -13.12 |

# Summary

- Volume exclusion due to chromatin may help speed-up the time required for regulatory proteins to find specific DNA binding sites.
  - For regions of low chromatin density (low fluorescence), weak to moderate volume exclusivity gives the fastest times.
  - For regions of sufficiently high density, very little or no speed-up is seen in comparison to a model without volume exclusion.
- Fastest binding times in potential model are similar to those where we threshold near the euchromatin peak.
  - At this threshold level, the “free space” in the nucleus has a box counting dimension of ~2.4-2.5

# Acknowledgements


- Collaborators:
  - Carolyn Larabell (UCSF/LBL), X-ray CT data.
  - David McQueen and Charles Peskin, Courant Institute, NYU, modeling and data analysis.
- Ravi Iyengar and SBCNY for support and helpful discussions.
- NSF and NIH for support.


Thank you for coming and inviting me!