The Molecular Dynamics Simulation Process

For textbooks see:


More at http://www.biomath.nyu.edu/index/course/99/textbooks.html

---

Classical Dynamics

*at 300K*

Energy function: \( U(\vec{r}_1, \vec{r}_2, \ldots \vec{r}_N) = U(\vec{R}) \)

used to determine the force on each atom:

\[
    m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i = -\vec{\nabla} U(\vec{R})
\]

yields a set of 3N coupled 2\textsuperscript{nd}-order differential equations that can be propagated forward (or backward) in time.

Initial coordinates obtained from crystal structure, velocities taken at random from Boltzmann distribution.

Maintain appropriate temperature by adjusting velocities.
Classical Dynamics

discretization in time for computing

\[ m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i = -\nabla U(\vec{R}) \]

Use positions and accelerations at time \( t \) and the positions from time \( t - \delta t \) to calculate new positions at time \( t + \delta t \).

\[
\begin{align*}
\mathbf{r}(t + \delta t) &\approx \mathbf{r}(t) + \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2 \\
\mathbf{r}(t - \delta t) &\approx \mathbf{r}(t) - \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2
\end{align*}
\]

“Verlet algorithm”

\[
\mathbf{r}(t + \delta t) \approx 2\mathbf{r}(t) - \mathbf{r}(t - \delta t) + \mathbf{a}(t)\delta t^2
\]

Potential Energy Function of Biopolymer

- Simple, fixed algebraic form for every type of interaction.
- Variable parameters depend on types of atoms involved.

\[
U(\vec{R}) = \sum_{\text{bonds}} k_i^{\text{bond}} (r_i - r_0)^2 + \sum_{\text{angles}} k_i^{\text{angle}} (\theta_i - \theta_0)^2 + \sum_{\text{dihedrals}} k_i^{\text{dihedral}} [1 + \cos (n_i \phi_i + \delta_i)] + \sum_i \sum_{j \neq i} 4\epsilon_{ij} \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 + \sum_i \sum_{j \neq i} q_i q_j \epsilon_{ij}
\]

heuristic from physics
Improving the Force Field

• Atomic polarizability increases computation by 2x…
• …but, the additional computations are perfectly suited to the GPU!
• For now, NAMD calculates atomic polarizability on CPUs only...soon we will also use GPUs

Molecular Dynamics Ensembles

Constant energy, constant number of particles (NE)

Constant energy, constant volume (NVE)

Constant temperature, constant volume (NVT)

Constant temperature, constant pressure (NPT)

Choose the ensemble that best fits your system and start the simulations, but use NE to check on accuracy of the simulation.
Langevin Dynamics
for temperature control

Langevin dynamics deals with each atom separately, balancing a small friction term with Gaussian noise to control temperature:

\[ m \ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r}) - \gamma m \dot{\mathbf{r}} + \mathbf{R}(t) \]

\[ \langle \mathbf{R}(t) \cdot \mathbf{R}(t') \rangle = 6k_B T \gamma \delta(t - t') \]

Langevin Dynamics
for pressure control

**Underlying Langevin-Hoover barostat equation for all atoms:**

Equations solved numerically in NAMD

\[ \frac{d^2 V(t)}{dt^2} = \frac{1}{W_{bs}} [P(t) - P_{\text{target}}] - \frac{1}{\tau_{bs}} \frac{dV(t)}{dt} + R_{bs}(t) \]

\[ P = \rho k_B T + \frac{1}{V_d} \sum_{i<j} \langle r_{ij} \frac{dU_{\text{tot}}(r_{ij})}{dr_{ij}} \rangle \quad d = \text{dimension} \]

\[ \langle R_{bs}(t) R_{bs}(t') \rangle = \frac{2k_B T_{\text{target}} \delta(t - t')}{W_{bs} \tau_{bs}} \quad W_{bs} = d N_{\text{atoms}} k_B T_{\text{target}} \tau_{\text{period}}^2 \]

\[ \dot{\mathbf{r}}_i = \mathbf{v}_i + s \mathbf{r}_i \quad \dot{\mathbf{v}}_i = \mathbf{F}_i / m_i - s \mathbf{v}_i \]

\[ \dot{V} = dV_s \quad \dot{s} = dV (P - P_{\text{target}}) / W - s / \tau_{bs} + R(t) \]

\[ d - \text{dimension} \]
Large is no problem. But …

Molecular dynamics simulation of alpha-hemolysin with about 300,000 atoms; 10 million atom simulations are routine today, 200 million atom simulations are possible.

But long is still a problem!

**Biomolecular timescale and timestep limits**

<table>
<thead>
<tr>
<th>Timescale</th>
<th>Timestep (steps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>$10^{15}$</td>
</tr>
<tr>
<td>ms</td>
<td>$10^{12}$</td>
</tr>
<tr>
<td>µs</td>
<td>$10^{9}$</td>
</tr>
<tr>
<td>ns</td>
<td>$10^{6}$</td>
</tr>
<tr>
<td>ps</td>
<td>$10^{3}$</td>
</tr>
<tr>
<td>fs</td>
<td>$10^{0}$</td>
</tr>
</tbody>
</table>

- Bond stretching
- Molecular dynamics timestep
- Rotation of buried sidechains
- Local denaturations
- Allosteric transitions
- Small protein folding
- Hinge bending
- Rotation of surface sidechains
- Elastic vibrations

**SPEED LIMIT**

$\delta t = 1 \text{ fs}$

(NSF center, Shaw Res.)
Preparing Your System for MD Solvation

Biological activity is the result of interactions between molecules and occurs at the interfaces between molecules (protein-protein, protein-DNA, protein-solvent, DNA-solvent, etc).

Why model solvation?
• many biological processes occur in aqueous solution
• solvation effects play a crucial role in determining molecular conformation, electronic properties, binding energies, etc

How to model solvation?
• explicit treatment: solvent molecules are added to the molecular system
• implicit treatment: solvent is modeled as a continuum dielectric or so-called implicit force field
Preparing Your System for MD Solvation

Biological activity is the result of interactions between molecules and occurs at the interfaces between molecules (protein-protein, protein-DNA, protein-solvent, DNA-solvent, etc).

Why model solvation?
• many biological processes occur in aqueous solution
• solvation effects play a crucial role in determining molecular conformation, electronic properties, binding energies, etc

How to model solvation?
• explicit treatment: solvent molecules are added to the molecular system
• implicit treatment: solvent is modeled as a continuum dielectric

From the Mountains to the Valleys

*how to actually describe a protein*

- Initial coordinates have bad contacts, causing high energies and forces (due to averaging in observation, crystal packing, or due to difference between theoretical and actual forces)

Minimization finds a nearby local minimum.

Heating and cooling or equilibration at fixed temperature permits biopolymer to escape local minima with

Initial dynamics samples thermally accessible states.
Longer dynamics access other intermediate states; one may apply external forces to access other available states in a more timely manner.

VMD-NAMD plugin
Keith