Simulating ribosome biogenesis in replicating whole cells

Tyler Earnest / PI: Zan Luthey-Schulten (UIUC)
NCSA Blue Waters Symposium for Petascale Science and Beyond: June 15, 2016
Multi-scale modeling in biology

Timestep:
- Femtoseconds
- Picoseconds
- Microseconds

MD: Potential-based, all atom

BD: Potential-based, coarse grained

RDME: Probability-based

Atomic interactions
Rigid body interactions
Diffusion probabilities
Reaction probabilities

Molecules to macromolecular assemblies
Whole cells and colonies
Multi-scale modeling in biology
Whole-cell modeling with Lattice Microbes
Whole-cell modeling with Lattice Microbes
Whole-cell modeling with Lattice Microbes

Discretize to lattice
Whole-cell modeling with Lattice Microbes

Discretize to lattice
Whole-cell modeling with Lattice Microbes

Discretize to lattice

Probability of the state of the cell:

Reaction-diffusion master equation

$$\frac{dP(x, t)}{dt} = \sum_v \sum_r \left[ -a_r(x_v)P(x_v, t) + a_r(x_v - S_r)P(x_v - S_r, t) \right] + \sum_v \sum_{\xi} \sum_{\alpha} \left[ -d^{\alpha}x^{\alpha}_v P(x, t) + d^{\alpha}(x^{\alpha}_{v+\xi} + 1)P(x + 1^{\alpha}_{v+\xi} - 1^{\alpha}_v, t) \right]$$
Discretize to lattice

\[ x(t) = \begin{pmatrix}
  x_{1,1,1}(t) \\
  x_{1,1,2}(t) \\
  \vdots \\
  x_{1,1,nz}(t) \\
  \vdots \\
  x_{1,ny,nz}(t) \\
  \vdots \\
  \vdots \\
  x_{nx,ny,nz}(t)
\end{pmatrix} \]

Rate of the cell:

\[ \frac{dP(x,t)}{dt} = V_x v_R x_r \cdot a_r(x_v)P(x_v,t) + a_r(x_v - S_r)P(x_v - S_r,t) \]

\[ + V_x v_{\hat{i},\hat{j},\hat{k}} \cdot \nabla P(x,t) + d_R(x_v \cdot \nabla + 1)P(x + 1_v + \xi - 1_v, t) \]

Whole-cell modeling with Lattice Microbes
Whole-cell modeling with Lattice Microbes

Discretize to lattice

State of the cell:

\[
x(t) = \begin{pmatrix}
x_{1,1,1}(t) \\
x_{1,1,2}(t) \\
\vdots \\
x_{1,1,n_z}(t) \\
\vdots \\
x_{1,n_y,n_z}(t) \\
\vdots \\
x_{n_x,n_y,n_z}(t)
\end{pmatrix}
\]

\[
x_{i,j,k}(t) = \begin{pmatrix}
x_{1}(t) \\
x_{2}(t) \\
\vdots \\
x_{n_{sp}}(t)
\end{pmatrix}
\]

\[
- d^{\alpha} x_{v}^\alpha P(x, t) + 1) P(x + 1_{v+\xi} - 1_{v}, t)
\]
Whole-cell modeling with Lattice Microbes

Designed for CUDA

Whole-cell modeling with Lattice Microbes

Designed for CUDA

Multiple levels of parallelism
- GPGPU  Roberts et al. *IPDPS* 2009
- MPI  (w.i.p.)
Whole-cell modeling with Lattice Microbes

Designed for CUDA

Multiple levels of parallelism
  GPGPU  Roberts et al. *IPDPS* 2009
  MPI  (w.i.p.)

Extensible through Python  Peterson et al. *PyHPC* 2013
Ribosome biogenesis in *Escherichia coli*

- Protein synthesis
Ribosome biogenesis in *Escherichia coli*

- Protein synthesis
- 1/4 of the cellular dry mass
Ribosome biogenesis in *Escherichia coli*

- Protein synthesis
- 1/4 of the cellular dry mass
- Requires understanding and modeling
  - Gene regulation
  - Transcription
  - Translation

Large subunit (LSU)

Small subunit (SSU)
Ribosome biogenesis in *Escherichia coli*

- Protein synthesis
- 1/4 of the cellular dry mass
- Requires understanding and modeling
  - Gene regulation
  - Transcription
  - Translation
- Model serves as the foundation for more complete cell models
Complexity of SSU assembly

16S rRNA + 20 r-protein

30S small subunit
Complexity of SSU assembly

16S rRNA + 20 r-protein

30S small subunit
Complexity of SSU assembly

$2^{20} \ (10^6)$ intermediate species; $20! \ (10^{18})$ reactions; $20! \ (10^{18})$ parameters
Kinetic model of small subunit assembly

Reactions

Rates
Kinetic model of small subunit assembly

Held, ..., Nomura. JBC (1974)
Kinetic model of small subunit assembly

Mulder, ..., Williamson Science (2010)
Kinetic model of small subunit assembly
Kinetic model of small subunit assembly

Earnest, ..., ZLS

Biophys. J. (2015)
# Reactions involved in ribosome biogenesis

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assembly</strong></td>
<td><strong>Sx + I_i → I_{i+1}</strong></td>
</tr>
<tr>
<td><strong>Transcription: rRNA</strong></td>
<td>DNA → DNA + rRNA</td>
</tr>
<tr>
<td><strong>Transcription: mRNA</strong></td>
<td>DNA → DNA + mRNA</td>
</tr>
<tr>
<td><strong>Degradation</strong></td>
<td>mRNA → ∅</td>
</tr>
<tr>
<td><strong>Translation</strong></td>
<td>30S + mRNA + 50S → 30S + mRNA + 50S + n Protein</td>
</tr>
<tr>
<td><strong>Diffusion</strong></td>
<td>$X_i(x) → X_i(x + δ_i)$</td>
</tr>
</tbody>
</table>
Operons are localized to sites within the nucleoid

rrn Operon
r-prot Operon
mRNA are emitted from r-protein operons
Particles diffuse by hopping to nearest neighbor lattice sites
Translation: mRNA associates with SSU
Translation: SSU:mRNA associates with LSU
Translation: This forms the translating ribosome.
Translation: Proteins are emitted from the ribosome sequentially using transcript derived rates
Translation: Ribosome dissociates into mRNA, SSU, and LSU in a single step.
Assembly: **16S rRNA** is produced at the site of **rRNA operons**
Assembly: proteins bind to the assembly intermediate in an order defined by the Nomura map
Cell division and DNA replication
Measure cell cycle parameters

- *Escherichia coli* K-12 MG1655 \( \Delta lac \)
  - 120 minute doubling time
Measure cell cycle parameters

- *Escherichia coli* K-12 MG1655 Δlac - 120 minute doubling time
- Construct 14 strains fluorescently labeling a different locus in the genome

Earnest, ..., Kuhlman, ZLS *Biopolymers* (2016) Accepted
Measure cell cycle parameters

- *Escherichia coli* K-12 MG1655 Δlac
  - 120 minute doubling time

- Construct 14 strains fluorescently labeling a different locus in the genome

- Measure cell lengths and count fluorescent foci

Earnest, …, Kuhlman, ZLS *Biopolymers* (2016) Accepted
Measure cell cycle parameters

- *Escherichia coli* K-12 MG1655 ∆lac
  - 120 minute doubling time

- Construct 14 strains fluorescently labeling a different locus in the genome

- Measure cell lengths and count fluorescent foci

- Fit to probabilistic model to get cell cycle parameters and mean cell size at division

\[
P(\ell, n|\text{gene}; \theta) = \begin{cases} 
\mu, \sigma & \text{length at division} \\
\mu, \sigma & \text{time to begin replication} \\
\mu, \sigma & \text{duration of replication} 
\end{cases}
\]

Earnest, …, Kuhlman, ZLS *Biopolymers* (2016) Accepted
Measure cell cycle parameters

- *Escherichia coli* K-12 MG1655 Δlac - 120 minute doubling time

- Construct [gene; ✓] labeling a different locus in the genome

- Measure cell lengths and count fluorescent foci

- Fit to probabilistic model to get cell cycle parameters and mean cell size at division

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{\text{len0}}$</td>
<td>Mean length at division</td>
<td>$4.772 \pm 0.021 , \mu m$</td>
</tr>
<tr>
<td>$\sigma_{\text{len0}}$</td>
<td>SD length at division</td>
<td>$0.744 \pm 0.024 , \mu m$</td>
</tr>
<tr>
<td>$\mu_{\text{trep}}$</td>
<td>Mean replication initiation time</td>
<td>$42.2 \pm 3.0 , \text{min}$</td>
</tr>
<tr>
<td>$\sigma_{\text{trep}}$</td>
<td>SD replication initiation time</td>
<td>$22.1 \pm 1.9 , \text{min}$</td>
</tr>
<tr>
<td>$T_{\text{rep}}$</td>
<td>Replication duration</td>
<td>$42.2 \pm 5.0 , \text{min}$</td>
</tr>
</tbody>
</table>

Earnest, …, Kuhlman, ZLS Biopolymers (2016) Accepted
Cell geometry

Earnest, ..., Kuhlman, ZLS Biopolymers (2016) Accepted
Cell geometry

Earnest, ..., Kuhlman, ZLS. *Biopolymers* (2016) Accepted

Cell geometry

Earnest, ..., Kuhlman, ZLS Biopolymers (2016) Accepted
Results

Earnest, ..., Kuhlman, ZLS Biopolymers (2016) Accepted
Results

Ribosome abundance

Earnest, ..., Kuhlman, ZLS Biopolymers (2016) Accepted
Results

Total protein abundance

Copy number

r-Protein operons

Concentration/10^{-6} M

Earnest, ..., Kuhlman, ZLS Biopolymers (2016) Accepted
Results

Earnest, …, Kuhlman, ZLS Biopolymers (2016) Accepted
Results

Earnest, ..., Kuhlman, ZLS Biopolymers (2016) Accepted
Results

All 30S intermediates

Earnest, …, Kuhlman, ZLS Biopolymers (2016) Accepted
Benchmarks: Lattice Microbes v2.3a

New reaction kernel: Hallock & ZLS. HiCOMB 2016
Benchmarks: Lattice Microbes v2.3a

Reaction/diffusion/growth

Walltime per timestep/ms

- Standard Rxn Kernel, 8 pps
- Standard Rxn Kernel, 16 pps
- Precomputed propensities, 8 pps
- Precomputed propensities, 16 pps

New reaction kernel: Hallock & ZLS. HiCOMB 2016
Benchmarks: Lattice Microbes v2.3a

Reaction/diffusion/growth

- Standard Rxn Kernel, 8 pps
- Standard Rxn Kernel, 16 pps
- Precomputed propensities, 8 pps
- Precomputed propensities, 16 pps

New reaction kernel: Hallock & ZLS. HiCOMB 2016
Outlook

E. coli whole-cell model
- Regulation of ribosomal gene expression
- Explicit treatment of DNA
- Metabolism
Outlook

**E. coli whole-cell model**
- Regulation of ribosomal gene expression
- Explicit treatment of DNA
- Metabolism

**Lattice Microbes**
- Multi-scale modeling: high concentration and/or fast diffusion rates
- MPI: Eukaryotic cells, bacterial colonies, tissues *(BW necessary!)*
- IPython/Jupyter notebook environment
Acknowledgements

Zan Luthey-Schulten
Tom Kuhlman
Mike Hallock
Joe Peterson
John Cole

Lattice Microbes: http://goo.gl/akkSyg