Part II - Applications of MultiSeq Evolution of Translation: Dynamics of Recognition in RNA:Protein Complexes

Part III – Towards in silico Cells: Simulating processes in entire cells

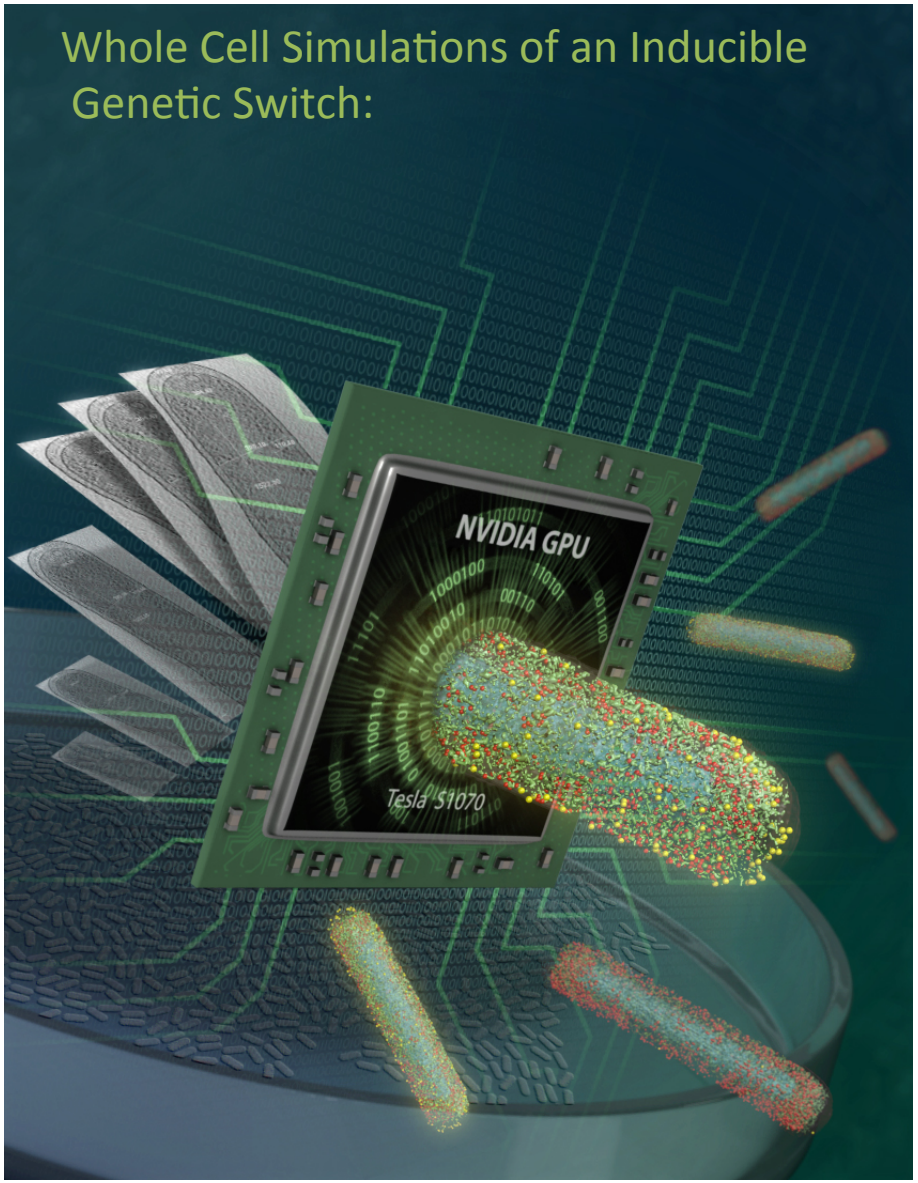
Zaida (Zan) Luthey-Schulten
Dept. Chemistry, Physics, Beckman Institute, Biophysics, Institute of Genomics

Biology

NIH Resource Macromolecular Modeling and Bioinformatics Workshop 2012

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Cellular Processes in Bacterial Cells



- Assemble cells for *in silico* studies with molecular crowding from CET & proteomics data
- Lac Genetic switch in E. coli
- Stochastic gene expression models
- Kinetic parameters from in vitro & SM experiments
- Compare solns for 2 cell models: fast and slow growing E. coli
- Multi-particle reaction-diffusion solns on a 3D lattice using GPUs

  for an entire cell cycle
- Progress towards eukaryal cells
- Connecting to systems biology

Crowding from CET, Proteomics, & Scell Data

Slow growing *E. coli cells with 3000 ribosomes*

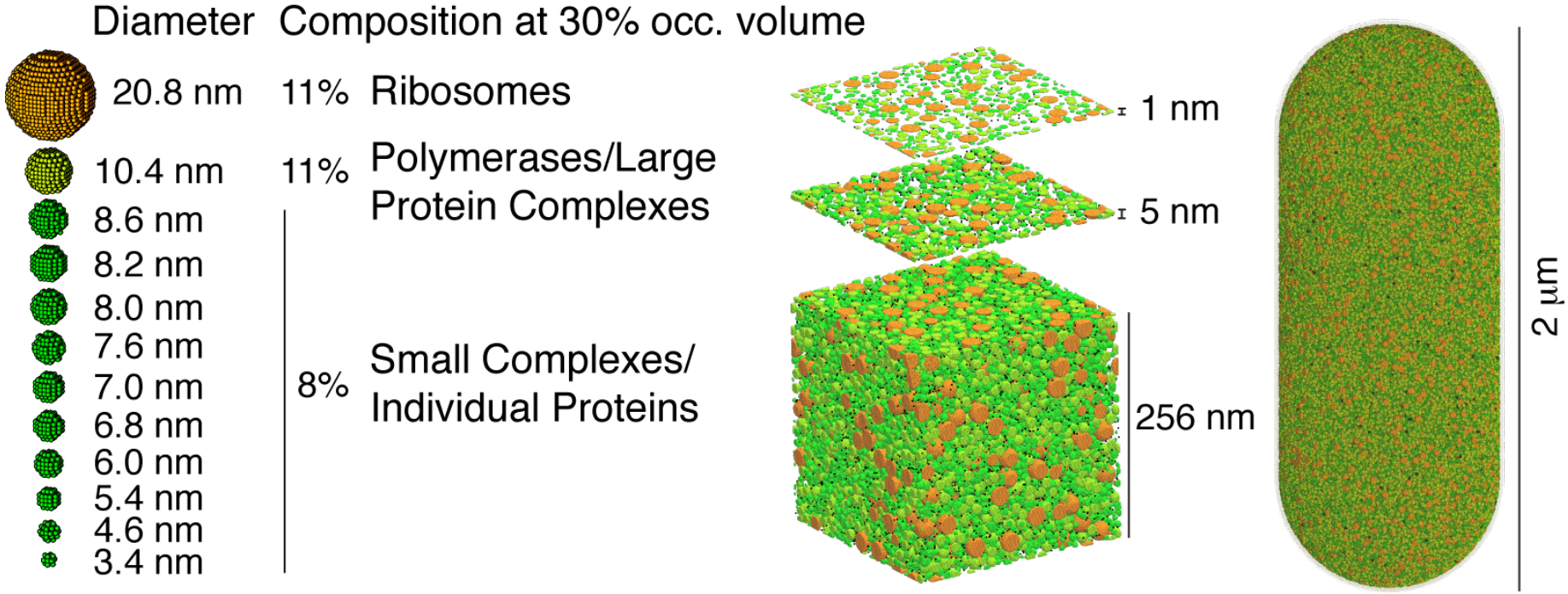
Cryo-electron tomography – ribosomes and membrane

Modeling – nucleoid and *in vivo* obstacles

Roberts, Magis, Ortiz,Baumeister, ZLS *Plos Comp. Biol* (2011), Cybercell DB, Xie et al. Science 2011

Ridgway et al., (2008) *Biophys J*
Roberts, Stone, Sepulveda, Hwu, Luthey-Schulten (2009) *IEEE Proc. HPCB - Italy*

In vivo Crowding in *E. coli* – *Fast Growing*



Immobile obstacle classes

Abundances from proteomics studies

20-30,000 ribosomes randomly placed

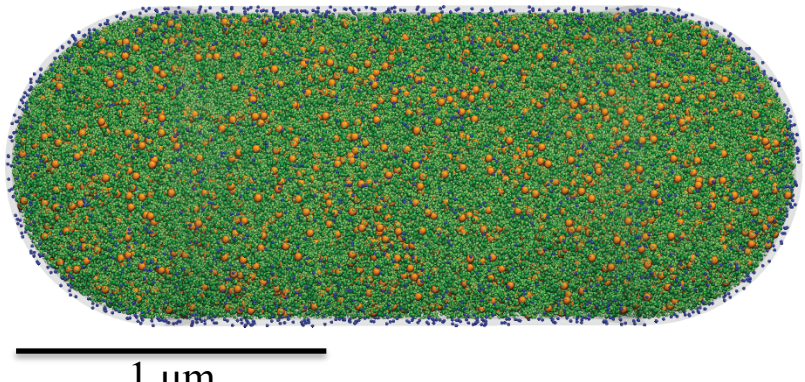
Fast growth phenotype (55 min) – uniformly distributed obstacles

Reaction diffusion master equation simulations on GPU

Lattice Microbes Software: A Brief Introduction

Software for simulating reactions within "in vivo" cells with molecular crowding

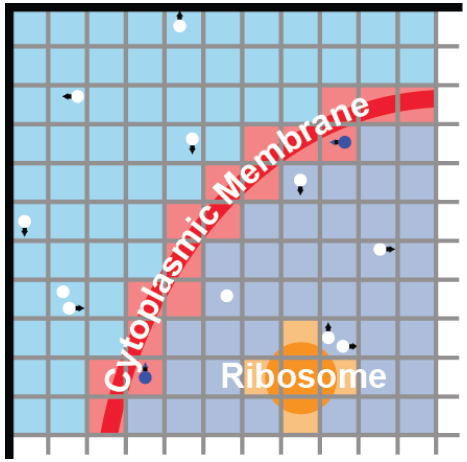
Integrating experimental data into computational cell models



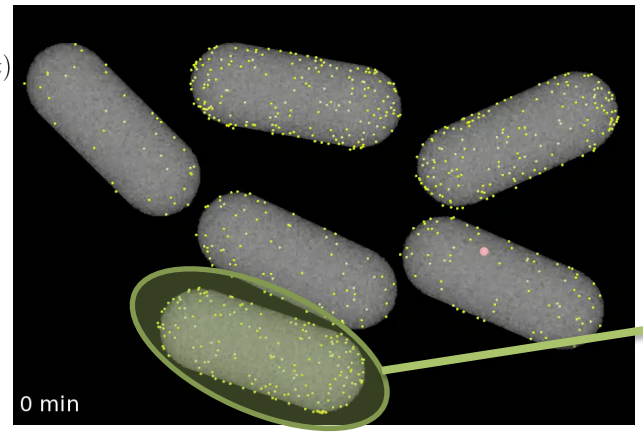
Solving the Reaction-Diffusion Master Equation (RDME) on a Lattice:

$$\frac{\partial P(x,t)}{\partial t} = \sum_{v\in V}\sum_{r=1}^{R} -a_r(x_v)P(x,t) + a_r(x_v - S_r)P(x - S_r 1_v, t)$$

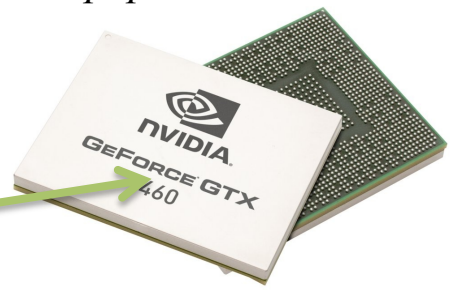
$$+ \sum_{i\in V}\sum_{j\in V}\sum_{\alpha=1}^{N} -d_{ij}^{\alpha} x_i^{\alpha} P(x,t) + d_{ji}^{\alpha}(x_j^{\alpha}+1)^{\alpha} P(x + 1_j^{\alpha} - 1_i^{\alpha}, t)$$



Lac Genetic Switch



GPUs enable modeling cells in a heterogeneous population

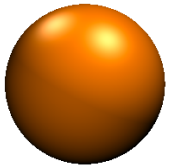


Roberts, Stone, Sepulveda, Hwu, Luthey-Schulten (2009) *IEEE IPDPS*
Roberts, Magis, Ortiz, Baumeister, Luthey-Schulten (2011) *PLoS Comp. Bio.*
Roberts, Stone, Luthey-Schulten (2012) *J. Comp. Chem.*

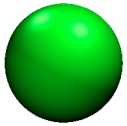
A Window into the Cell with VMD

John Stone

John Cole

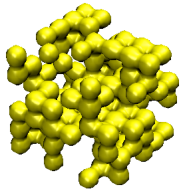


Ribosome



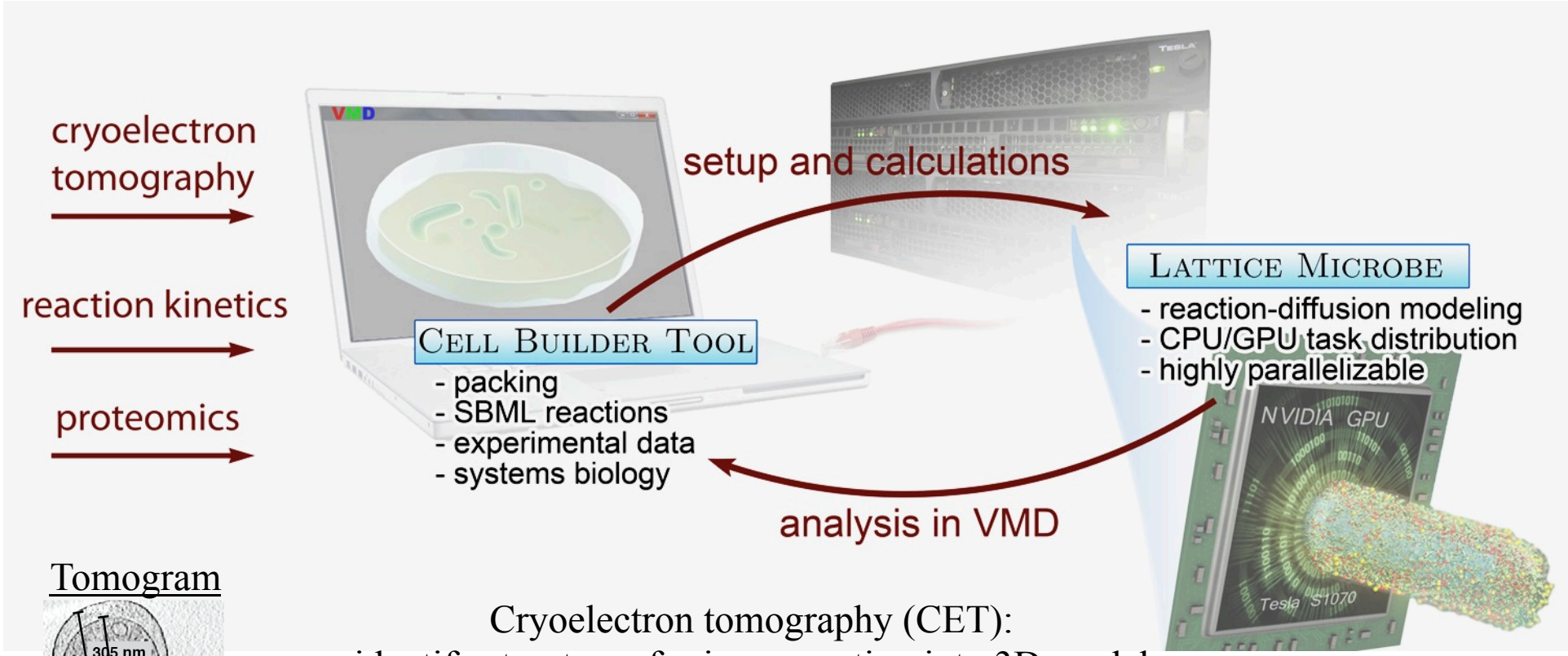
Polymerases/
Large Complexes

Small Complexes/
Proteins

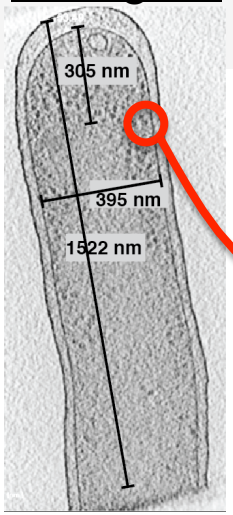


Lattice sites

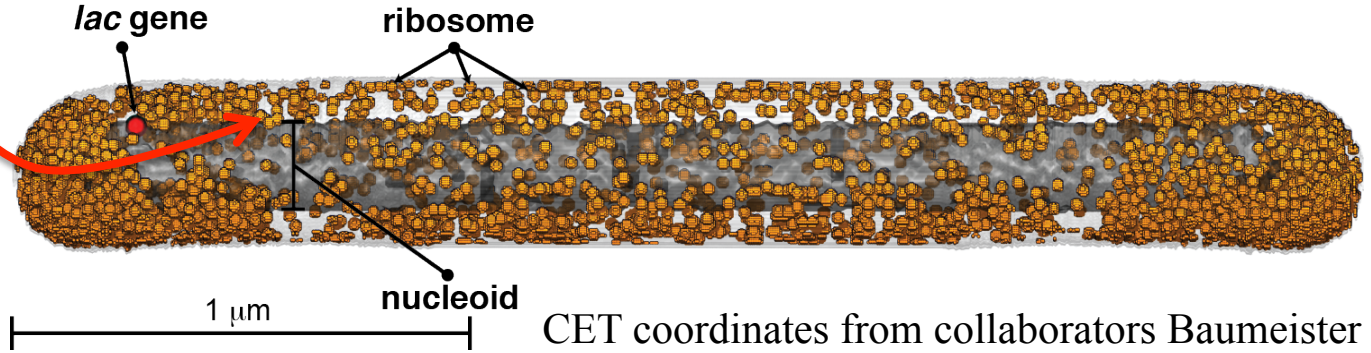
Overview of Cell Simulation Tools



Tomogram

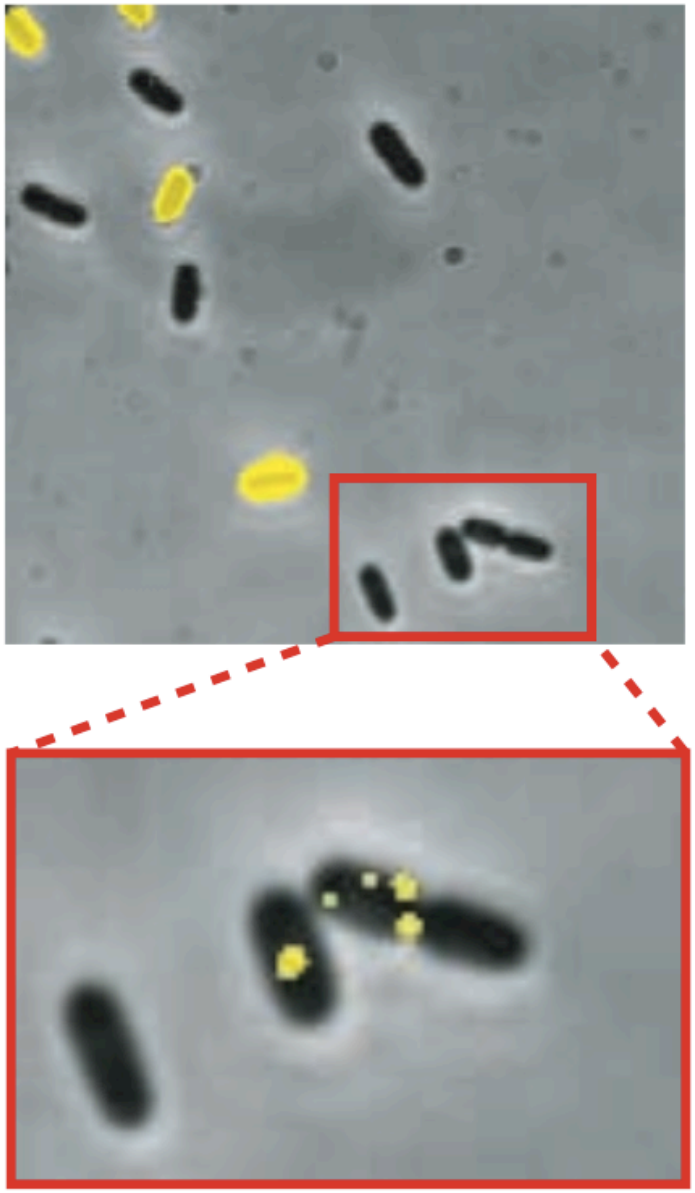


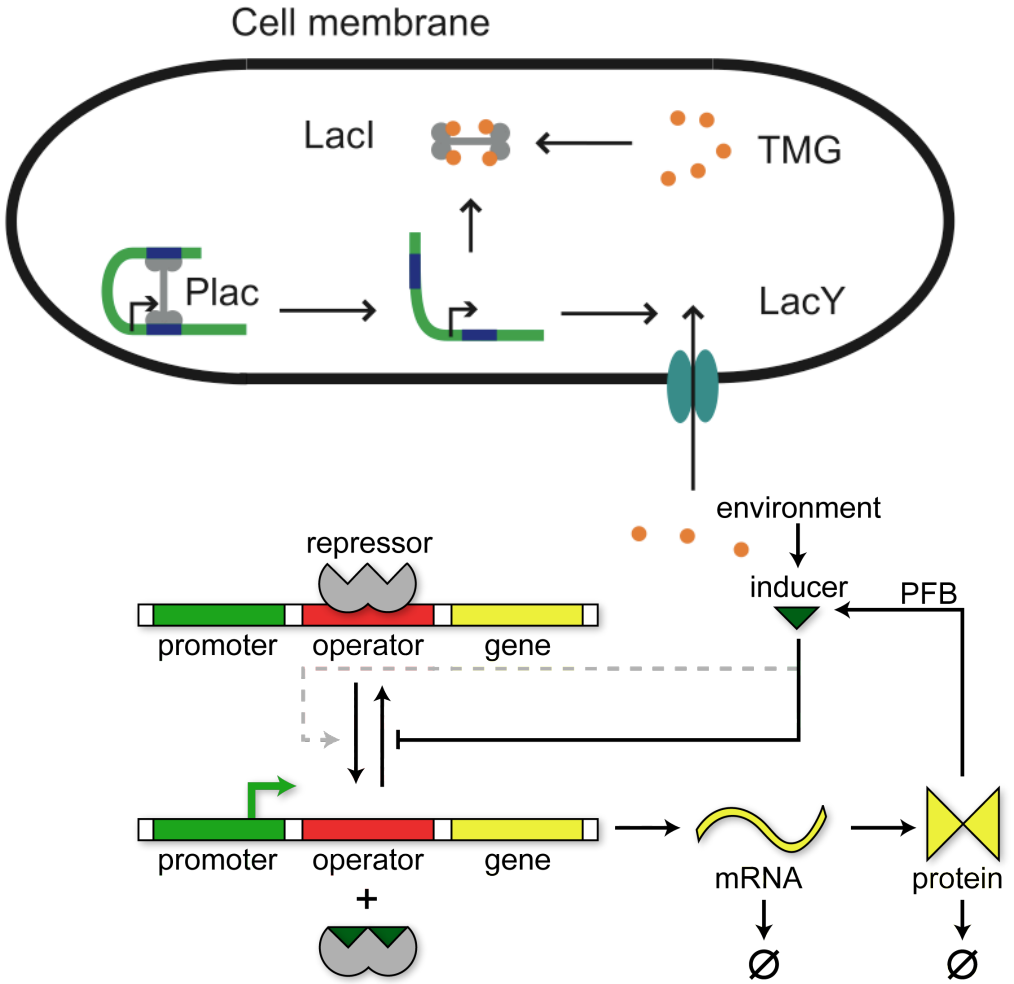
Cryoelectron tomography (CET):
identify structures for incorporation into 3D model



CET coordinates from collaborators Baumeister & Ortiz
Future: STED, Confocal Microscope, Virtual Cell models, …

Lac genetic switch in E. coli





- Kinetic model for lac regulatory circuit
- Stochasticity & population heterogeneity?

SM experiments (Xie, Science 2007, 2008):

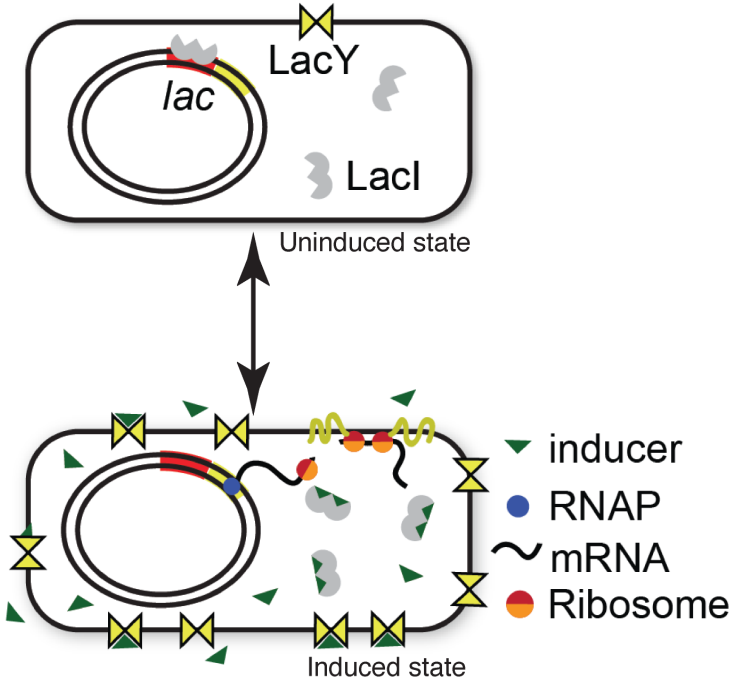
Kinetic Model of *lac* System

Reaction	Param	Stochastic Rate		Units	Source[a]	
Lac operon regulation						
$R_2 + O \rightarrow R_2O$	k_{ron}	2.43e+06		$M^{-1}s^{-1}$	M	
$IR_2 + O \rightarrow IR_2O$	k_{iron}	1.21e+06		$M^{-1}s^{-1}$	M	
$I_2R_2 + O \rightarrow I_2R_2O$	k_{i2ron}	2.43e+04		$M^{-1}s^{-1}$	M	
$R_2O \rightarrow R_2 + O$	k_{roff}	6.30e-04		s^{-1}	S	
$IR_2O \rightarrow IR_2 + O$	k_{iroff}	6.30e-04		s^{-1}	S	
$I_2R_2O \rightarrow I_2R_2 + O$	k_{i2roff}	3.15e-01		s^{-1}	M	
Transcription, translation, and degredation						
$O \rightarrow O + mY$	k_{tr}	1.26e-01		s^{-1}	M	
$mY \rightarrow mY + Y$	k_{tn}	4.44e-02		s^{-1}	S	
$mY \rightarrow \varnothing$	k_{degm}	1.11e-02		s^{-1}	S	
$Y \rightarrow \varnothing$	k_{degp}	2.10e-04		s^{-1}	M	
Lac inducer–repressor interactions		TMG	IPTG		TMG	IPTG
$I + R_2 \rightarrow IR_2$	k_{ion}	2.27e+04	9.71e+04	$M^{-1}s^{-1}$	M	K
$I + IR_2 \rightarrow I_2R_2$	k_{i2on}	1.14e+04	4.85e+04	$M^{-1}s^{-1}$	M	K
$I + R_2O \rightarrow IR_2O$	k_{iopon}	6.67e+02	2.24e+04	$M^{-1}s^{-1}$	M	K
$I + IR_2O \rightarrow I_2R_2O$	k_{i2opon}	3.33e+02	1.12e+04	$M^{-1}s^{-1}$	M	K
$IR_2 \rightarrow I + R_2$	k_{ioff}	2.00e-01		s^{-1}		K
$I_2R_2 \rightarrow I + IR_2$	k_{i2off}	4.00e-01		s^{-1}		K
$IR_2O \rightarrow I + R_2O$	k_{iopoff}	1.00e+00		s^{-1}		K
$I_2R_2O \rightarrow I + IR_2O$	$k_{i2opoff}$	2.00e+00		s^{-1}		K
Inducer transport						
$I_{ex} \rightarrow I$	k_{id}	2.33e-03		s^{-1}		K
$I \rightarrow I_{ex}$	k_{id}	2.33e-03		s^{-1}		K
$Y + I_{ex} \rightarrow YI$	k_{yion}	3.03e+04		$M^{-1}s^{-1}$		K
$YI \rightarrow Y + I_{ex}$	k_{yioff}	1.20e-01		s^{-1}		K
$YI \rightarrow Y + I$	k_{it}	1.20e+01		s^{-1}		K

K – in vitro kinetic experiment

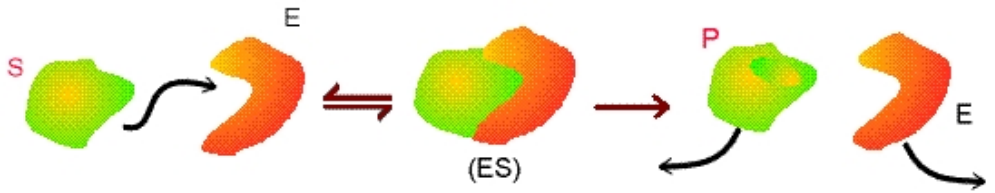
S – single molecule experiment

M – model parameter fit to single-molecule distributions



Analysis of biochemical reactions

$$E + S \xrightarrow{k_1} ES$$
$$ES \xrightarrow{k_2} E + S$$
$$ES \xrightarrow{k_3} E + P$$

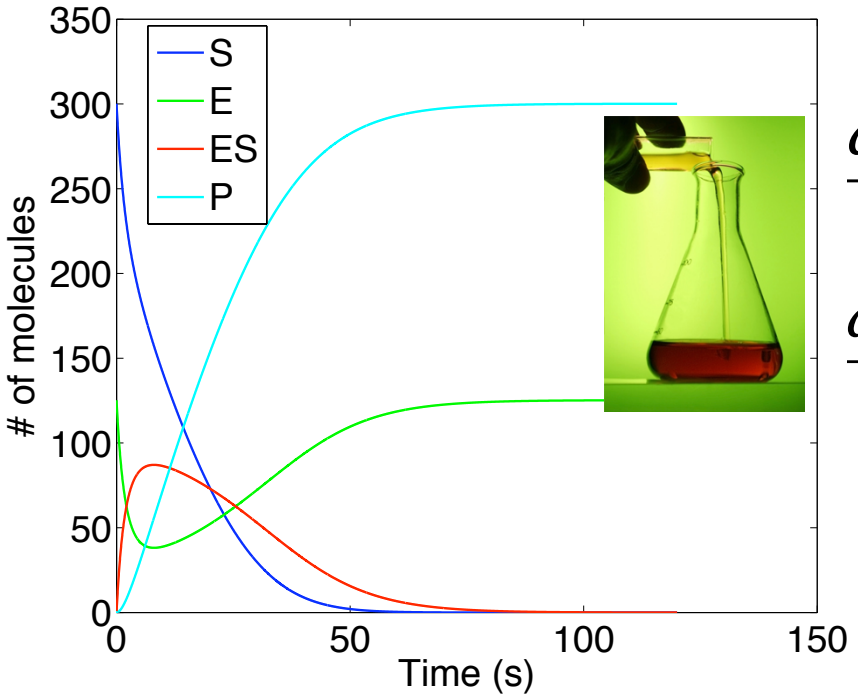


$$\frac{d[S]}{dt} = k_2[ES] - k_1[E][S]$$

$$\frac{d[E]}{dt} = k_2[ES] + k_3[ES] - k_1[E][S]$$

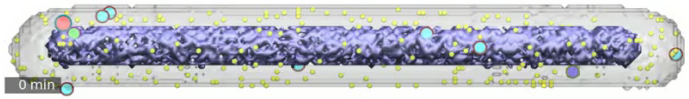
$$\frac{d[ES]}{dt} = k_1[E][S] - k_2[ES] - k_3[ES]$$

$$\frac{d[P]}{dt} = k_3[ES]$$



In cells hundreds of such reactions exist, some Involving only a few molecules (9 repressors, 1DNA, …) others have thousands of reactants.

Stochastic *vs*. Deterministic Solutions



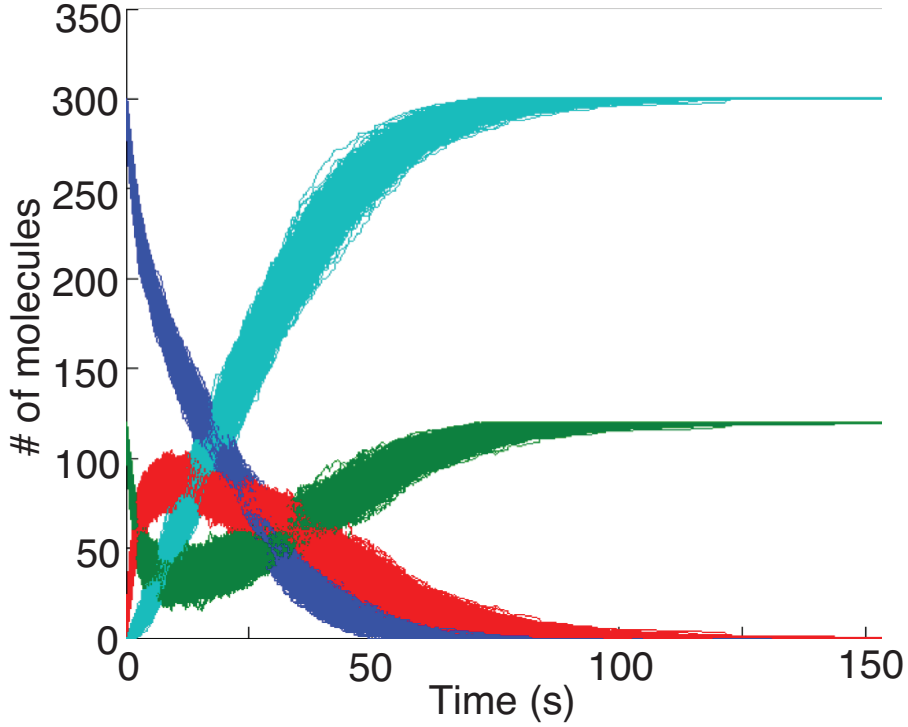
Cell is shaken, but not well-stirred

Gillespie's SSA on GPU (2000 threads)

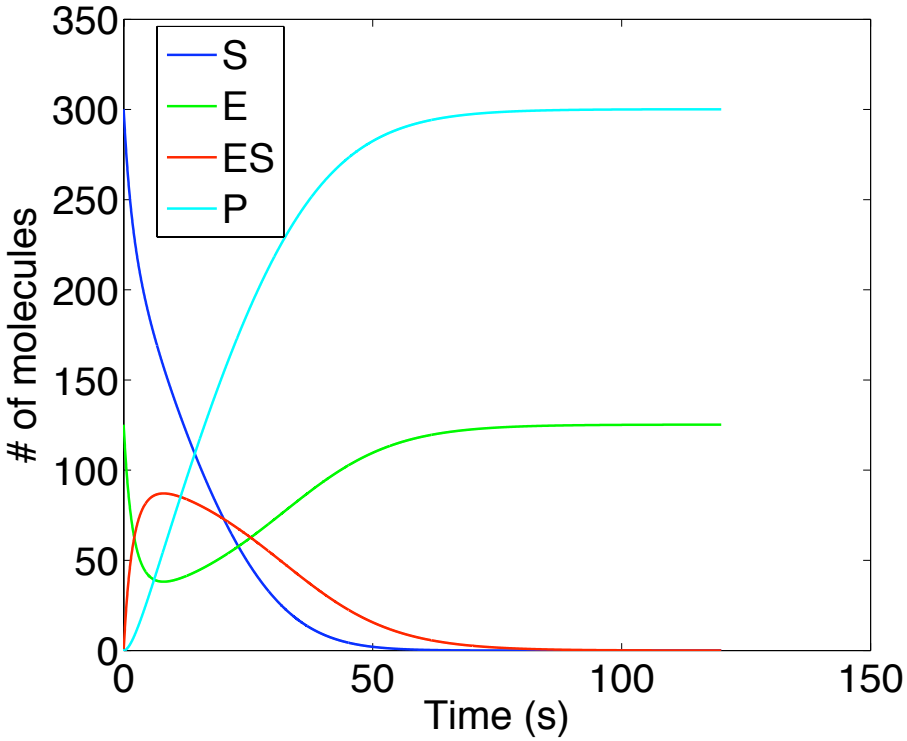
$$E + S \xrightarrow{k_1} ES$$

$$ES \xrightarrow{k_2} E + S$$

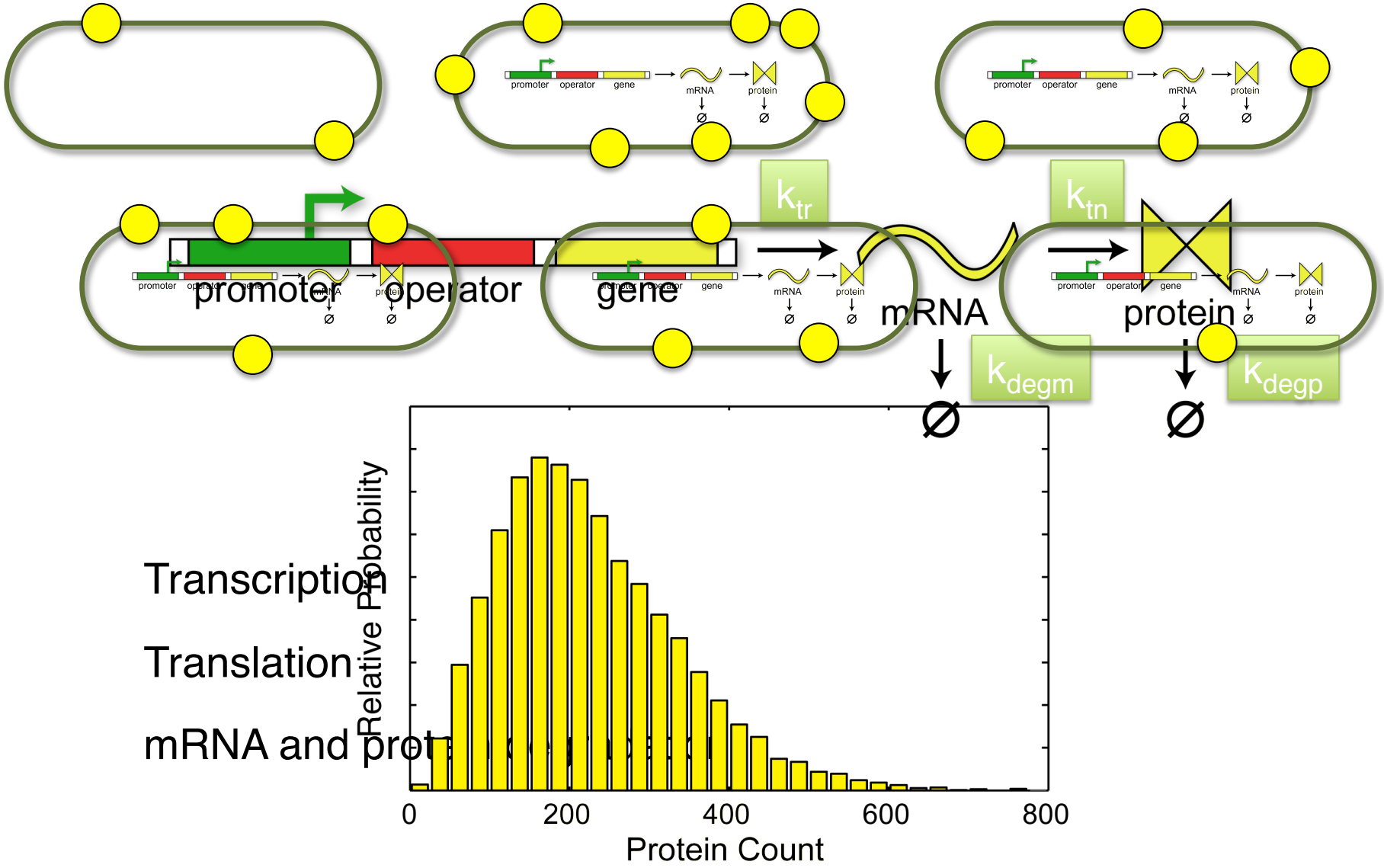
$$ES \xrightarrow{k_3} E + P$$



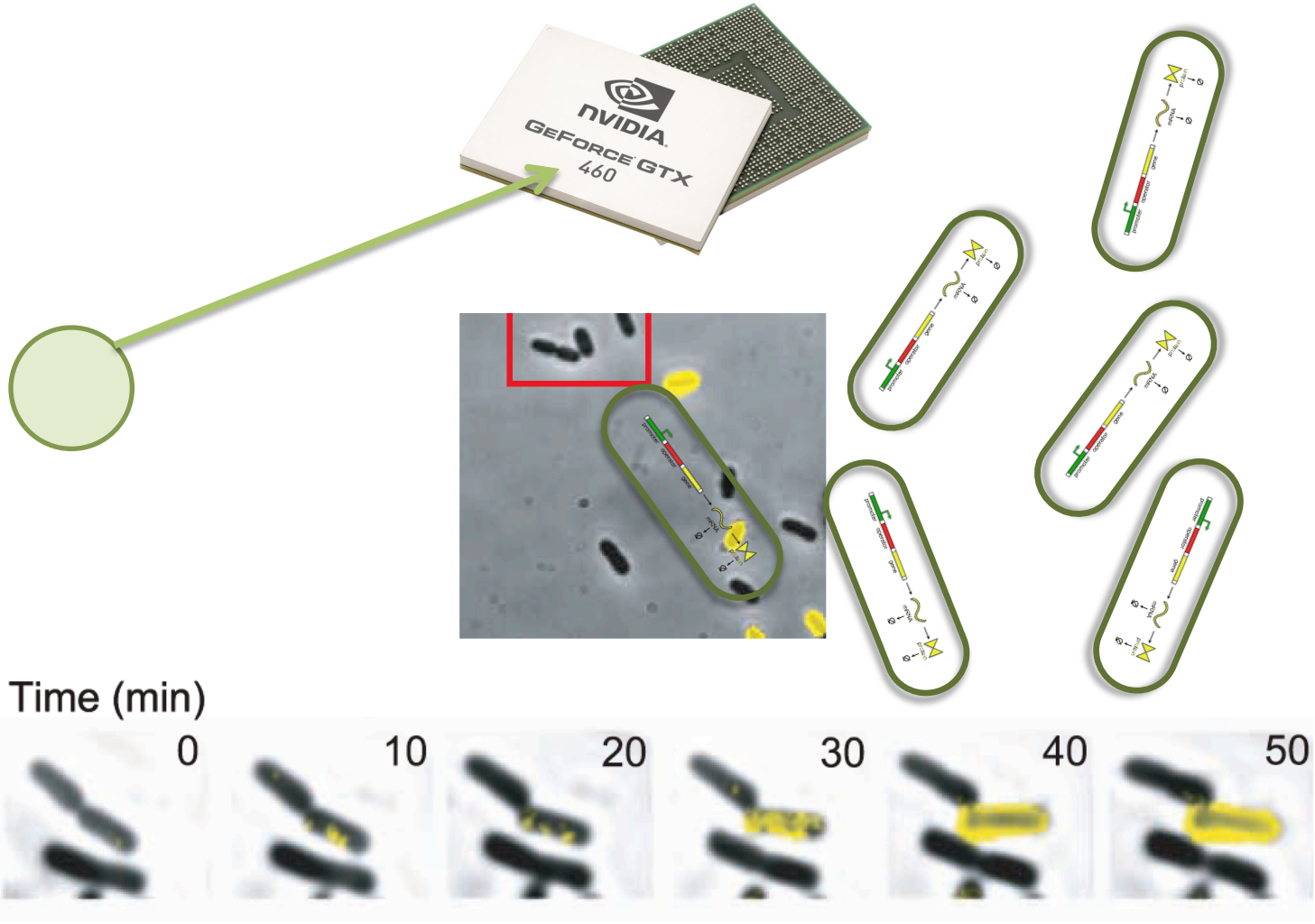
Well-stirred reactions



Un-induced state – low inducer

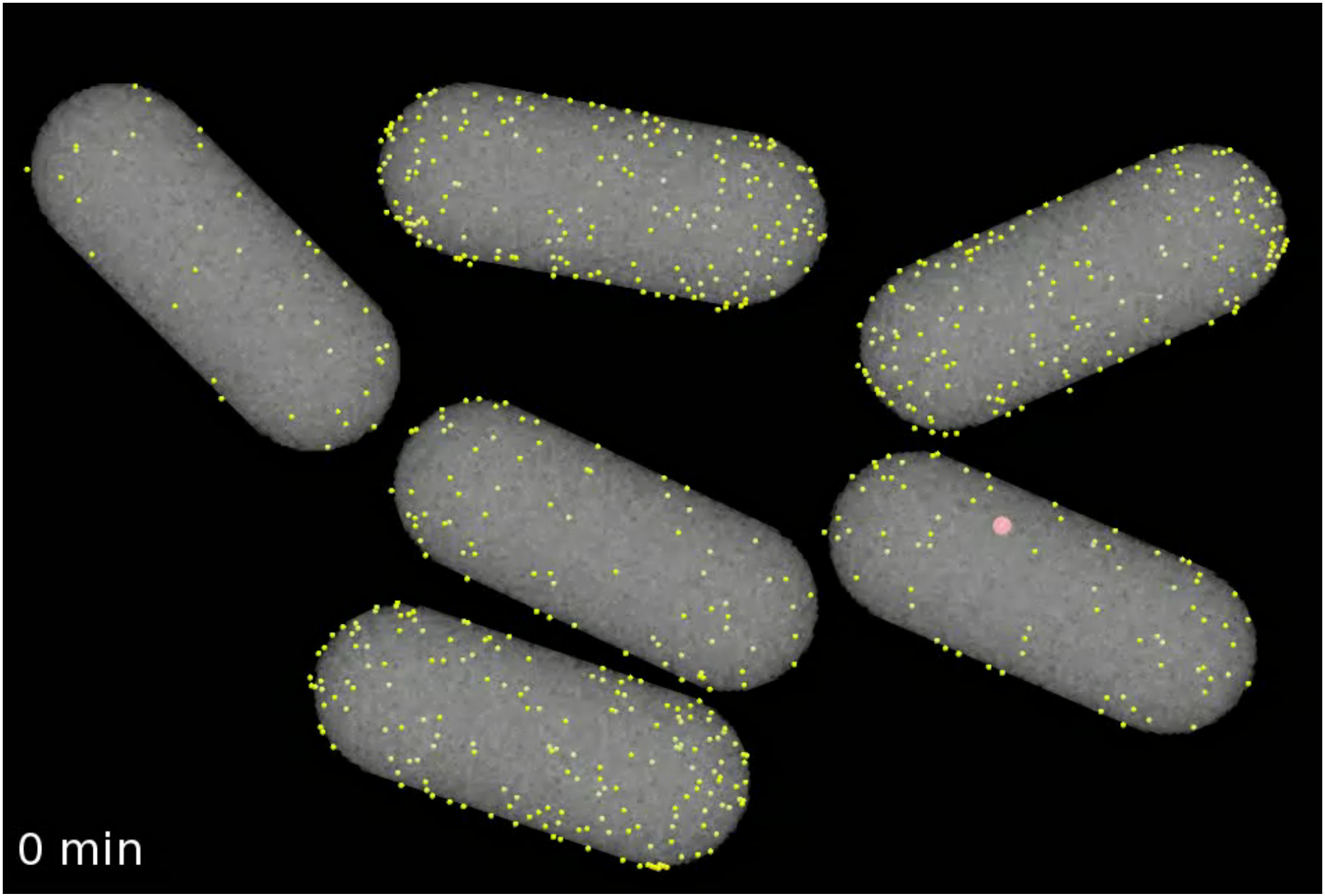


Motivation: Capture Timescale and Fraction of Cells Undergoing Phenotypic Switching

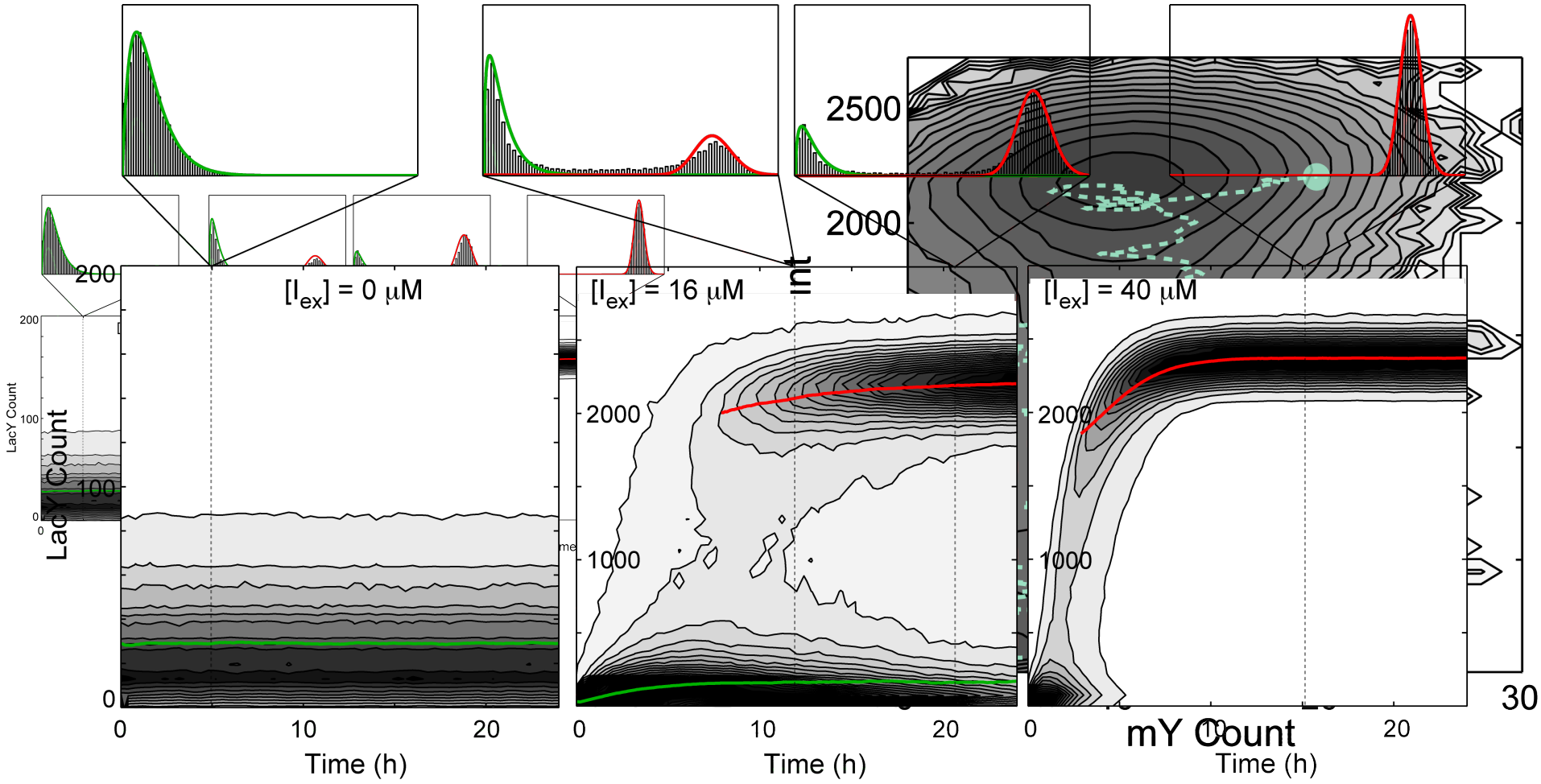


Choi, Cai, Frieda, Xie (2008) *Science*

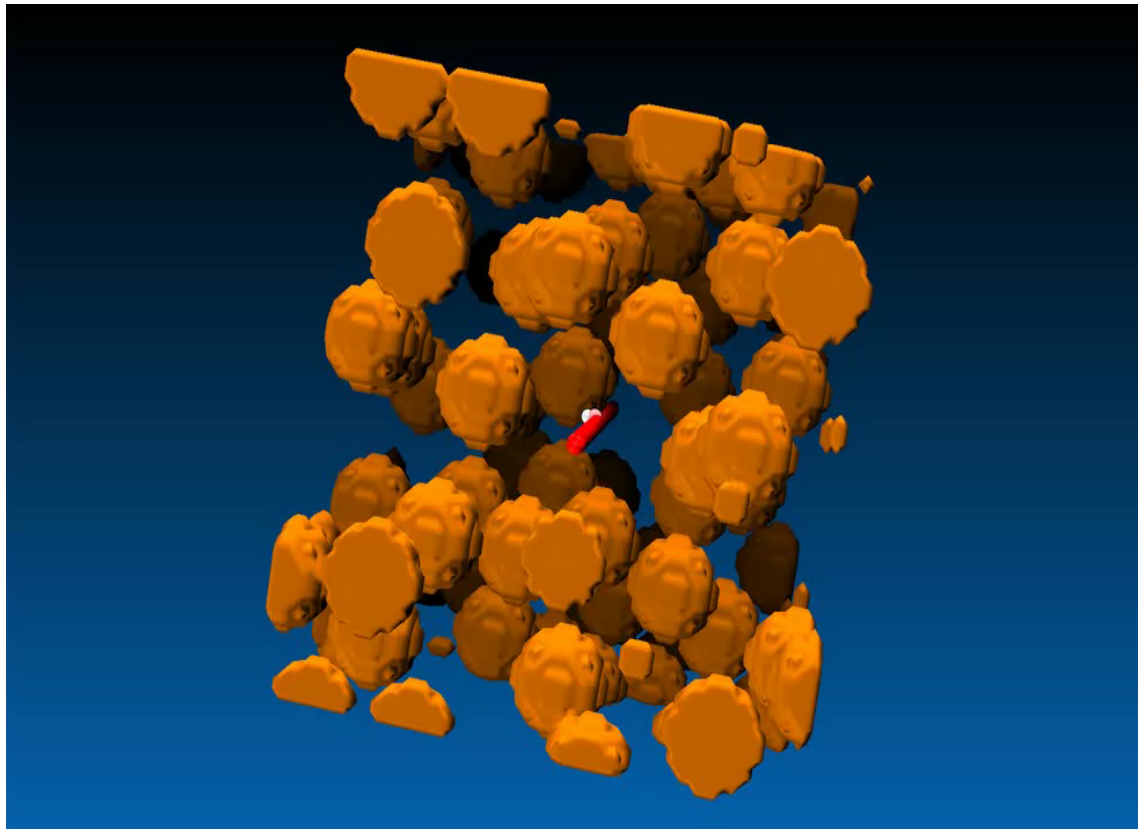
Switching in Fast Growing E. coli Cells – Bursting of mRNA

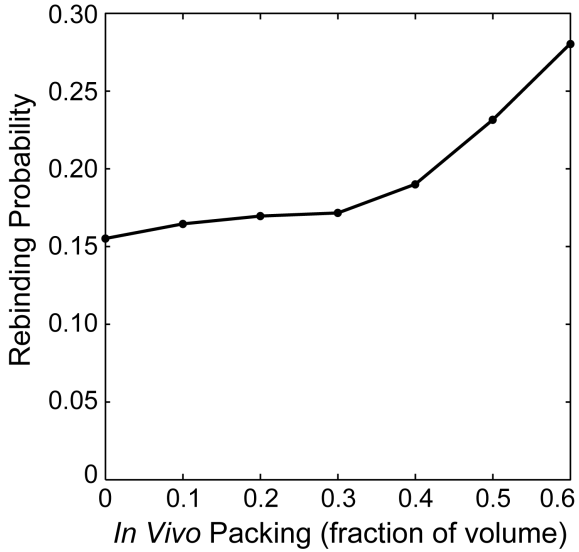


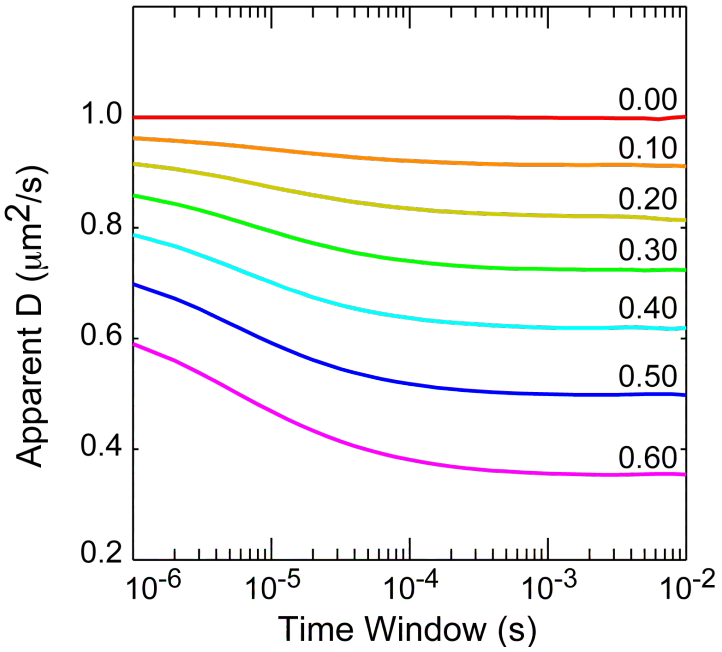
Switching Behavior – LacY/mRNA Distributions



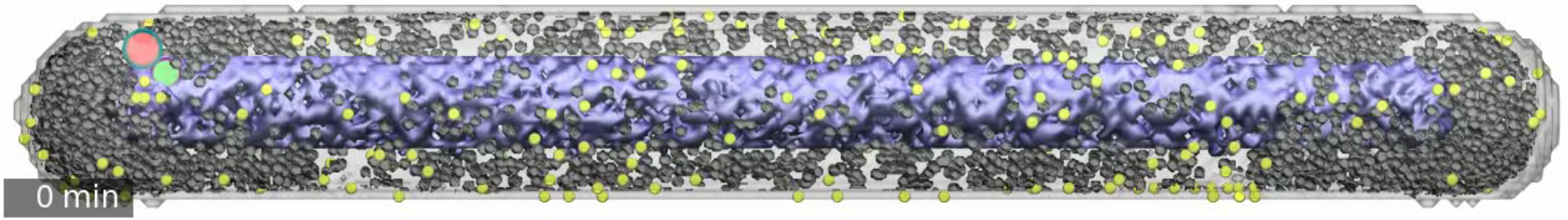
Effect of *in vivo* crowding on repressor re-binding



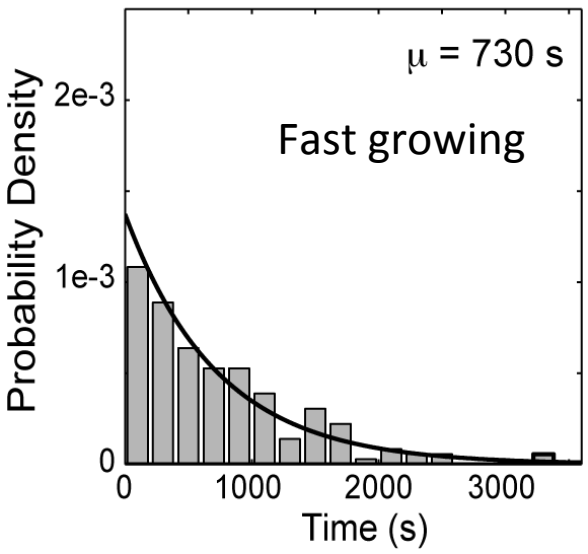


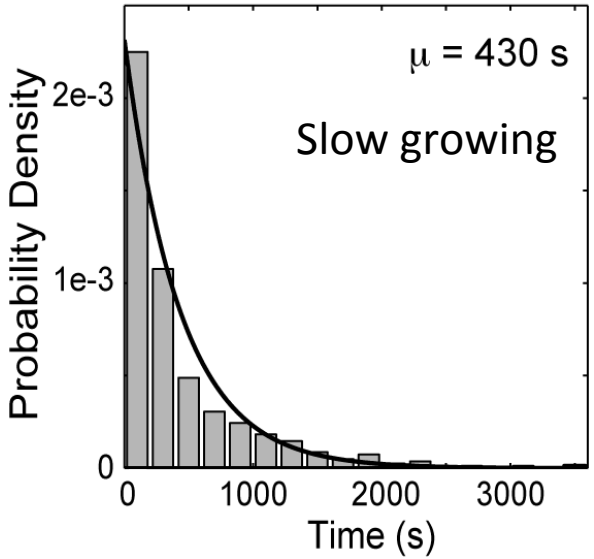


In vivo – Slow Growing *E. coli*

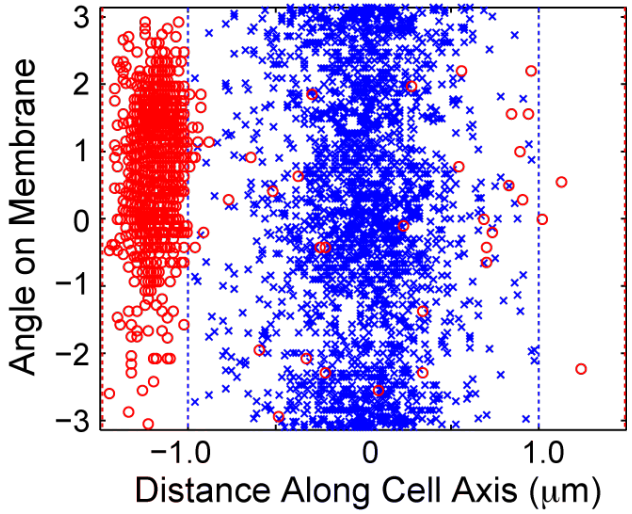


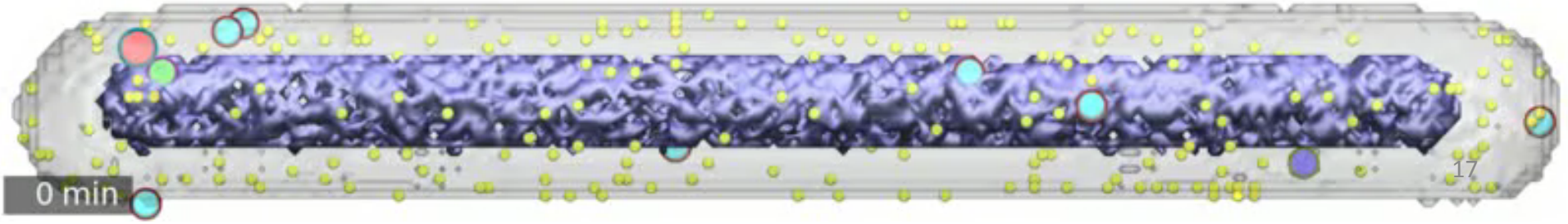
Lifetimes Repressor-Operator Complexes





Localization of mRNA



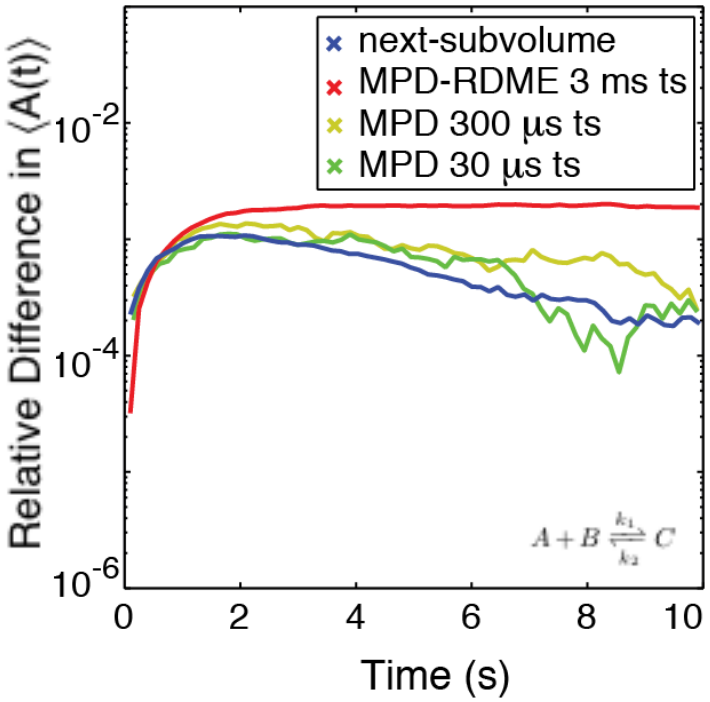


Lattice Microbes Software Release

Single GPU version available since January 2012

http://www.scs.uiuc.edu/schulten/lm

Validation of GPU-based approximate algorithm (MPD-RDME) compared to the exact next-subvolume method



Lattice Microbes: as much as *360 times faster*

Method	Lac Switch, [I] = 5μM rxns/sec	wall clock/sim hr	Lac Switch, [I] = 40μM rxns/sec	wall clock/sim hr
CME				
Direct GPU	7.3×10^6	0.018 sec	7.4×10^6	4.0 sec
Direct No-GPU	5.0×10^6	0.026 sec	5.0×10^6	5.9 sec
Next React GPU	5.5×10^6	0.024 sec	6.2×10^6	4.8 sec
Next React No-GPU	3.9×10^6	0.032 sec	4.5×10^6	6.5 sec
Stochkit SSA	4.2×10^6	0.031 sec	4.2×10^6	7.2 sec
COPASI	2.7×10^6	0.048 sec	2.8×10^6	11 sec
RDME				
MPD-RDME GPU	0.44	1.89 days	0.1200	0.191 yrs
Next Subvol GPU	0.10	8.33 days	0.0004	59 yrs
Next Subvol No-GPU	0.10	8.67 days	0.0004	59 yrs
MesoRD	0.01	83 days	0.0003	68 yrs

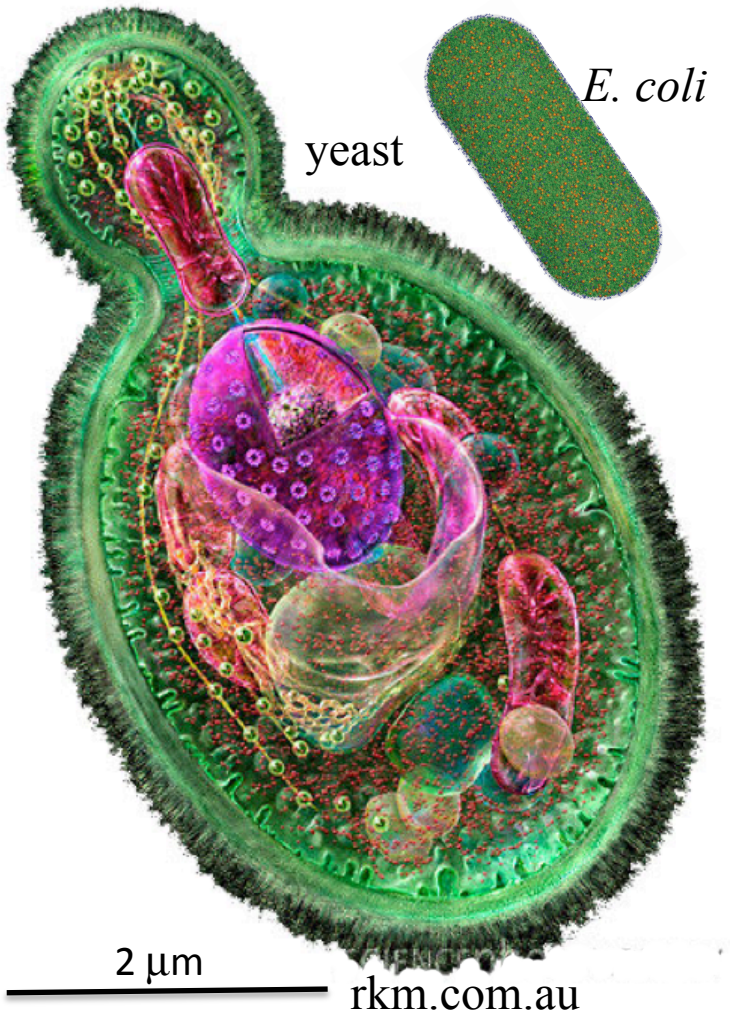
Roberts, Stone, Luthey-Schulten (2012) JCC

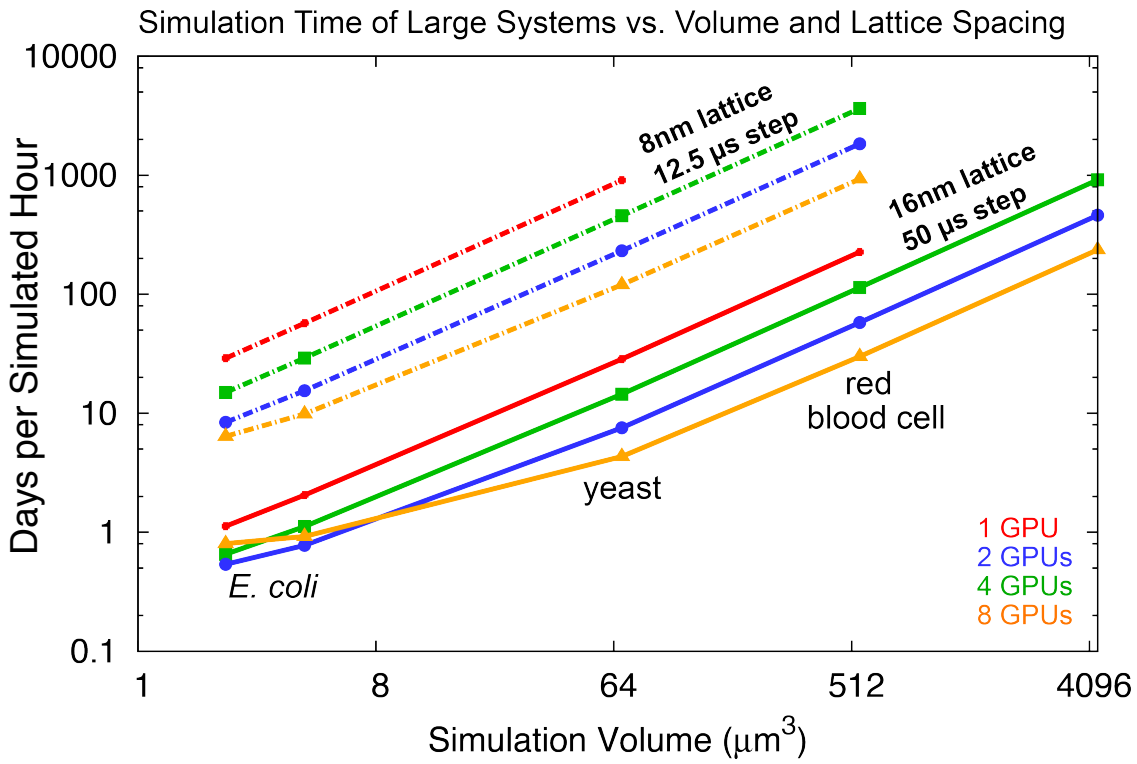
National Center for Research Resources

Ongoing Development of **Lattice Microbes**

*Expanding **Lattice Microbes** software to eukaryotic-sized cells*

- Progress: Added support for using **multiple GPUs** in a single node with good scaling
- Looking Forward: Distribute simulations across multiple nodes in a cluster





Hallock, Stone, Fry, Roberts, Luthey-Schulten (2012)