Integration of Spatial Data in Modeling

Ádám M Halász (West Virginia University)

Q-bio summer school 2015

Dynamical modeling of molecular processes in living cells

- Seeks to **recapitulate higher functionality** by emulating the dynamics of components
- **Utility:** Verification of hypotheses, insight into mechanism of emerging properties, virtual experimentation
- Several **levels of detail**: parts of molecules to tissues; each with their own **abstractions**
- Preceding lectures discussed **Rule Based Models**, **Kinetic Monte Carlo**, **Molecular Dynamics**

Topics

- Summary of dynamical modeling / simulation methods
  - ODE (continuous and deterministic)
  - well mixed stochastic simulations (Gillespie)
  - random motion (Brownian or not)
  - reaction-diffusion systems (agent based, stochastic)
- Membrane bound receptors and signaling
  - impact of spatial organization and movement of receptors on the cell membrane
  - anomalous diffusion and membrane landscape
- Connecting simulations and microscopic data
  - clustering of receptors & anomalous diffusion
  - a unifying hypothesis & connection to models

Molecular Processes in Cells

- A living cell may be regarded as a collection of complex molecules, organized in a specific spatial pattern
  - a **type of molecule** in a **specific location** defines a **species**
  - the state of the system is defined by the amount (concentration, copy number, …) of each species
- The molecular transformations and even the movement of these molecular species can be described using terminology from chemistry

  (Chemical Reaction Network - **CRN** models)

$$A + B \rightleftharpoons C \; ; \; A + E \rightarrow A + D \; ; \cdots$$

ODE models

- The amount of each species is represented by a **concentration**, proportional to the (copy) number of molecules per unit volume

  $1\mathrm{nM} \approx 6\times10^{23-9} = 6\times10^{14}$ molecules per liter

- The **state** of the system is the collection of all concentrations $X = ([A],[B],\cdots)^T$
- Evolves as an ODE system:

$$\frac{dX}{dt} = S\cdot\Phi(X)$$

ODE models

- The **reaction rate** is the rate of change of the concentrations of the participating species

  e.g. $$\Phi_{A\to B} = -\left(\frac{d[A]}{dt}\right)_{A\to B} = \left(\frac{d[B]}{dt}\right)_{A\to B}$$

- With **mass action**, the rate is the product of all incoming concentrations and a rate constant

  e.g. $$\Phi_{A\to B} = k_1[A]; \Phi_{A+C\to D} = k_2[A][C]$$

ODE models

- Using concentrations and reaction rates, we have a framework that culminates in sets of ordinary differential equations (ODE)

$$A + B \underset{k_1'}{\overset{k_1}{\rightleftharpoons}} C \quad \Rightarrow \quad \Phi_1 = k_1[A]\cdot[B] - k_1'\cdot[C]$$

$$A + E \xrightarrow{k_2,\, K_m} D + E \Rightarrow \quad \Phi_2 = k_2 \cdot [E] \cdot \frac{[A]}{[A] + K_m}$$

$$\frac{d}{dt}\begin{bmatrix}[A]\\ [B]\\ [C]\\ [D]\\ [E]\end{bmatrix} = \begin{bmatrix} -1 & -1\\ -1 & 0\\ 1 & 0\\ 0 & 1\\ 0 & 0\end{bmatrix} \cdot \begin{bmatrix}\Phi_1\\ \Phi_2\end{bmatrix} \quad \Leftrightarrow \quad \boxed{\frac{dX}{dt} = F(X) = S\cdot\Phi(X)}$$

$$X = ([A],[B],\cdots)^T$$

Issues…

- How many molecules are there again?

  example: **1 nM** ≈ 6 x 10^{14} molecules per liter..

  ..in **one cubic micron** (1 liter = $(10^{-1}/10^{-6})^3=10^{15}$ µm^3)

  ..works out to **0.6 molecules**

- How does one measure those rate constants?

  In vivo on cells (e.g. flow cytometry)

  Quantitative *in vitro* assay

  Molecule by molecule (e.g. SPT)

  The resulting values are likely **very different***

  *due to different **physical circumstances** / **spatial scales**

" Well-mixed" molecular simulations

- The amount of substance is **not** continuous
  - species represented by a **molecule number**: $N_A, N_B, \cdots$
- Instances of molecular transformations are **random events**, triggered by **Poisson processes**:
  - Probability of firing in $(t, t+\Delta t)$, if did not fire until t:

$$\Delta p(t, t+\Delta t) = \lambda \cdot \Delta t$$

  - Probability density function (pdf) of the **firing time** τ:

$$f(\tau) = \lambda \cdot e^{-\lambda \tau}$$

  - the constant λ is the **rate** of the process
  - rates of concurrent, independent processes are additive (…)

" Well-mixed" molecular simulations

The system is a continuous time Markov chain

- States defined by copy numbers: $(N_A, N_B, \ldots)$
- Reactions correspond to discrete state transitions:

  $A \rightarrow B$ means $(N_A, N_B) \rightarrow (N_A - 1, N_B + 1)$
- The Poisson rate for a reaction is its <u>propensity</u> (eg. $\gamma_{A \rightarrow B}$)
- Multiple transformations may originate in a given state
  - independent Poisson processes run concurrently
  - when one of them fires, the state changes accordingly
  - new Poisson processes start for feasible transformations

" Well-mixed" molecular simulations

Propensities typically proportional to the number of possible instances (similar to mass action kinetics)

- for first order reactions (A→B, A→B+C,..), the propensity is the number of molecules times the macroscopic reaction rate constant: $\gamma_{A\rightarrow B} = k_{A\rightarrow B} N_A$
- for bi-molecular reactions (A+B→…), there is a volume factor $\gamma_{A+B\rightarrow C} = k_{A+B\rightarrow C} N_A N_B / (V_0 N_{Avogadro})$

Many algorithms to **simulate** these stochastic models, often called Gillespie simulations:

- direct method, first reaction method, tau leaping,…

Slide 11

2 Expan on this, maybe make it a separate slide
Adam Halasz, 7/22/2015

Brownian motion (BM) and diffusion

- Bio-molecules are generally **not** uniformly distributed
  - the spatial pattern impacts the rates of **some** reactions
- If no other factors are present, molecules move randomly
- Ideal BM is equivalent to Fickian diffusion
  - displacement of a particle over a time interval is normally distributed, with standard deviation proportional to the time
  - the localization density of a set of BM particles follows a diffusion equation
- BM emerges as the large spatial (or time) scale limit of a broad class of random walks
  - a consequence of the Central Limit Theorem
  - effective diffusion coefficient given by variance / time

Simulation of molecular movement

- **Agent based** algorithms: follow the position (and chemical state) of individual molecules
- **Brownian-motion** – use the definition of BM
  - continuous (X,Y) coordinates
  - positions updated with random displacements taken from a normal distribution with $\sigma = 2\ D\ t$
- **Lattice** – spatial discretization (grid of sites)
  - hops to adjacent sites triggered by Poisson processes
  - rate determined by grid spacing and diffusion coeff.

Brownian Motion Simulations

- Conceptually very straightforward; implements the definition of the PDF:
  - let the position at time t be (x,y)
  - then the position (x',y') at time t' is a RV that follows

    $f(x',y',t') = \exp(-((x'-x)^2+(y'-y)^2)/(4D(t'-t)))/(4\pi D(t'-t))$
- For a single particle, we simply generate displacements in the X and Y directions, using normally distributed random numbers of the appropriate variance ($\sigma_x=\sigma_y=2D(t'-t)$)
  - The two displacements are independent
  - So are the displacements of other particles in the system
  - For N particles, we can generate a 2N dimensional vector of displacements and add it to the 2N dimensional position vector
  - Takes about 10 minutes to implement in Matlab ;-)
  - Eminently vectorizable
- No need to simulate intermediate states
  - the time step can be anything
  - unless we care about collisions and incursions into special regions

Boundaries…

- Most of the complexity in spatial simulations stems from dealing with boundaries and collisions
  - Particles will “wander off” from the region of interest if we don’t keep them in there somehow – **virtual** boundaries
  - Membrane landscape – barriers that may be more or less permeable, and may favor crossing in one direction
- Simulation space defined by the system to be modeled
  - a patch of membrane, a small window into a large world
  - there may be real obstacles to the movement of our bio-molecules: linear barriers, attractive or forbidden regions, point obstacles (=other particles), traps

Boundaries…

- In the open-system situation, an easy trick is **periodic boundary conditions**
  - Particles that leave through the right edge re-enter through the left
  - The position vectors are “modulo” the simulation boundaries
- Physical boundaries are usually modeled with **reflecting boundary conditions**
  - An attempted move beyond the boundary results in a position equivalent to a reflection of the displacement vector by the boundary
  - Partially reflecting boundaries – allow crossing with a probability <1

Lattice-based Simulations

- There are many technical benefits of discretizing space in a simulation
  - Easy identification of "neighboring" particles, avoiding unphysical overlap, identification of collisions and other particle-particle interactions
  - Possibility of a joint "species - position" state space and a uniform CTMC type simulation engine
- Idea: positions are restricted to a (uniform) grid of coordinates

  $(x,y) \rightarrow (i,j)$; $x=0, \pm\Delta x, \pm 2\Delta x, \ldots$; $y=0, \pm\Delta y, \pm 2\Delta y, \ldots$
- Particles "hop" between lattice sites
  - One could consider an implementation similar to BM where there is a finite probability of jumping into any site
  - The more common (and practical) approach is to only allow hops into neighboring sites, with a certain probability per unit time ("hopping rate" δ)
- Correspondence with physical parameters
  - The hopping rate should be consistent with the physical diffusion rate, so that the means square displacement after a finite time $T >> 1/\delta$ is the same

Lattice Simulation of Diffusion

- Less latitude in choosing the time step then in BM simulations, but there are still some choices
  - Use a fixed simulation time step Δt
    - Probability of a hop $\approx \delta \Delta t$
    - It must be small enough to justify neglecting the possibility of multiple hops: $\Delta t << 1 / \delta$
  - Trigger hops with Poisson processes
    - Recall that for a Poisson process, the average firing time is $\tau = 1 / \lambda$; if we trigger hops with Poisson rate equal to δ, the average hopping rate will be exactly δ
    - We may simulate multiple particles this way, following the Gillespie protocol; only one move per update eliminates contradictions (more than one particle per site)
- Boundaries and particle overlaps
  - Keep track of "allowed" moves for each particle
  - In a rectangular lattice, a particle can move to any of its four neighbors
    - The hopping propensity is typically defined per individual move
  - If a neighboring position is occupied, that move is disallowed, i.e. its propensity set to zero
  - Directions that take the particle beyond the boundary may be excluded or implemented using specials rule for reflecting or periodic boundary conditions

Reaction-Diffusion Systems

- The combined simulation of molecular transformations and diffusion
  - necessary to reproduce important phenomena
- First-order reactions* (A→B) are not influenced by spatial aspects and also do not complicate spatial simulations*
- By contrast, bi-molecular reactions (A+B→C or A+B→C+D) require an entire new layer of analysis and simulation machinery

Uni-Molecular Reactions in Space

- Consider several molecular species A,B,… which are subject only to simple reactions of the form A→B
- These molecules also diffuse in two or more dimensions
- If we have a separate simulation of the diffusion of the molecules and another one that simulates their chemical transformations, we could run them side by side without the need to transfer information between them
- How would we combine the two simulations though?

BM + “Gillespie”

- Assume we have a maximum simulation time step Δt so that we need to report the state of the system at times 0,Δt,2Δt,…
- We also have an initial distribution of the particles and we know their chemical species; we want that information, consistently with the rules of the reactions and diffusion
- The variables are the 2 spatial coordinates and the chemical species (=chemical state) of each molecule; we initialize them according to the data received
- The Gillespie module normally just gives us the time and type of the next transformation;
  - We can make it “agent based” by adding a step where we choose the actual molecule that undergoes that transformation
  - Now we know all the elements of the next chemical transformation
- We evolve the system to the next time, which is whatever comes first between the next Gillespie event and the next recording time
  - If the recording step is shorter, we generate new positions for all the particles and move on
  - If the Gillespie step is the shorter, we generate new positions for all the particles and also perform the chemical update: (1) change the chemical state of the particle that reacts (2) generate the next Gillespie event
- This sounds simple (and it is) but already requires some construction in terms of state variables

Lattice Diffusion + “Gillespie”

- The lattice based diffusion algorithm we discussed was specifically designed to be similar to a(n agent-based) Gillespie type simulation
- Particles will have two discrete spatial indices and a third to indicate their chemical state
- There are two types of transformations
  - Chemical – characterized by a (per-particle and reaction type) elementary rate and an overall propensity
  - Spatial – characterized by a hopping rate (per particle and direction); the sum of the propensities of all allowed moves give the total diffusion propensity
- We can treat all transformations the same way, i.e. put their propensities together and then pull out a single elementary transformation from the bag
- It is a bit more elegant to first choose whether the next event is a move or a chemical transformation, then implement each the way we did before
- This approach will result in much shorter time steps compared to the BM version – basically is evolves the system one particle at a time; the slowdown is a price paid for a spatial simulation that avoids particles “moving through” each other.

Bi-molecular Reactions in Space

- Two molecules can interact only if they are close enough to each other
  - the "well-mixed" assumption implies that molecules diffuse so fast that whatever inhomogeneities exist, they are washed away on a time scale well below that of a typical two-molecule interaction
- The main type of interaction that needs to be addressed is one that results from a binary collision
  - Complex (dimer) formation: $A+B \rightarrow C$
  - Mutual transformation: $A+B \rightarrow A'+B'$
  - Dissociation, $A \rightarrow B+C$ presents its own issues

Bi-molecular Reactions in Space

Focusing on complex formation, the main issue to be addressed is to establish whether a specific pair of (A,B) molecules may interact

- they need to be “close enough” to each other;
- exact dynamical details may actually be known or could be simulated in a molecular dynamics setting
- However, our focus is on the reaction in the context of a larger system, so we rely on approximations developed for the two types of simulations, lattice based and BM

Bi-molecular Reactions on a Lattice

- An implicit assumption behind lattice based spatial simulations is that one lattice site is about the size of a single molecule of interest
- Then, two molecules are "close enough" if they occupy neighboring lattice sites.
- If we are only interested in A+B→A'+B', that is, the molecules may transform but do not form a complex we only need to add neighborhood based rules to the Lattice+Gillespie machinery
- If complexes are formed, there is the question of how many lattice sites will the complex occupy?
  - In the important case of RTK receptors, the only thing we need to worry about are receptor dimers (all membrane bound species are singleton receptors or dimers in various binding and activation states)
- In this particular case, dimers are described as two molecules that keep their identity and footprint (lattice site) that diffuse together
  - Diffusion rules are modified such that the elements of the dimer can only hop simultaneously, in a way that keeps them adjacent
- Otherwise the implementation adds
  - A mechanism to track possible dimerizations
  - Dissociation reactions
- There are artifacts that result from this way of describing dimers

Bi-molecular Reactions in BM

- The main ingredient is the interaction radius
  - A bi-molecular reaction for a specific pair of molecules is considered during an update, if the distance between them is below a pre-determined value
  - The reaction radius is calculated semi-empirically so that the macroscopic reaction rate is matched
- The time evolution of the system is done using a fixed time step
  - The step is chosen to be small so that the update is done for one particle at a time
  - For the chosen particle and the respective update time step, finite probabilities are calculated for unimolecular reactions
  - Binary reactions occur based on the reaction radius
  - The position of the particle is updated using a Brownian PDF
- The added computational cost (due to the one particle at a time update) is comparable to that of lattice based simulation methods
- Approaches that would provide a Gillespie-like propensity for the occurrence of bi-molecular reactions are not practical (to my knowledge)

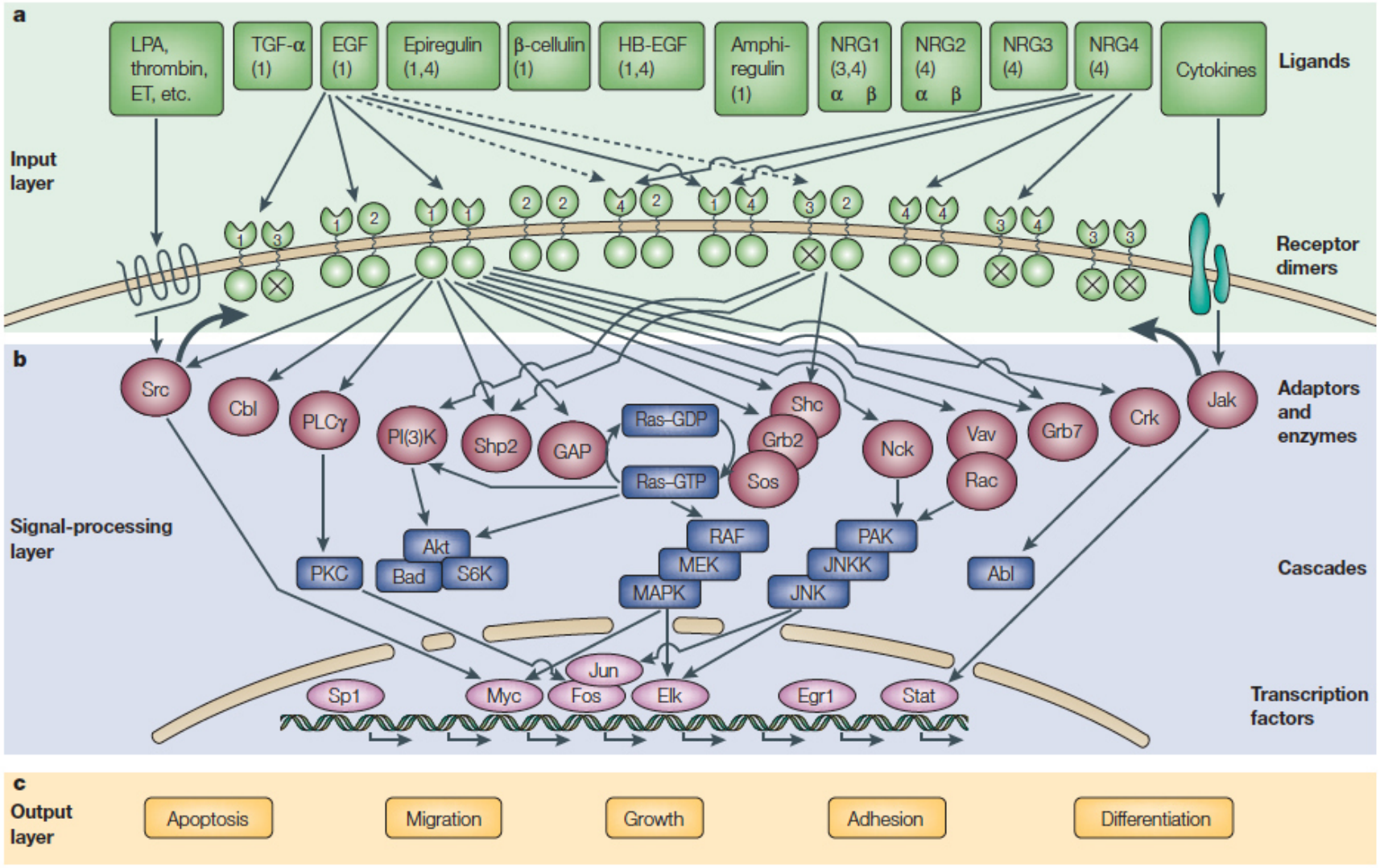
Reaction-Diffusion Simulations

- Agent-based – molecules have an “identity”
- Movement is random, often approximated by Brownian motion
- Second order reactions triggered by collisions
- BM (continuous space) or lattice (discretized space) approaches
- **Rates and collision / dissociation distances calibrated to match required values**

Molecular Processes in Cells

- The resulting dynamical systems are complex and solving them is impractical
  - even if the details are often not available, this picture is behind most modeling approaches to cell dynamics
- To avoid dealing with everything in the same time, focus on specific types of processes or functions
  - metabolism, sensing, growth, …
- One particularly important aspect is **cell signaling**

Signaling



EGF signaling network

Yarden, Y. and Sliwkowski, M. X. (2001) *Nature Reviews* 2: 127-137

Molecular Processes in Cells

- A cell interacts with the rest of the organism largely through signaling
  - Complex, interconnected ~~pathways~~ networks
- Changes in a signaling network may have far reaching implications [1]
  - EGFR (Her2) mutations involved in cancers via overexpression, changes in activity
- Signaling is a favorite target of therapies
  - suppress VEGF to block tumor angiogenesis

Slide 30

1 Check with Bridget

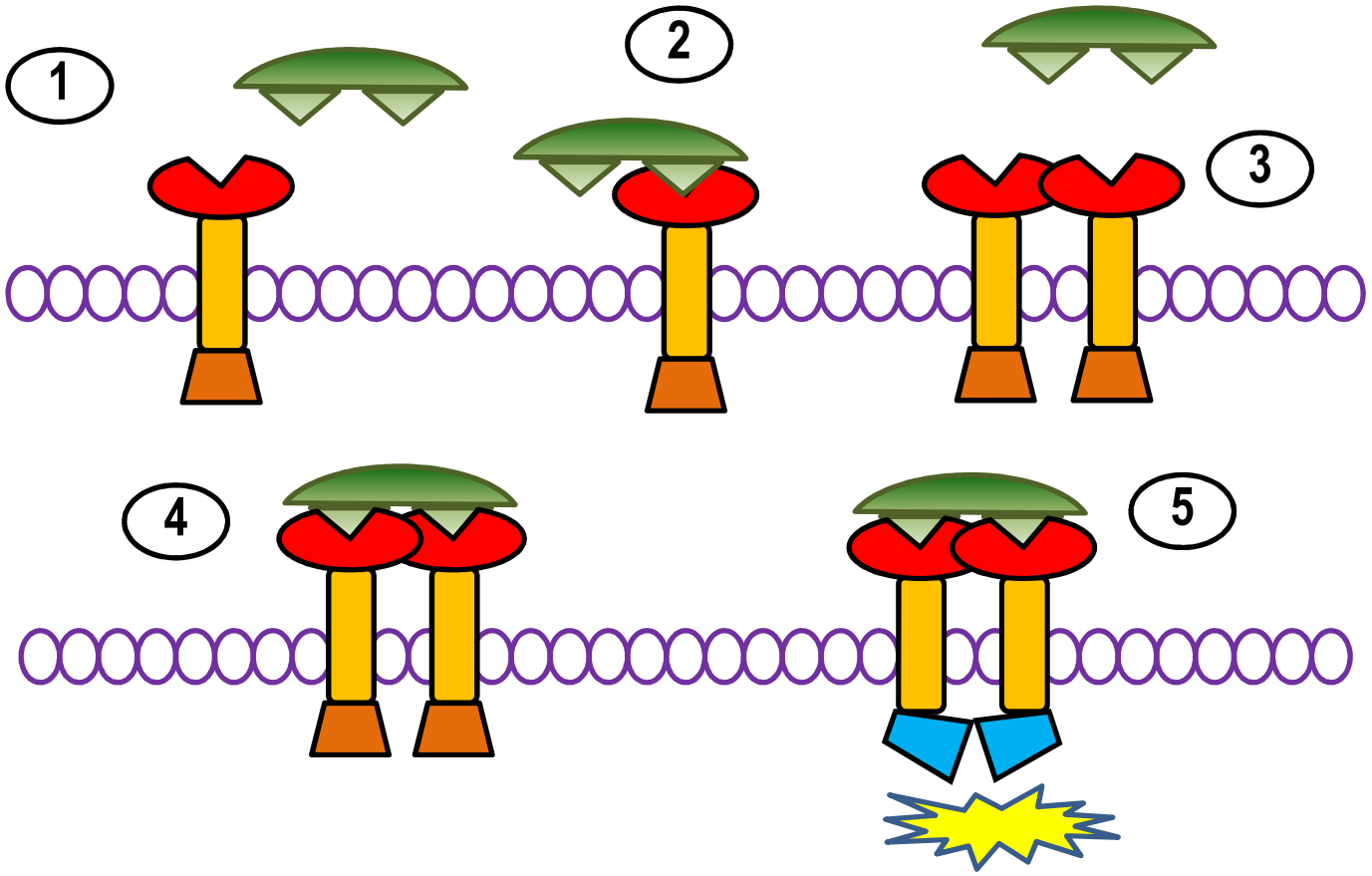
Adam Halasz, 7/22/2015

Example: VEGF Signal Initiation

- Vascular Endothelial Growth Factor is involved in the growth of blood vessels, relevant to cancer, diabetes,..
- A somewhat typical example, VEGFR is a receptor tyrosine kinase (RTK), just like the EGF receptor family
- **Initiation** refers to processes that occur from the appearance of the ligand to the activation of the intracellular domain of the receptor
- Complexity:
  - 6 species and 7 reactions for one receptor and ligand type;
  - there are several types of VEGF, also quite a few (at least two important) VEGFR types, plus soluble versions, etc.

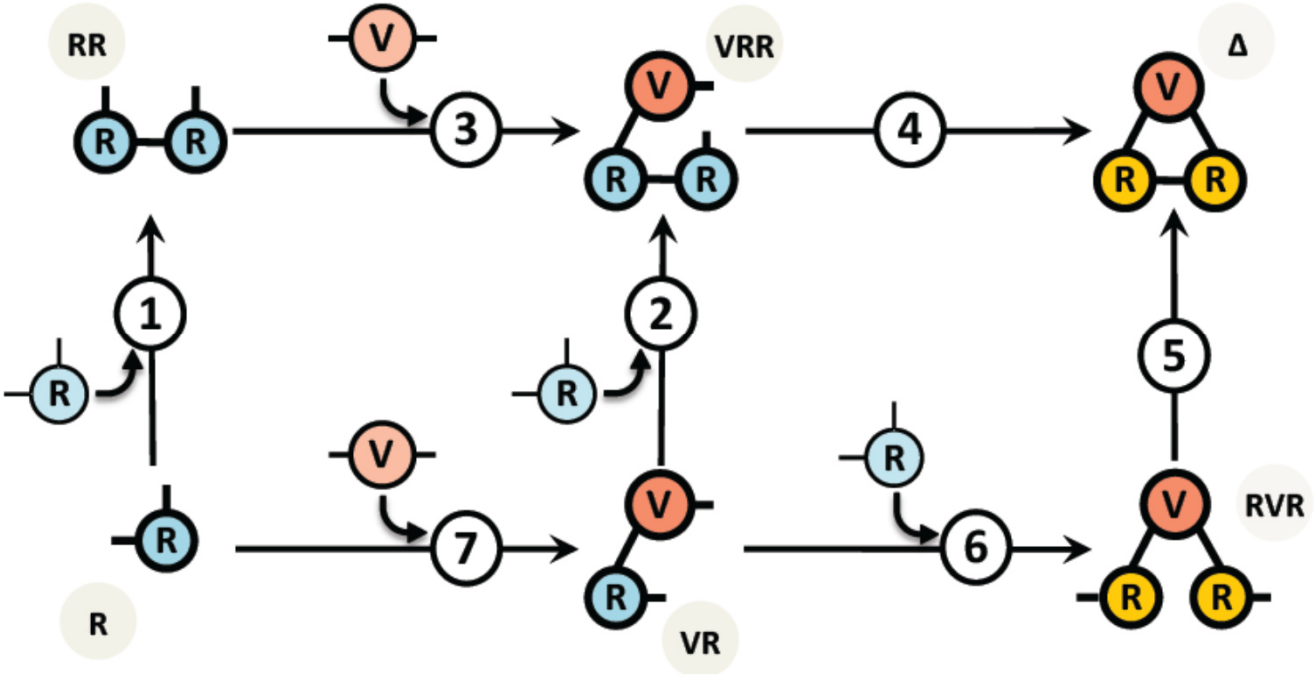
VEGF Signal Initiation

Monovalent receptors straddle the cell membrane; the VEGF **ligand is bivalent**; receptor dimerization is **required** for signal initiation



VEGF Signal Initiation

- VEGF receptors are membrane bound; all processes take place on the cell membrane
- (Ligand-supported) dimerization is a necessary step for signal initiation (similar to EGF)



Importance of Membrane Dynamics

- **Dimerization** of membrane bound receptors is a common feature of RTK and other signaling machineries
- Dimerization occurs between receptors that move semi-freely along the cell membrane
- It is not clear how susceptible the signaling mechanism is to variations in the kinetics of dimerization;
- We do know that cancer-related mutations involve receptor overexpression, or molecular changes that result in modified receptor activity

The Experimental Picture

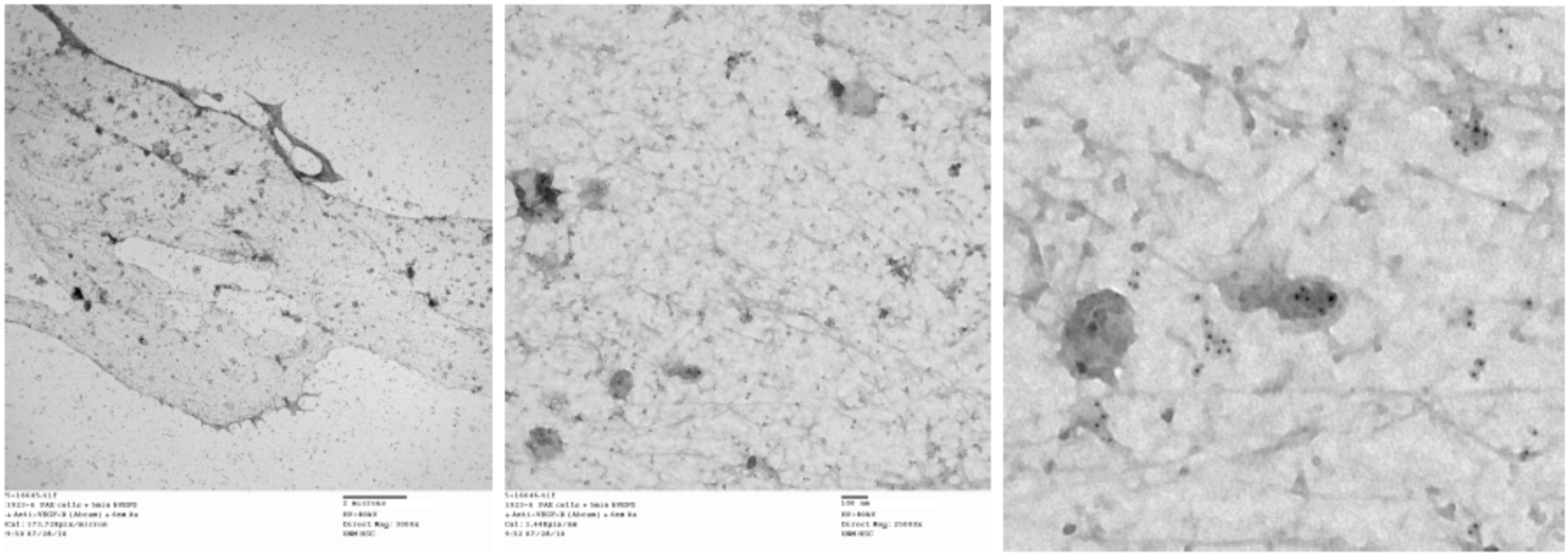
- Labeling / imaging techniques allow us to interrogate cells on the level of individual molecules
- **Flow cytometry** – quantifies the amount of several molecular species in individual cells in a population
- **Light microscopy** with fluorescent tags – movies of individual molecules within a cell
- **Electron microscopy** with metal beads – snapshots of individual molecules with high spatial resolution

The Experimental Picture

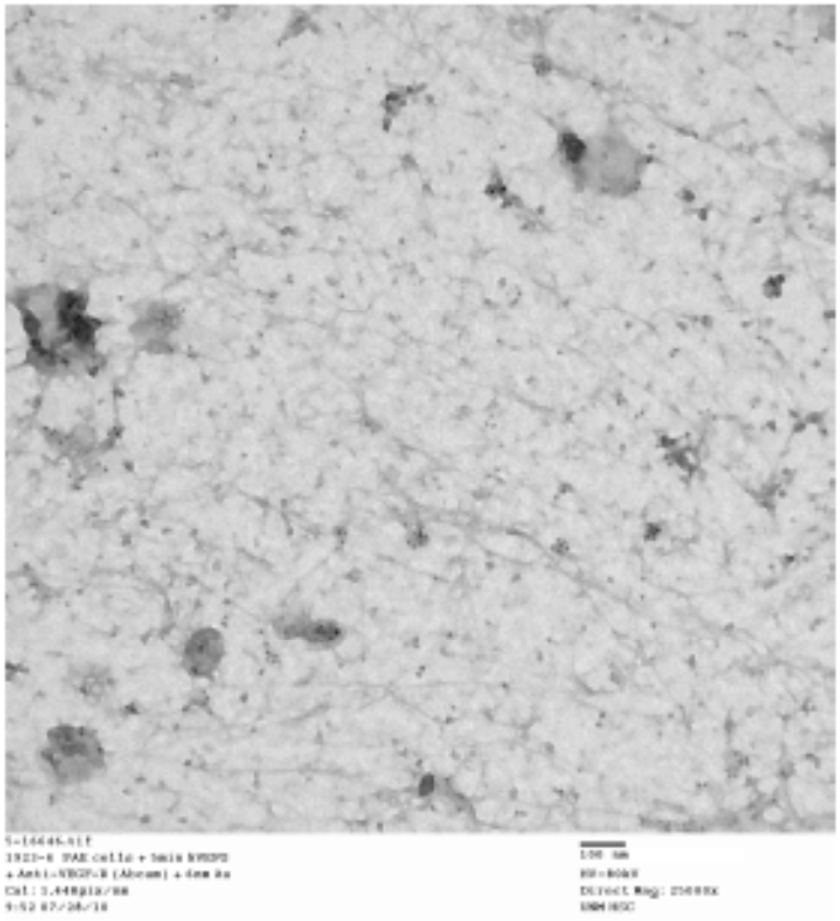
- The cell membrane has a **landscape** of **domains** (concentrations of lipids and / or proteins), and **barriers** (elements of the cytoskeleton)
- Receptors such as VEGFR or EGFR tend to localize in **clusters**, that are sometimes enhanced in the presence of ligand
- Dynamic imaging indicates **co-confinement** of receptor pairs (correlated movement that is not consistent with bond formation)

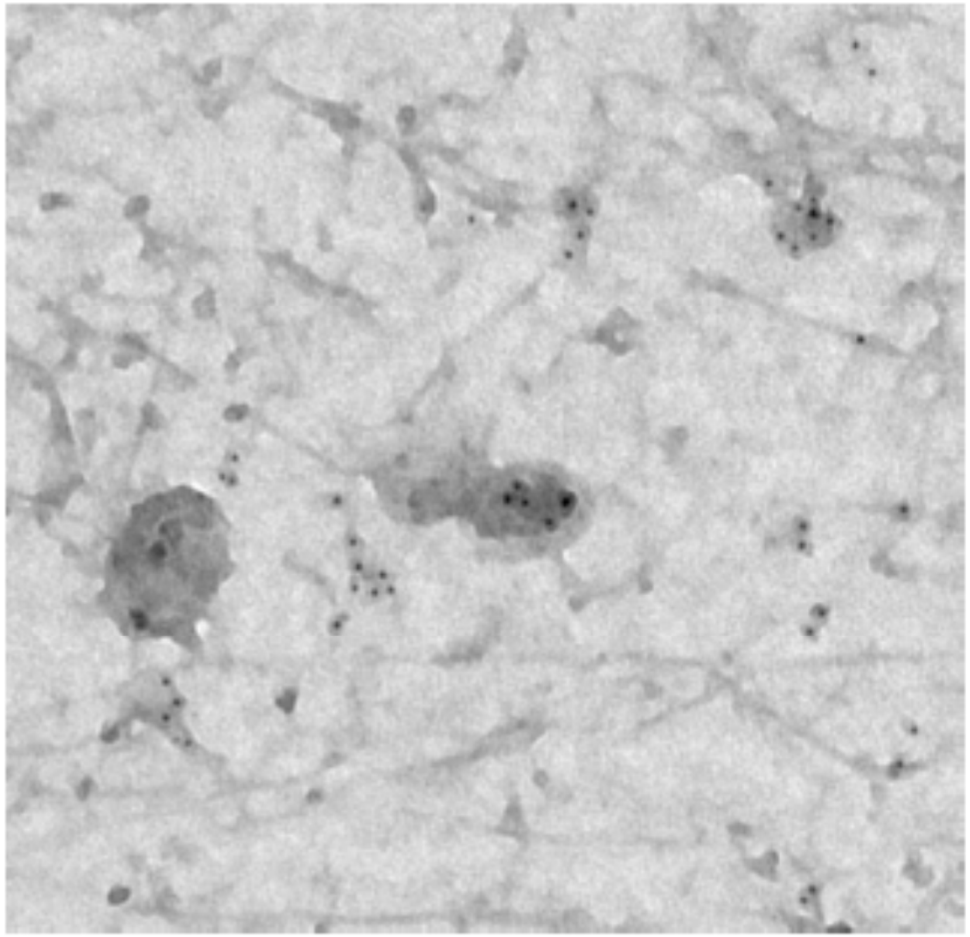
TEM Images of Receptors

- VEGF receptors labeled with gold particles
- Receptors tend to localize in **clusters**, that are sometimes enhanced in the presence of ligand



TEM Images of Receptors





Data Analysis: Clusters

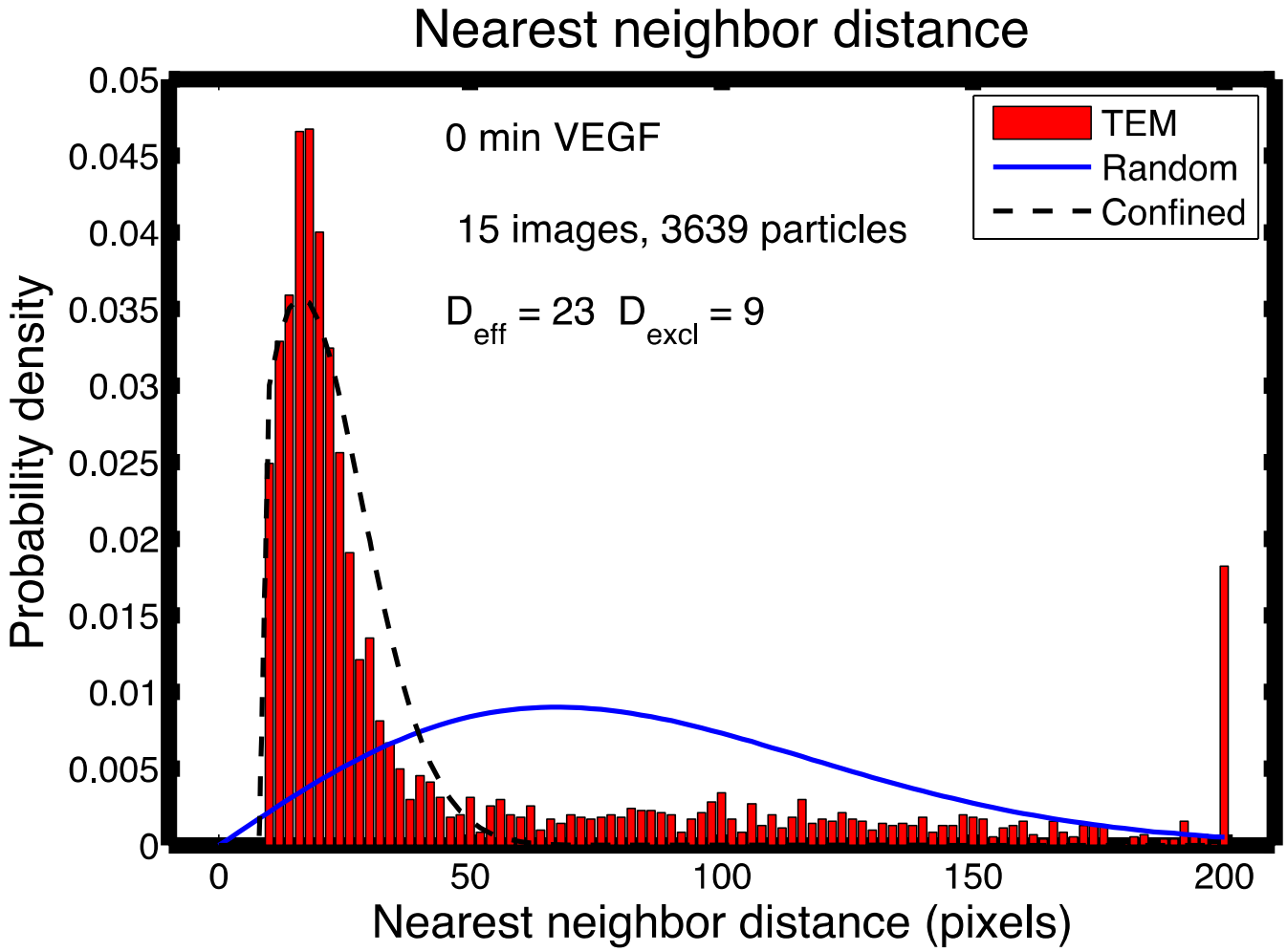
- Transmission Electron Microscopy of static receptors
  - Membrane sheets taken from [HUVEC] cells
  - VEGF receptors labeled with specific antibodies
  - Tags are 6-10 nm diameter gold particles, which appear as dark dots on TEM micrographs
- Preliminary analysis
  - Individual images cover ≈2x2 μm^2 (1-2% of the membrane)
  - A few tens to a few hundreds of receptors per image
  - Receptors identified by a semi-manual procedure (ImageJ)
- We derive the distribution of points for a set of images

Data Analysis: Clusters

How to quantify these receptor distributions?

- There are several commonly used measures, originally used for the study of trees
  - nearest-neighbor distance, Hopkins, …
  - all of these indicate that the distributions are not uniformly random
- Beyond that,
  - Identify the clusters (distance based clustering)
  - Seek to identify a basic (proximal) mechanism

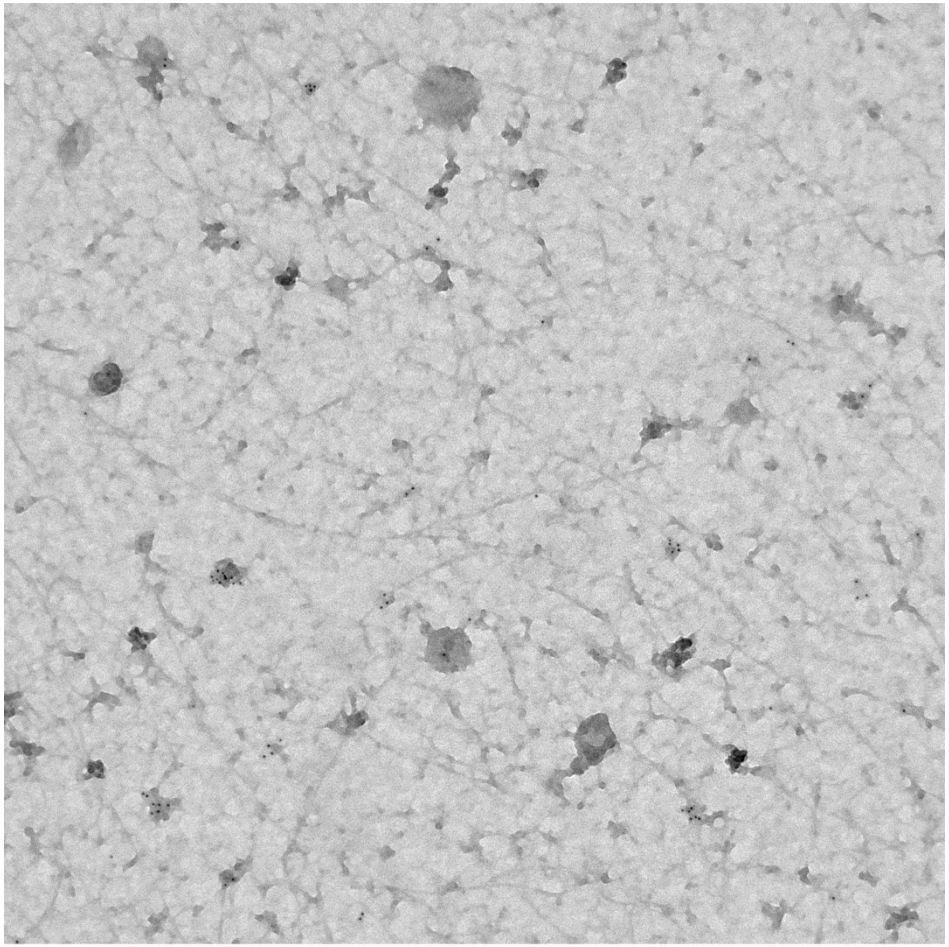
Data Analysis: Clusters



Data Analysis: Clusters

- Hierarchic distance based clustering
  (Espinoza et al, Bull. Math. Biol. 2013)
- Define clusters by comparing the mutual distance between points to a fixed distance parameter (L)
- Two points A,B are in the same cluster if either:
  - their distance is less than L: $d(A,B)<L$
  - there is a point C in the same cluster with A and B
- For a given configuration of points and length parameter L, we obtain a unique partition
- The issue of the optimal L remains…

Data Analysis: Clusters



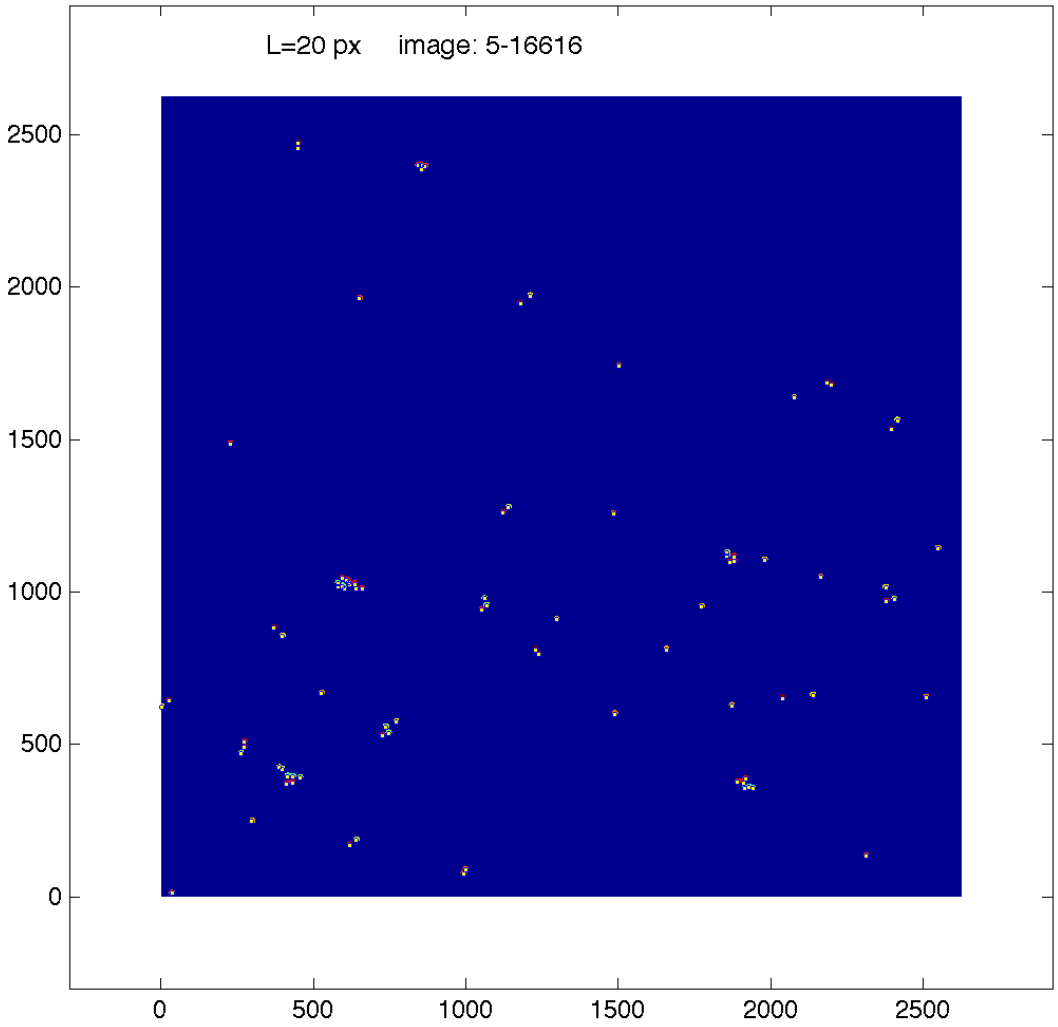
What is the right distance?

Shorter distance

- more clusters

Larger distance

- fewer, larger clusters



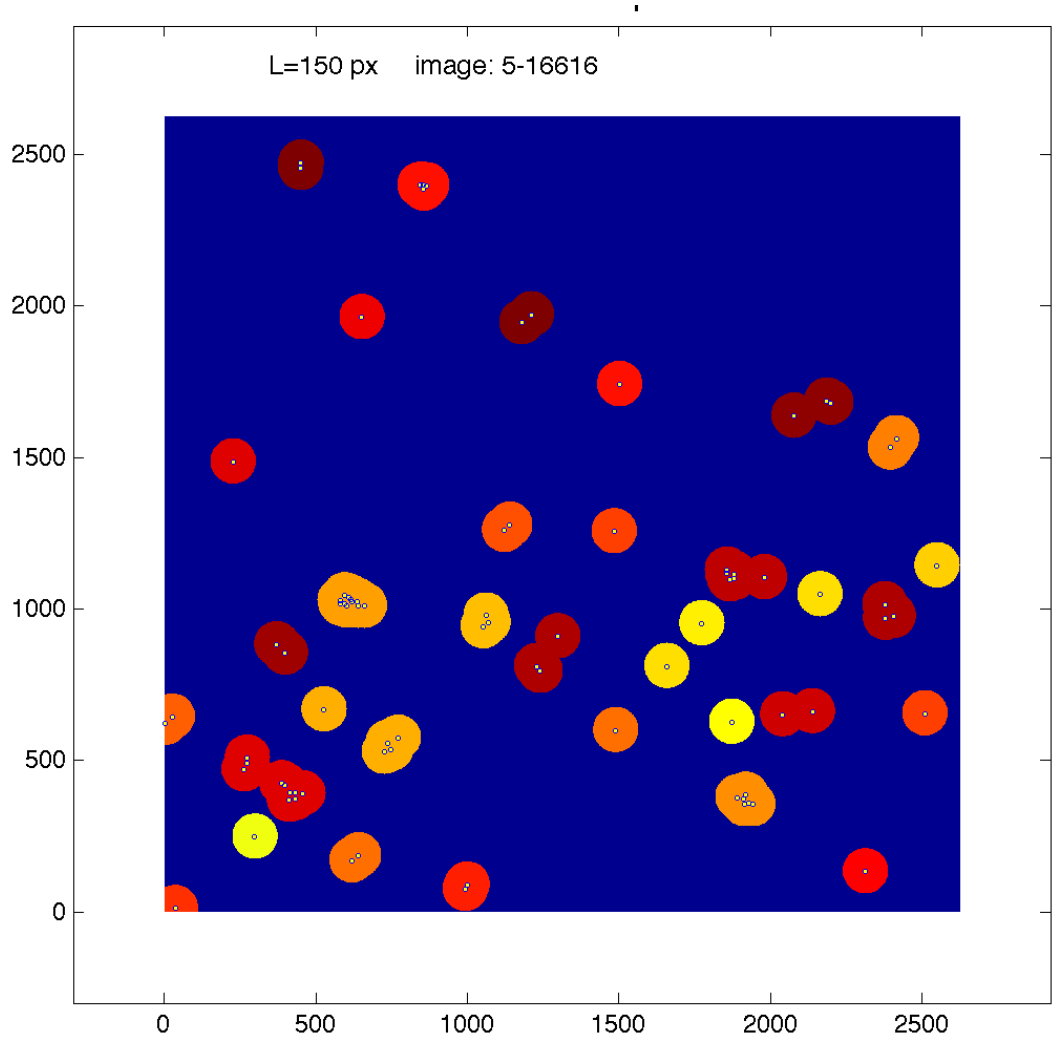
What is the right distance?

Shorter distance

- more clusters

Larger distance

- fewer, larger clusters



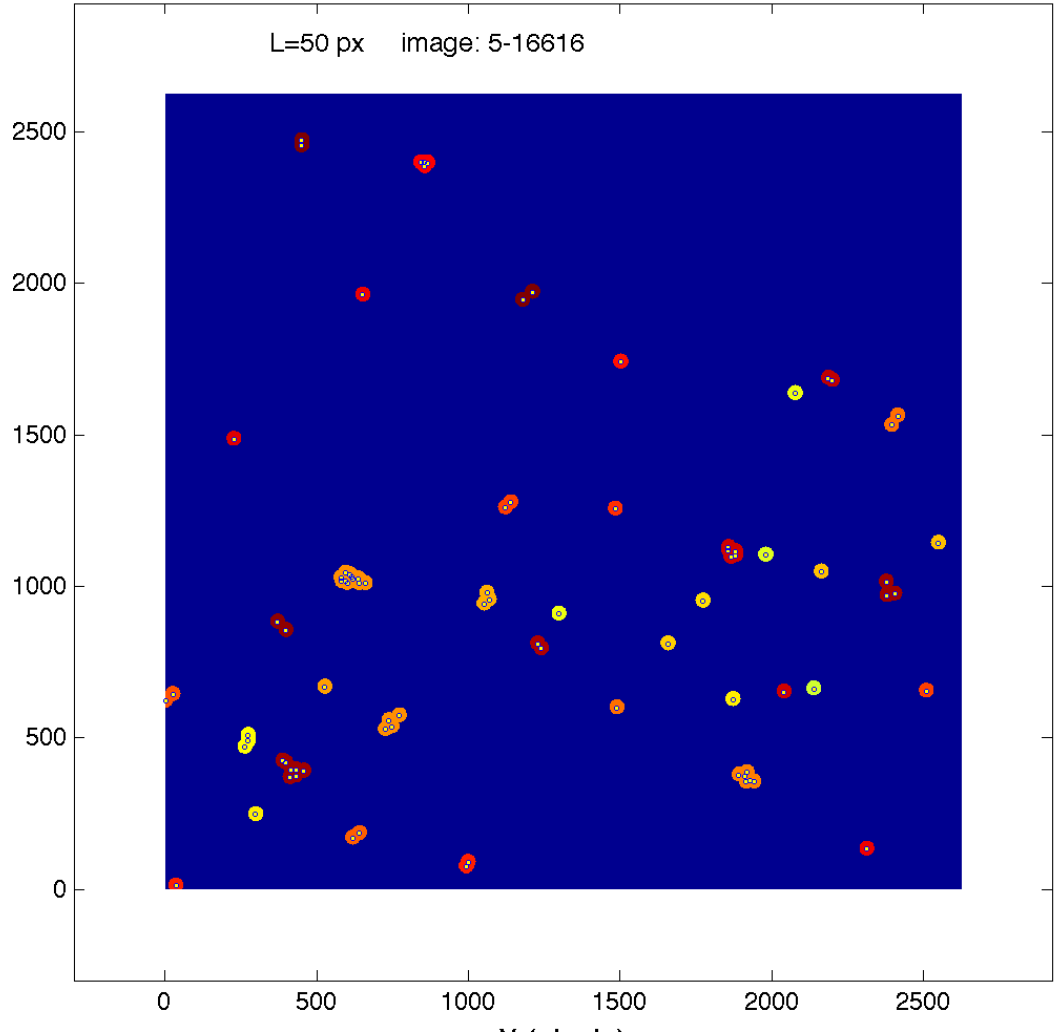
What is the right distance?

Shorter distance

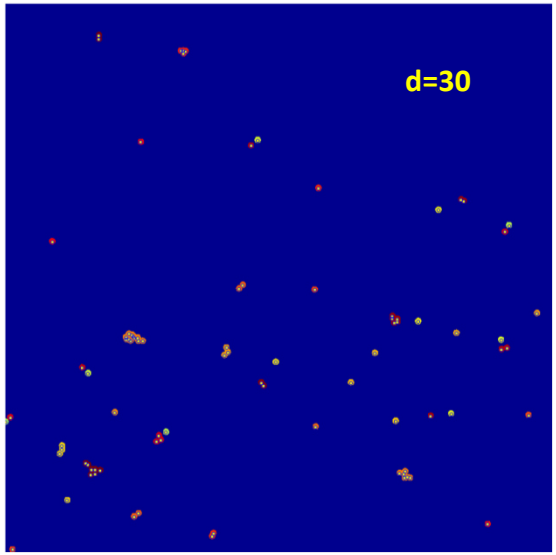
- more clusters

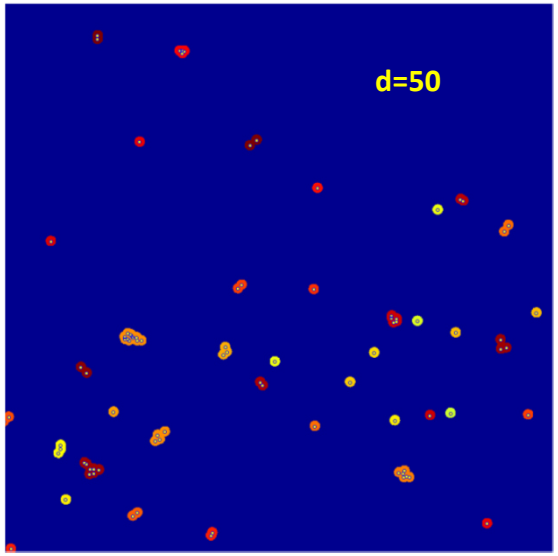
Larger distance

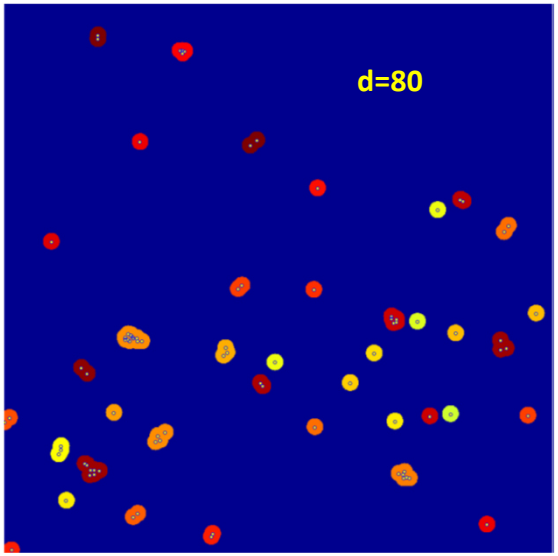
- fewer, larger clusters

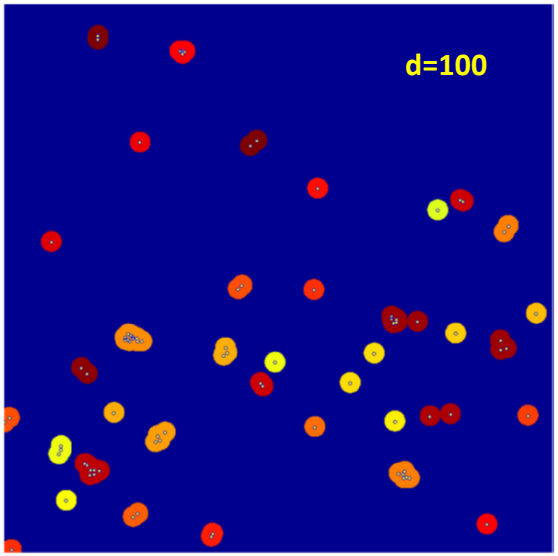


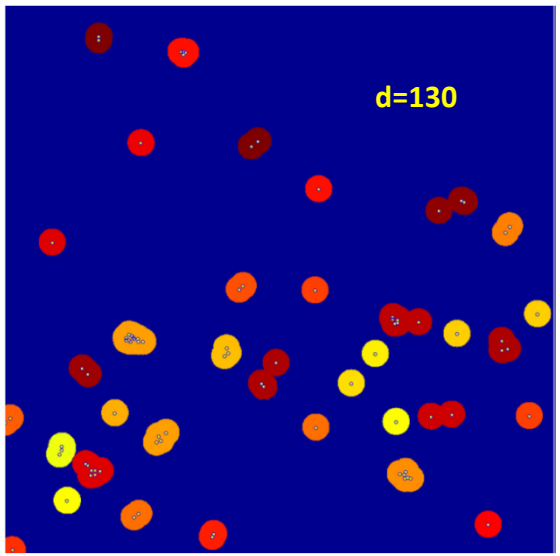
What is the right distance?

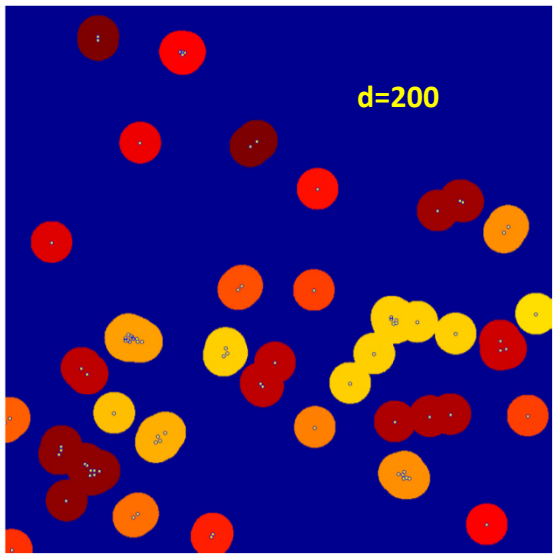




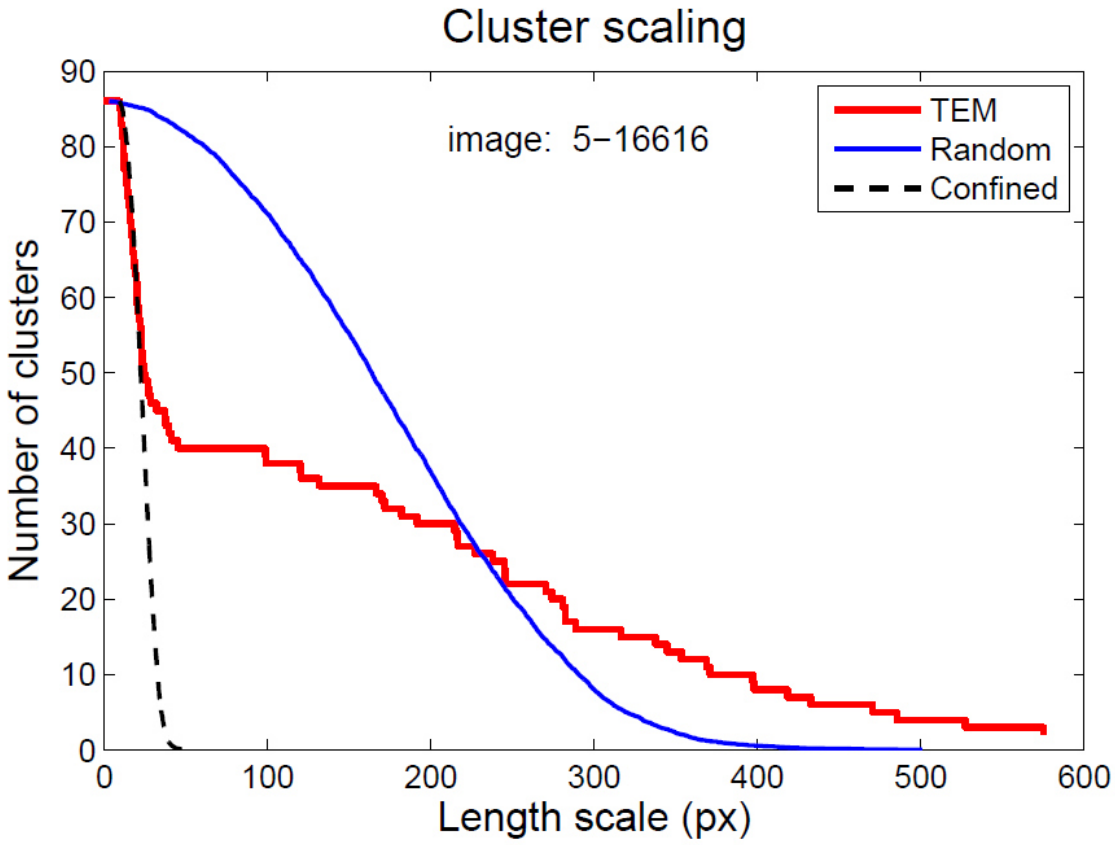




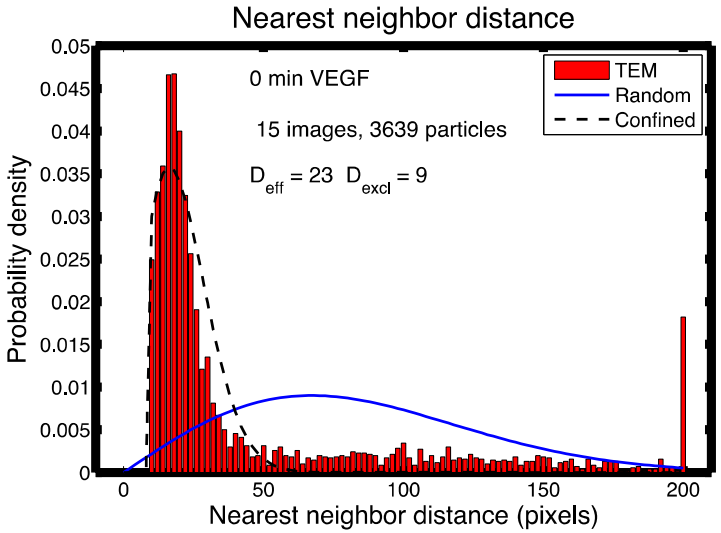


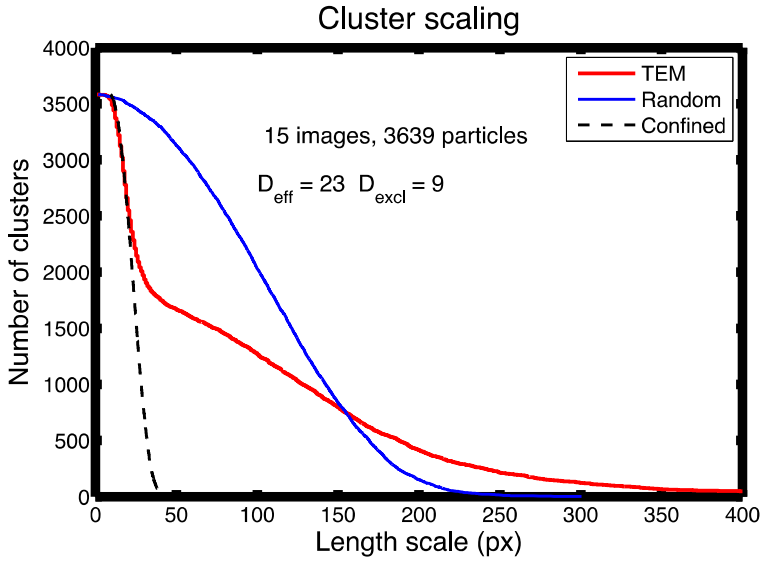


Cluster Number versus Length Scale

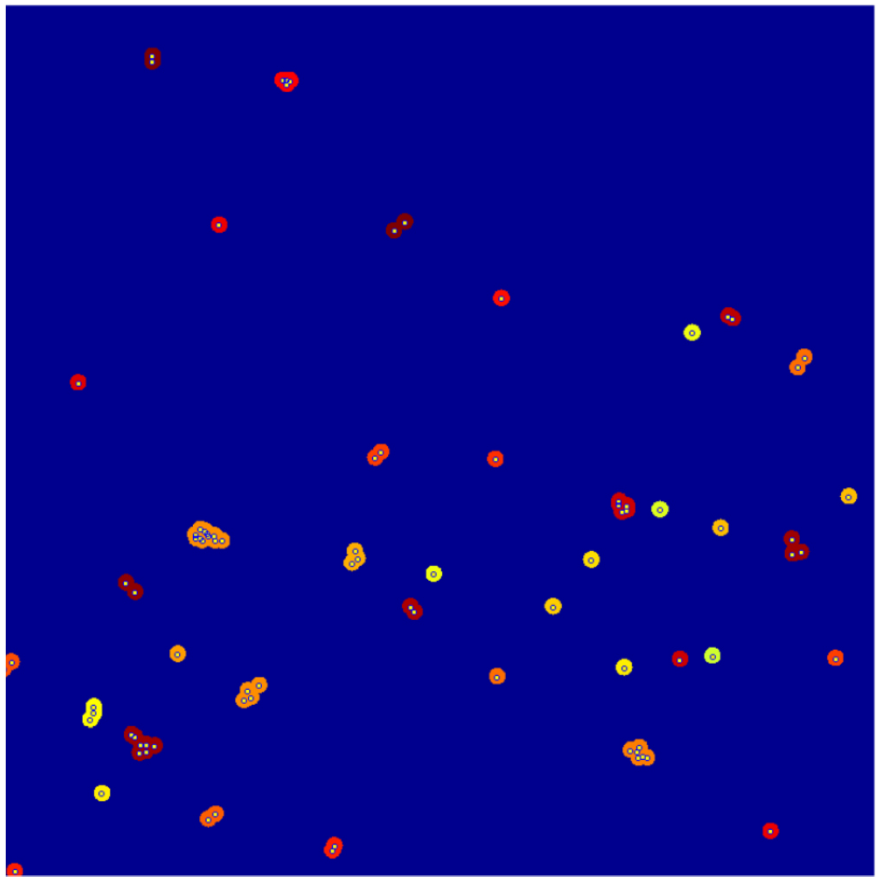


Clustering: Data Analysis





Particles distributed randomly within randomly distributed clusters?



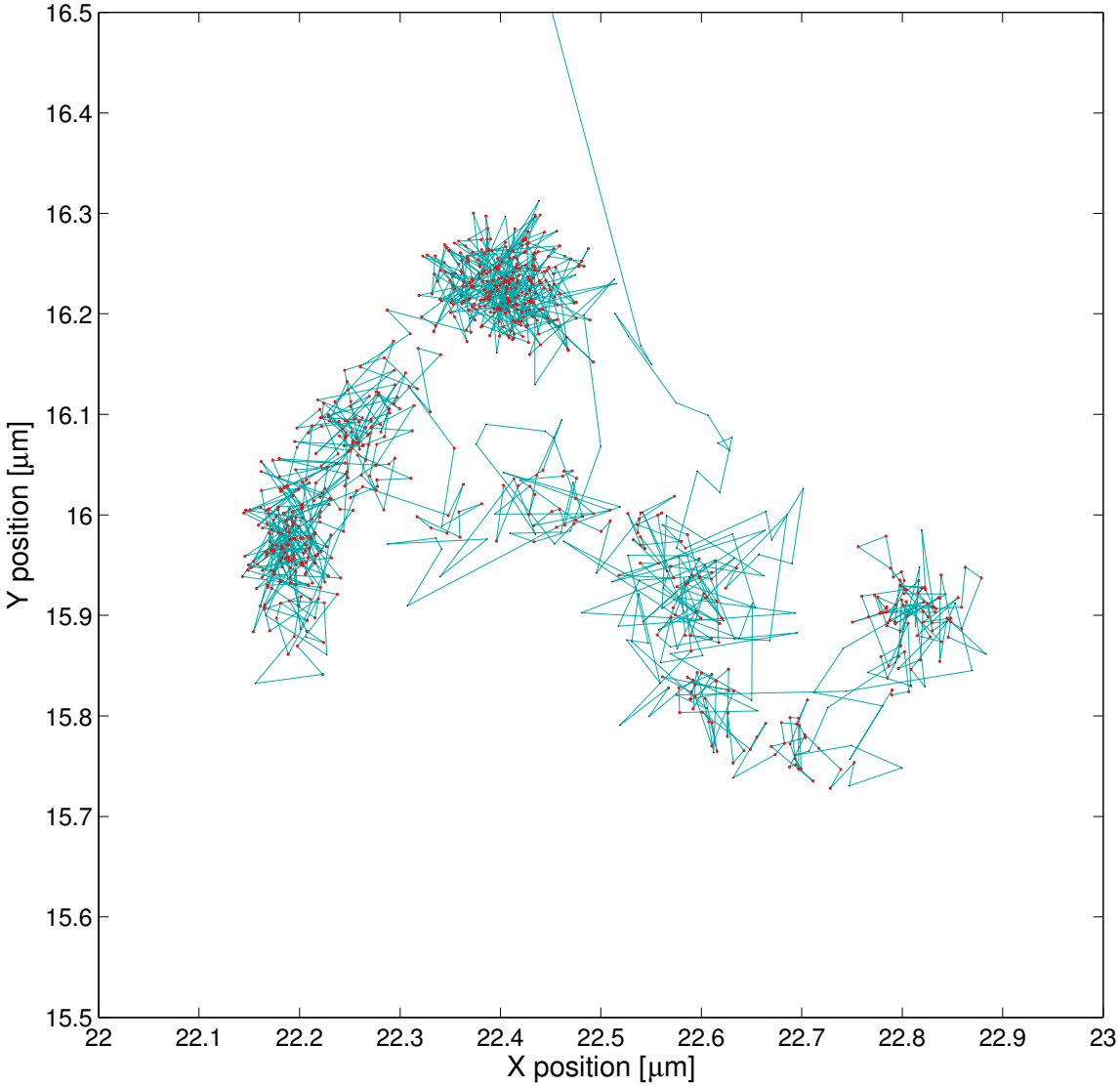
Data Analysis: Diffusion

- In the classic picture, proteins move freely along the membrane (in two dimensions), and this movement is uniform and random
- Dynamic imaging reveals that the random motion is modulated by **domains** that transiently trap receptors
- These domains are part of a varied landscape; the mechanism of confinement is not well understood

Data Analysis: Diffusion

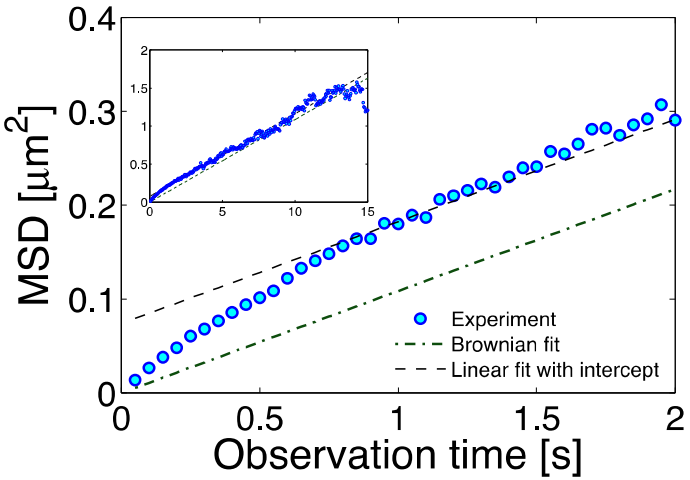
Single Particle Tracking (SPT) trajectory exhibiting transient confinement

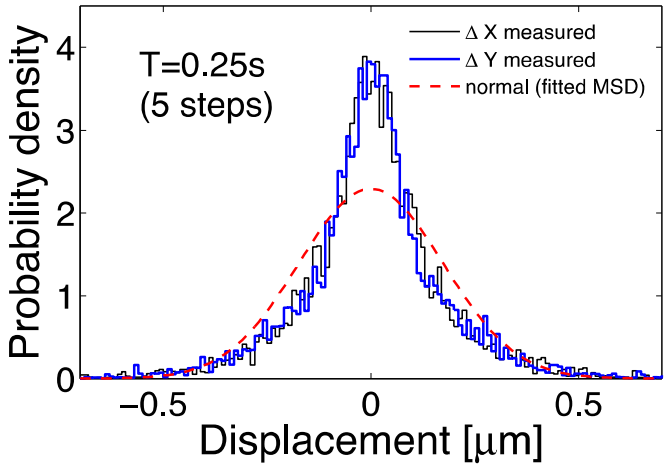
- the particle spends time exploring a compact area
- then proceeds to an adjacent or nearby patch and moves randomly in that one
- points labeled in red correspond to slower movement



Data Analysis: Diffusion

- The random movement of receptors is **anomalous diffusion**
  - similar to Brownian motion but the distribution of jump sizes deviates from the standard
- Square displacements over a fixed time should be distributed exponentially;
  - the data is consistent with two or more exponential components
- The mean square displacement (MSD) does not grow linearly with time;
  - rather, it tends to slow down



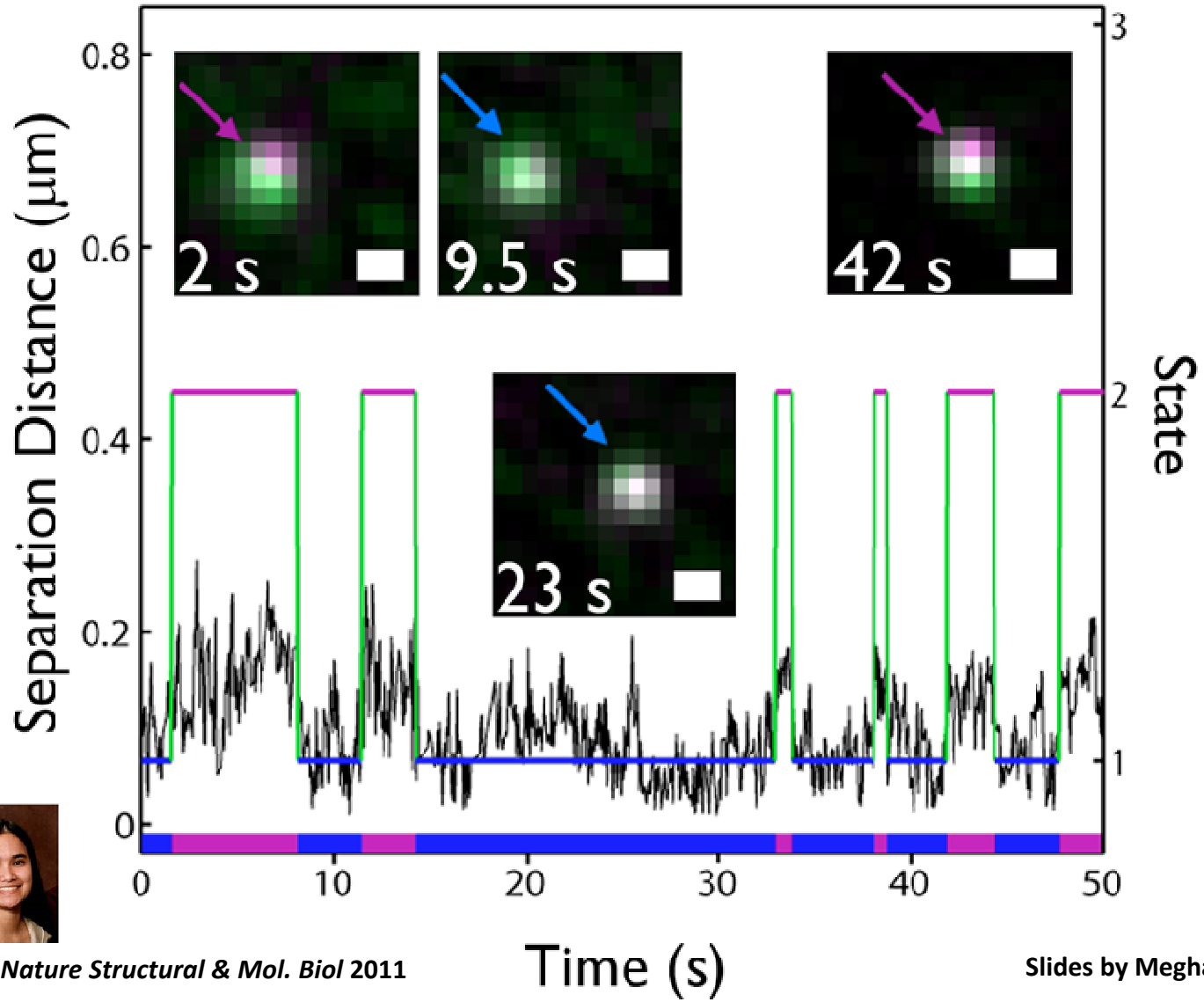


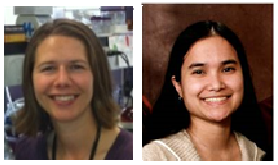
Hypothesis: Confining Domains

- The features of the anomalous diffusion are consistent with the hypothesis of **confining domains**:
  - which are up to a few hundred nm in size and represent a small fraction (10%) of the membrane
  - whose physical properties cause receptors to preferentially localize there during their otherwise random movement
- Transient confinement has been observed or hypothesized in many imaging studies;
  - classic work of Kusumi and coworkers
  - a recent example is **co-confinement of receptor pairs**[*]

[*] Low-Nam et al., *Nature Structural & Mol. Biol* 2011

Remarkable observation: Same "pair" could repeatedly dimerize. What can stochastic modeling tell us about this?



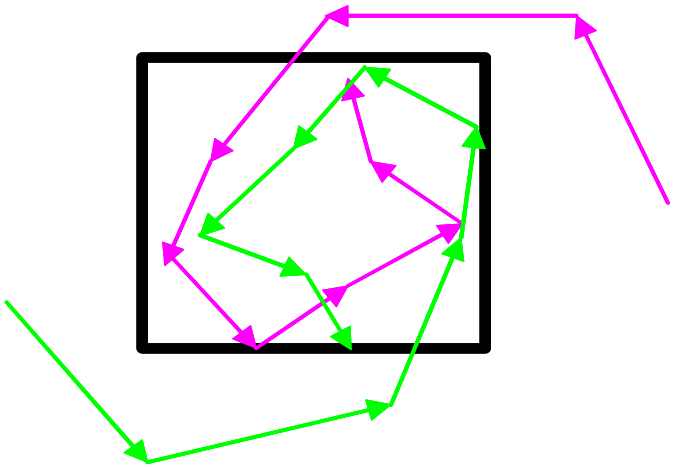


Low-Nam et al., *Nature Structural & Mol. Biol* 2011

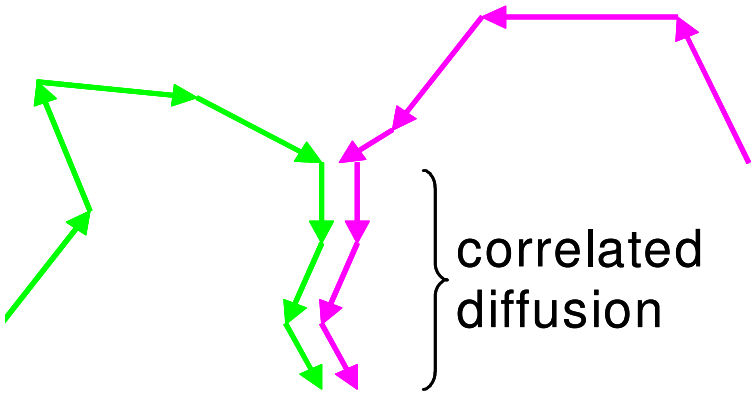
Slides by Meghan McCabe Pryor

In a complex … or just close?

co-confinement



complex



Slides by Meghan McCabe Pryor

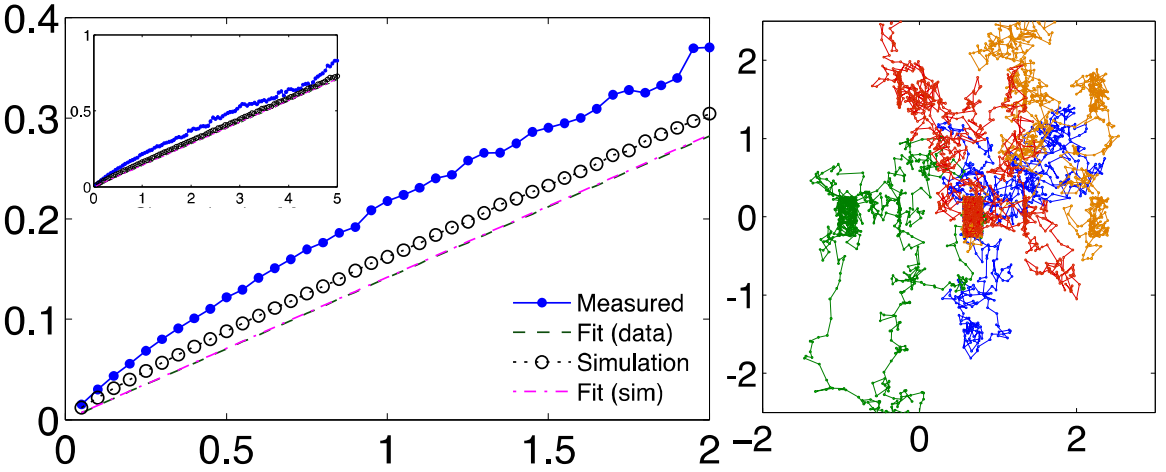
Modeling: Diffusion

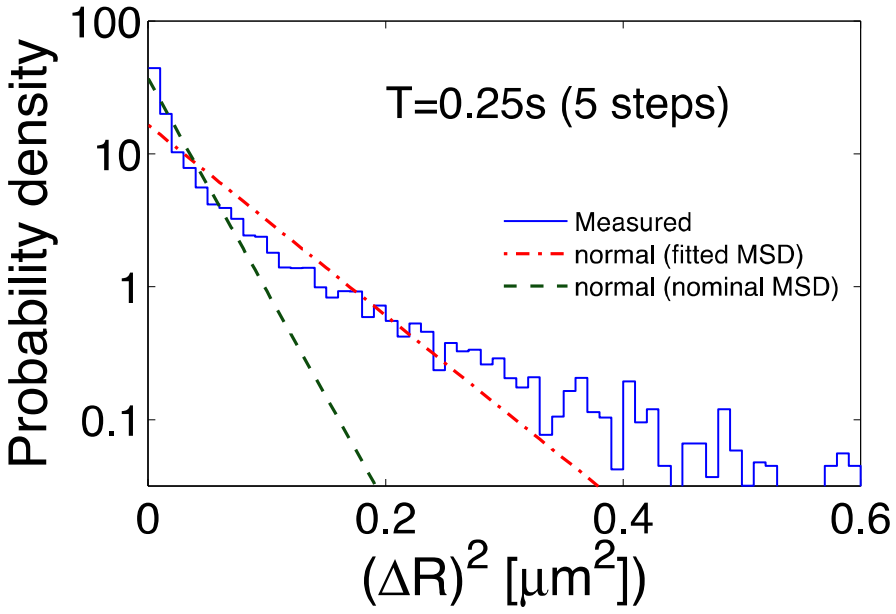
- Simulations in a landscape of confining domains reproduce the qualitative features of anomalous diffusion
  - Particles alternate between stretches of confinement and free movement
  - The **mean** square displacement grows for short times but slows down as the confined particles are blocked at the domain boundary
  - **Individual** displacements over **fixed time** intervals are the overlap of a free distribution that is close to Brownian and one corresponding to confined particles

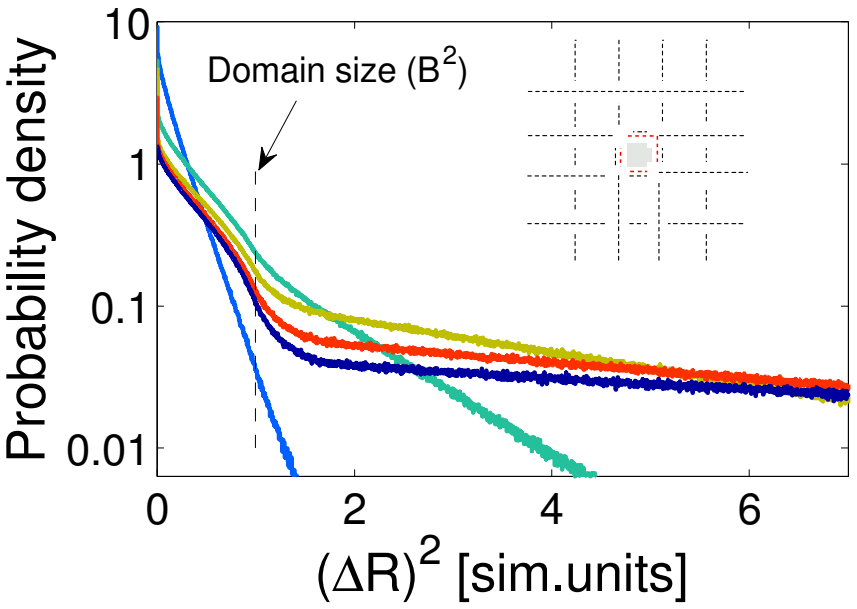
Modeling: Diffusion

$$P(r^2) = \frac{1}{2\sigma_{xy}^2} \exp\left(-\frac{r^2}{2\sigma_{xy}^2}\right)$$

$$\langle r^2 \rangle = 2\sigma_{xy}^2(t) = 4Dt$$







Modeling: Clusters

- Confining domains also provide a plausible explanation for the static clusters:
  - Larger clusters are simply the confining domains that accumulate the majority of receptors
  - Receptors that are outside a confining domain are singletons or form truly random, small clusters
- Given the size and number of confining domains
  - we can predict the size distribution of the clusters
  - compare images that contain different numbers of particles with the same model

Confining Domains: Just a Theory

- **Alternative hypothesis** (not the only one)
  - there are no confining domains per se; the membrane is partitioned into comparable domains, separated by barriers (actin filaments or cytoskeletal elements)
  - some feedback mechanism (e.g. crowding), leads to ever increasing occupancy of domains that acquired more receptors through a statistical fluctuation
- There is no known modality to image the domains directly; we must **validate** or falsify the hypothesis using the data we have

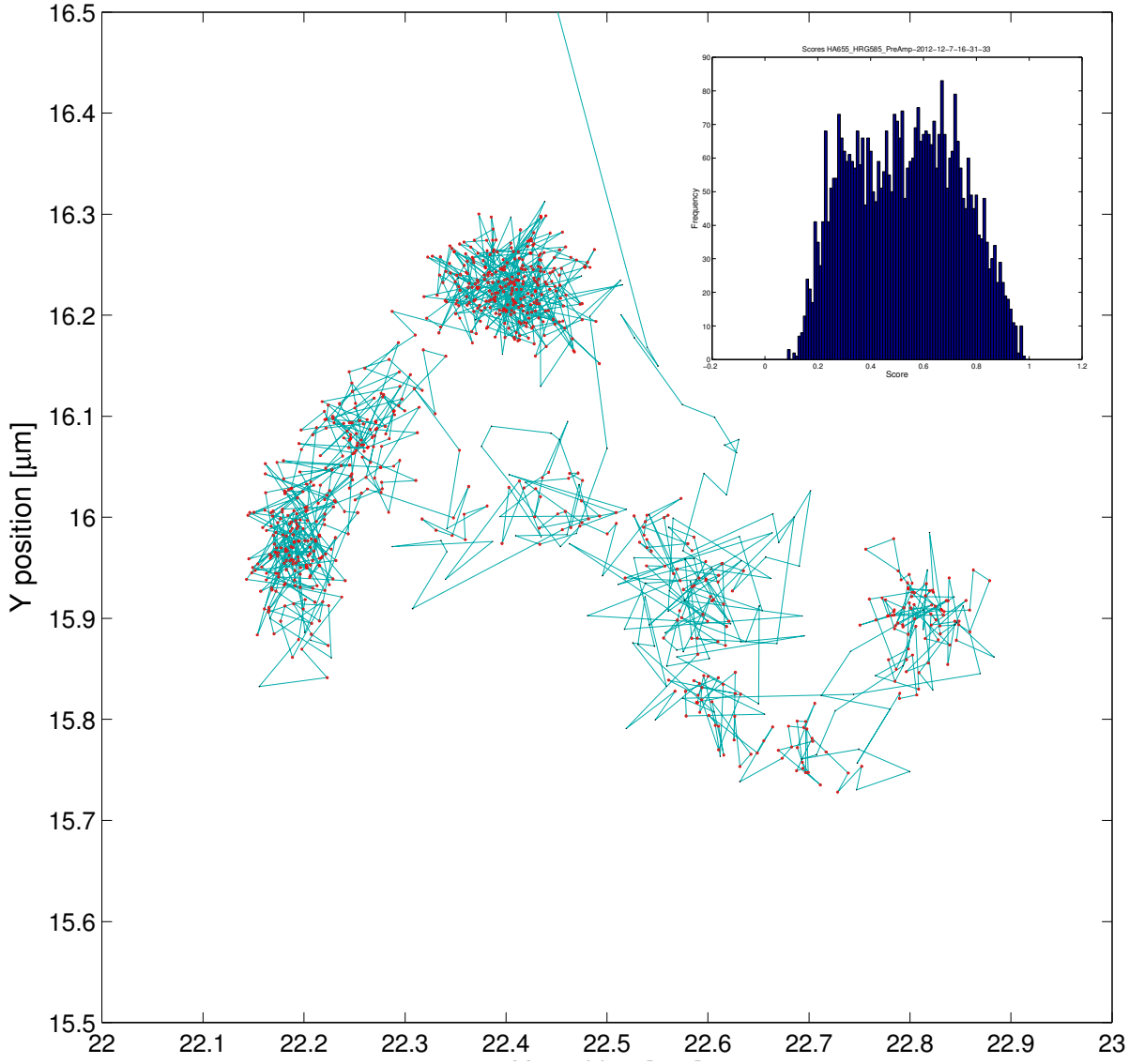
Domain Reconstruction: Clusters

- Can we recapitulate the clustering and diffusion data using the same confining domain model / parameters?
- Static imaging of clusters combined with a **statistical model for cluster sizes*** provides a range of domain size and attractiveness
- Using a typical density of particles within a domain and the actual images, we may reconstruct the outline of the domains

Domain Reconstruction: Diffusion

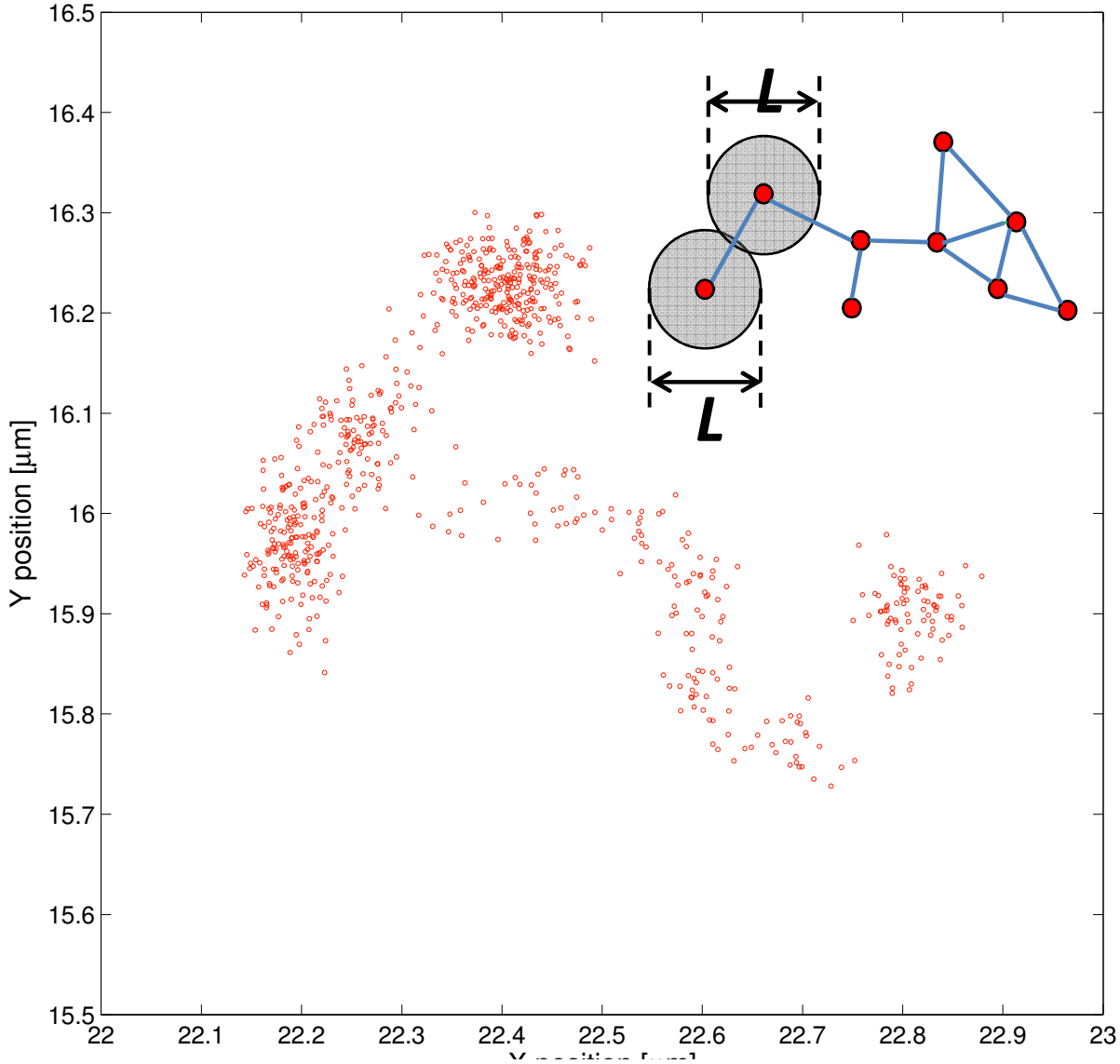
- Can we recapitulate the clustering and diffusion data using the same confining domain model / parameters?
- Within diffusion data, we can estimate the likelihood that a given point in a trajectory is part of the confined population or not
- Using variations in mobility, we can attempt to reconstruct the confining domains that modulate the movement of the particles.

Domain Reconstruction: Diffusion



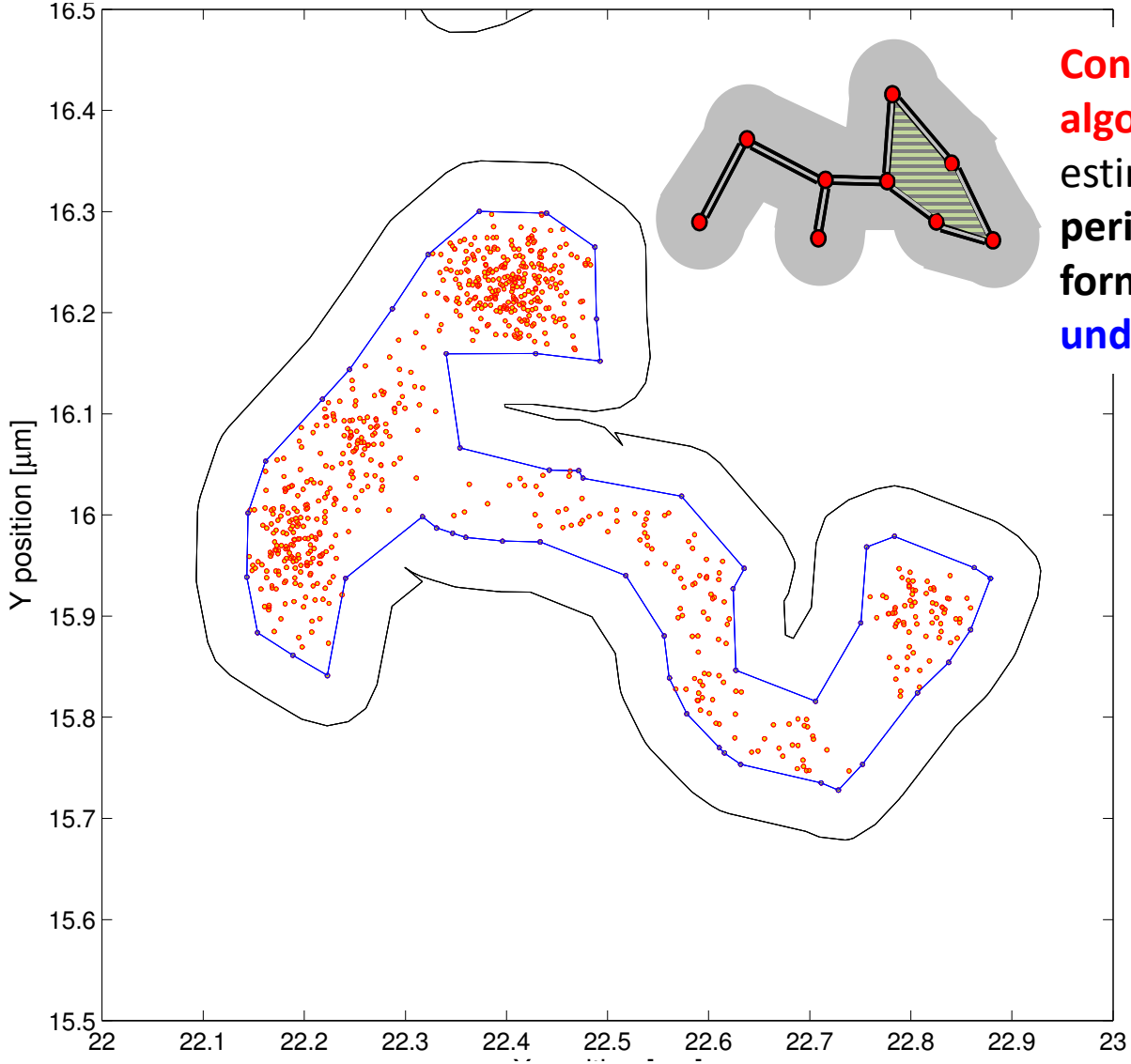
Slow points identified from a combined score based on jump sizes over several frame intervals

Domain Reconstruction: Diffusion



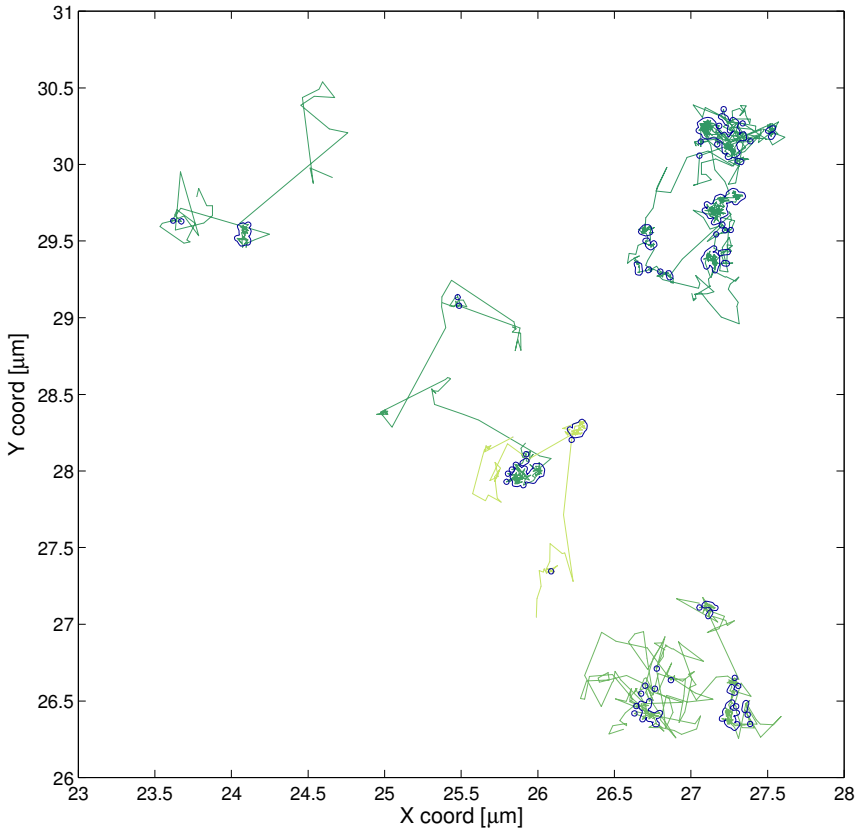
Clusters of points identified via distance based clustering

Domain Reconstruction: Diffusion

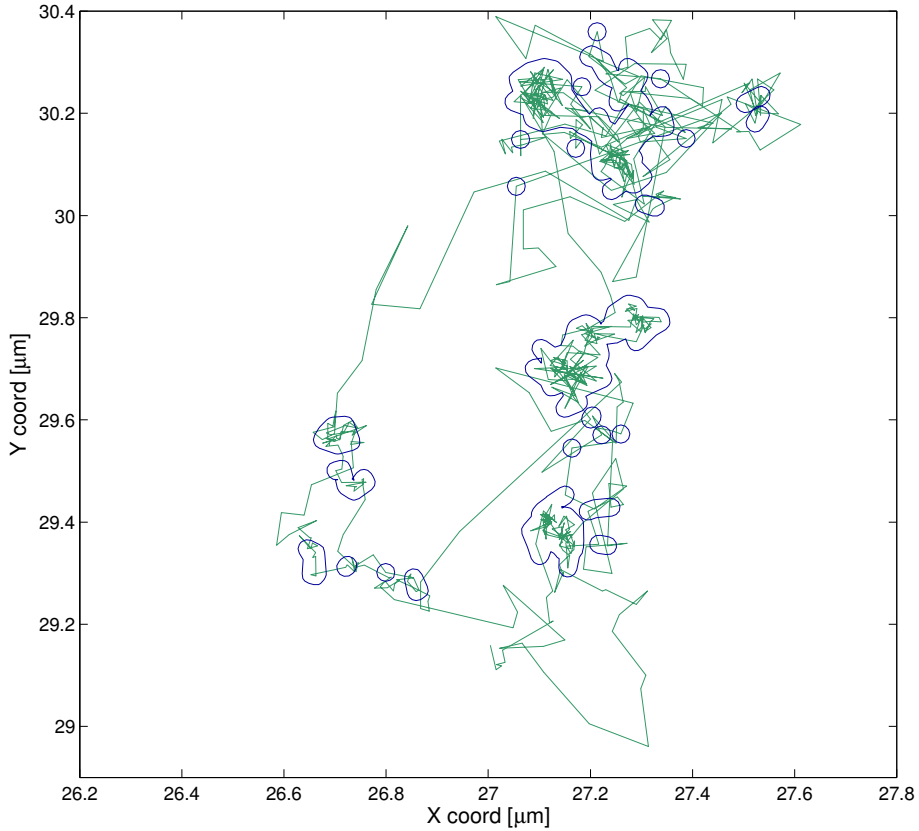


Contour building algorithm - estimates **area**, **perimeter**, and **form factor of underlying domains**

Domains from contours



Domains from contours

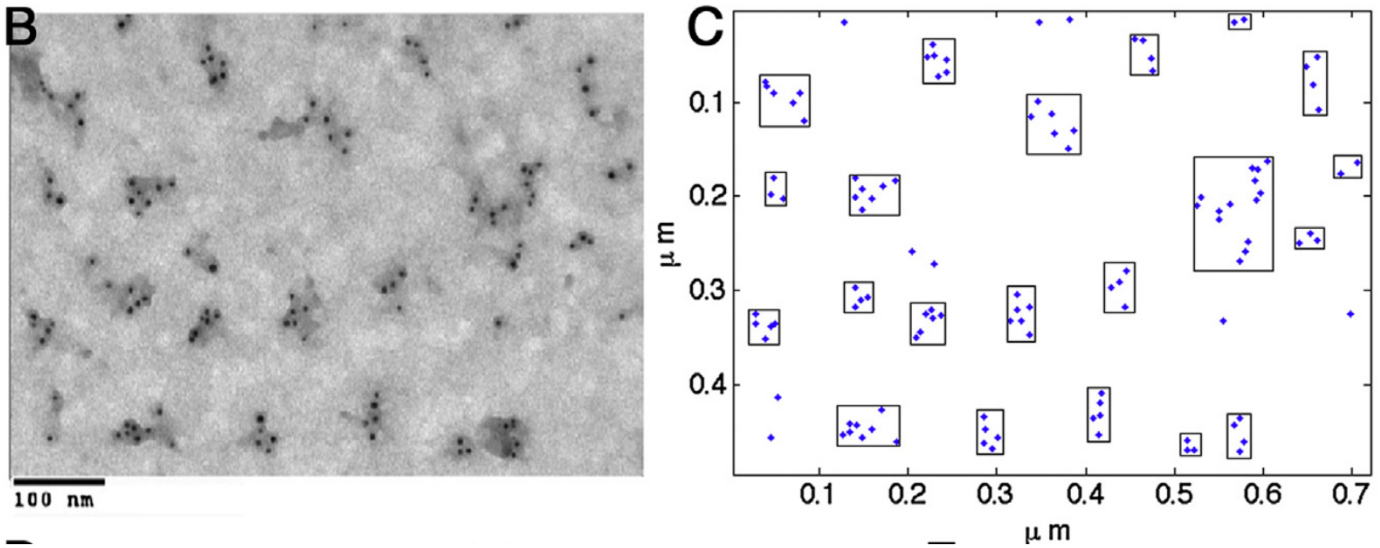


Modeling: Impact on Signaling

- This confining domain hypothesis is used in detailed models (simulations) of EGF signaling
- “agent based” simulations, that follow the movement and interactions of individual molecules (Smoldyn*)
  - Brownian motion for the movement part
  - First order reactions triggered by Poisson processes
  - Second order reactions triggered by geometric proximity (a collision / dissociation radius)
- The movement is limited to confining domains, taken from TEM images (Pryor et al Biophys J. 2013)
- Study detailed molecular mechanisms in this context

*Andrews & Bray, Phys. Biol. 2004

Modeling: Impact on Signaling



Fully spatial simulations in the ErbB1 system

- Movement is limited to confining domains, taken from TEM images
- Kinetic parameters inferred directly from SPT

This model is currently used to investigate specific molecular structure aspects

Pryor et al Biophys J. 105(6) 1533-1543 (2013)

Steinkamp et al. (in preparation)

Modeling: Impact on Signaling

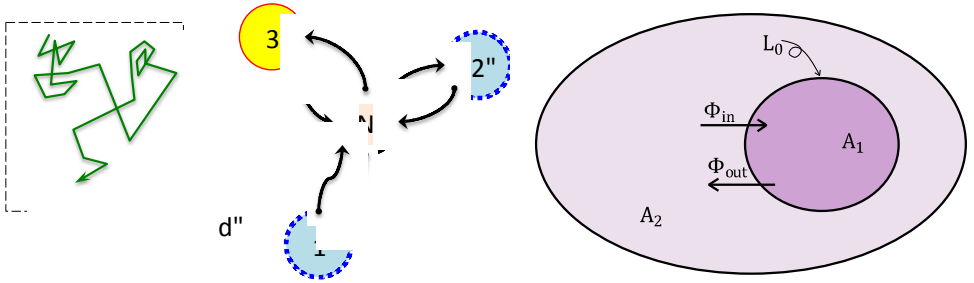
- A continuing problem is that fully detailed spatial reaction-diffusion simulations can not be scaled up the level of the entire cell
- One approach is to model the domains as well-mixed compartments that communicate with the rest of the membrane
- Back to CRN or well-mixed stochastic (Gillespie) models that can offer insight into how **confining domains** fit into the machinery of signaling

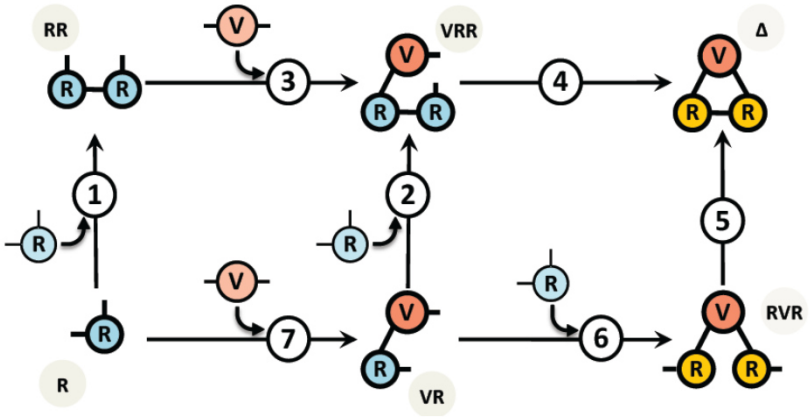
Modeling: Impact on Signaling

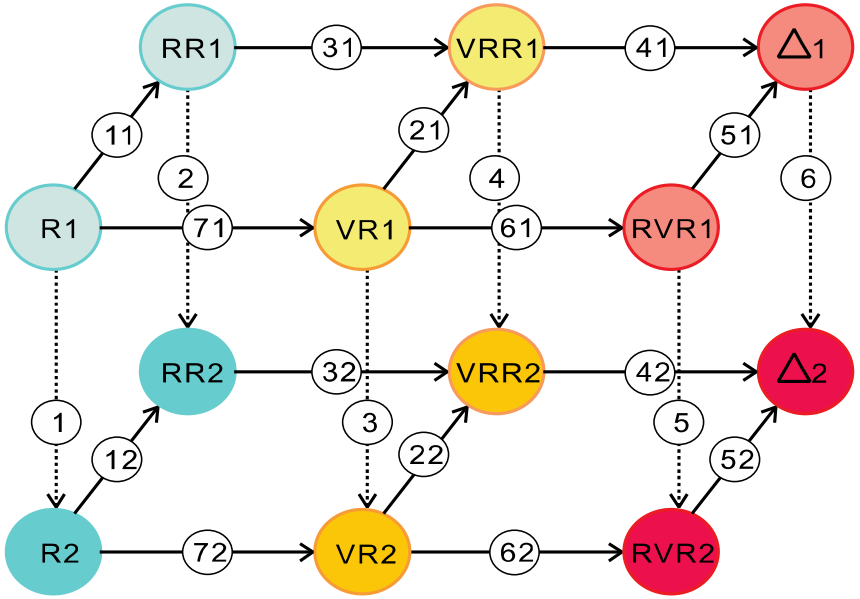
- A well-mixed model of signal initiation with domains
  - a compartment for the “normal” part of the membrane and one or more representing confining domains
  - each compartment has an instance of the basic CRN; species may transfer between compartments consistently with the attractiveness and size parameters
- “Cartesian product” of the basic CRN model and a spatial network of domains

Modeling: Impact on Signaling

- Abstract the domains into well-mixed compartments
- Place a copy of the CRN into each compartment







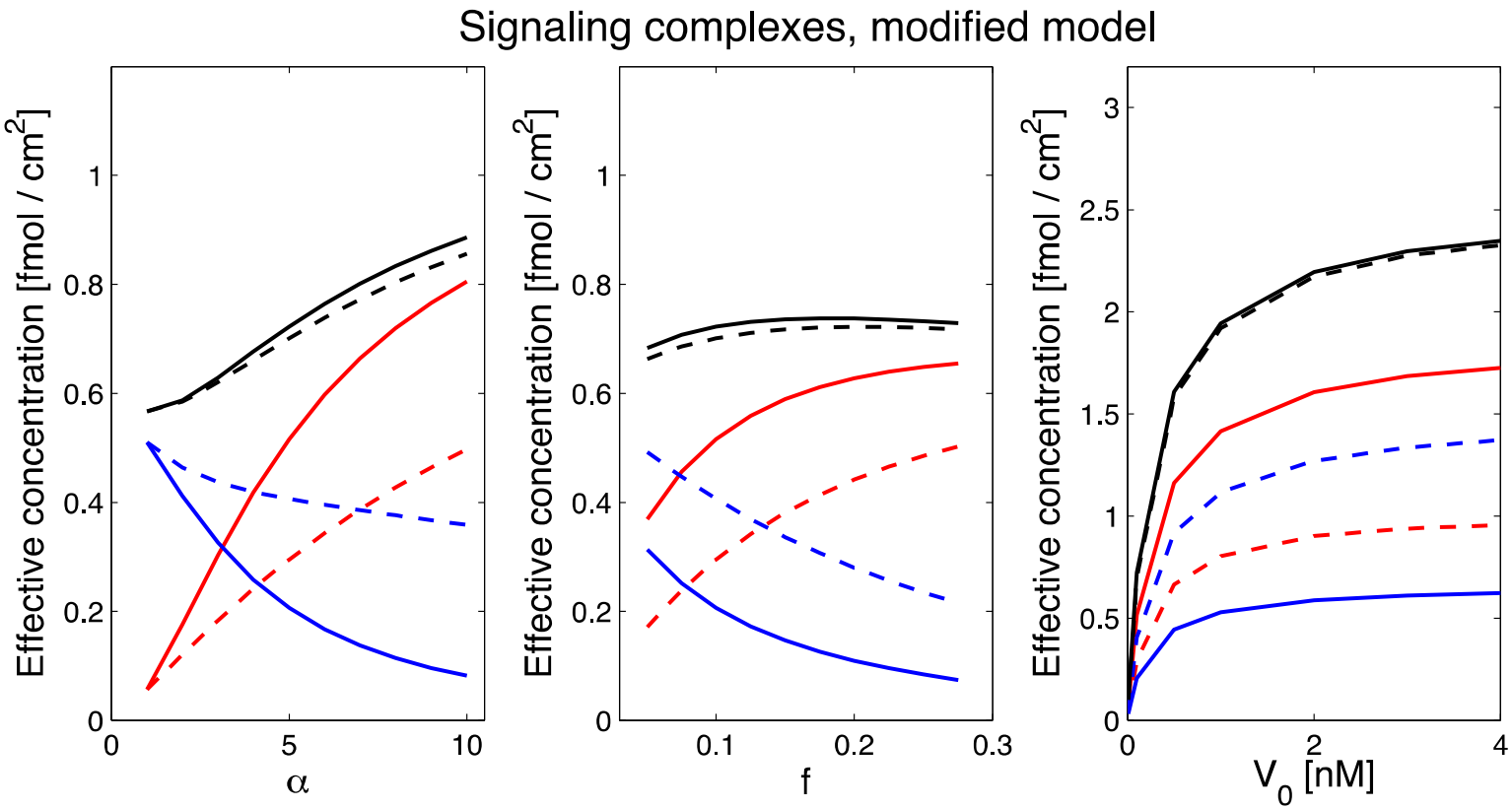
Chen et al, EPTCS Vol.125 37-52 (2013) – Proc. of HSB 2013

Q-bio 2015 → posters of C. Short and E. Güven

Modeling: Impact on Signaling

- Results:
  - The presence of VEGF ligand directly increases dimerization; confining domains accumulate more receptors in the presence of VEGF
  - Conversely, confining domains may increase the observed dimerization several fold; this enhances the intensity and speed of signal initiation
- The domain story can be connected directly with signaling by estimating effective kinetic parameters

Modeling: Impact on Signaling



Confining domains may increase the observed dimerization several fold; this enhances the intensity and speed of signal initiation

Chen et al, EPTCS Vol.125 37-52 (2013) – Proc. of HSB 2013

Role of Models

- There are solid methods to model reaction-diffusion on a molecular level
- Use modeling in conjunction with microscopy data to characterize and explain these “nano”-level phenomena
- Need to make connection to data on the cellular level
- Not to mention the emerging wealth of genomic data – mutations, markers, etc.

Role of Models

- But signaling is already complicated. It is challenging to gain insight from the signaling models as it is.
- We also need to **simplify**, summarize what we learn on the molecular level in a way that fits into a higher level description

Thanks

University of New Mexico (STMC)

- Jeremy Edwards
- Flor Espinoza
- Diane Lidke
- Janet Oliver
- Meghan McCabe Pryor
- Krishnan Radhakrishnan
- Stanly Steinberg
- Mara Steinkamp
- Bridget Wilson

West Virginia University (Math)

Ye Chen

Emine Güven

Christopher Short

NIH (NCI, NIGMS)

In case you were wondering…