Modeling Molecular Motions

david.wishart@ualberta.ca

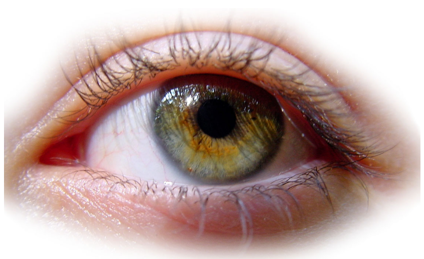
3-41 Athabasca Hall

Oct. 7, 2013

Why Should I Care?

- Most of today's largest computers were purchased or developed to perform accurate dynamic simulations (weather, nuclear explosions, molecular modeling)
- Parallel computing, cloud computing, GPU computing have all been driven by needs of chemists, physicists and engineers to do simulation
- High-performance computing (HPC): lots of resources, lots of research opportunities, lots of jobs
- Molecular simulation is used in physics, chemistry, nanotechnology, and biology
- Same ideas are often used in game development and special effects for movies/games

Our Seeing Limits (and Limitations)

Free	$5	$5000	$500,000
			
1 m	1×10^{-3} m	1×10^{-6} m	1×10^{-9} m
Live, moving	Live, moving	Live or stained	Fixed, stained

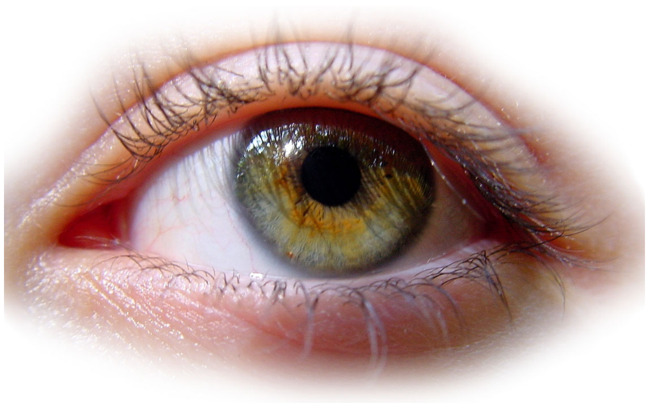
Our Seeing Limits (and Limitations)

$5,000,000 **$500,000,000**

$1x10^{-10}$ m **$1x10^{-12}$ m**

Extracted, crystallized **Atomized, vaporized**

What We See

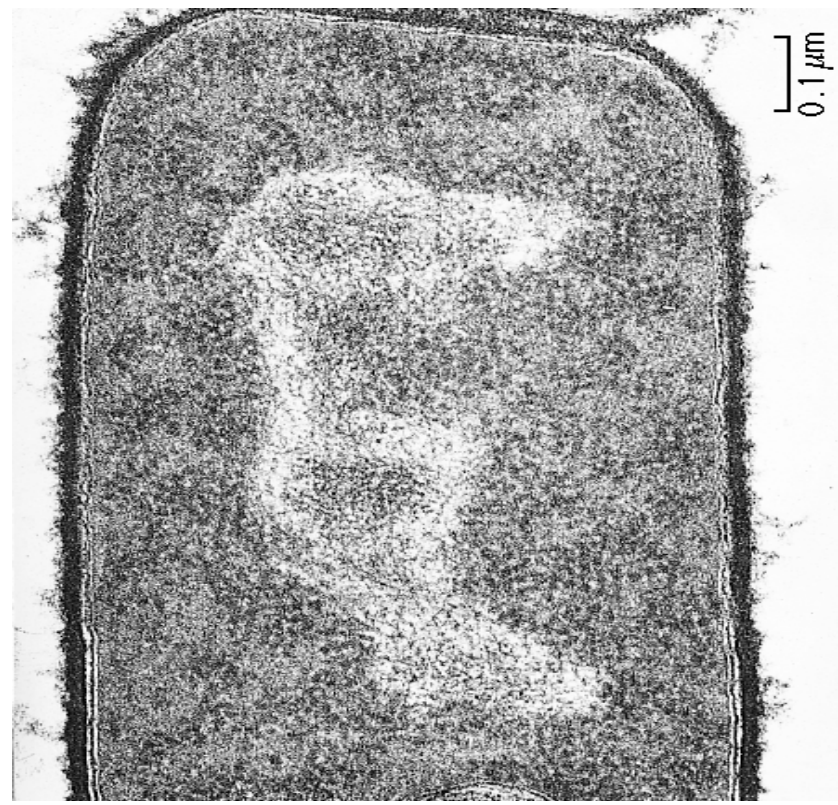


A Skier Jumping

Optical Microscopy

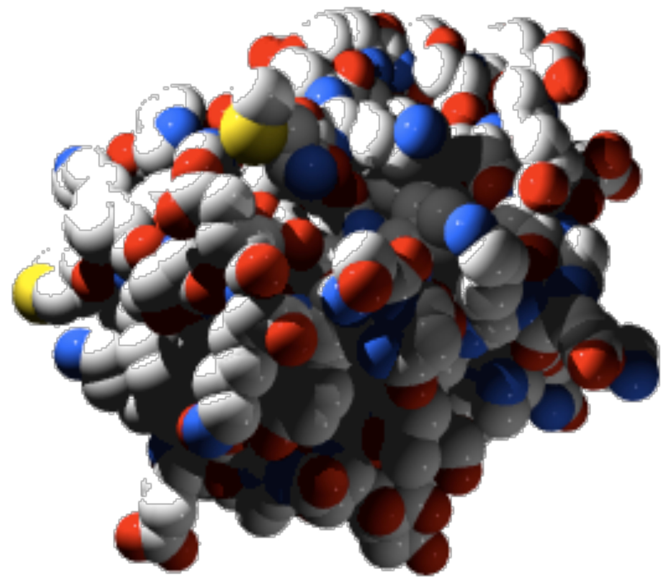
Cells Dividing

Electron Microscopy



Rigid structure, no movement

X-ray or NMR

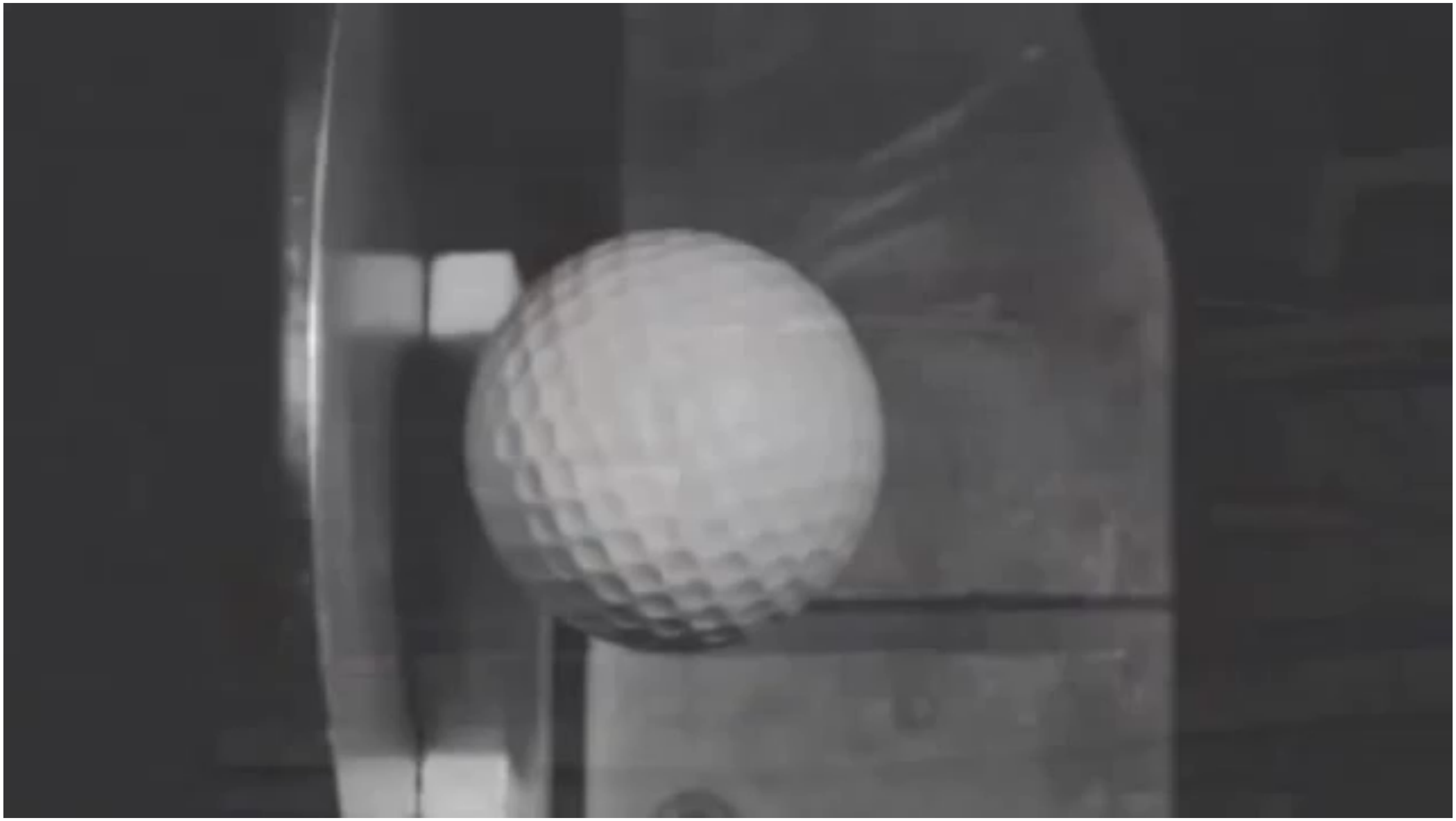


Rigid structure, no movement

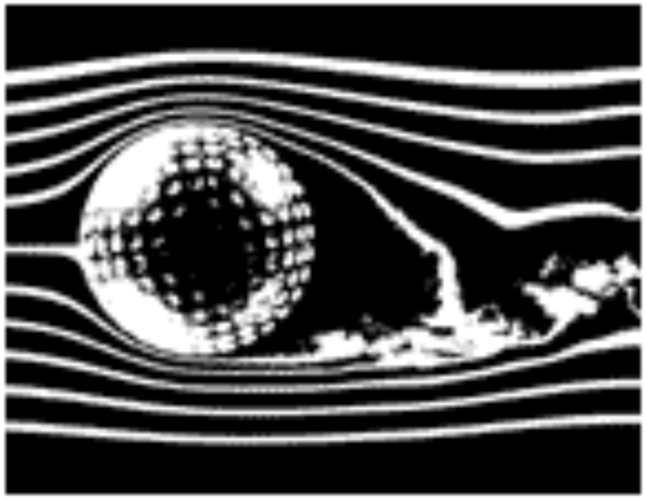
Problems with Visualization

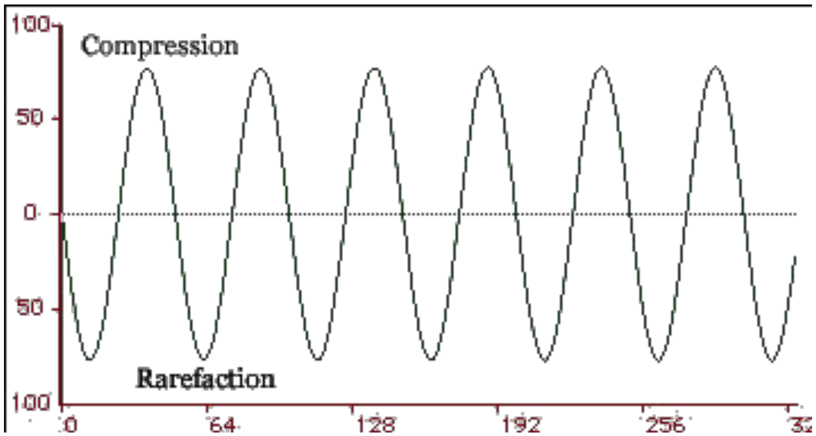
- What if the event happens too fast to see?
  - High speed photography
  - Measuring bulk kinetic variables (pressure, rate constants, heat gradients, enthalpy)
  - Computer simulation
- What if the system is too small to see?
  - Measuring bulk kinetic variables (pressure, rate constants, heat gradients, enthalpy)
  - Computer simulation

High Speed Photography

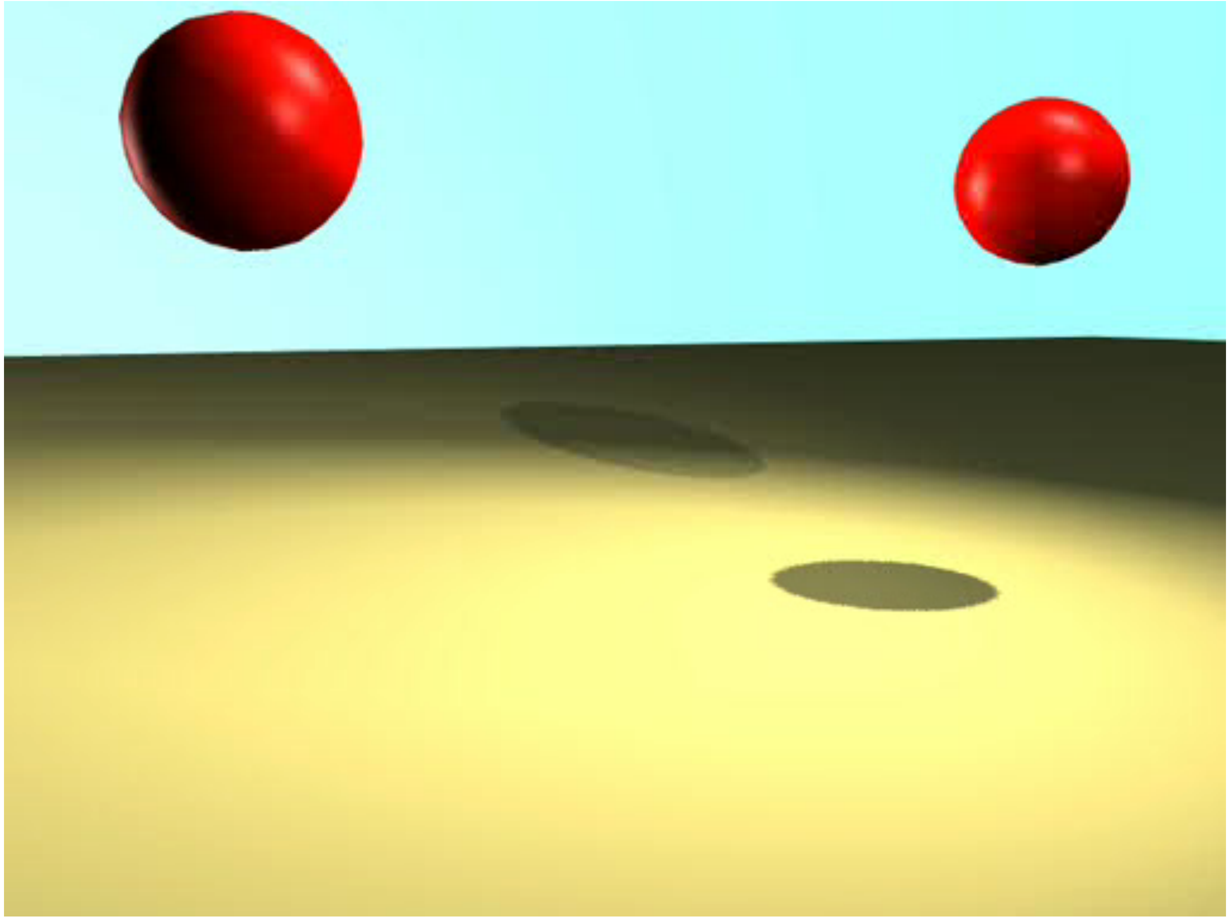


Measuring Bulk Properties





Computer Simulations



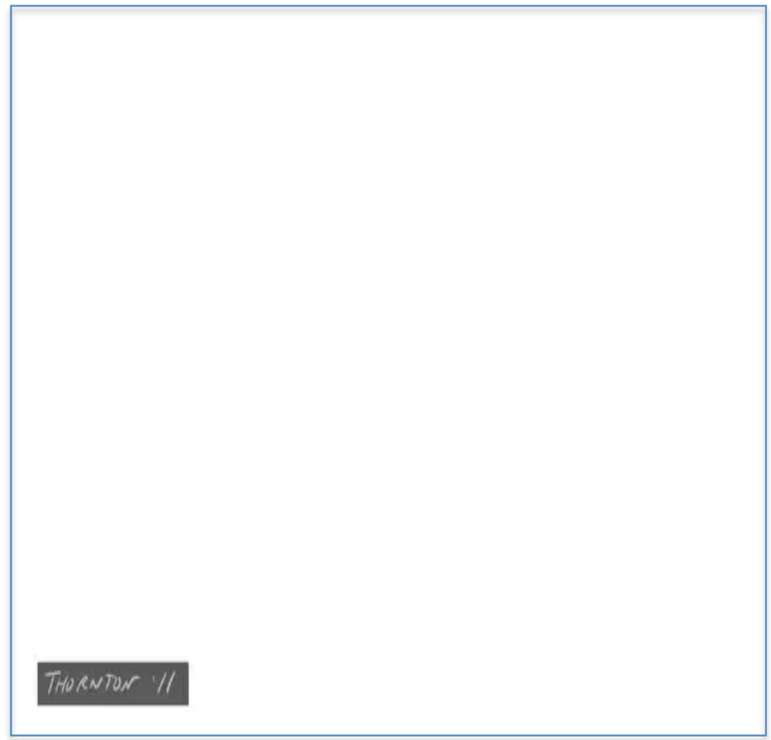
Deformation modeled by modal analysis

Simulations are Different than Animations

- Animation is art, simulation is science
- Animation tries to simulate reality but only approximately (animator's intuition)
- Better the animation, the closer it is to simulation – or reality (high quality gaming involves high quality animation and some even involve simulation)
- Simulation games (pong, bowling, golf, some team sport games, SimCity)

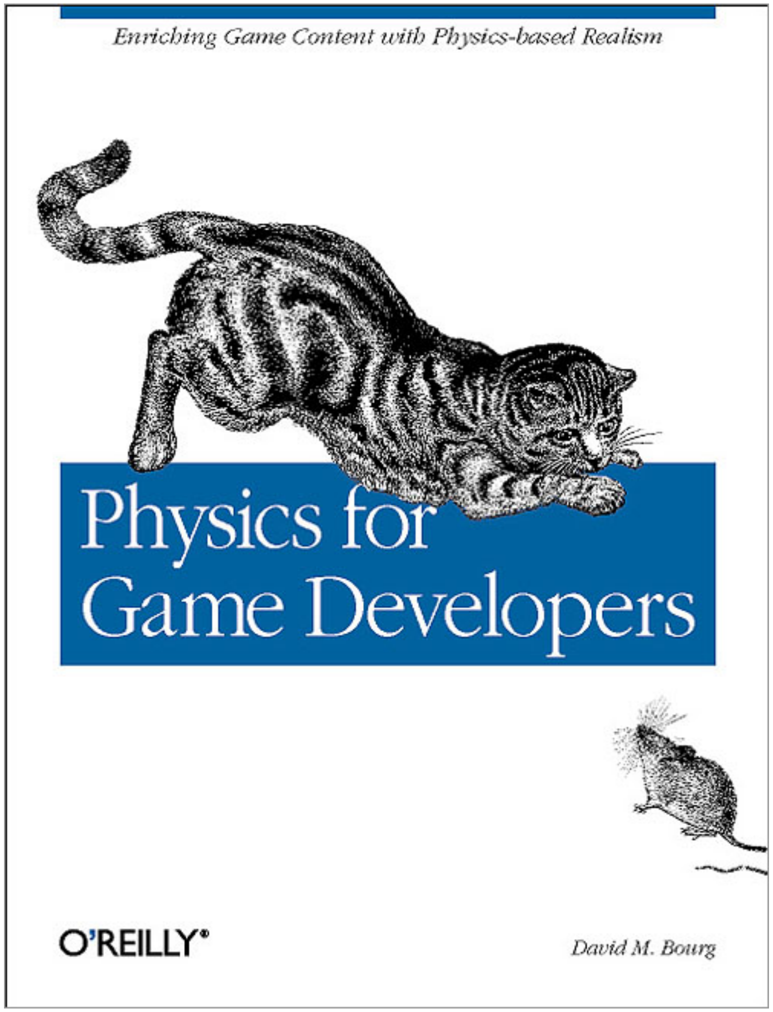
Simulation vs. Animation

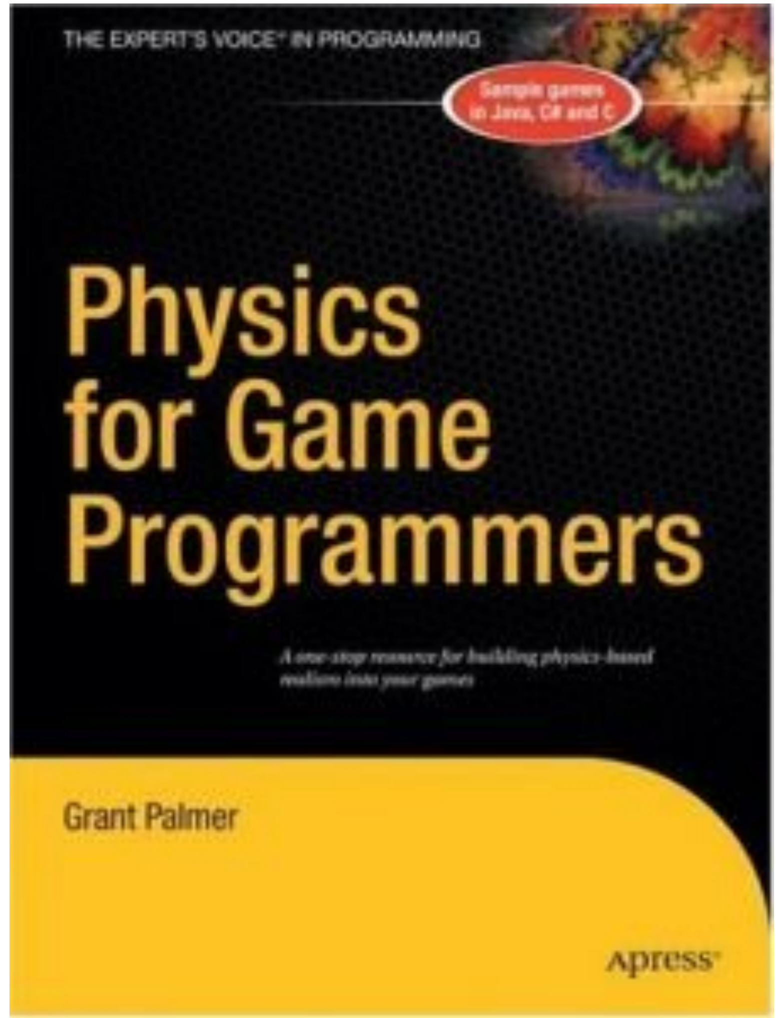
Simulation with Blender 2.5



Animation

The "Bibles" for Game Developers

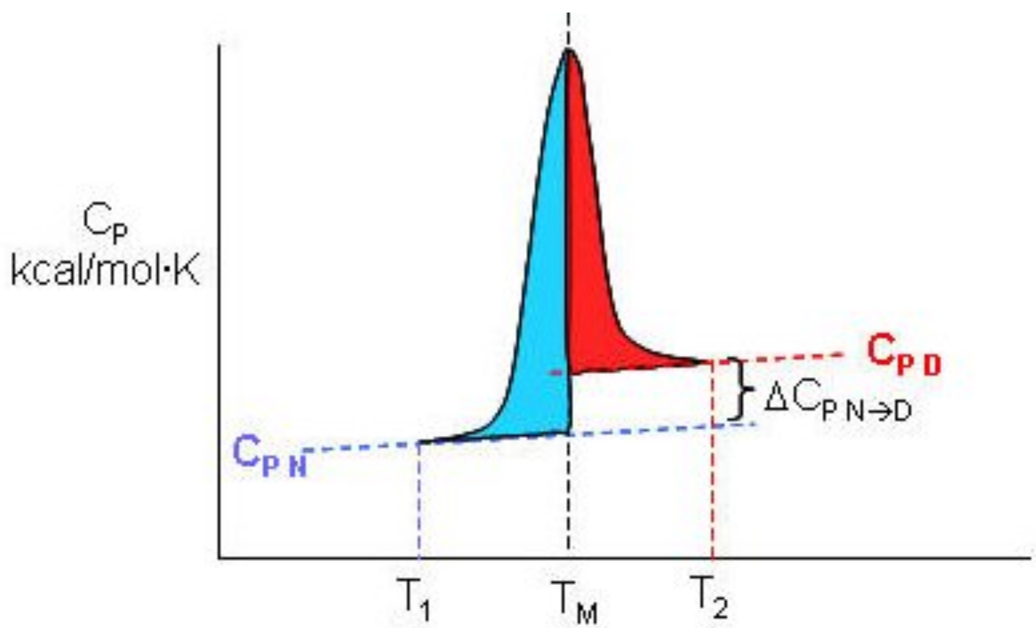




Problems with Visualization

- What if the event happens too fast to see?
  - High speed photography
  - Measuring bulk kinetic variables (pressure, rate constants, heat gradients, enthalpy)
  - Computer simulation
- What if the system is too small to see?
  - Measuring bulk kinetic variables (pressure, rate constants, heat gradients, enthalpy)
  - Computer simulation

Measuring Bulk Properties



Heat Capacity Curve

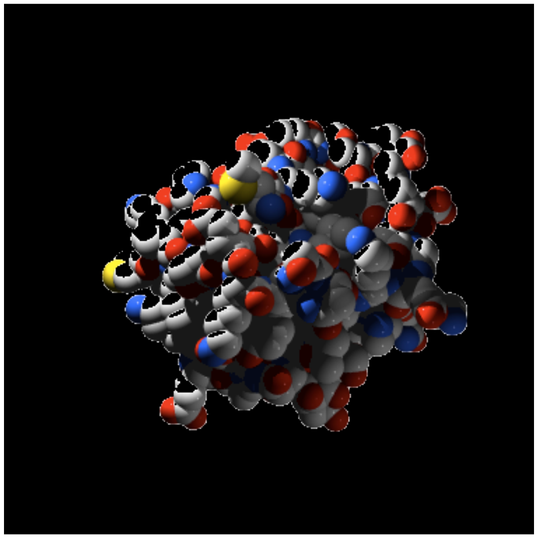
Isothermal Calorimeter

Instruments used to measure kinetics and thermodynamics of molecules

Cellular & Molecular Time Scales

Cell division	20 minutes
Lifetime of average mRNA	1 minute
Time to synthesize 1 protein	1 second
Time for protein to move across cell	1 millisecond
Time for 1 protein to fold	1 microsecond
Time for enzyme catalysis	1 nanosecond
Time for helix bending	1 picosecond
Time for bond vibration	1 femtosecond

3 Ways to Simulate

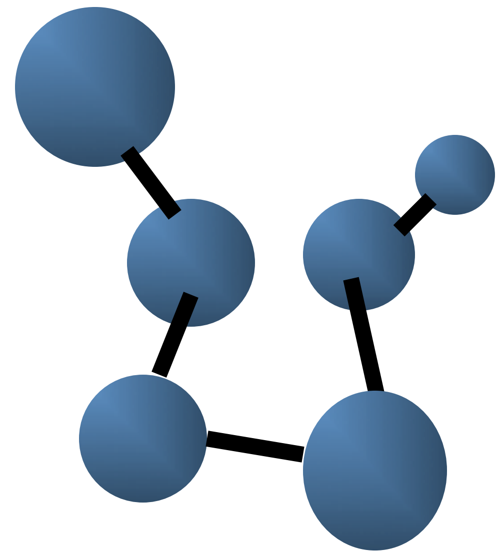


Atomic Scale
0.1 - 1.0 nm
Coordinate data
Dynamic data
0.1 - 10 ns
Molecular dynamics

Meso Scale
1.0 - 10 nm
Interaction data
Kon, Koff, Kd
10 ns - 10 ms
Mesodynamics

Continuum Model
10 - 100 nm
Concentrations
Diffusion rates
10 ms - 1000 s
Fluid dynamics

Molecular Dynamics



Newton's Equation

$$\vec{f}_i = m_i \vec{a}_i$$

Differential Equation

$$-\frac{dU}{d\vec{r}_i} = m_i \frac{d^2\vec{r}_i}{dt^2}$$

Leapfrog Verlet Algorithm

$$\vec{v}(t+1/2\Delta t) = \vec{v}(t-1/2\Delta t) + \Delta t \vec{a}(t)$$

$$\vec{r}(t+\Delta t) = \vec{r}(t) + \Delta t \vec{v}(t+1/2\Delta t)$$

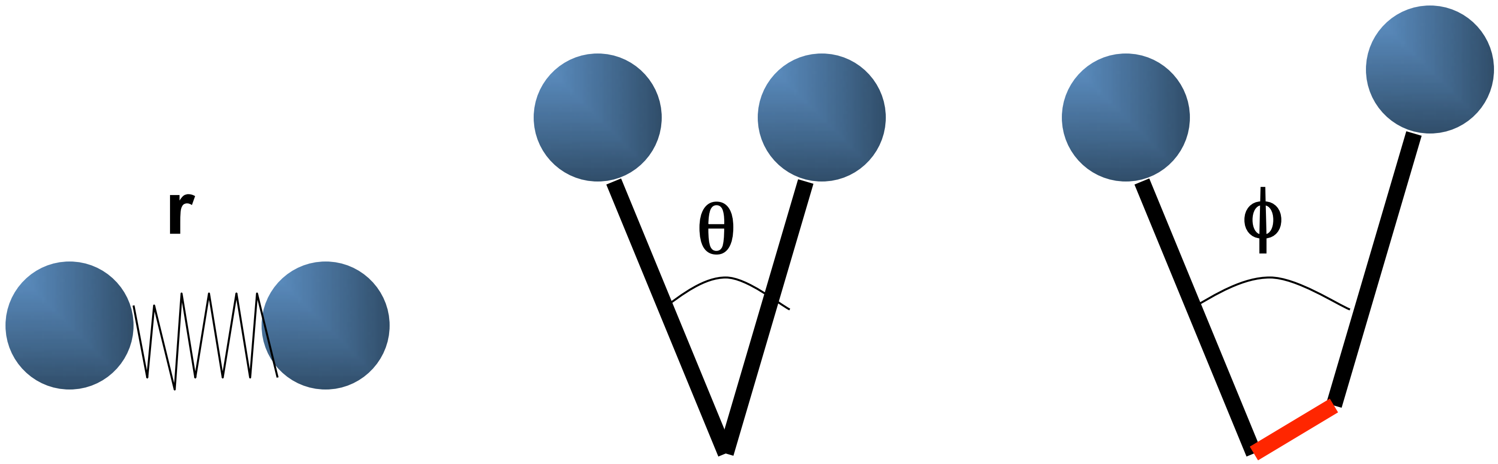
$$\vec{a}(t+\Delta t) = \frac{\vec{f}(t+\Delta t)}{m}$$

Standard Energy Function

$U =$	$K_r(r_i - r_j)^2 +$	Bond length
	$K_\theta(\theta_i - \theta_j)^2 +$	Bond bending
	$K_\phi(1-\cos(n\phi_j))^2 +$	Bond torsion
	$q_i q_j / 4\pi\varepsilon r_{ij} +$	Coulomb
	$A_{ij}/r^6 - B_{ij}/r^{12} +$	van der Waals
	$C_{ij}/r^{10} - D_{ij}/r^{12}$	H-bond

$$-\frac{dU}{d\vec{r}_i} = m_i \frac{d^2\vec{r}_i}{dt^2}$$

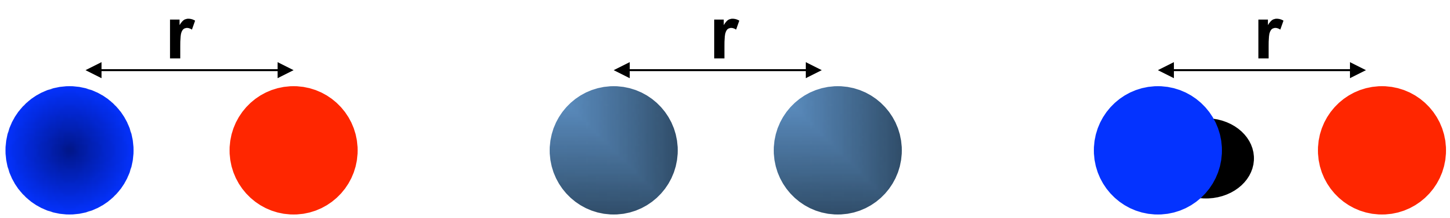
Energy Terms



$K_r(r_i - r_j)^2$ $K_\theta(\theta_i - \theta_j)^2$ $K_\phi(1-\cos(n\phi_j))^2$

Stretching **Bending** **Torsional**

Energy Terms

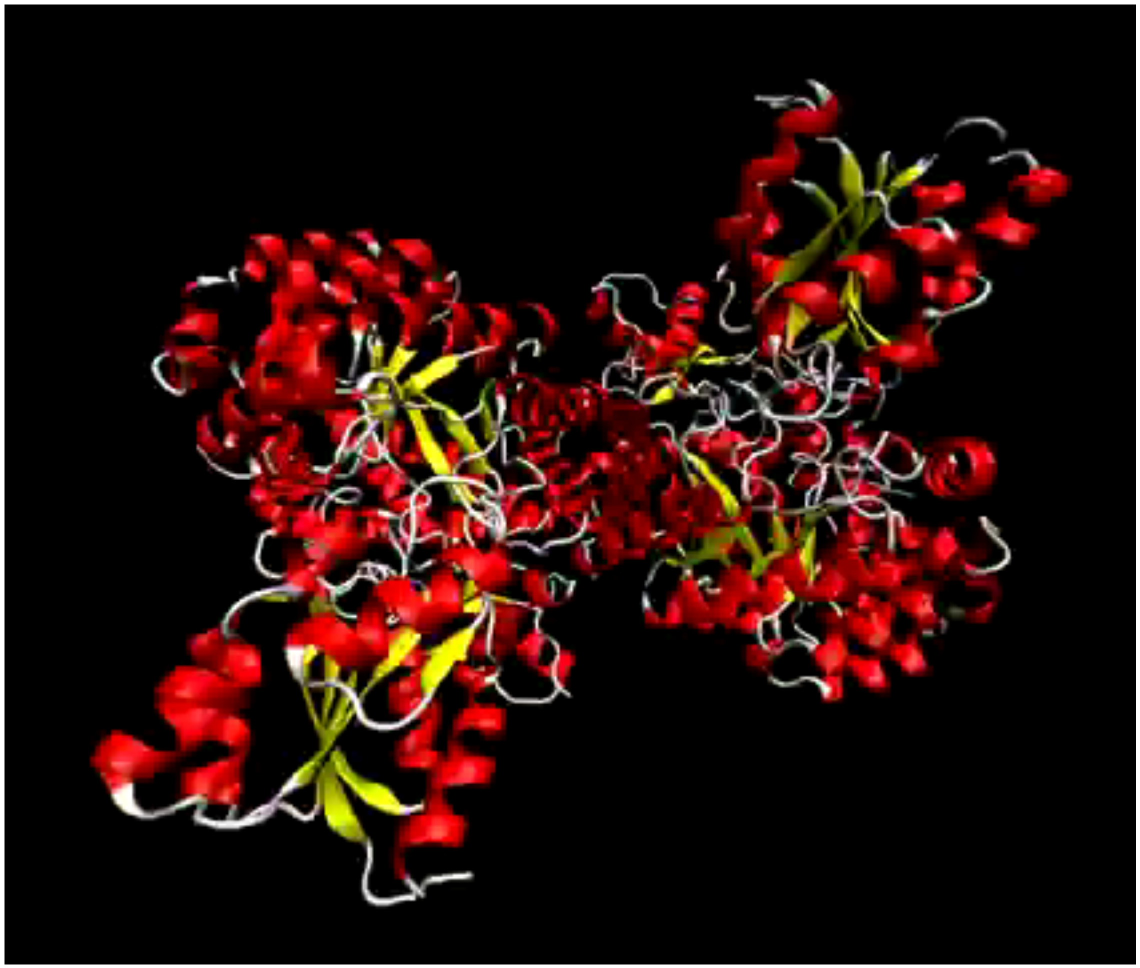


$q_i q_j / 4\pi\varepsilon r_{ij}$ — Coulomb

$A_{ij}/r^6 - B_{ij}/r^{12}$ — van der Waals

$C_{ij}/r^{10} - D_{ij}/r^{12}$ — H-bond

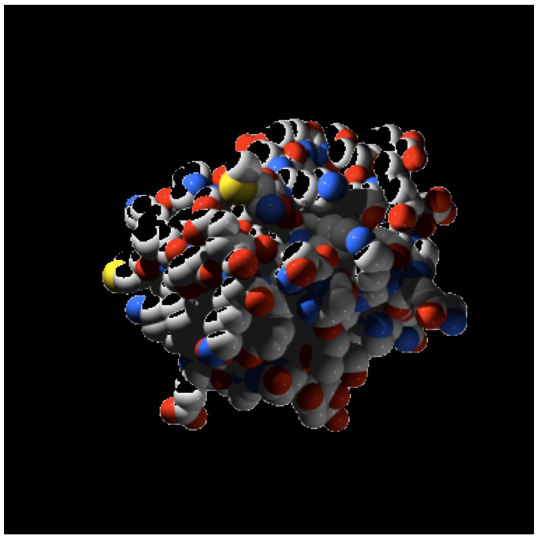
2 ns MD Simulation of A Large Protein



What Can MD Do?

- Allows scientists to visualize motions that cannot be seen with experimental methods
- MD methods are becoming very accurate and allow calculations of micro-scale and macro-scale properties – agreement with experimental data is very good
- More and more scientists are relying on MD simulations to predict or to determine properties that are not available via expt.

3 Ways to Simulate

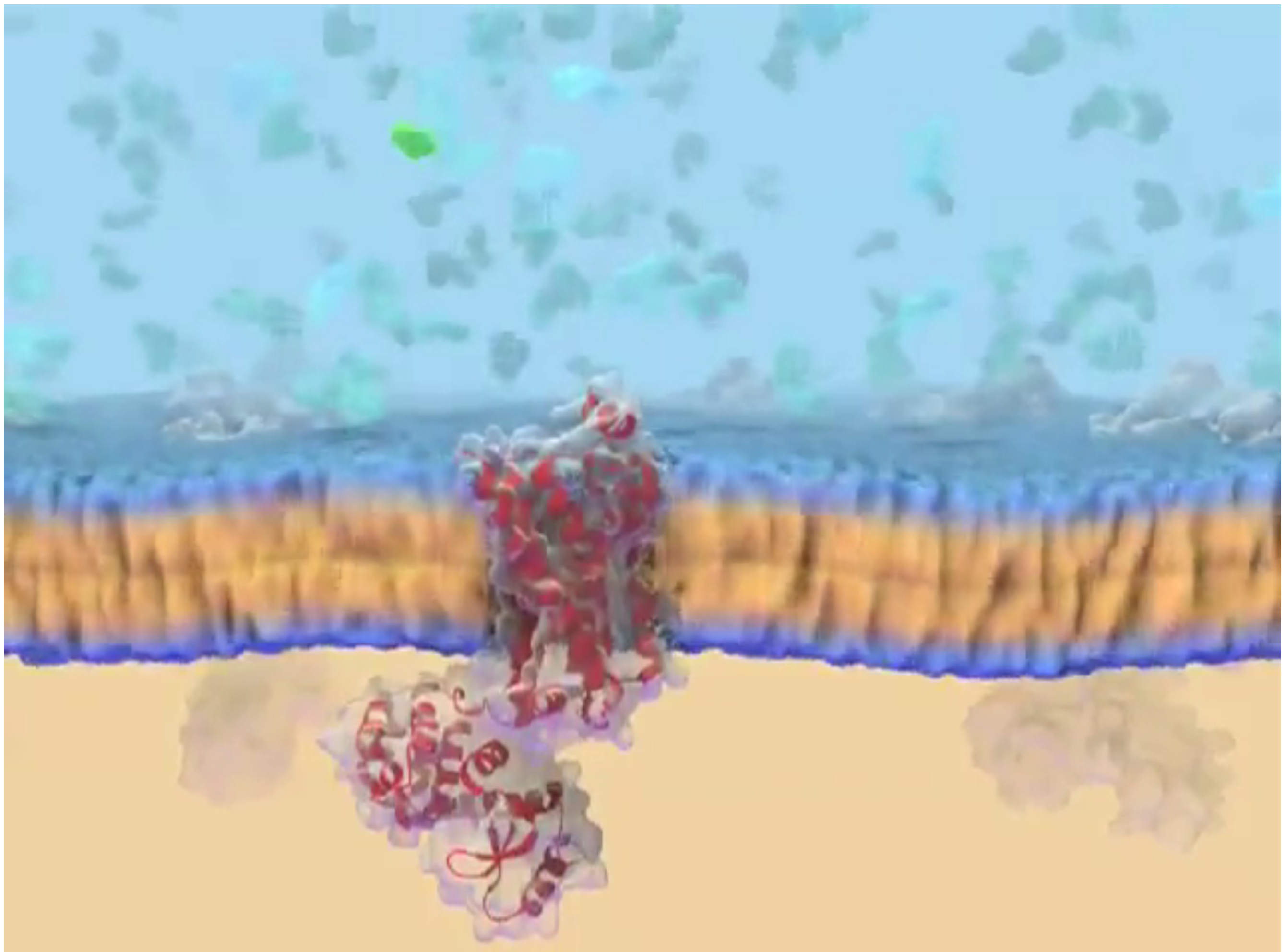


Atomic Scale
0.1 - 1.0 nm
Coordinate data
Dynamic data
0.1 - 10 ns
Molecular dynamics

Meso Scale
1.0 - 10 nm
Interaction data
Kon, Koff, Kd
10 ns - 10 ms
Mesodynamics

Continuum Model
10 - 100 nm
Concentrations
Diffusion rates
10 ms - 1000 s
Fluid dynamics

Mesoscale Simulation



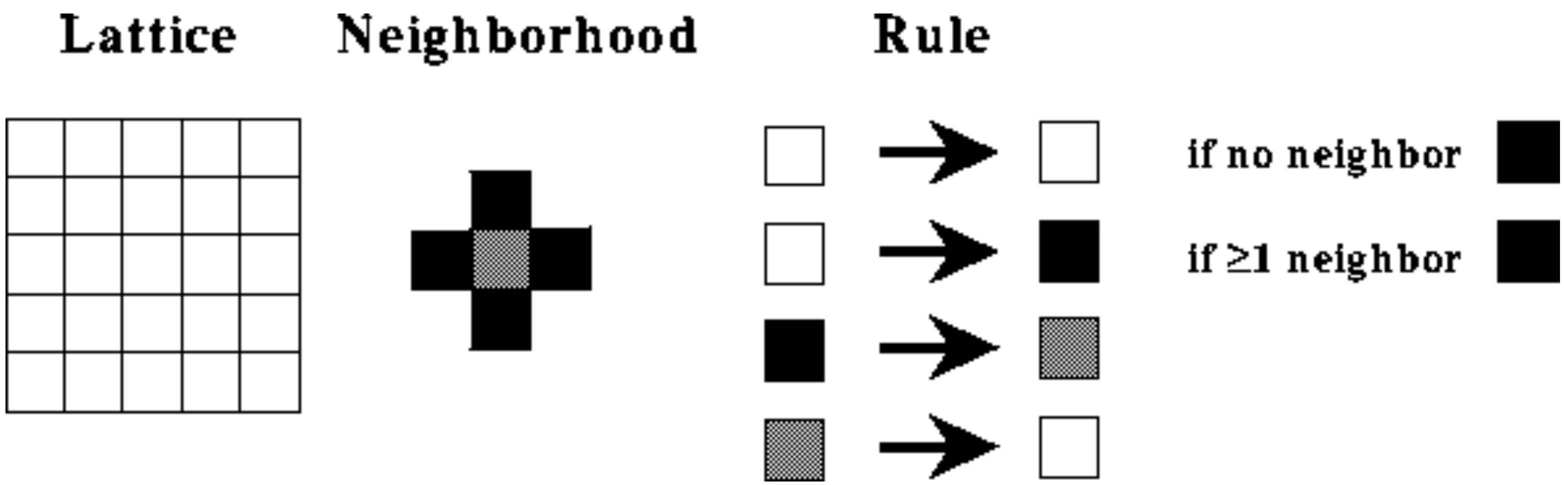
MesoScale Simulation is Tough

- Components are not “atomistic” but are “blobs”
- Blobs don’t have well defined physics or efficient ways of rendering
- Blobs are always diffusing, spinning and rotating randomly through Brownian motion – lots of randomness to motions with very different diffusion constants for different sized blobs
- Solving DE and PDE equations with stochastic effects is very difficult and compute intensive

Cellular Automata

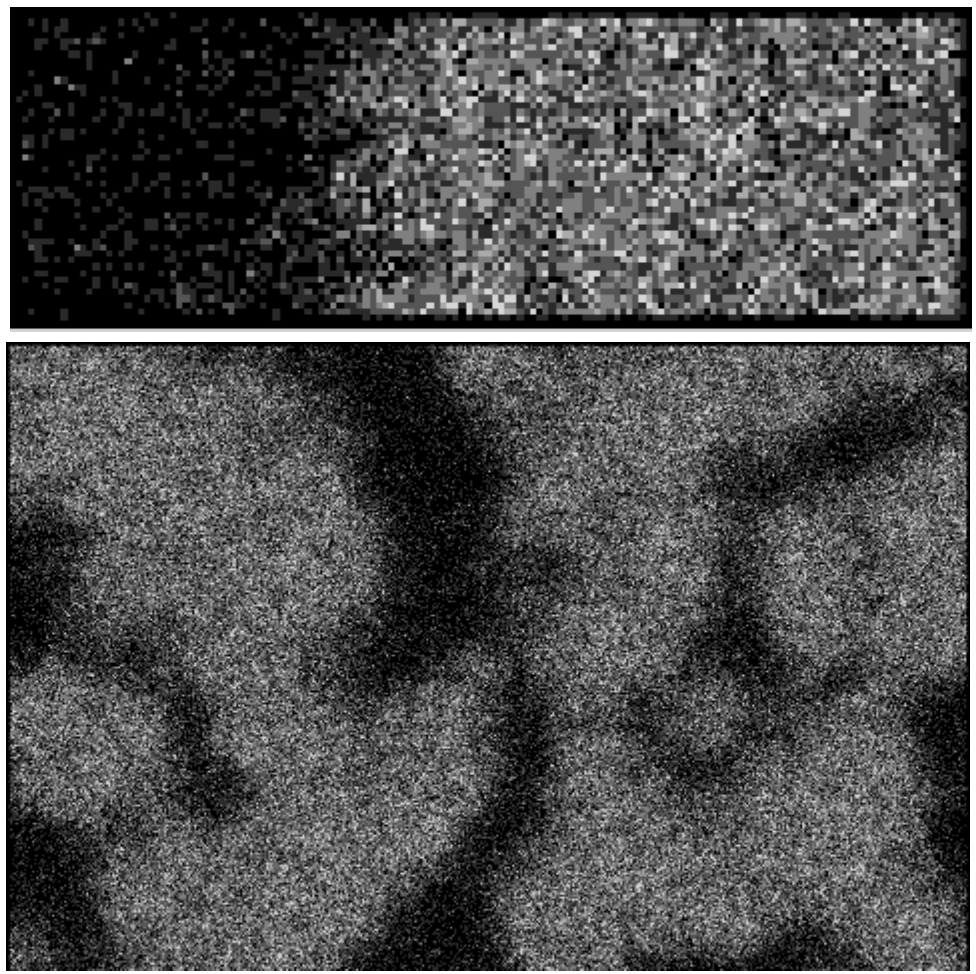
- Computer modelling method that uses lattices and discrete state “rules” to model time dependent processes – a way to animate things
- No differential equations to solve, easy to calculate, more phenomenological
- Simple unit behavior -> complex group behavior
- Used to model fluid flow, percolation, reaction + diffusion, traffic flow, pheromone tracking, predator-prey models, ecology, social nets
- *Scales from 10^{-12} to 10^{+12}*

Cellular Automata

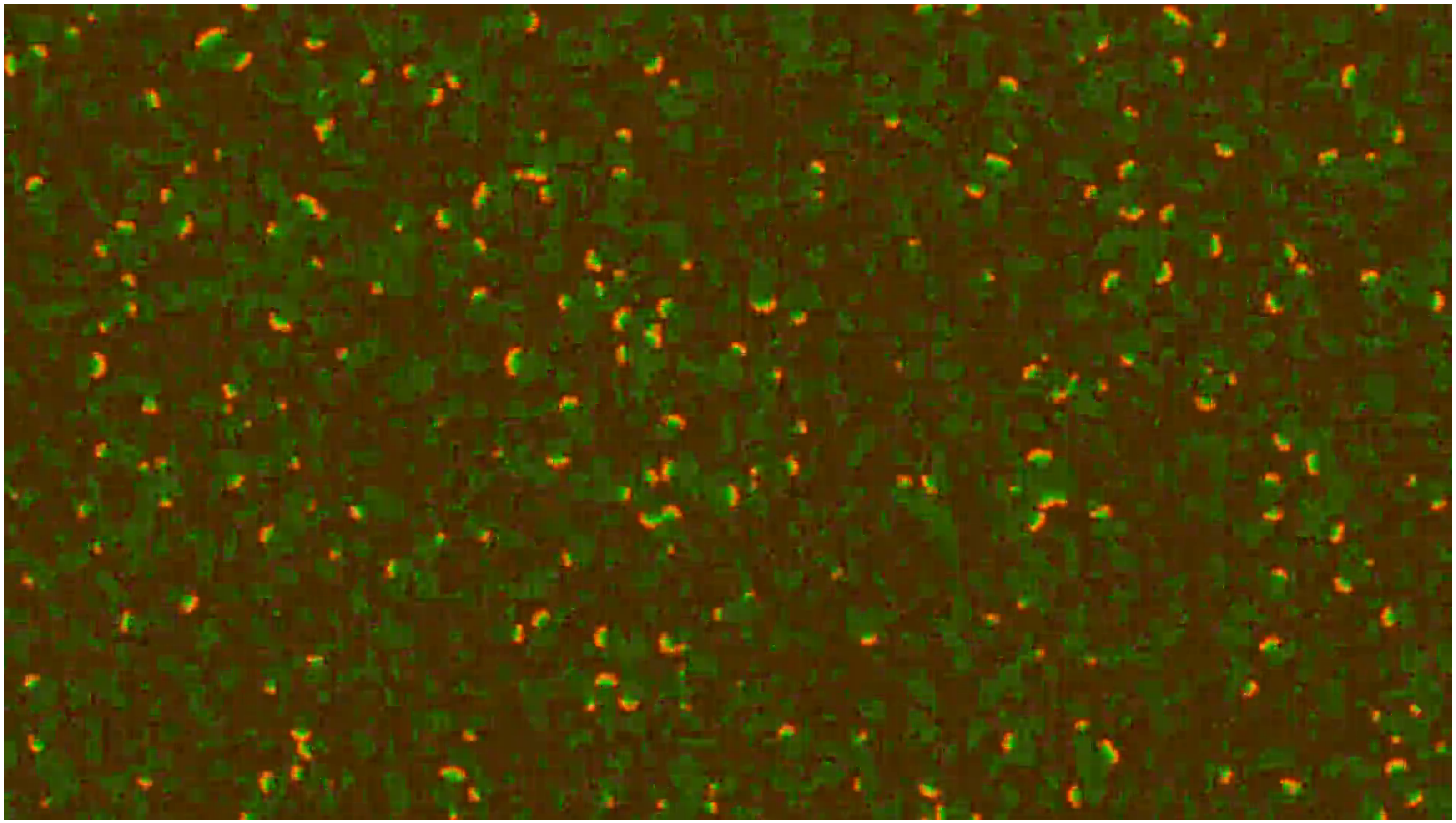


Can be extended to 3D lattice

Reaction/Diffusion with Cellular Automata



Reaction Diffusion Cellular Automata



CA Methods in Games

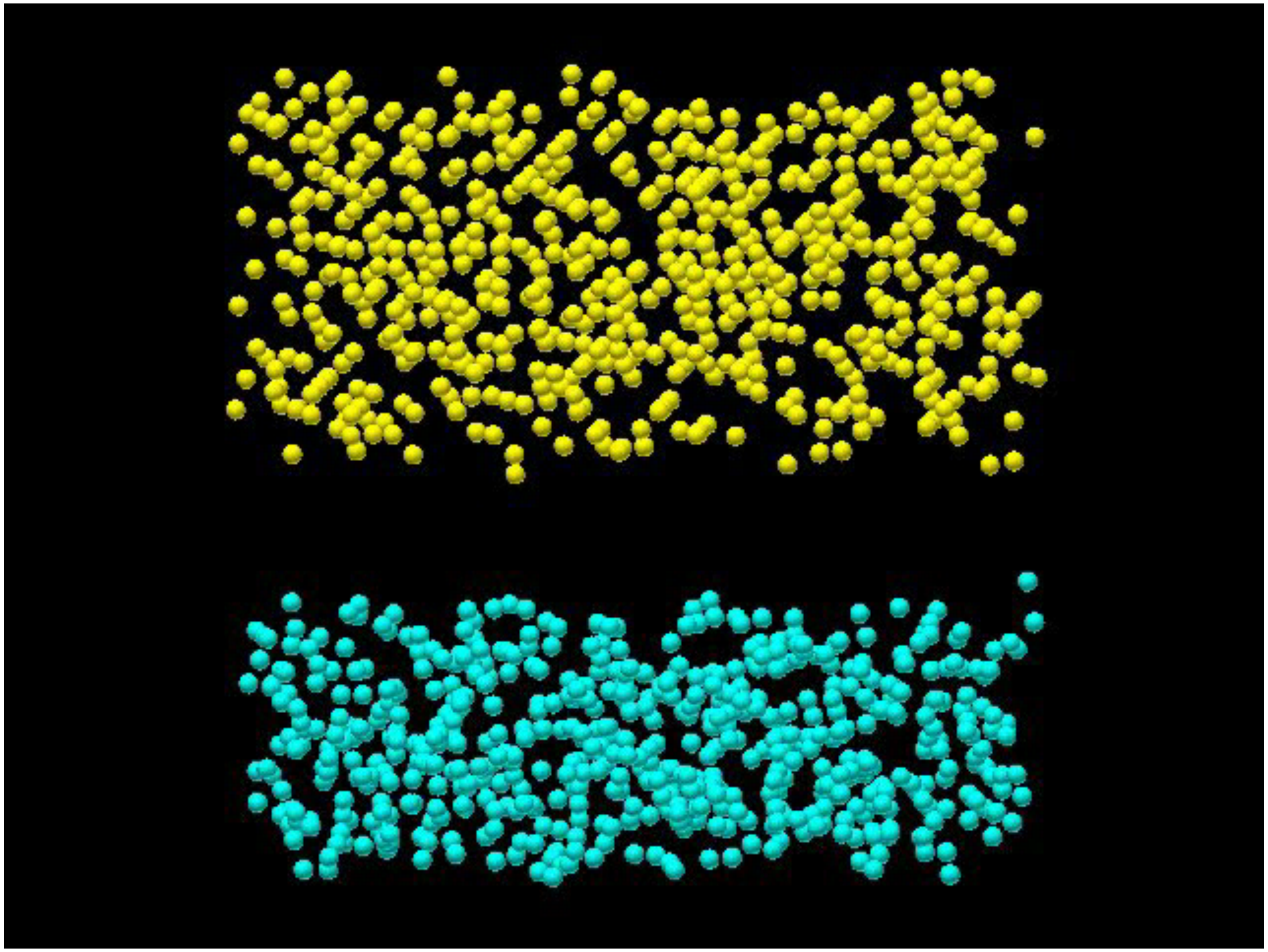
SimCity 2000

The SIMS

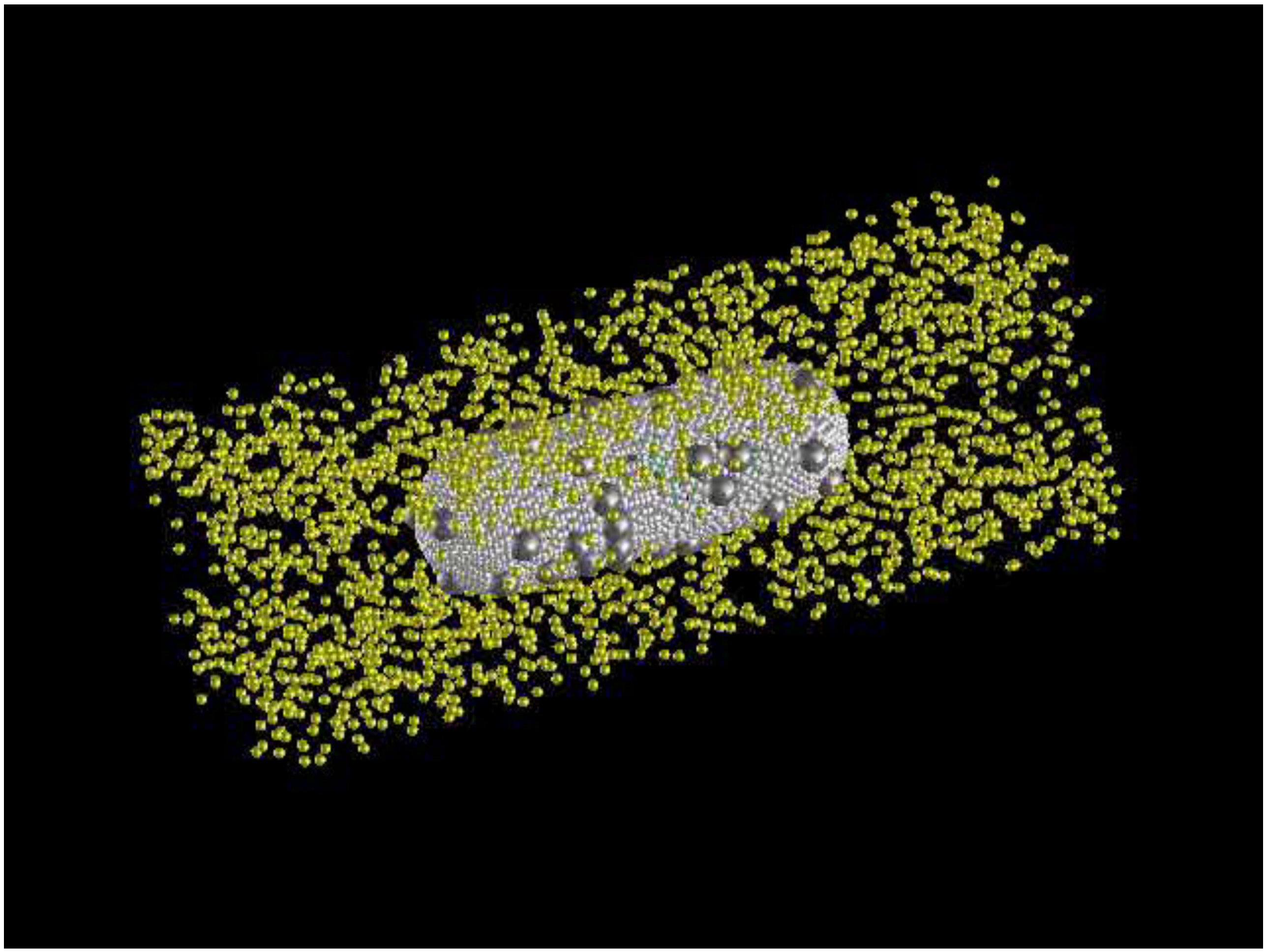
Dynamic Cellular Automata

- A novel method to apply Brownian motion to objects in the Cellular Automata lattice (mimics collisions)
- Takes advantage of the scale-free nature of Brownian motion and the scale-free nature of heterogeneous mixtures to allow simulations to span many orders of time (nanosec to hours) and space (nanometers to meters)

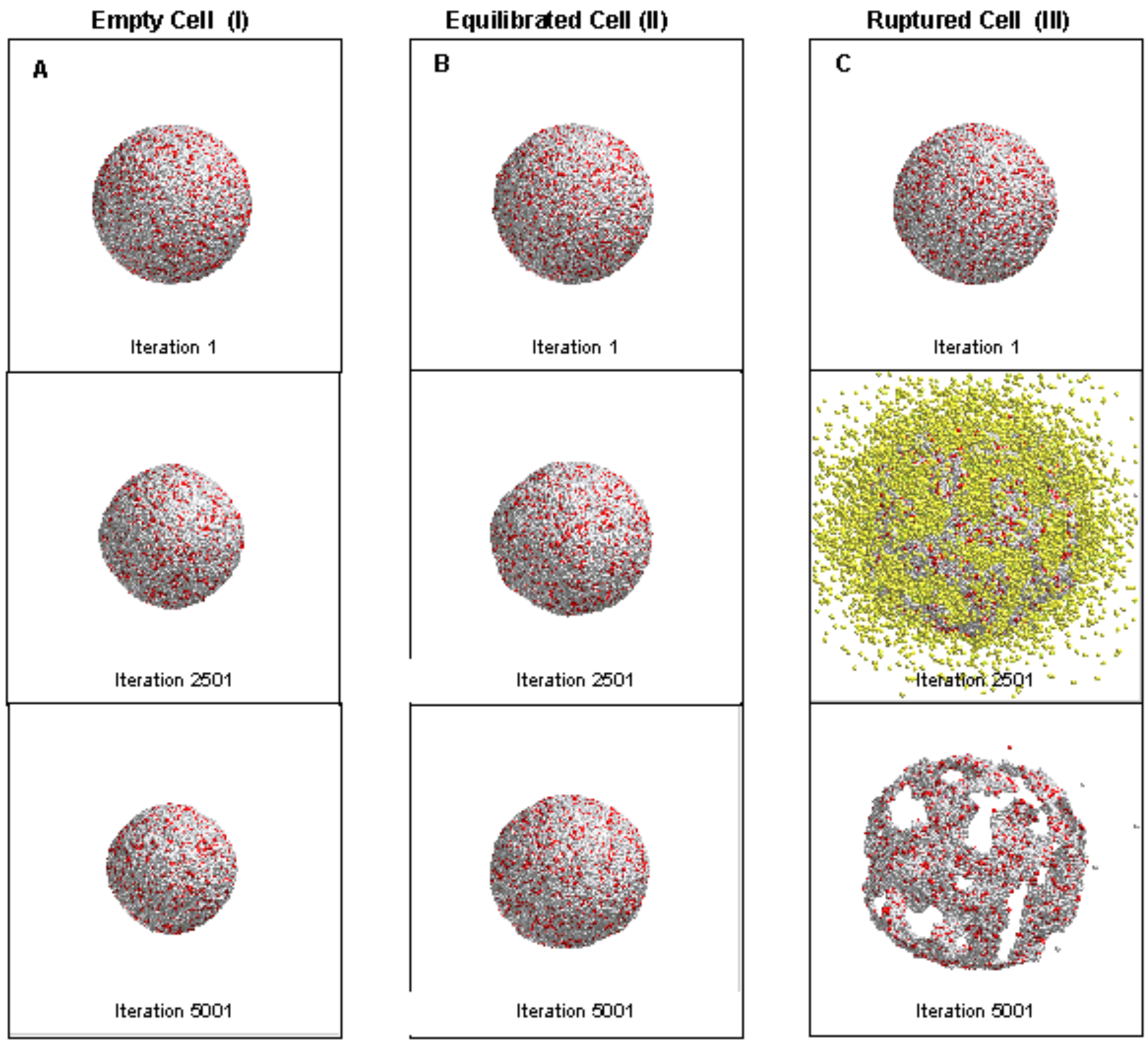
3-D CA of Diffusion + Reaction



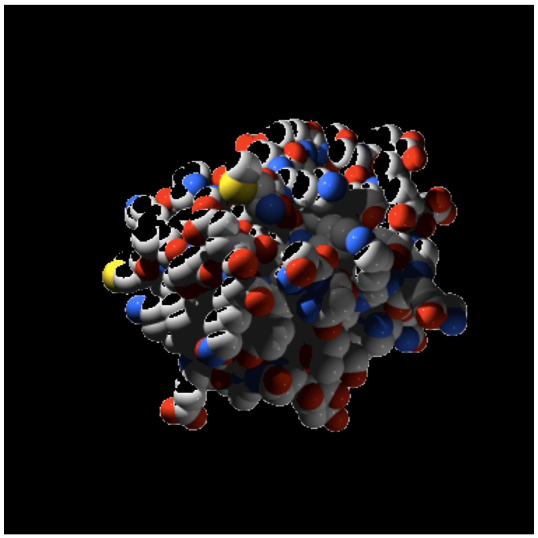
3D-CA Simulations of Transport



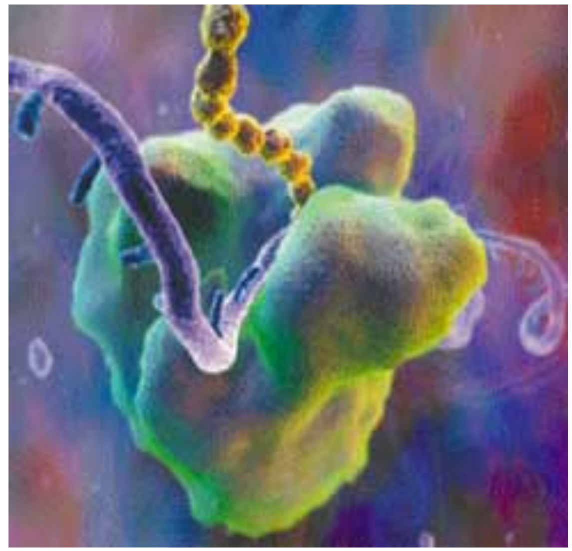
Simulating Membranes & Osmotic Shock



3 Ways to Simulate



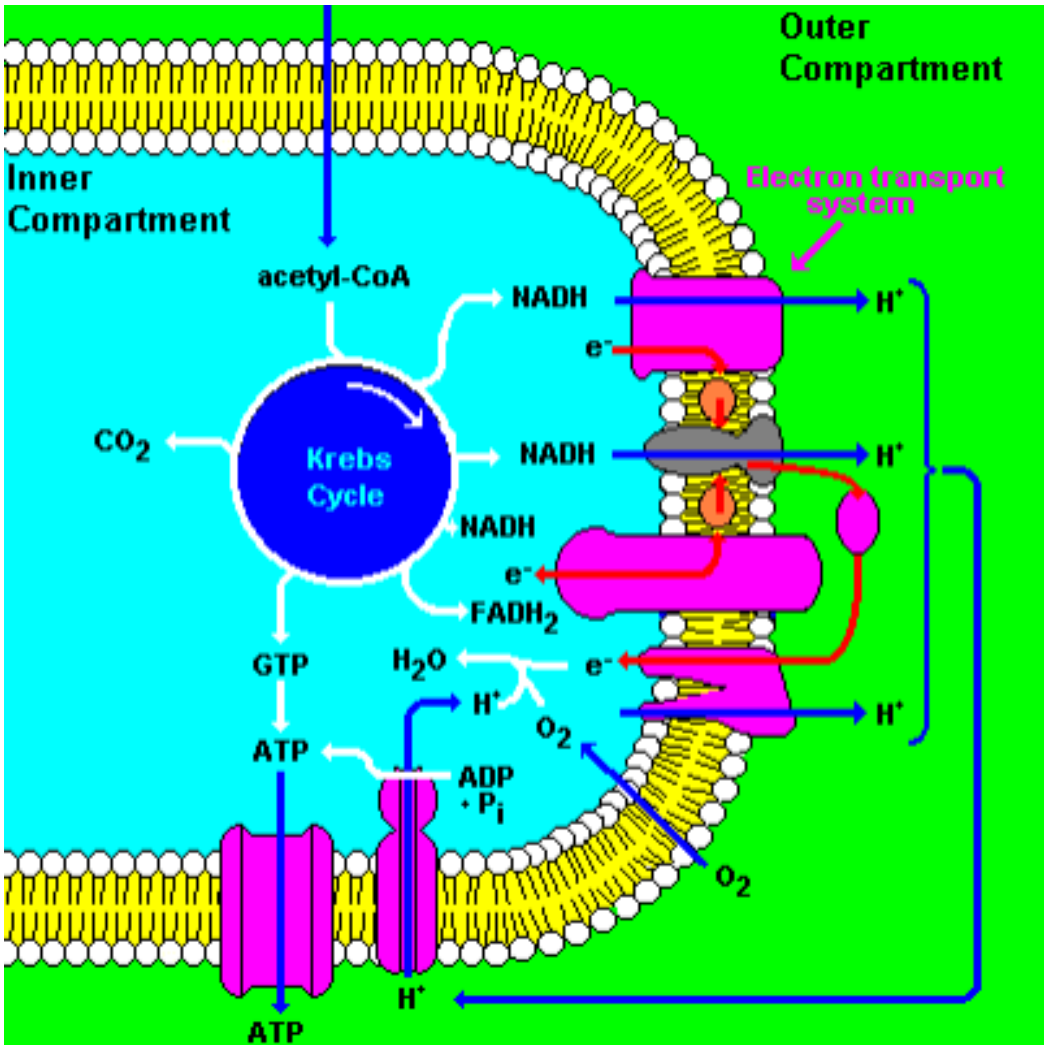
Atomic Scale
0.1 - 1.0 nm
Coordinate data
Dynamic data
0.1 - 10 ns
Molecular dynamics



Meso Scale
1.0 - 10 nm
Interaction data
Kon, Koff, Kd
10 ns - 10 ms
Mesodynamics

Continuum Model
10 - 100 nm
Concentrations
Diffusion rates
10 ms - 1000 s
Fluid dynamics

Cell Simulation with DEs



$$\frac{dx_1}{dt} = k_{11}x_1 + k_{21}x_2 + k_{31}x_3 + \ldots$$

$$\frac{dx_2}{dt} = k_{12}x_1 + k_{22}x_2 + k_{32}x_3 + \ldots$$

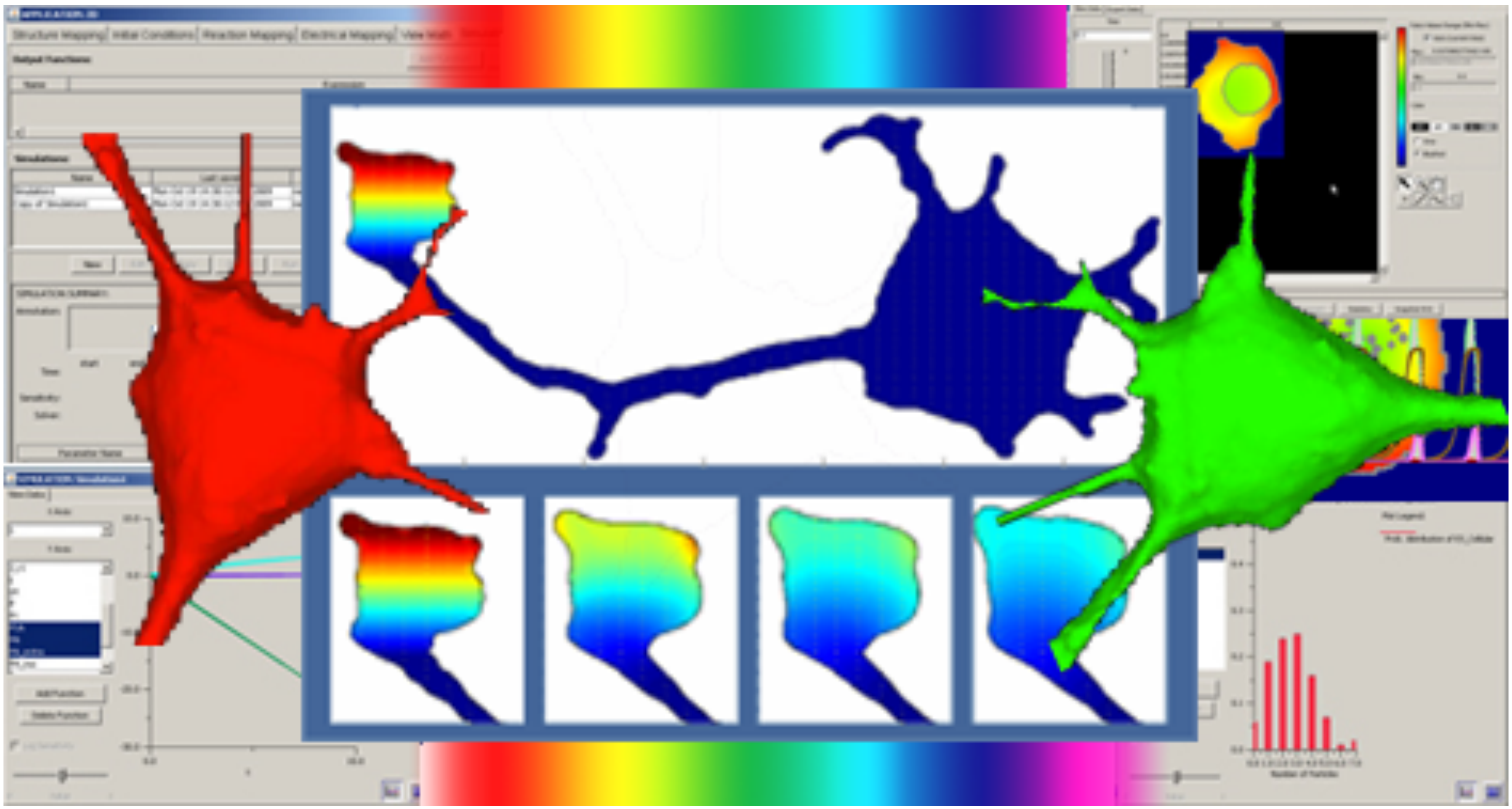
$$\frac{dx_3}{dt} = k_{13}x_1 + k_{23}x_2 + k_{33}x_3 + \ldots$$

$$\frac{dx_4}{dt} = k_{13}x_1 + k_{24}x_2 + k_{34}x_3 + \ldots$$

Continuum Modelling

- Desire to simulate events spatially and temporally (to make movies)
- Use a combination of ordinary differential equations to simulate kinetics and partial differential equations to simulate spatial movements
- Use numerical solvers to solve equations and generate simulation
- Requires user to provide measured parameters from real cells, real metabolites, proteins

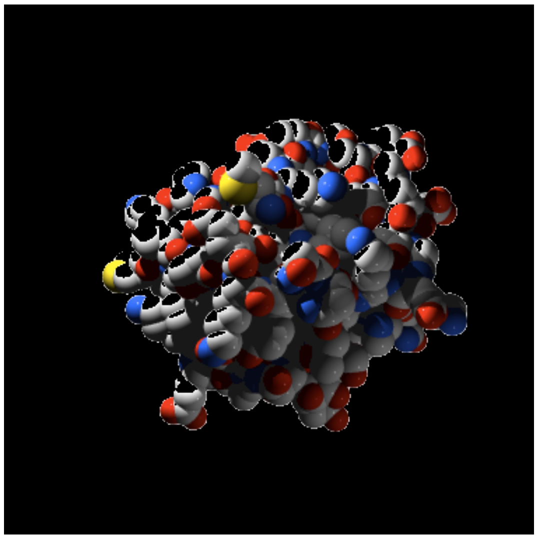
VCell



http://vcell.org

Computer Needs

Atomic



Petaflop computer
Shared Memory or
Grid Computing
Parallelized MD code
HQ 3D Graphics
VR Environment

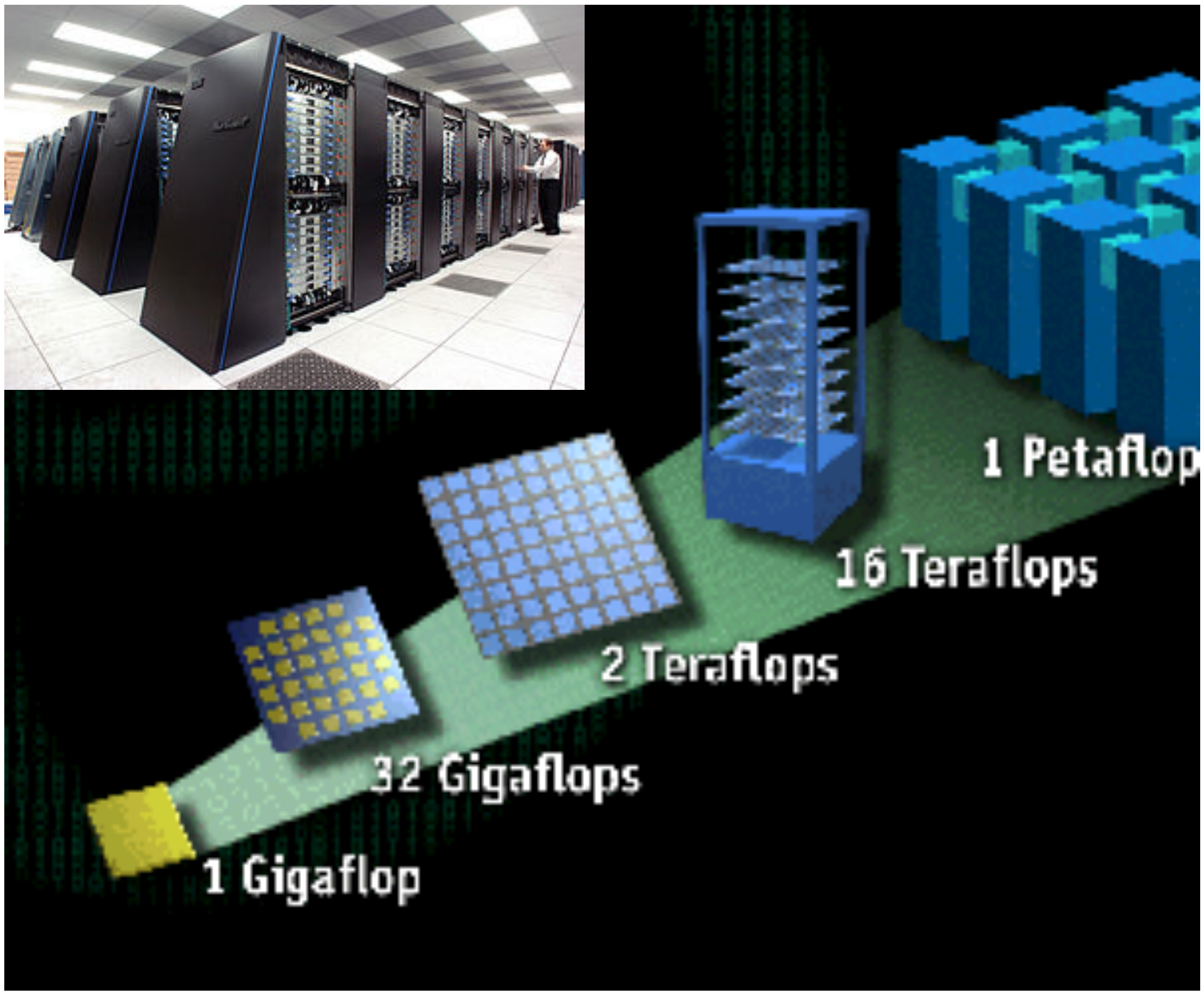
Meso-scale

Teraflop computer
Shared Memory or
Grid Computing
Parallelized MD code
HQ 3D Graphics
VR Environment

Continuum

Gigaflop computer
Shared Memory
Gbytes of RAM
HQ 3D Graphics
VR Environment

BlueGene (600 Teraflops)



Conclusions

- Simulation and modeling is critical to visualize dynamic events that are either too fast or too small to see
- Good simulations are very accurate and very predictive
- Multiple routes to performing molecular or molecule scale modeling – each has their benefits and drawbacks
- Computer scientists often work hand-in-hand with other scientists to perform these challenging tasks