Molecular mechanics

Molecular mechanics deals with the static features of biomolecular systems at \( T=0 \) K, that is, particles do not have any kinetic energy. [N.B. in molecular dynamics simulations, particles have an average kinetic energy of \((3/2)kT\), which is substantial at room temp., \( T=300 \) K] Thus the energy is given solely by the potential energy of the system.

Two important applications of molecular mechanics are:

1. Energy minimization (geometry optimization):
   Find the coordinates for the configuration that minimizes the potential energy (local or absolute minimum).

2. Normal mode analysis:
   Find the vibrational excitations of the system built on the absolute minimum using the harmonic approximation.
Energy minimization

The initial configuration of a biomolecule — whether experimentally determined or otherwise — does not usually correspond to the minimum of the potential energy function. It may have large strains in the backbone that would take a long time to equilibrate in MD simulations, or worse, it may have bad contacts (i.e. overlapping van der Waals radii), which would crash the MD simulation in the first step! To avoid such mishaps, it is a standard practice to minimize the energy before starting MD simulations.

Three popular methods for analytical forms:

• Steepest descent (first derivative)
• Conjugate gradient (first derivative)
• Newton-Raphson (second derivative)
Steepest descent:

Follows the gradient of the potential energy function \( U(r_1, \ldots, r_N) \) at each step of the iteration

\[
\mathbf{r}_i^{n+1} = \mathbf{r}_i^n + \lambda_i \mathbf{f}_i^n, \quad \mathbf{f}_i^n = -\nabla_i U(\mathbf{r}_1, \ldots, \mathbf{r}_N)|_n
\]

where \( \lambda_i \) is the step size. The step size can be adjusted until the minimum of the energy along the line is found. If this is expensive, a single step is used in each iteration, whose size is adjusted for faster convergence.

- Works best when the gradient is large (far from a minimum), but tends to have poor convergence as a minimum is approached because the gradient becomes smaller.
- Successive steps are always mutually perpendicular, which can lead to oscillations and backtracking.
A simple illustration of steepest descent with a fixed step size in a 2D energy surface (contour plots)
Conjugate gradient:

Similar to steepest descent but the gradients calculated from previous steps are used to help reduce oscillations and backtracking

\[ r_i^{n+1} = r_i^n + \lambda_i \delta_i^n, \quad \delta_i^n = f_i^n - \delta_i^{n-1} \frac{|f_i^n|^2}{|f_i^{n-1}|^2}, \quad f_i^n = -\nabla_i U(r_1, \ldots, r_N)|_n \]

(For the first step, \( \delta \) is set to zero)

• Generally one finds a minimum in fewer steps than steepest descent, e.g. it takes 2 steps for the 2D quadratic function, and \( \sim n \) steps for nD.
• But conjugate gradient may have problems when the initial conformation is far from a minimum in a complex energy surface.
• Thus a better strategy is to start with steepest descent and switch to conjugate gradient near the minimum.
Newton-Raphson:

Requires the second derivatives (Hessian) besides the first.

Predicts the location of a minimum, and heads in that direction.

To see how it works in a simple situation, consider the quadratic 1D case

\[ f(x) = ax^2 + bx + c, \quad f' = 0 \quad \rightarrow \quad x_{\text{min}} = -\frac{b}{2a} \]

\[ x_{\text{min}} = x_0 - \left. \frac{f'}{f''} \right|_{x_0} = x_0 - \frac{2ax_0 + b}{2a} = -\frac{b}{2a} \]

In general

\[ r_{i}^{n+1} = r_{i}^{n} - [H_{i}^{n}]^{-1} f_{i}^{n}, \quad [H_{i}^{n}]_{kl} = \left. \frac{\partial^2 U(r_1, \ldots, r_N)}{\partial x_k \partial x_l} \right|_{n} \]

For a quadratic energy surface, this method will find the minimum in one step starting from any configuration.
• Construction and inversion of the 3Nx3N Hessian matrix is computationally demanding for large systems (N>100).
• It will find a minimum in fewer steps than the gradient-only methods in the vicinity of the minimum.
• But it may encounter serious problems when the initial conformation is far from a minimum.
• A good strategy is to start with steepest descent and then switch to alternate methods as the calculations progress, so that each algorithm operates in the regime for which it was designed.

Using the above methods, one can only find a local minimum. To search for an absolute minimum, Monte Carlo methods are more suitable. Alternatively, one can heat the system in MD simulations, which will facilitate transitions to other minima.
Normal mode analysis

Assume that the minimum energy of the system is given by the 3N coordinates, \( \{r^0_i\} \). Expanding the potential energy around the equilibrium configuration gives

\[
U(r_1, \ldots, r_N) = U(\{r^0_i\}) - \sum_{i=1}^{N} (r_i - r^0_i) \cdot \nabla U_{\{r^0_i\}} \rightarrow 0
\]

\[
+ \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (r_i - r^0_i) \cdot \nabla U_{\{r^0_i\}} \cdot (r_j - r^0_j) + \cdots
\]

Ignoring the constant energy, the leading term is that of a system of coupled harmonic oscillators. In a normal mode, all the particles in the system oscillate with the same frequency \( \omega \). To find the normal modes, first express the 3N coordinates as \( \{x_i, \ i=1, \ldots, 3N\} \).
The potential energy becomes

\[ U(x_1, \ldots, x_{3N}) = \frac{1}{2} \sum_{i=1}^{3N} \sum_{j=1}^{3N} (x_i - x_i^0)K_{ij}(x_j - x_j^0) \]

where the spring constants are given by the Hessian as

\[ K_{ij} = \left. \frac{\partial^2 U}{\partial x_i \partial x_j} \right|_{\{x_i^0\}} \]

Introducing the 3Nx3N diagonal mass matrix \( M \)

\[ M = \text{diag}(m_1, m_1, m_1, m_2, m_2, m_2, \ldots, m_N, m_N, m_N) \]

The secular equation for the normal modes is given by

\[ \left| M^{-1/2} K M^{-1/2} - \omega^2 I \right| = 0 \]
For a \(3N\times3N\) matrix, solution of the secular equation will yield \(3N\) eigenvalues, \(\omega_i\), and the corresponding eigenvectors, \(\alpha_i\). Of these, 3 correspond to translations and 3 to rotations of the system. Thus there are \(3N-6\) intrinsic vibrational modes of the system.

At a given temperature \(T\), the motion of the \(i\)'th coordinate is given by

\[
x_i(t) = \sum_{j=1}^{3N} \left( \frac{kT}{m_j \omega_j^2} \right)^{1/2} \alpha_{ij} \cos(\omega_j t + \phi_j)
\]

The mean square displacement of the coordinates and atoms:

\[
\langle (\Delta x_i)^2 \rangle = \frac{1}{2} \sum_{j=7}^{3N} \frac{kT}{m_j \omega_j^2} \alpha_{ij}^2
\]

\[
\langle (\Delta r_n)^2 \rangle = \langle (\Delta x_{3n})^2 \rangle + \langle (\Delta x_{3n+1})^2 \rangle + \langle (\Delta x_{3n+2})^2 \rangle
\]
A simple example: normal modes of water molecule

Water molecule has $9-6=3$ intrinsic vibrations, which correspond to symmetric and antisymmetric stretching of H atoms and bending. Because of H-bonding, water molecules in water cannot freely rotate but rather librate (wag or twist).

Wave numbers: $3652 \text{ cm}^{-1}$  $3756 \text{ cm}^{-1}$  $1595 \text{ cm}^{-1}$  $(200 \text{ cm}^{-1} \sim 1 \text{ kT})$

Excitation energies $>> kT$, which justifies the use of a rigid model for water.
Normal modes of a small protein BPTI (Bovine pancreatic tyripsin inhibitor)

Some characteristic frequencies (in cm\(^{-1}\))

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Frequency Range (cm(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stretching</td>
<td>H-N</td>
<td>3400-3500</td>
</tr>
<tr>
<td></td>
<td>H-C</td>
<td>2900-3000</td>
</tr>
<tr>
<td></td>
<td>C=C, C=O</td>
<td>1700-1800</td>
</tr>
<tr>
<td></td>
<td>C-C, C-N</td>
<td>1000-1250</td>
</tr>
<tr>
<td>Bending</td>
<td>H-C-H</td>
<td>1500</td>
</tr>
<tr>
<td></td>
<td>H-N-H</td>
<td>1500</td>
</tr>
<tr>
<td></td>
<td>C-C=O</td>
<td>500</td>
</tr>
<tr>
<td>Torsion</td>
<td>C=C</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>C-O</td>
<td>300-600</td>
</tr>
<tr>
<td></td>
<td>C-C</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>S-S-C</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>C-S</td>
<td>200</td>
</tr>
</tbody>
</table>