

Event based modeling & Multilevel modeling in CA

Course Bioinformatic Processes 2018/2019; Paulien Hogeweg;
Theoretical Biology and Bioinformatics Grp Utrecht University

modeling formalisms (last)

LAST TIME:

Network models: Gene regulation

themes

Multiple attractors, domain of attraction,
divergent trajectories

(double) knockouts and “buffering”

Multiple models for predefined behaviour (cf robustness):
no reversed engineering; not simplest one ‘best’ (evolved)

TODAY:

- Event based models
data intensive modeling of specific systems
- Multilevel modeling: predefined multiple levels.
counterintuitive behavior
debugging experimental inferences
debugging models

EVENT based models: continuous time, discrete events

Gillespie algorithm

1: seen als stochastic ODE

Example: logistic stochastic population growth

$$dN/dt = aN - bN^2 + \text{noise}$$

EVENT based

all events (birth + death) :

$$e_0 = (a_1 + a_2)N - b_1 N^2 + b_2 N^2$$

$$\tau = 1/e_0 \ln(1/\text{rand1}); T = T + \tau$$

N=N+1 if $(a_1 N - b_1 N^2) < \text{rand2} * e_0$

else N=N-1;

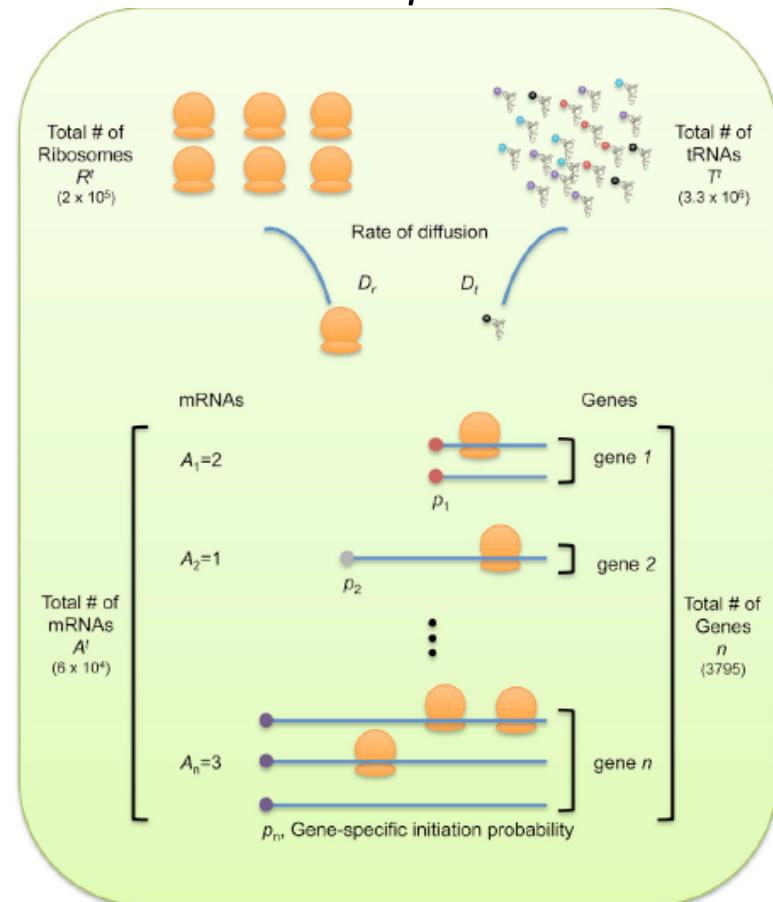
EVENT based models: continuous time, discrete events

Gillespie algorithm

seen as multi-entity - multistate decomposition

Example

Rate-Limiting Steps in
Yeast Protein Translation
Premal Shah et al Cell 2013



Data, states, events

DATA

- fasta file of yeast mRNA + number of mol/cell
- yeast tRNA's (41) + number in cell + wobble
- number of ribosomes
- initiation prob of all mRNA types
- size of ribosome/tRNA's yeast cell
- diffusion constant ribosomes, tRNA's
- -- > characteristic times

STATES

- number of free ribosomes/tRNA's(of every type)
- Position of each bound ribosomes/tRNA's on each individual mRNA

EVENTS

- Initiation (binding of ribosome at free 5'end of mRNA)
- Elongation (change position, free - bind tRNA)

Yeast data on cell content

Table 1. Summary of Model Parameters

Parameter	Description	Value or Range of Values	References
R^t	number of ribosomes	2×10^5	(Warner, 1999; von der Haar, 2008)
A^t	number of mRNAs	6×10^4	(Zenklusen et al., 2008)
T^t	number of tRNAs	3.3×10^6	(Waldron and Lacroute, 1975)
T_n	number of types of tRNAs	41	(Chan and Lowe, 2009)
T_j	number of tRNAs of type j	$\sim 12,000\text{--}190,000$	(Chan and Lowe, 2009)
A_i	number of mRNAs of type i	1–1,254	(Ingolia et al., 2009)
p_i	gene-specific initiation probability	$\sim 3.5 \times 10^{-6}\text{--}0.115$	(Experimental Procedures)
n	number of genes	3,795	(Ingolia et al., 2009)
D_r	diffusion coefficient of ribosomes	$3 \times 10^{-13} \text{ m}^2/\text{s}$	(Politz et al., 2003)
D_t	diffusion coefficient of tRNAs	$8.42 \times 10^{-11} \text{ m}^2/\text{s}$	(Werner, 2011)
C_r	size of ribosome footprint in codons	10	(Ingolia et al., 2009)
s	tRNA competition coefficient	7.78×10^{-4}	(Experimental Procedures)
V	volume of the cell	$4.2 \times 10^{-17} \text{ m}^3$	(Siwiak and Zielenkiewicz, 2010)

Algorithm (pseudocode)

while $time < t$ (*total simulation time*) **do**

 Calculate

 Fraction of mRNAs of gene i that are *initiable*, f_i - i.e., those mRNAs with first 10 codons unbound.

 Number of *elongatable* ribosomes waiting at codon j , $R^b(j)$ - ribosomes with next 10 codons unbound.

 Rates of all possible events (see Table S2)

$$\text{Total initiation rate: } \rho^t = \sum_{i=1}^n \frac{R^f f_i A_i p_i}{\tau_r N_r}$$

$$\text{Total elongation rate: } \epsilon^t = \sum_{j=1}^{61} \frac{R^b(j) T_{\phi(j)}^f w_j s}{\tau_t N_t}$$

 Probability of each possible event (see Table S2)

 Randomly select an event based on its probability of occurrence (see Table S2)

 Update the changes in the state of the cell (see $\Delta State$ in Table S2)

 Increment $time$ by $\frac{1}{\rho^t + \epsilon^t}$

 Update the number of free ribosomes, R^f

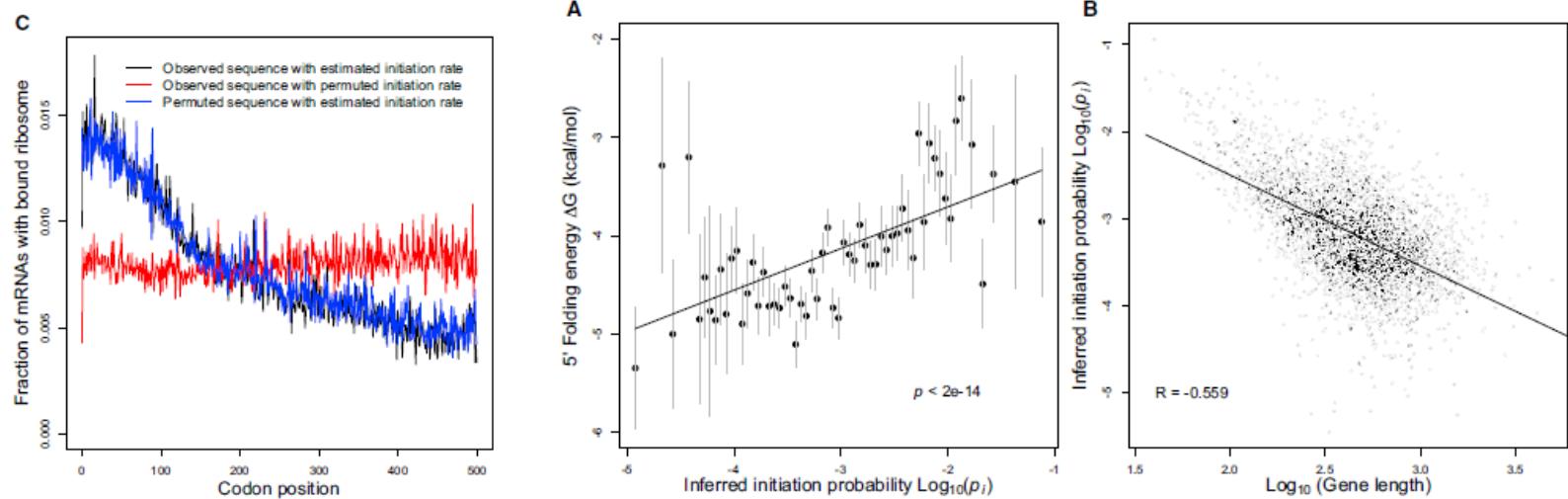
 Update the number of free tRNAs of type $\phi(j)$, $T_{\phi(j)}^f$

end

Table S2. Markov States and Transition Rates, Related to Figure 1

Initiation	Gene #	mRNA # of gene			Initiation rate	Event probability	Δ State
	1	1			0	0	N/A
	1	2			0	0	N/A
	1	3			$\frac{R^f p_1}{\tau_r N_r}$	$\frac{R^f p_1}{\tau_r N_r (\rho^t + \epsilon^t)}$	$R^f \rightarrow R^f - 1$
	1	A_1			$\frac{R^f p_1}{\tau_r N_r}$	$\frac{R^f p_1}{\tau_r N_r (\rho^t + \epsilon^t)}$	$R^f \rightarrow R^f - 1$
	2	1			$\frac{R^f p_2}{\tau_r N_r}$	$\frac{R^f p_2}{\tau_r N_r (\rho^t + \epsilon^t)}$	$R^f \rightarrow R^f - 1$
	2	2			$\frac{R^f p_2}{\tau_r N_r}$	$\frac{R^f p_2}{\tau_r N_r (\rho^t + \epsilon^t)}$	$R^f \rightarrow R^f - 1$
2
2	A_2				0	0	N/A
.
n
	n	A_n			$\frac{R^f p_n}{\tau_r N_r}$	$\frac{R^f p_n}{\tau_r N_r (\rho^t + \epsilon^t)}$	$R^f \rightarrow R^f - 1$
			Total initiation rate	$\rho^t = \sum_{i=1}^n \frac{R^f f_i A_i p_i}{\tau_r N_r}$			
Elongation	Gene #	mRNA # of gene	Codon position	Ribosome bound	Elongation rate	Event probability	Δ State
	1	1	1	N	0	0	N/A
	1	1	2	Y	$\frac{T^f_{\phi(1,2)W_{1,2}S}}{\tau_t N_t}$	$\frac{T^f_{\phi(1,2)W_{1,2}S}}{\tau_t N_t (\rho^t + \epsilon^t)}$	Ribosome bound at codon 2 → N Ribosome bound at codon 3 → Y
	1	1	3	N	0	0	N/A
	1	1
	1	1	L_1	N	0	0	N/A
	1	2	1	Y	0	0	N/A
	1	2
	1	2	11	Y	$\frac{T^f_{\phi(1,11)W_{1,11}S}}{\tau_t N_t}$	$\frac{T^f_{\phi(1,11)W_{1,11}S}}{\tau_t N_t (\rho^t + \epsilon^t)}$	Ribosome bound at codon 11 → N Ribosome bound at codon 12 → Y
	1	2
	1	2	L_1	Y	$\frac{T^f_{\phi(1,L_1)W_{1,L_1}S}}{\tau_t N_t}$	$\frac{T^f_{\phi(1,L_1)W_{1,L_1}S}}{\tau_t N_t (\rho^t + \epsilon^t)}$	$R^f \rightarrow R^f + 1$
	1
	1	A_1	L_1
	2
	2	A_2	L_2
.
n	A_n	L_n
			Total elongation rate	$\epsilon^t = \sum_{j=1}^{61} \frac{R^b(j) T^f_{\phi(j)W_j S}}{\tau_t N_t}$			

Is protein production initiation or elongation limited in exponential growing yeast populations?



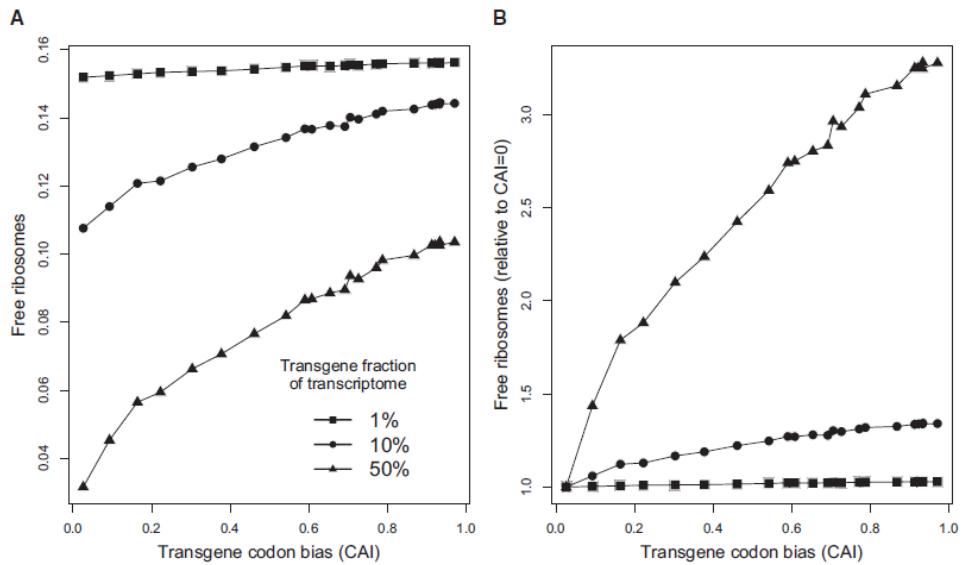
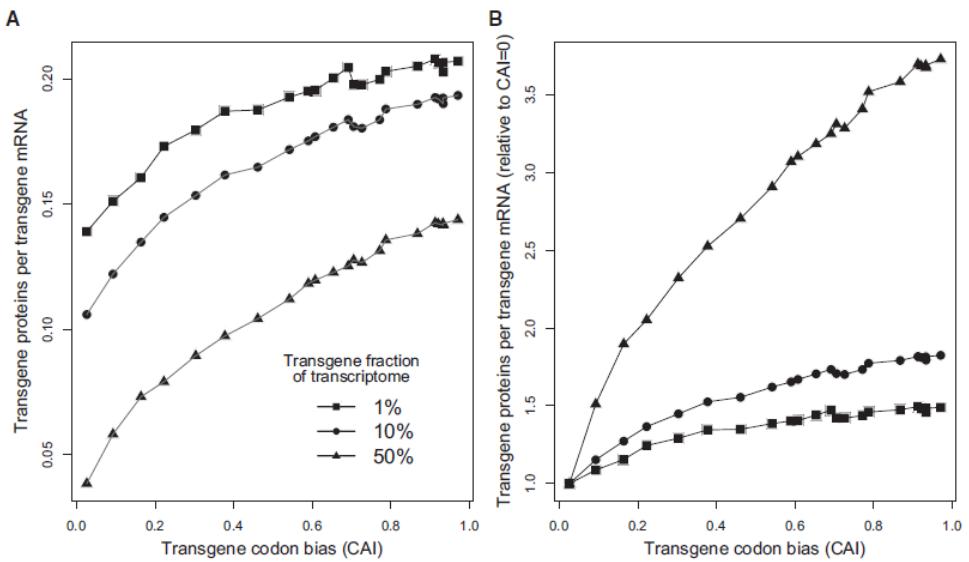
more ribosomes at 5' end BUT due to >> initiation prob. on short genes

initiation limited

debugging of wrong inference from exp. data

Optimizing codon usage
of transgenes
only useful
at very high dosage

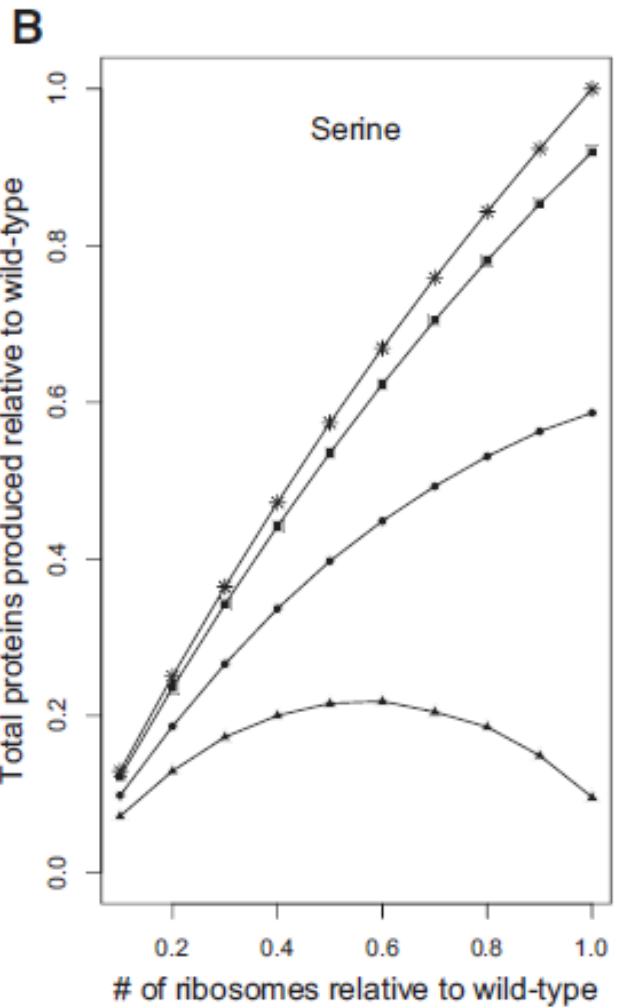
due to freeing of ribosor



Under amino-acid starvation down regulating ribosomes can increase protein production

*because translation becomes
elongation limited
reducing Ribosomes increases
free tRNA'*

- * Wild-type amino-acid abundance
- 2-fold amino-acid starvation
- 5-fold amino-acid starvation
- ▲- 10-fold amino-acid starvation



conclusions event based modeling of stochastic reaction kinetics

Data intensive modeling

Quantitative conclusions

Upscaling to “whole cell modeling” feasible

But note simplifications:

space but no spatial structure

fixed number of molecules

fixed conditions

....

McGuffee & Elcock 2010

Molecular packing within E coli cell

macromolecules in the system (85% of mass)

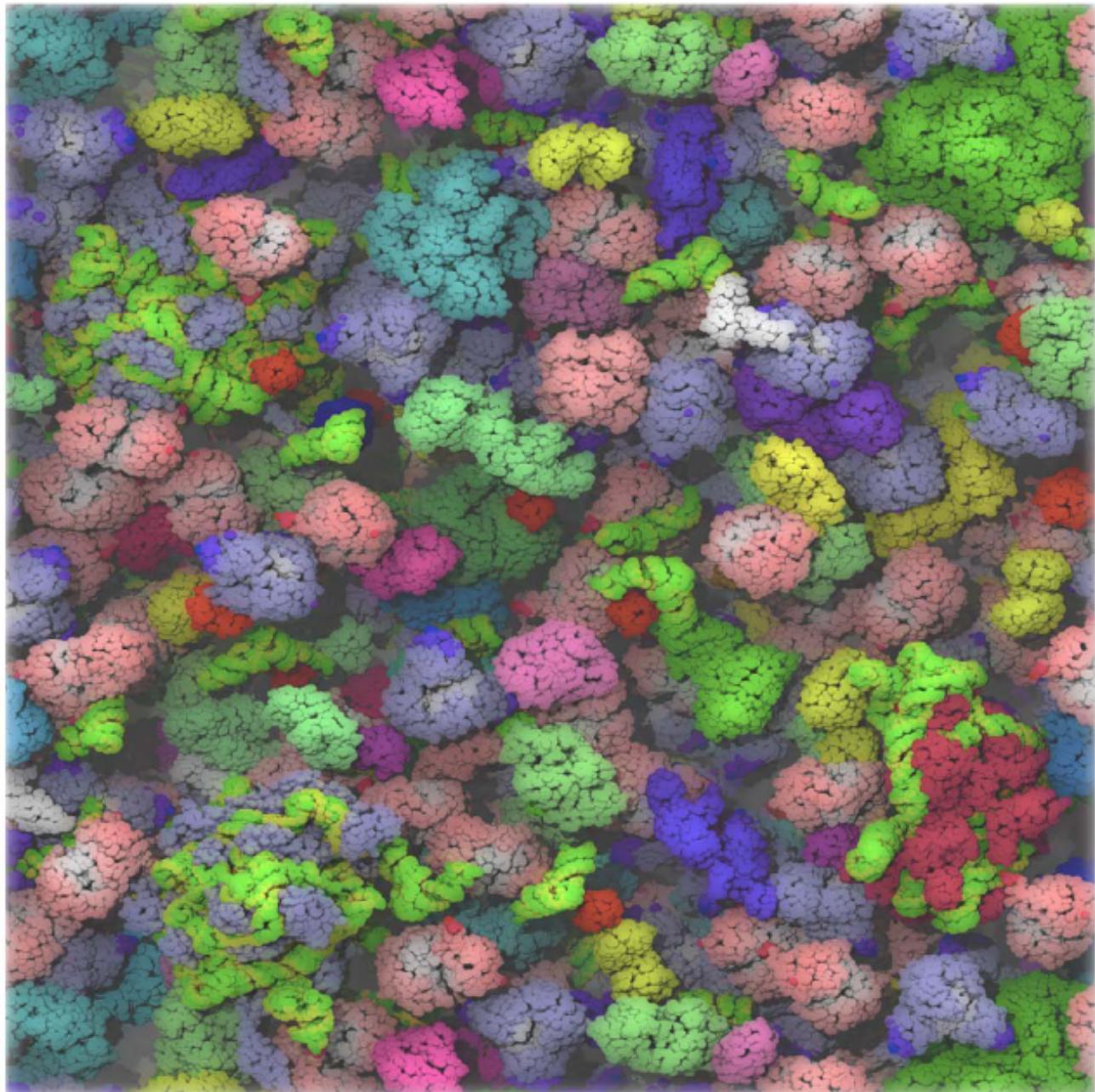
A

	Name	Mw	#
	Adk	24	14
	AhpC	187	7
	Asd	80	4
	Bcp	11	8
	CspC	7	72
	CysK	64	13
	DapA	125	2
	DnaK	41	11
	Efp	20	14
	Eno	91	18
	Fba	78	6
	Frr	21	7
	FusA	69	22

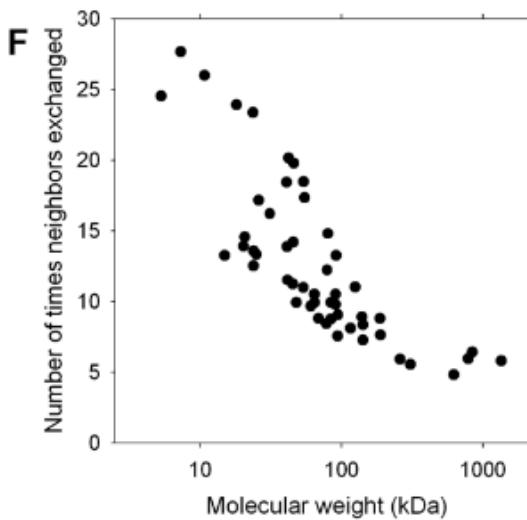
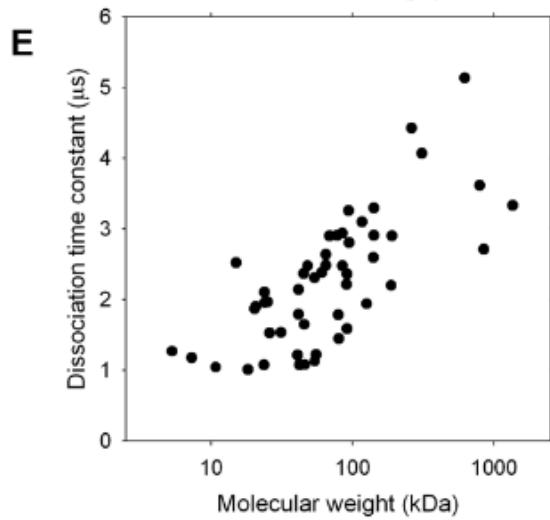
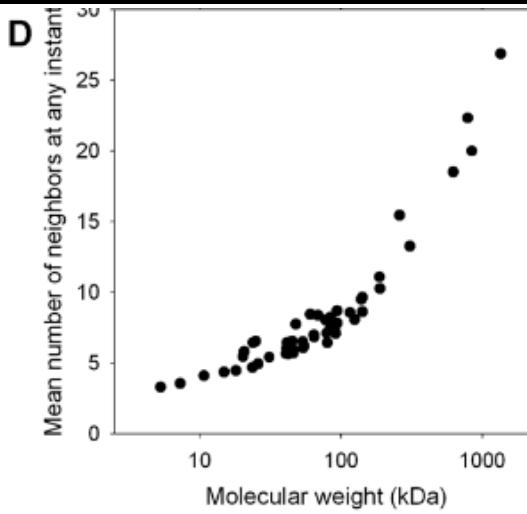
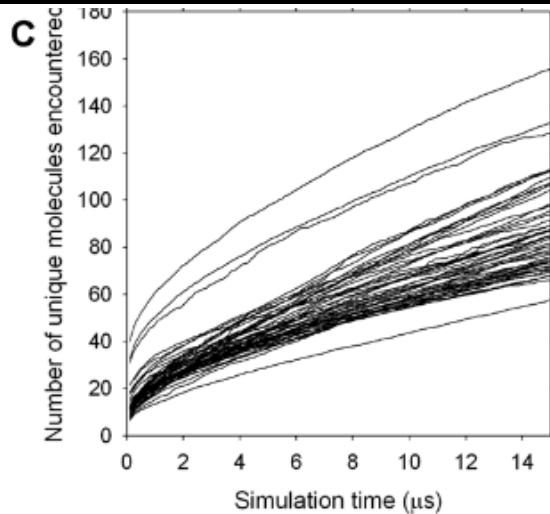
	Name	Mw	#
	GapA	142	10
	GlnA	621	1
	GltD	94	3
	GlyA	91	15
	GpmA	55	4
	Hns	5	7
	Hup	15	12
	IcdA	92	43
	IlvC	54	18
	Mdh	65	13
	MetE	84	213
	Mop	845	2

	Name	Mw	#
	PanB	140	2
	Pgk	41	26
	Pnp	190	3
	Ppa	116	9
	PpiB	18	7
	PurA	94	4
	PurC	42	7
	Pyr	308	3
	RpiA	46	3
	Rpo	260	4
	SerC	79	11
	SodA	46	13
	SodB	42	9

	Name	Mw	#
	Suc	142	4
	Tig	48	9
	TpiA	54	5
	Tsf	61	12
	TufA	84	181
	Upp	45	11
	UspA	31	7
	50S	1,355	10
	30S	788	10
	tRNA-C	24	37
	tRNA-Q	24	37
	tRNA-F	25	37
	GFP	26	8



Who meets whom? $15\mu s$



conclusions

It is crowded!

Nevertheless fairly much movement

'diffusion' ca 10% from diluted environment

changes neighbours frequently even in $15\mu s$

stabilization/destabilization dependent on protein

..... many observations possible

spectacular that it is possible

but 1 year real simulation time /run of $20\mu s$

modeling biotic systems as multilevel systems

Previously:

EMERGENT MESOSCALE ENTITIES:

- discovery and description
- modeling these entities
- variable number of 'entities',
- mean field approximation

Now:

PREDEFINED MULTIPLE LEVEL

- e.g. predefined cells as mesoscale
- multiple timescales of information transfer
- multiple scales of interaction

example of cell movement

How to represent a cell?

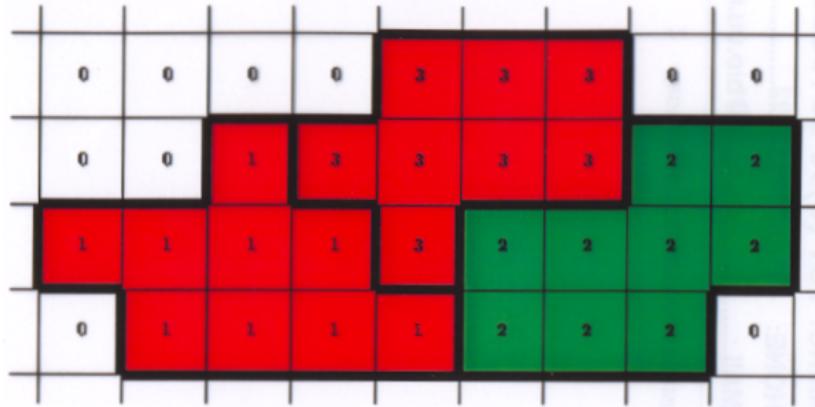
cell basic unit in single celled and multi-cellular organisms

- cell as a dimensionless point: PDE
- cell as occupation of a patch of space: CA
NB particle conservation!
- cell as a “homunculus” IBM
- cell as a ball being moved by external forces (finite element models)
- **Cells are deformable highly viscous objects, behaviour determined by internal state (gen expression) and external interactions operating in subcellular scale**

How to model? Multilevel model formalism (CPM)

Glazier-Graner 2-Scale CA model

CPM



- A 'biotic' cell consists of many lattice sites in same 'state' (= cell identity)
- Cells have a type τ , volume v (and...)
- Between cells: free energy $bod J_{ij}$ where i and j are the types of the cells
- *dynamics*: Free energy minimization with volume conservation:

$$H = \sum \frac{J_{ij}}{2} + \sum J_{im} + \lambda(v - V)^2$$

- Copy state of neighbouring cell with probability:

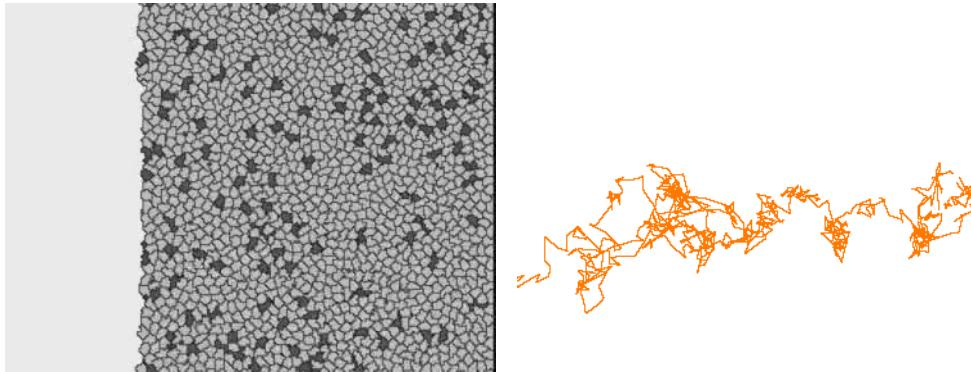
$$P_{i \rightarrow j} = 1 \text{ iff } \Delta H < -\beta ; \quad P_{i \rightarrow j} = e^{-(\Delta H + \beta)/M} \text{ iff } \Delta H \geq -\beta$$

	
Initial configuration	
Cell Sorting $J_{white,white} = J_{grey,grey} < J_{white,grey}$	
Cell Mixing $J_{white,white} = J_{grey,grey} > J_{white,grey}$	
Engulfment $J_{white,grey} < J_{white,medium}$ $J_{grey,medium} < J_{white,medium}$	
No cell cell adhesion $J_{cell,cell} > 2J_{cell,medium}$	

Table 2.2: A list of cell sorting behaviours in the Glazier and Graner model

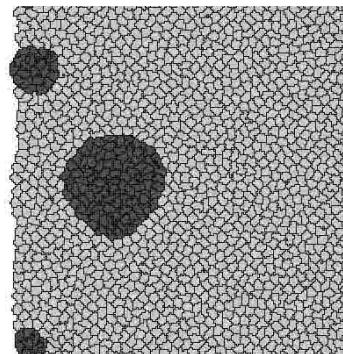
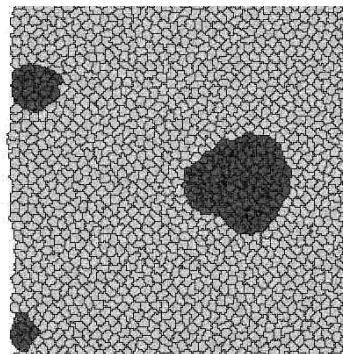
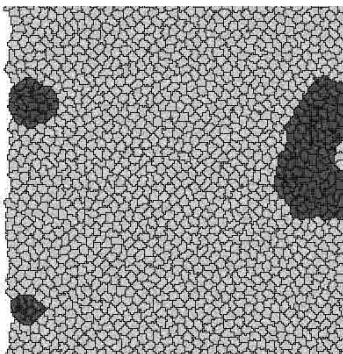
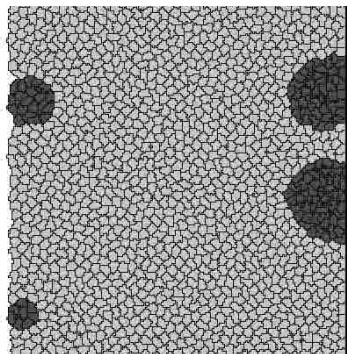
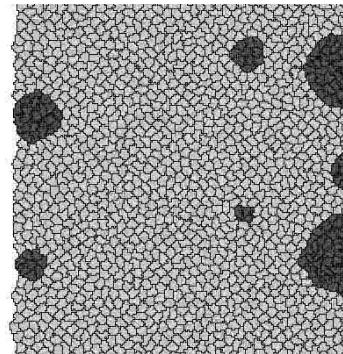
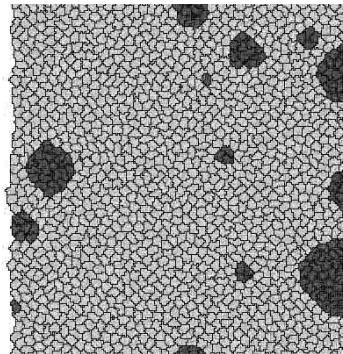
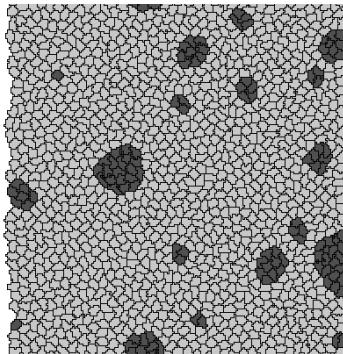
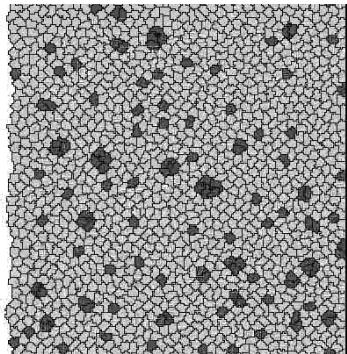
from cells to tissues (and beyond)

- cell sorting by differential adhesion
- Individual cells 'wiggle' through cell mass



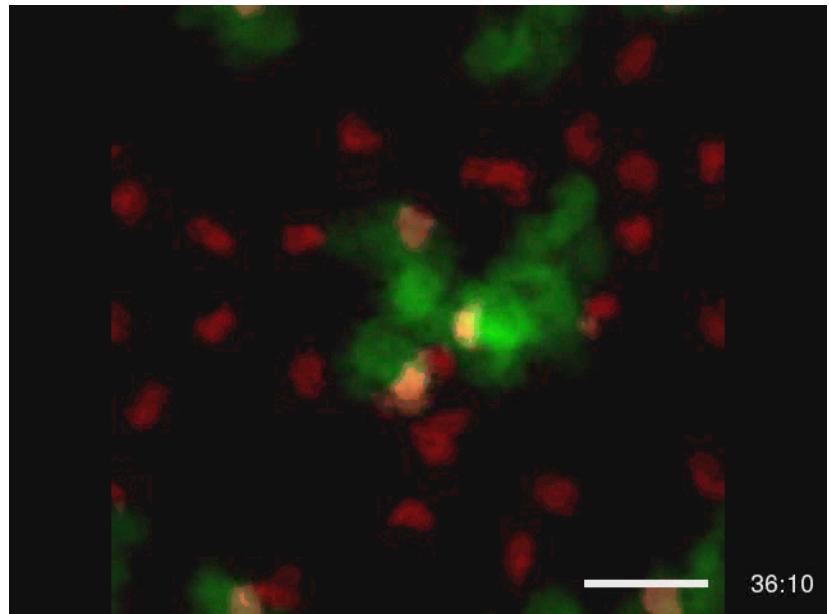
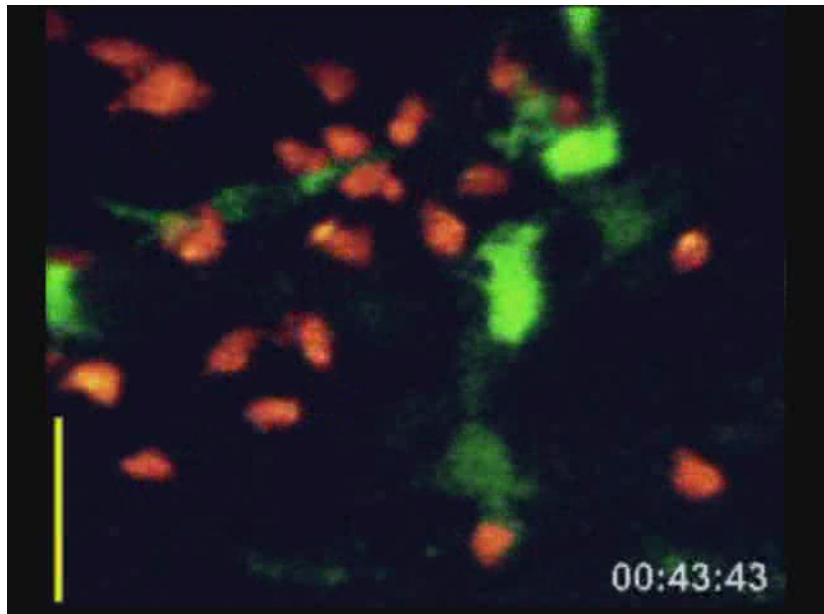
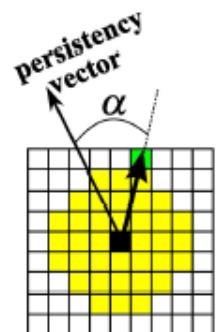
- Individual cells can 'move against the flow'
e.g. by being larger; being in the minority, adhesion

**change direction without changing
cell or environmental properties**



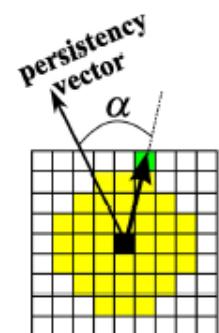
--> chenotaxis

Cell movement in Lymphnode: stop and go (Beltman et al 2007)

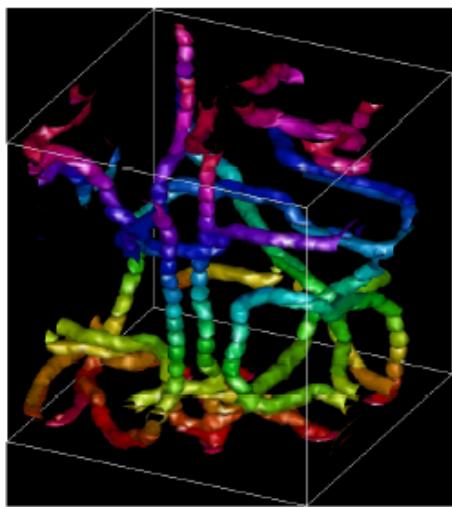


Cell movement in empty Lymphnode

Beltman et al 2007

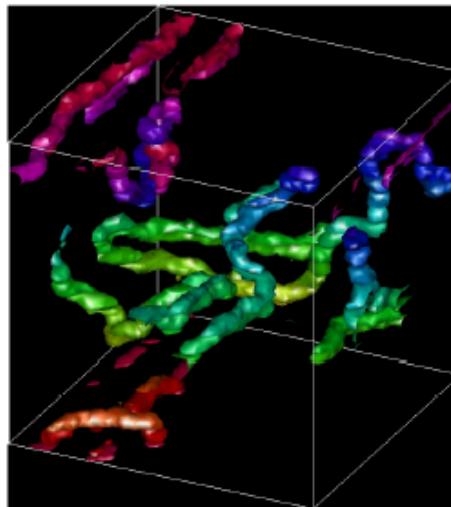


Cell track of individual lone cell

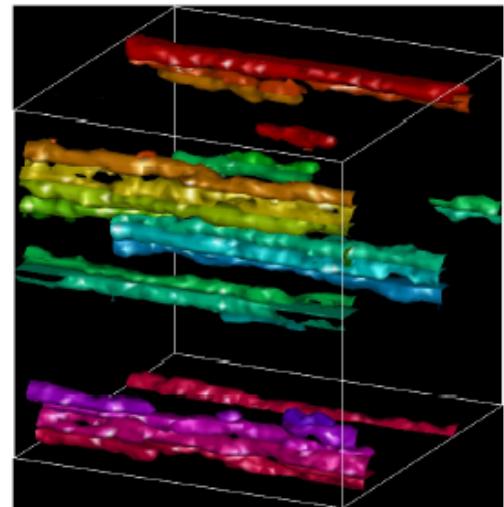


b

Cell track of 5 cells in an environment with many cells



c



d

Chemotaxis in lymphnodes?

beware of modeling artifacts

Weak chemotaxis hard to detect by cell tracking. Augment by modeling: what would the effect of chemotaxis be?

2 recent models: opposite conclusions

Riggs et al JTB 2008: "A comparison of random vs. chemotaxis-driven contacts of T cells with dendritic cells during repertoire scanning".

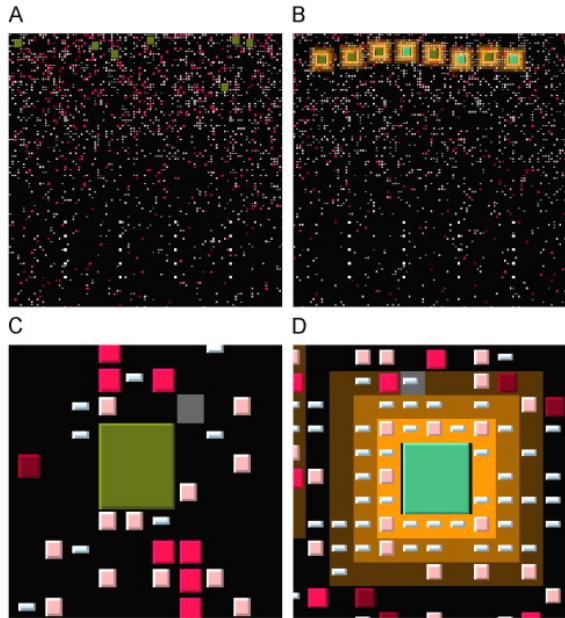
Our [CA modeling results] suggest that, within a LN T-zone, a random search strategy is optimal for a rare cognate T cell to find its DC match and maximize production of activated T cells "

Vroomans et al 2012: "Chemotactic Migration of T Cells towards Dendritic Cells Promotes the Detection of Rare Antigens"

Our [CPM] simulations show that chemoattraction of T cells enhances the DC scanning efficiency, leading to an increased probability that rare antigen-specific T cells find DCs carrying cognate antigen.

Models incorporate very similar biological assumptions
Difference in modeling formalism

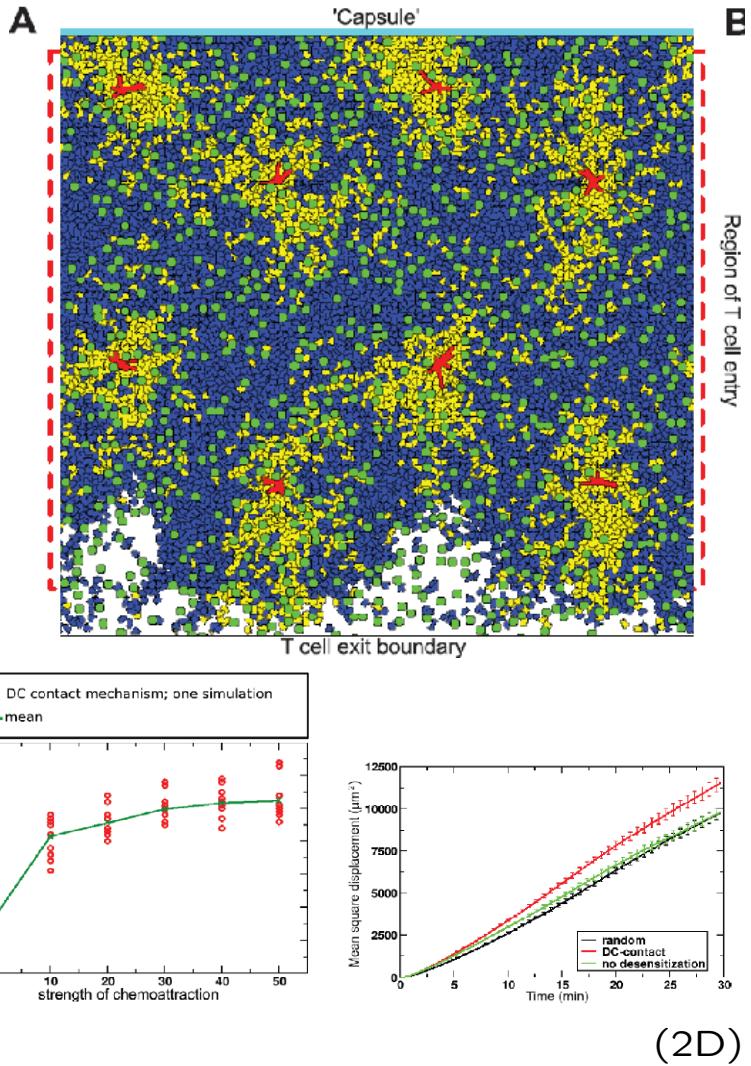
CA model of Riggs et al



RANDOM CHEMOTAXIS
(3D)

Vroomans: sensitive T cells (blue), insensitive T cells (yellow), DCs (red), reticular

CPM model of Vroomans et al



network (green)

Spatial modeling formalisms

space / time / var.	formalism
ccc (ddc)	partial differential equations (PDE) reaction diffusion systems
ddc	map lattices
ddd	CA
ccd	individual oriented models off lattice particle systems / event-based
dcc	meta-population models
c/dc (d+c)	hybrid models

note: translations