The modelling and simulation of Cellular Kinetics

Kevin Burrage
Federation Fellow of the ARC – kb@maths.uq.edu.au
UQ, Brisbane, Australia

Dresden/UQ July 2005

Contents

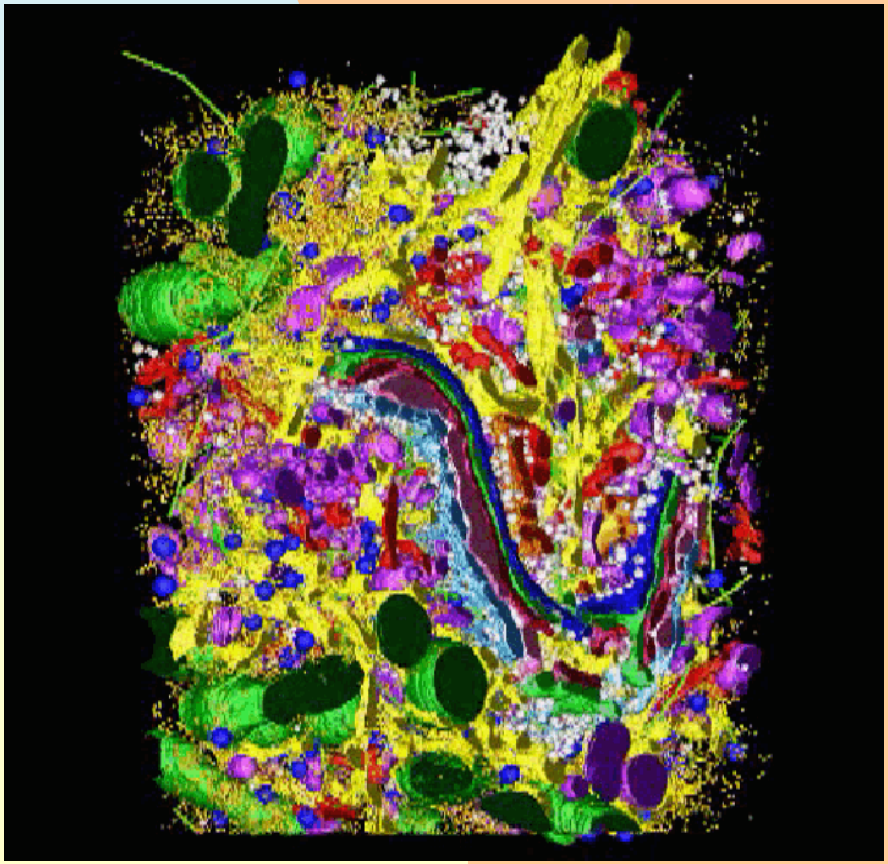
- Cell Biology, Noise and Genetic Regulation
- Discrete stochastic techniques
- Modelling and Simulation Issues
- Modelling and Simulation of Delays in GR
- External Noise
- Space and time
- Challenges

Joint work with

- Margherita Carletti (Urbino)
- Tian Tianhai, Pamela Burrage, Dan Nicolau Jr (UQ)
- Peter Hunter (Auckland)
- Grant Lythe (Leeds)

Cell Biology, Noise and Genetic Regulation

- Cells are complex with a wide variety of ultrastructure.
- Many different types of dynamics processes and transports
- Single Particle Tracking in a living cell.
- Magnify a cell a million times -Water molecule: full stop; Protein: ping pong ball; Ribosome: soccer ball; Mitochondion: person; Nucleus: car; Cell membrane: <1 cm thick!

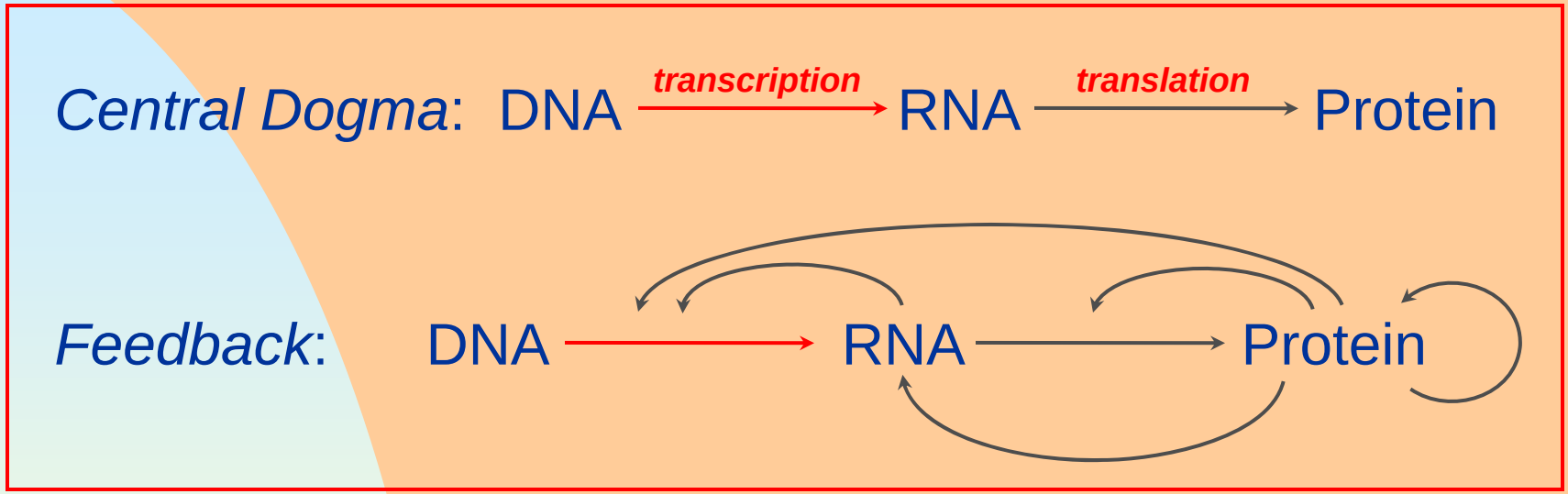


Three D EM image of a mammalian insulin secreting cell (Marsh)

Biological Evidence of noise

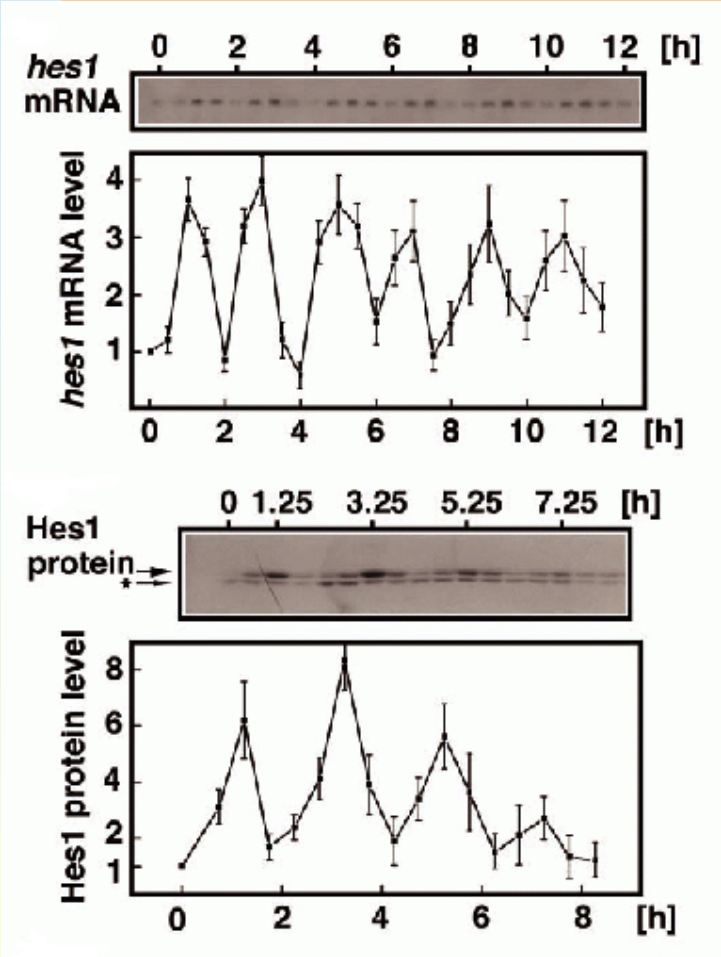
- “Stochasticity is evident in all biological processes … the proliferation of both noise and noise reduction systems is a hallmark of organismal evolution” – Federoff et al.(2002).
- “Transcription in higher eukaryotes occurs with a relatively low frequency in biologic time and is regulated in a probabilistic manner” – Hume (2000).
- “Gene regulation is a noisy business” – Mcadams et al. (1999).
- “Initiation of gene transcription is a discrete process in which individual protein-coding genes in an off state can be stochastically switched on, resulting in sporadic pulses of mRNA production” – Sano 2001.

Elements of Genetic Regulatory Networks

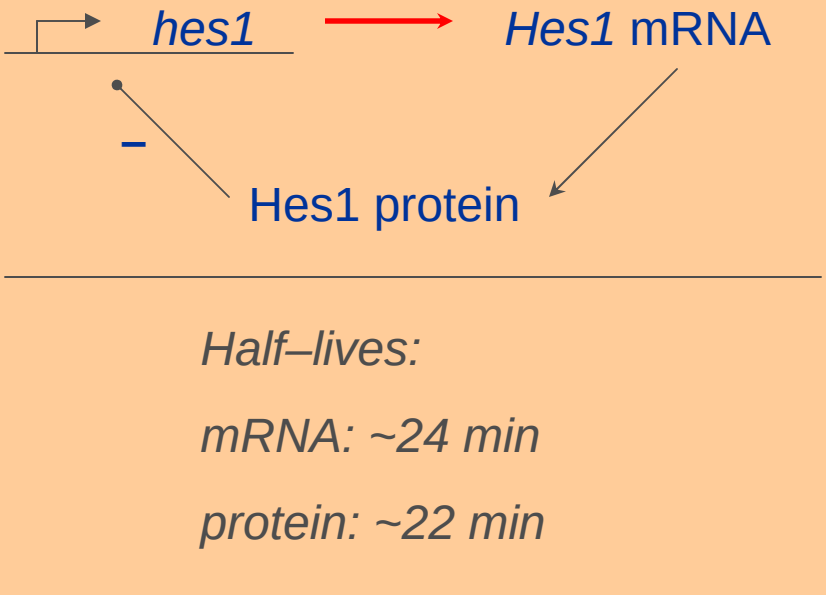


- **Transcription factors** bind to DNA sequences in regulatory regions of genes.
- Binding regulates the rate at which transcripts (polymerase) of the gene are **initiated**.
- Protein is made off mature mRNA transcripts by **translation** (ribososomes).

Hes1 oscillates in cultured mouse cells



Hirata *et al*., *Science* **298**, 840–843 (2002).



… but a non-delayed (ODE) model cannot oscillate without additional variables

Hirata *et al*. predicted extra components in the feedback loop.

Discrete Stochastic Techniques

Michaelis – Menten Reaction

$$S_1 + S_2 \xrightarrow{k_1} S_3 \qquad a_1(X) = k_1 S_1 S_2$$

$$S_3 \xrightarrow{k_2} S_1 + S_2 \qquad a_2(X) = k_2 S_3$$

$$S_3 \xrightarrow{k_3} P + S_1 \qquad a_3(X) = k_3 S_3$$

The stoichiometric vectors are and the Law of Mass Action gives

$$\nu_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \quad \nu_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \quad \nu_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$X'(t) = \sum_{j=1}^{m} \upsilon_j a_j(X(t))$$

Stochastic Simulation Algorithm

- Simulates the time evolution of a well stirred chemical reacting system by taking proper account of inherent randomness. Small Numbers of Molecules
- Well-stirred mixture
- N molecular species $S_1,...S_N$
- Constant temperature, fixed volume Ω
- M reaction channels $R_1,\ldots R_M$
- Dynamical state $X(t)=(X_1(t),...,X_N(t))$

where $X_i(t)$ is the number of S_i molecules in the system

- Propensity function $a_j(x)dt =$ the probability, given $X(t)=X$, that one R_j reaction will occur somewhere inside Ω in the next infinitesimal time interval $[t,t+dt]$

- Reaction pdf $P(\tau, k \mid X(t))$
- $P(\tau, k \mid X(t))d\tau$ Probability kth reaction in $(t+\tau, t+\tau+d\tau)$

$$P(\tau, k \mid X) = a_k(X)\exp(-\tau a_0(X)),\ \ a_0(X) = \sum_{j=1}^{m} a_j(X)$$

- SS Algorithm

  While t < T

  2 samplings of U(0,1) $\tau = \frac{1}{a_0(X(t))} Ln(\frac{1}{U_1})$

  Choose j $[0, \frac{a_1}{a_0}], (\frac{a_1}{a_0}, \frac{a_1 + a_2}{a_0}], \cdots (\frac{a_1 + \ldots a_{m-1}}{a_0}, 1]$

  Update $X \leftarrow X + \upsilon_j$

Modelling and Simulation Issues

Modelling Regimes

- **Discrete and stochastic** – Small numbers of molecules. Exact description via Stochastic Simulation Algorithm (SSA) - *Gillespie*.

  Large computational time.

- **Continuous and stochastic** - A bridge connecting discrete and continuous models.

  Described by SDEs – The Chemical Langevin Equation.

- **Continuous and deterministic** – Law of Mass Action. The Reaction Rate equations. Described by ordinary differential equations. Not valid if molecular populations of some critical reactant species are small.

- From moments of CME

$$\mu(X) = \sum_{j=1}^{m} \upsilon_j a_j(X)$$

$$B^2(X) = \upsilon \, Diag\left(a_1(X), \cdots, a_m(X)\right) \upsilon^T$$

- Chemical Langevin Equation - SDE

$$dX = \sum_{j=1}^{m} \upsilon_j a_j(X) dt + \sum_{j=1}^{N} b_j(X) dW_j$$

$$B = \left[b_1, \cdots, b_N\right]$$

Important Issues I

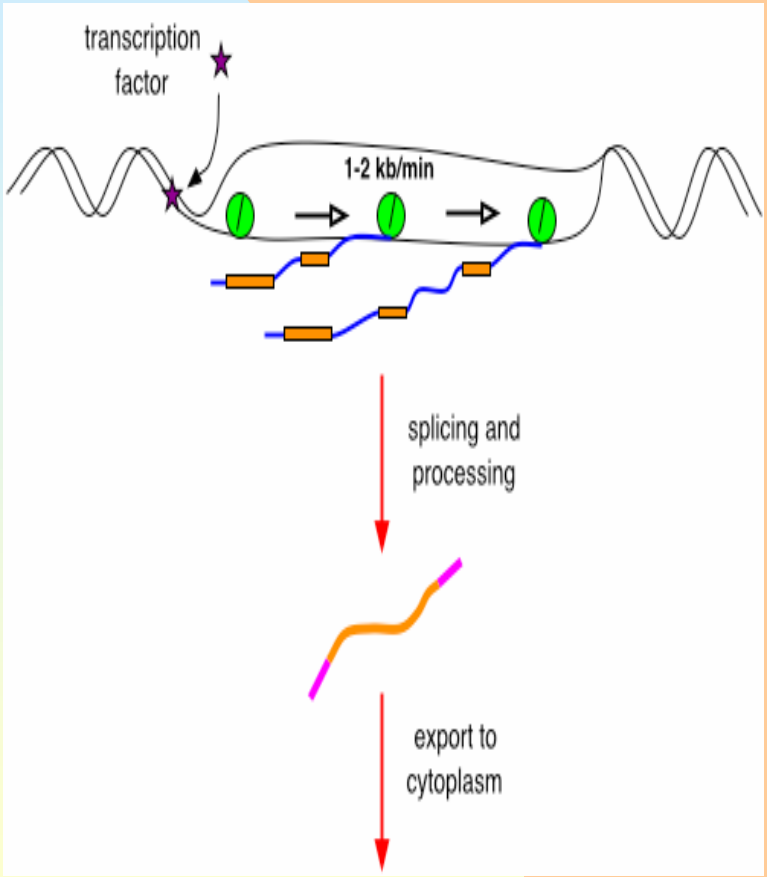
- Need effective stochastic simulation techniques that can move seamlessly across the 3 modelling regimes – multi-scale simulations.
- How do we represent external noise?
- When and how should we represent delays?
  - Should we lump transcriptional and translational delays?
  - Should we sample delays from a distribution? How do we estimate them from genomic information?
  - How do we use the Hill Function? – co-operativity;
  - Need a stochastic delay simulation algorithm.

Important Issues II

- How do we incorporate time and space into our models and simulations?
  - How do we represent anomalous diffusion?
  - Are there more effective approaches than just Monte-Carlo simulations and is there a difference between small and large numbers of molecules?
  - Should we have time dependent rate constants?
  - Can we represent spatial effects by temporal simulations?

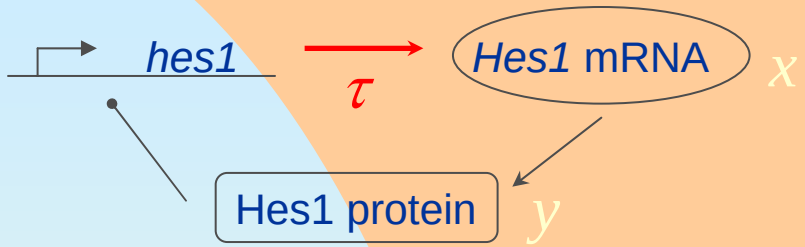
Modelling and Simulation of Delays in GRNs

Eukaryotic transcription and time delays



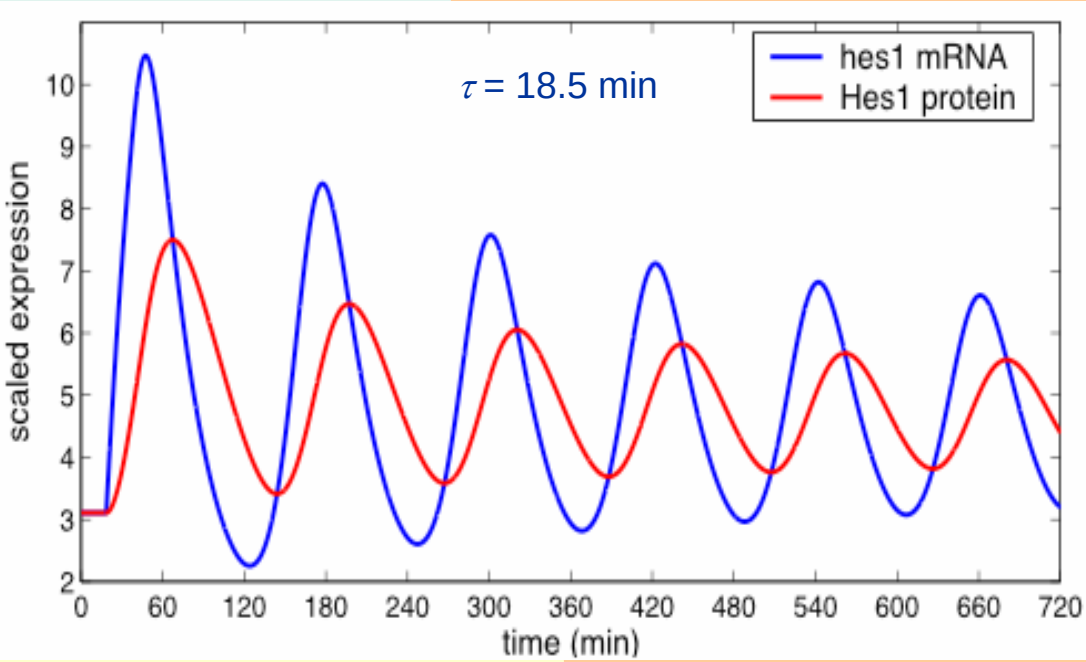
- There is an **irreducible** delay of ~15–20 min from initiation of a transcript to appearance of **functional mRNA** in the cytoplasm
- The delay can be much longer (>16 hrs for human *dystrophin*)
- **Delay equations** should be used to model transcription

Monk model for the Hes1 feedback loop



$$dx/dt = -\mu x(t) + pg[y(t-\tau)]$$

$$dy/dt = -\nu y(t) + qx(t)$$



g is a hill function with

$n > 4$ – co-operativity

$$g(y) = 1/\left(1 + \left(y(t-\tau)/y_0\right)^n\right)$$

The transcriptional delay has now been observed directly for Hes7

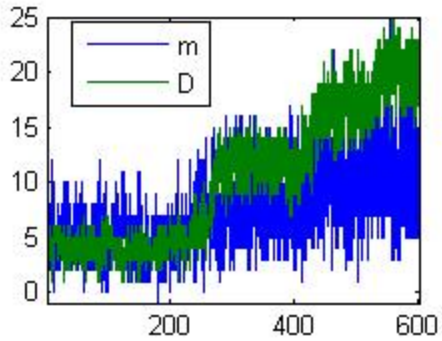
Bessho *et al. Genes & Dev.* 17, 1451 (2003).

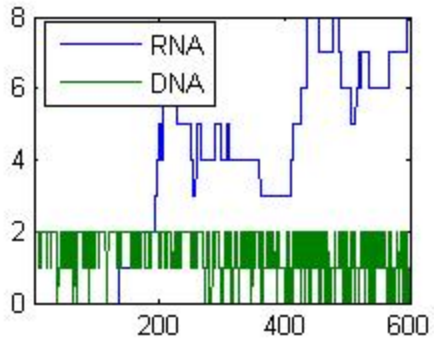
A regulatory model

- 10 reactions, 6 unknowns.
- *m* (monomer protein)
  D (dimer transcription factor)
  RNA (mRNA produced by transcription)
  DNA (free of dimers)
  DNA1 (bound by *D* at binding site R1)
  DNA2 (bound at sites R1 and R2).
- *X = (m, D, RNA, DNA, DNA1, DNA2).*
- *t_end = 600; X_initial=[2 6 0 2 0 0]';*
- *theta_C : [mC-aC,mC+aC] is the transcriptional delay*
- *theta_L: [mL-aL,mL+aL] is the translational delay*
- *facc=10; mC=12/facc; aC=1/facc; mL=2.8/facc; aL=1/facc;*

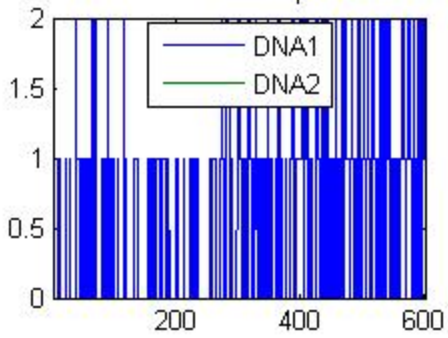
- Transcription to mRNA when *D* occupies R1;
  mRNA translated into proteins (both can decay);
  protein dimerises to transcription factor, *D;*
  binding of *D* at R1 activates transcription of *m;*
  binding of *D* at R2 excludes RNA polymerase from binding and transcription repressed.

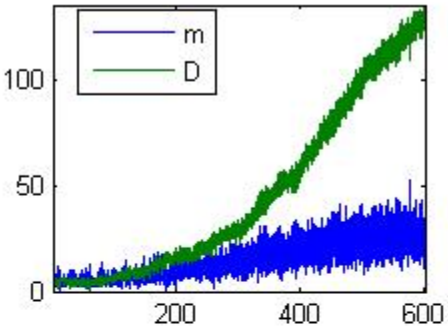
$$DNA \rightarrow RNA + DNA1 \quad - \quad a_1(X_5(t-\tau_1))$$
$$RNA \rightarrow RNA + m \quad - \quad a_2(X_3(t-\tau_2))$$
$$m \rightarrow 0$$
$$RNA \rightarrow 0$$
$$D \underset{\leftarrow}{\overset{\rightarrow}{}} m + m$$
$$DNA + D \underset{\leftarrow}{\overset{\rightarrow}{}} DNA1$$
$$DNA1 + D \underset{\leftarrow}{\overset{\rightarrow}{}} DNA2$$

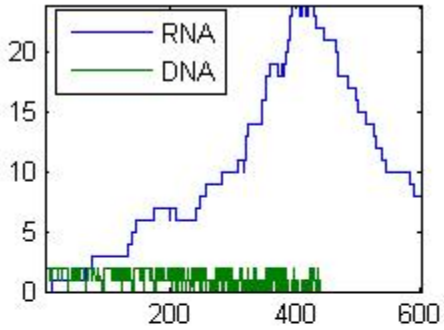




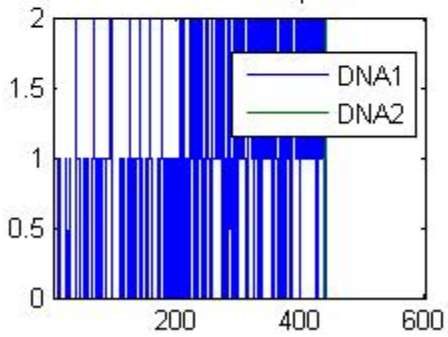
SSA with distributed transcription and translation delay, facc=0.1







SSA with distributed transcription and translation delay, facc=10



- Suppose delays associated with reactions 1 and 2 (transcription and translation)
- Delay SSA:
  - Reaction not updated until delay time has passed
  - Delays can be distributed.
- DDE model

$$\frac{dX}{dt} = \upsilon_1 a_1(X_5(t-\tau_1)) + \upsilon_2 a_2(X_3(t-\tau_2)) + \sum_{j>2}^{m} \upsilon_j a_j(X(t))$$
$$\equiv \mu\big(X(t), X(t-\tau_1), X(t-\tau_2)\big)$$

- SDDE model

$$dX = \mu\big(X(t), X(t-\tau_1), X(t-\tau_2)\big)dt + B\, dW$$
$$B^2 = \upsilon\, Diag\big(a_1(X_5(t-\tau_1)), a_2(X_3(t-\tau_2)), \cdots, a_m(X(t))\big)\upsilon^T$$

External Noise

- Assume external noise occurs in transcription
- Suppose this affects just the k^{th} reaction

$$\bar{\upsilon}_k = \begin{cases} \upsilon_k, & prob\ \theta \\ 0, & prob\ 1-\theta \end{cases}$$

- Modified SSA:
  - If reaction *k* selected, generate $u \in U(0,1)$
  - If $u \in [0,\theta)$ $\bar{\upsilon}_k = \upsilon_k$ $else$ $\bar{\upsilon}_k = 0$
- ODE Model

$$dX = \sum_{j \neq k} \upsilon_j a_j(X)dt + \theta \upsilon_k a_k(X)dt$$

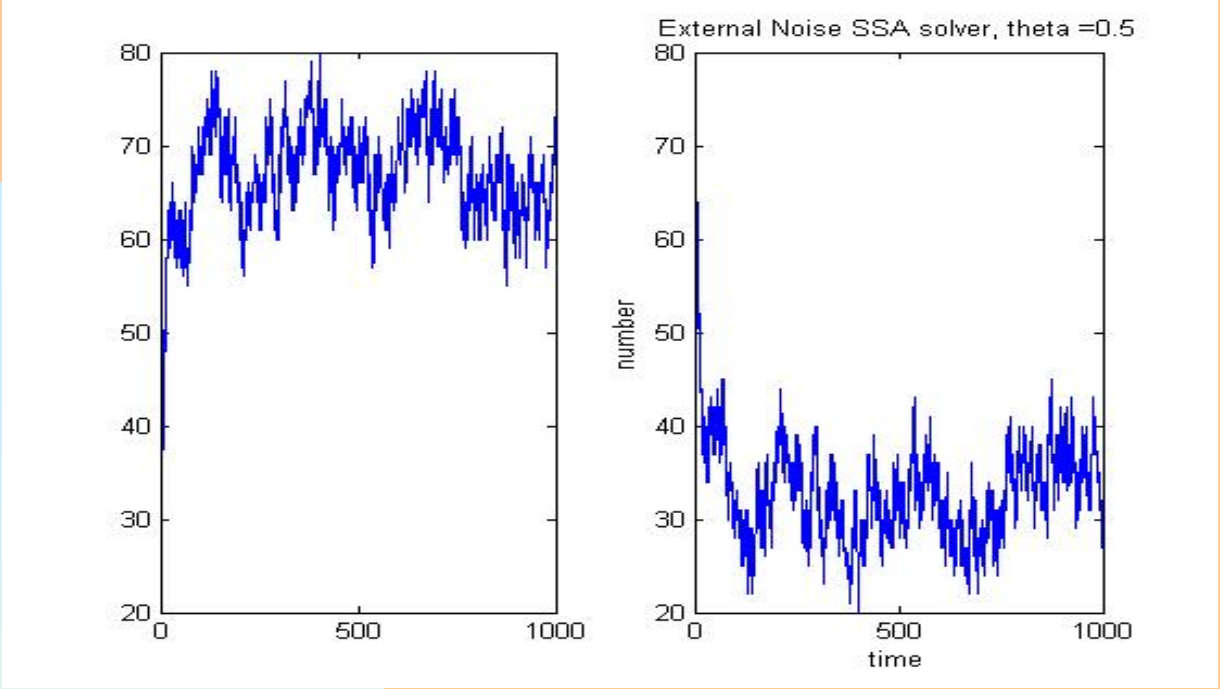
- Example:

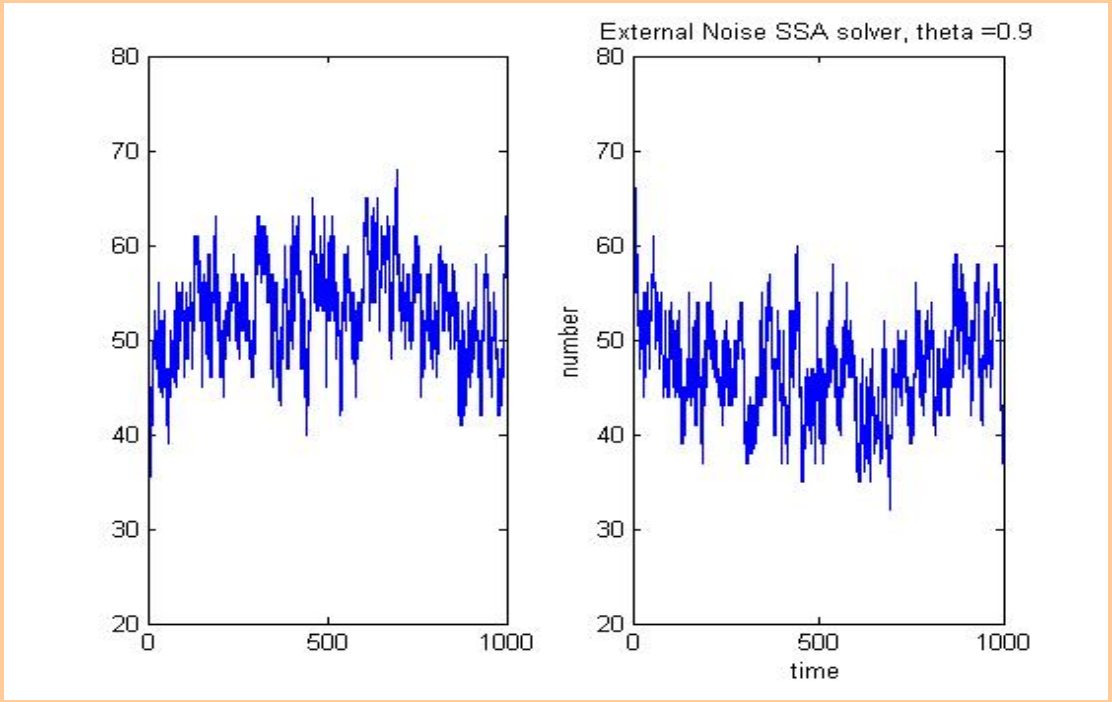
$$C \to S, \quad S \to C, \quad C(0)+S(0)=N, \quad k_1 = k_2$$

External Noise in first reaction (ODE)

$$C_{eq} = N\frac{1}{1+\theta}, \quad S_{eq} = N\frac{\theta}{1+\theta}$$

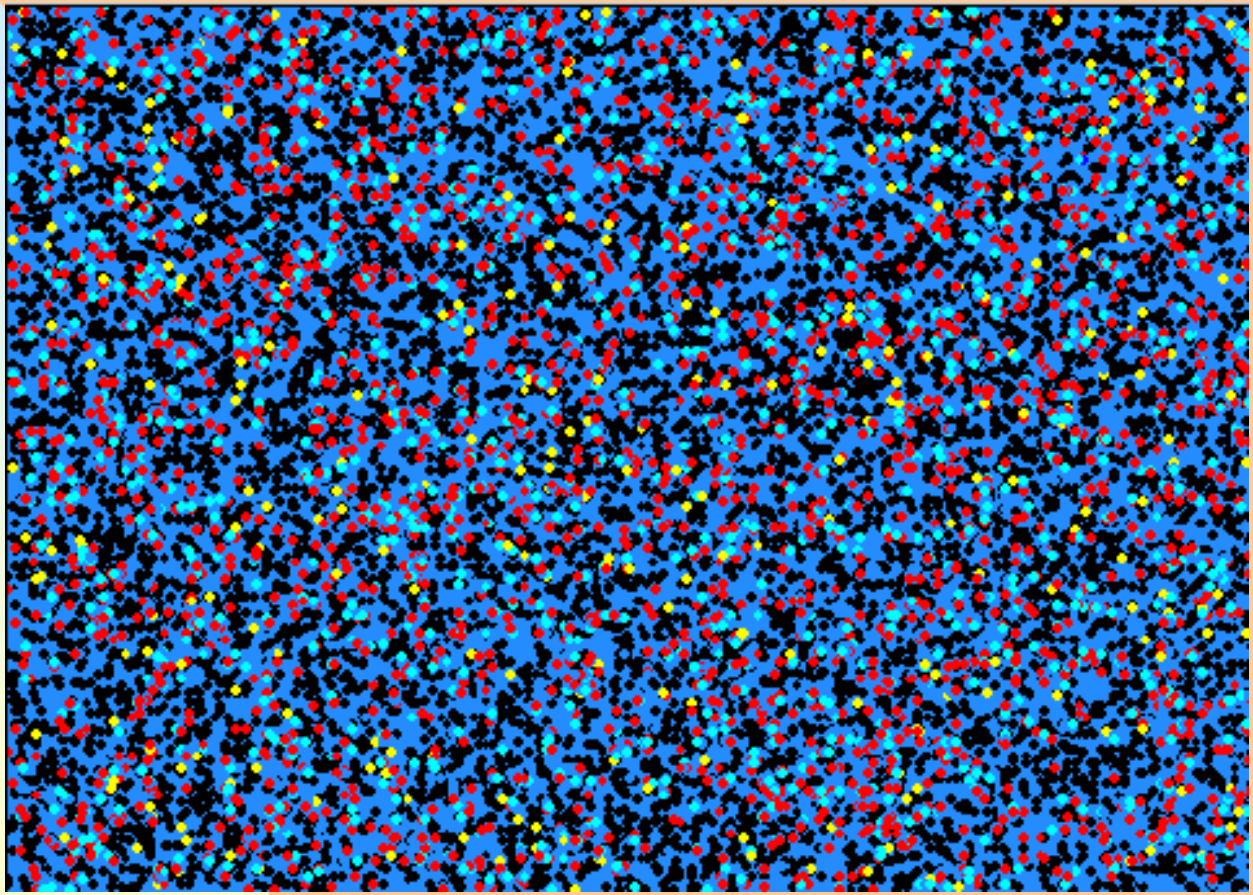
$$N = 100, \quad \theta = \tfrac{1}{2} \quad (66), \quad \theta = 0.9 \quad (52)$$


External Noise SSA solver, theta =0.5
80
70
60
50
40
30
20
0
500
1000
number
time


External Noise SSA solver, theta =0.9
80
70
60
50
40
30
20
0
500
1000
number
time

Space and Time

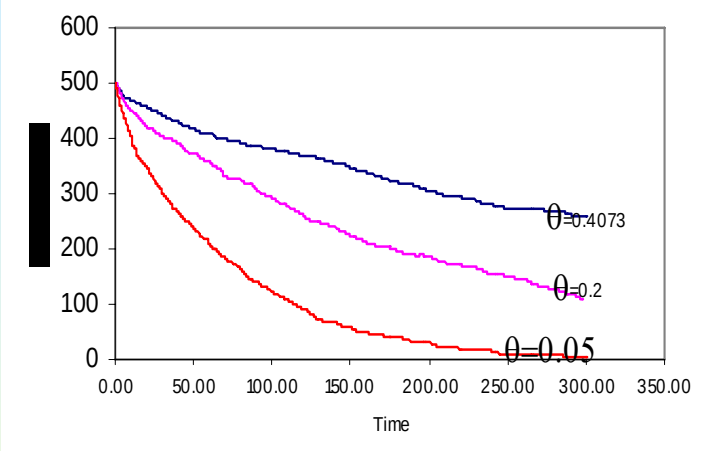
Kinetics with obstacles

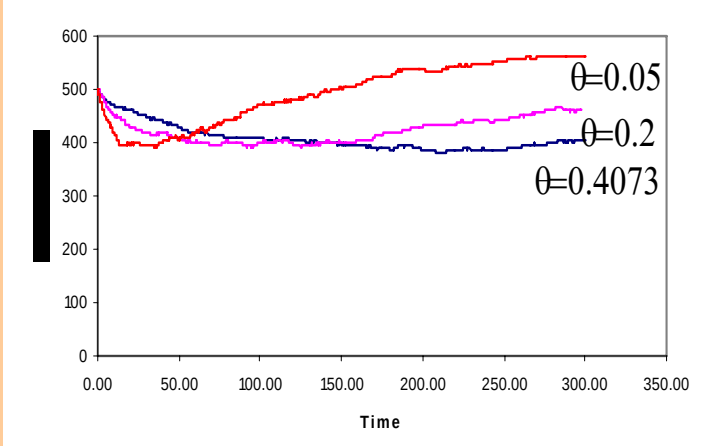


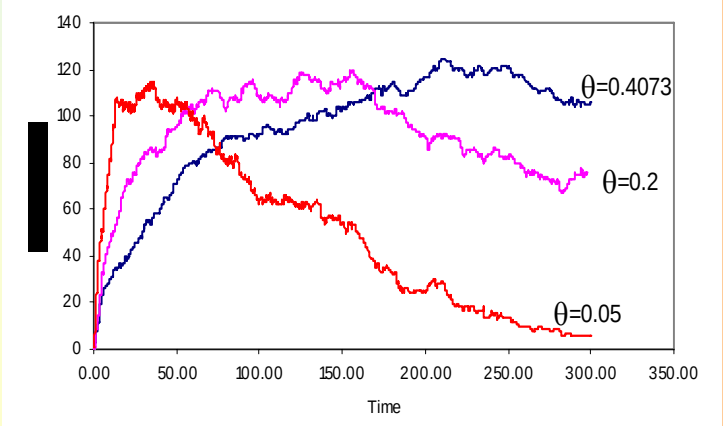
Black: obstacles; Blue: product; Yellow: complex; Red: substrate

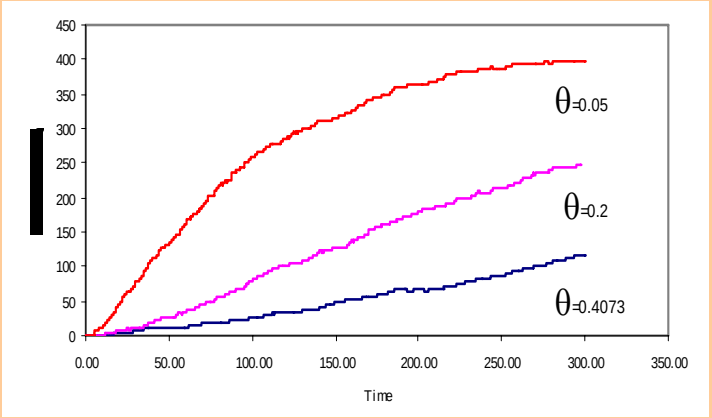
Kinetics with and without obstacles

$$E+S \leftrightarrow C \leftrightarrow E+P$$









Clockwise from top: Substrate, Enzyme, Complex, Product

Biological issues

- Percolation thresholds (obstacle density: 2D = 40%).
- Relate obstacle density to anomalous parameter.
- Measurements

  Diffusion of proteins in nucleus α =0.87

  Percolation threshold α =.54

Density is a long way from the threshold. But

  Diffusion on a 2D cell membrane α =0.74

  Percolation threshold α =0.69

- Kopelman (1986)

  Reactions $A + B \rightarrow 0$ in crowded environments (crystalline alloys) exhibit fractal-like kinetics $k(t) \approx t^{-\alpha}$

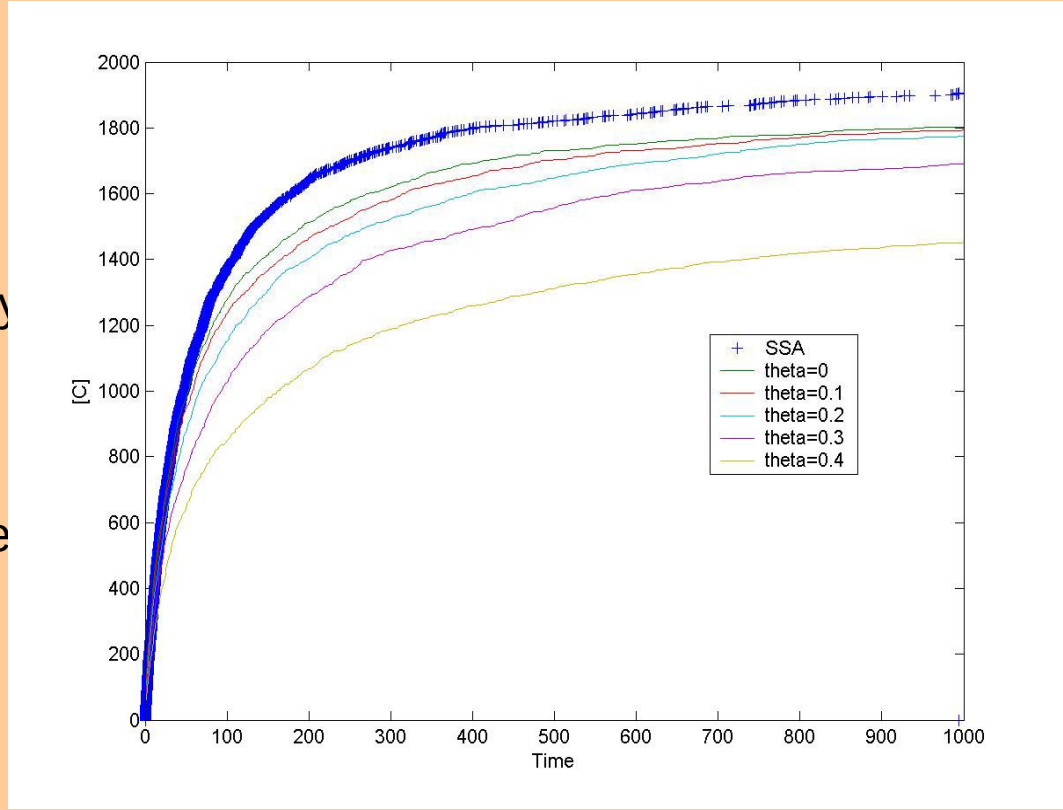
Spatial Simulation

- SSA ignores space and assumes the mixture is well stirred.

- Inhomogeneity, compartments, anomalous diffusion are hallmarks of biological systems, SSA applicability to cellular environments is questioned

- We simulated (by Monte Carlo) the evolution of A+B→C on 2D membrane in the presence of fixed obstacles.

- Initial conditions are A(0)=2000, B(0)=2000 and the system is relatively well mixed (D=5% of membrane size).



- Even if the obstacle concentration is 0, we do not observe perfect agreement between SSA and spatial simulation.
- As the obstacle concentration increases, the departure of SSA from Monte Carlo can become significant.

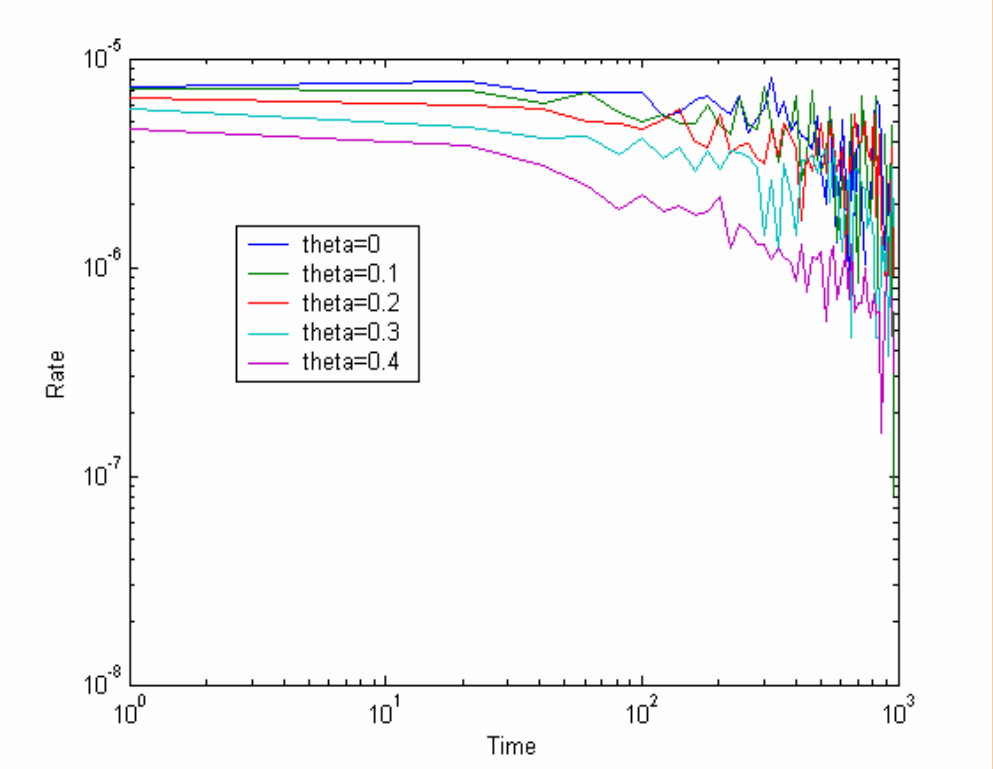
Anomalous SSA

- Non-constancy of the reaction rate at long reaction times.

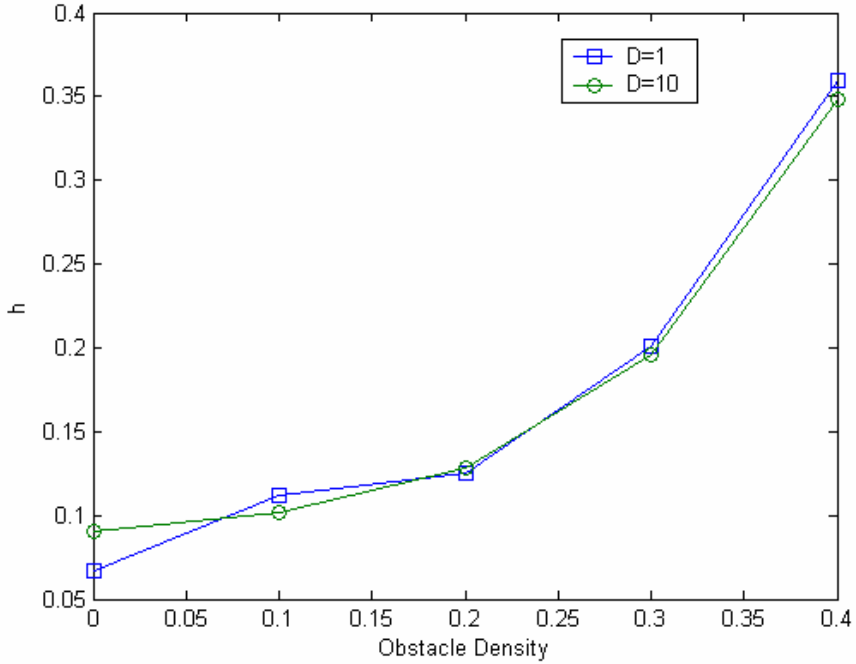
- Consider a power-law dependence of the rate on time T $k(t)=k(0)T^{-h}$ and sample from the “anomalous” waiting time (t) exponential distribution

$$w(t) = a(T)\exp\{-a(T)t\}$$

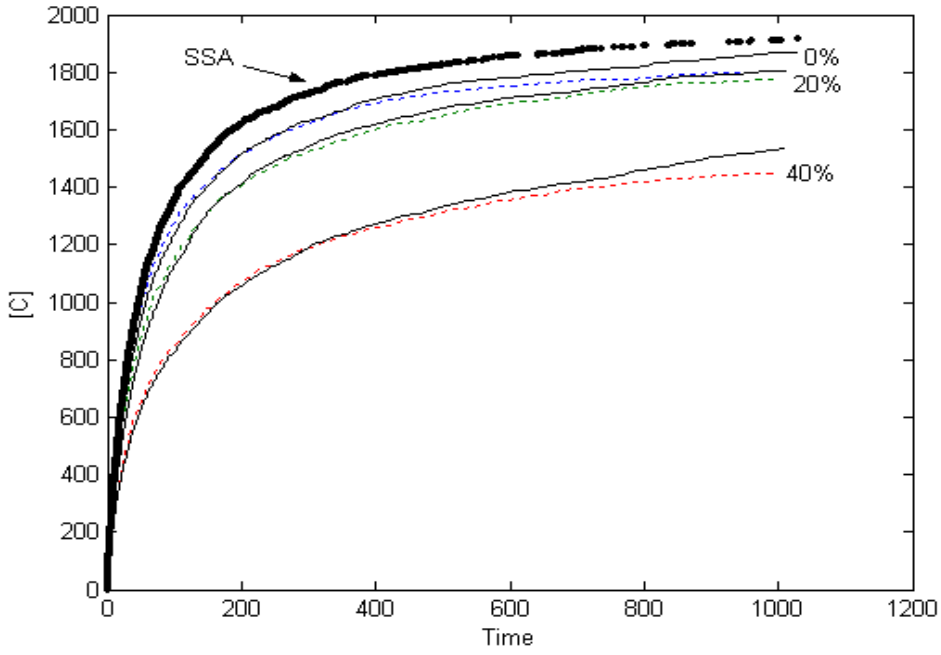
a(T)=k(0)A(T)B(T)T^(-h)



- Estimate *h* empirically from MC simulations at known obstacle concentrations and diffusion rates.

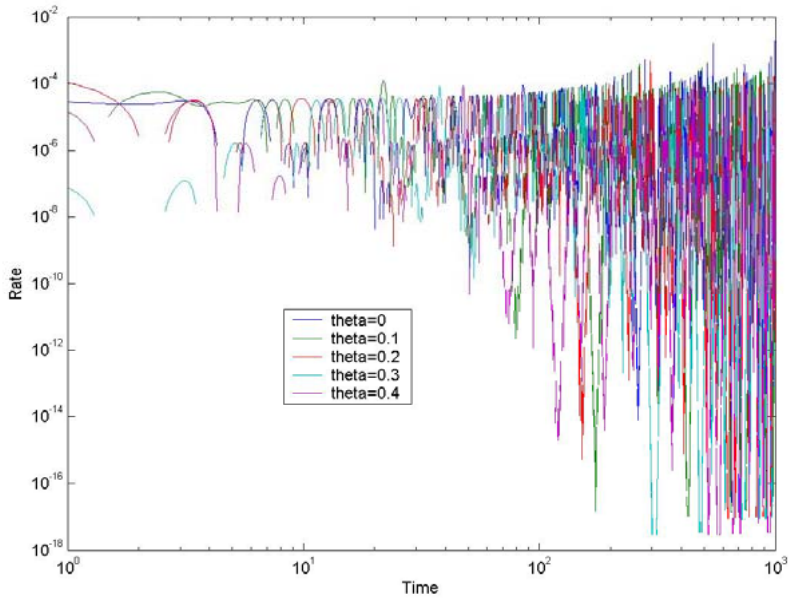


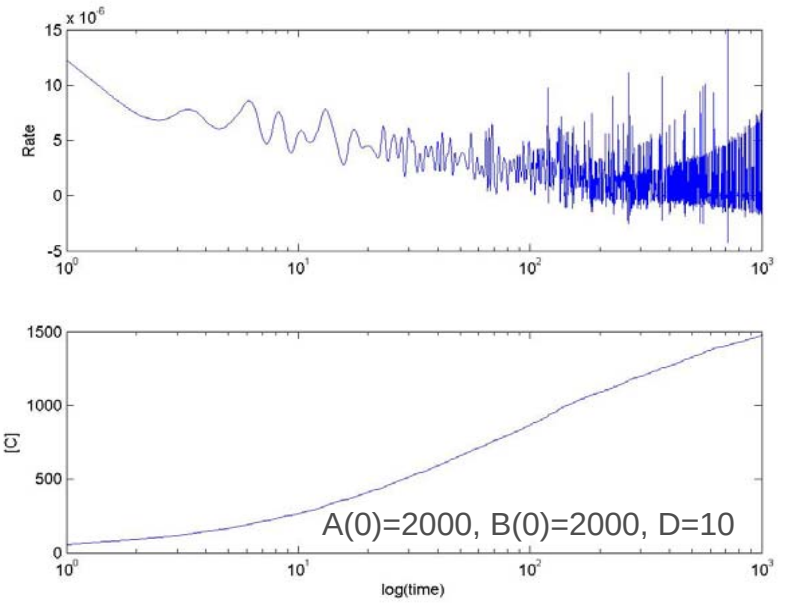
- Anomalous SSA meets with good success at different obstacle densities.



Rate Breakdown -Low Molecule with Obstacles

- As molecular numbers become smaller, the concept of a rate begins to break down due to spatial effects.
- More pronounced if there are obstacles.
- Even if diffusive jumps are large, the rate is noisy for most of the simulation time.
- The rate can behave stochastically even while the concentrations of reactants appear relatively smooth.





Challenges

- Temporal Simulation methods that move seamlessly between three modelling regimes.
- A multi-scaled approaches for hybrid systems – E.g. Burrage, Tian & Burrage – Progress in Biophysics and Molecular Biology, 2004.
- Role of delay and discrete models in GRN.
- A rigorous mathematical and simulation framework for chemical kinetics and anomalous diffusion in crowded environments.
- Convergence issues in time and space.
- Efficient methods for Fractional DEs.

Thank you