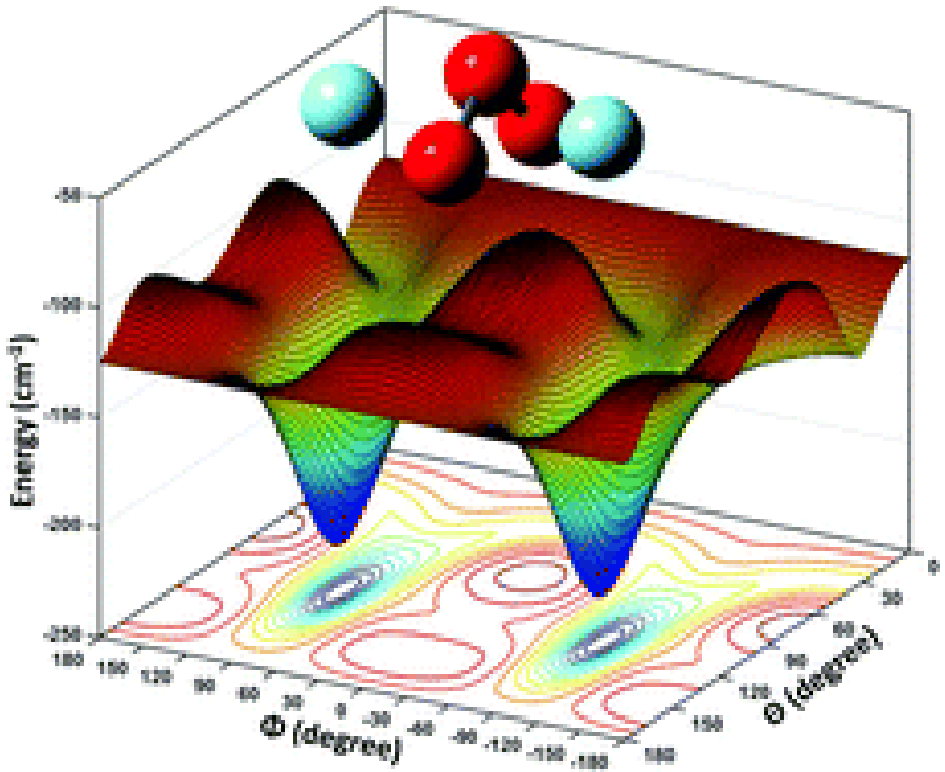


Conformational Search

Ms. Afreen Khan

Bombay College of Pharmacy.

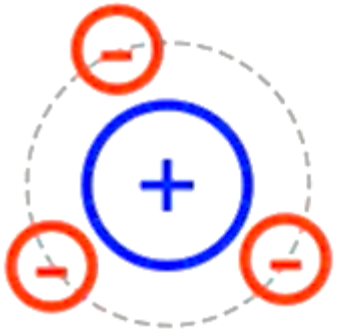
Introduction

- Potential energy
- Force fields
- Concept of torsion rotations
- Potential energy surface

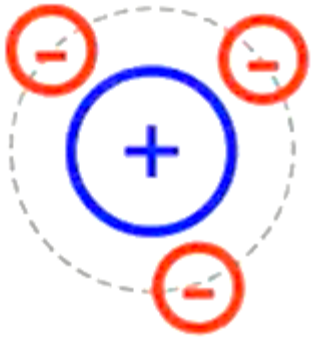
Potential Energy

- The potential energy dictates the behavior of the system.
- The potential energy is the interaction energy among all the particles of the system.

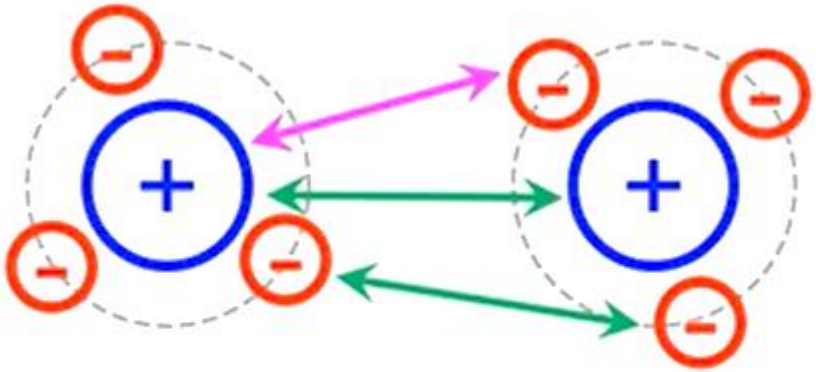
$$V = V_{NN} + V_{ee} + V_{eN}$$



A



B

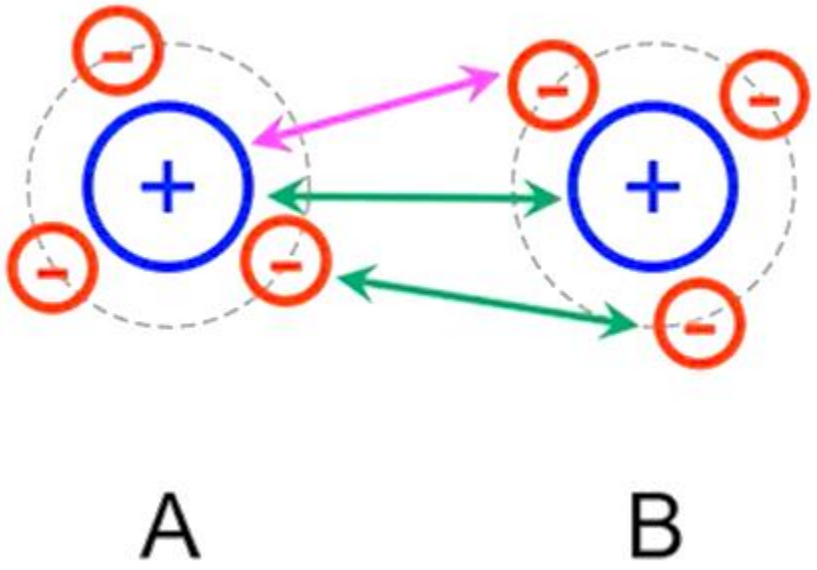


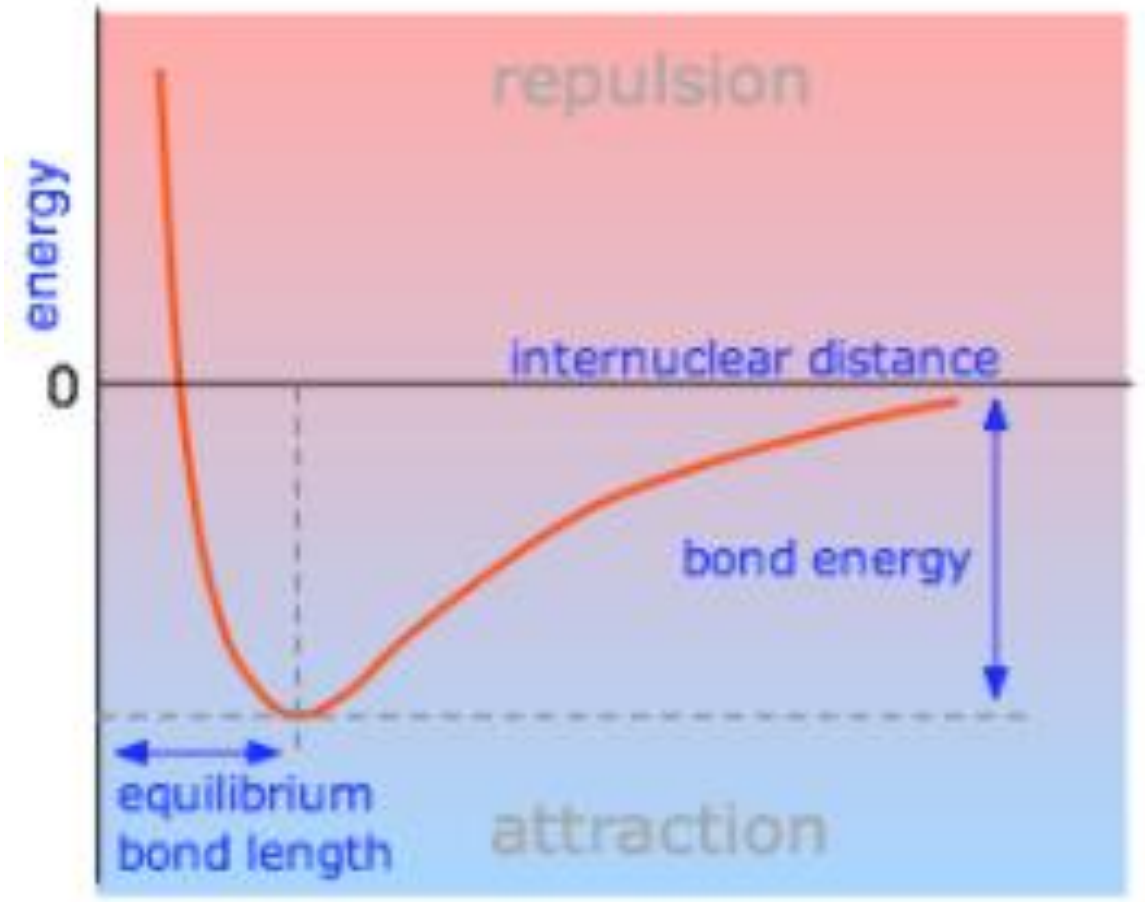
A B

Potential Energy

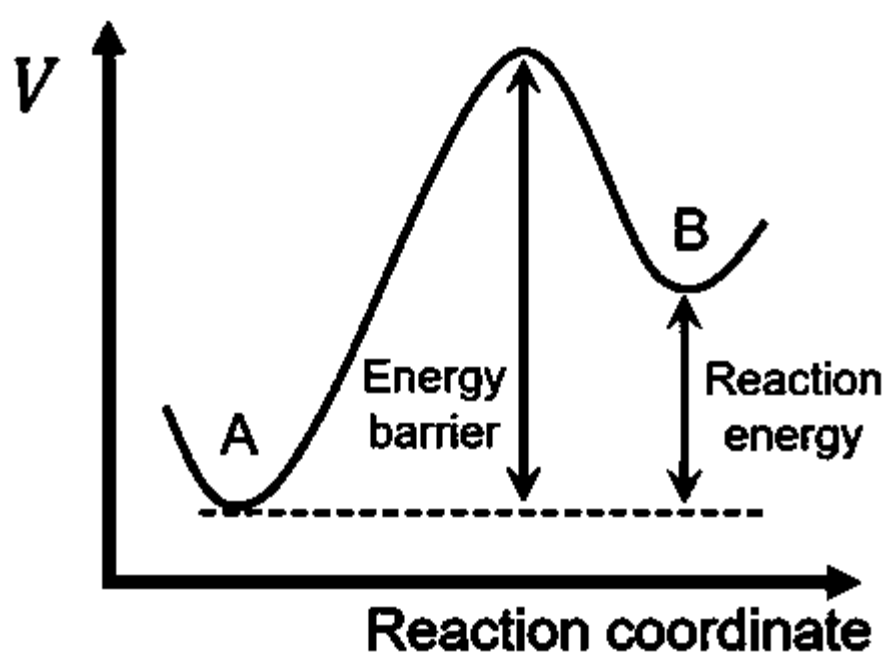
$$V = V_{NN} + V_{ee} + V_{eN}$$

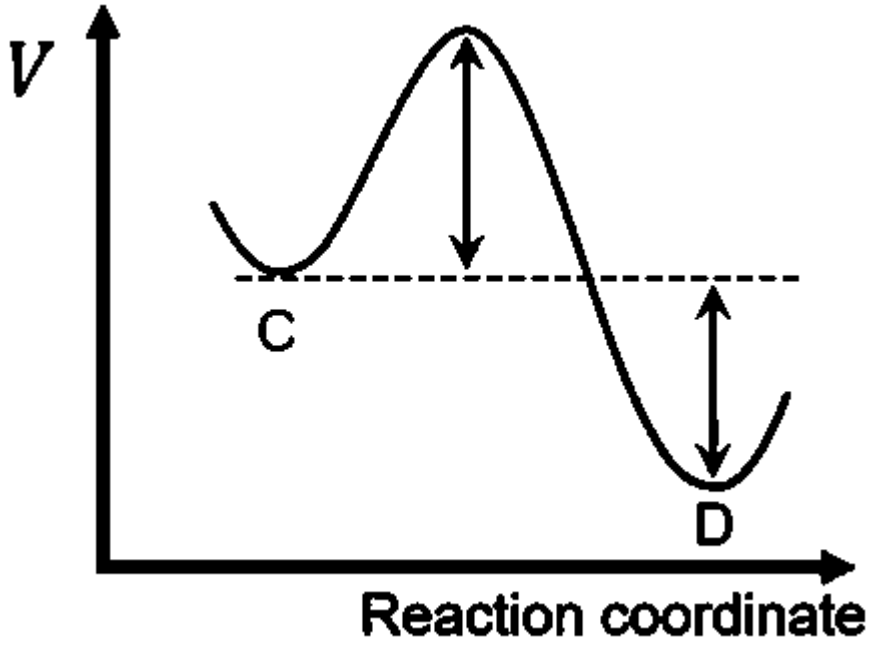
Repulsive (positive) Attractive (negative)





Potential Energy



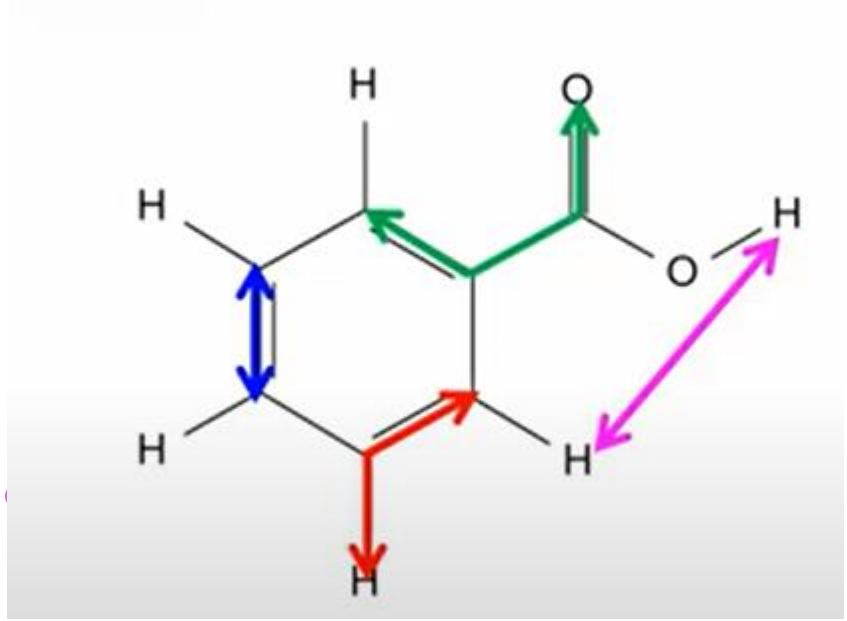


- **Large energy barrier → slow reaction**
- **Positive reaction energy → Unfavourable**

- **Small energy barrier → fast reaction**
- **Negative reaction energy → Favourable**

Force Fields

- A force field is a simple equation that relates the PE of the system with its internal coordinates (bond distances, bond angles,…)
- In most used force fields, the PE is split into bonded and non-bonded interactions.



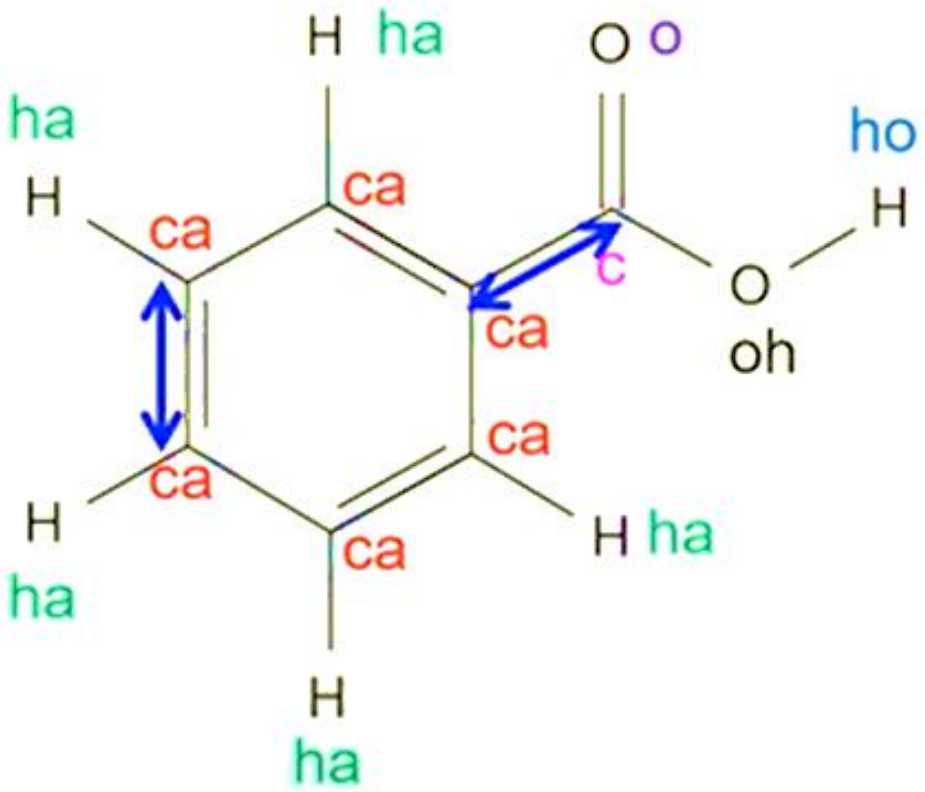
$$E_{total} = E_{bonded} + E_{nonbonded}$$

$$E_{bonded} = E_{bond} + E_{angle} + E_{dihedral}$$

$$E_{nonbonded} = E_{electrostatic} + E_{van\ der\ waals}$$

Force Fields

The atoms of the molecules are classified in different atom types to distinguish interactions between the same chemical class of atoms.



ca: sp^2 C in aromatic system

ha: H bonded to aromatic C

c: sp^2 C of carbonyl group

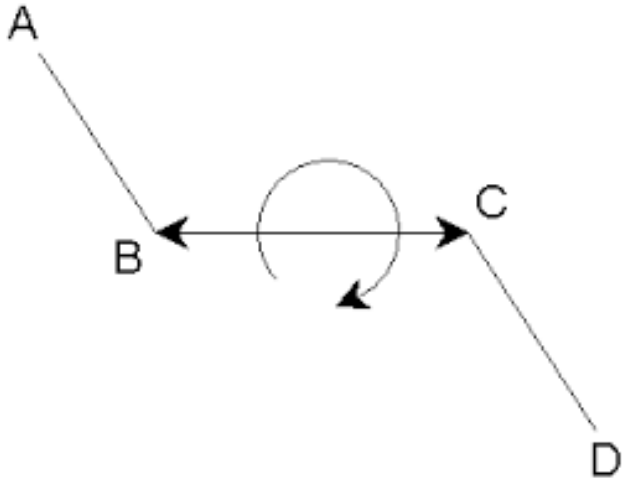
o: O with one connected atom

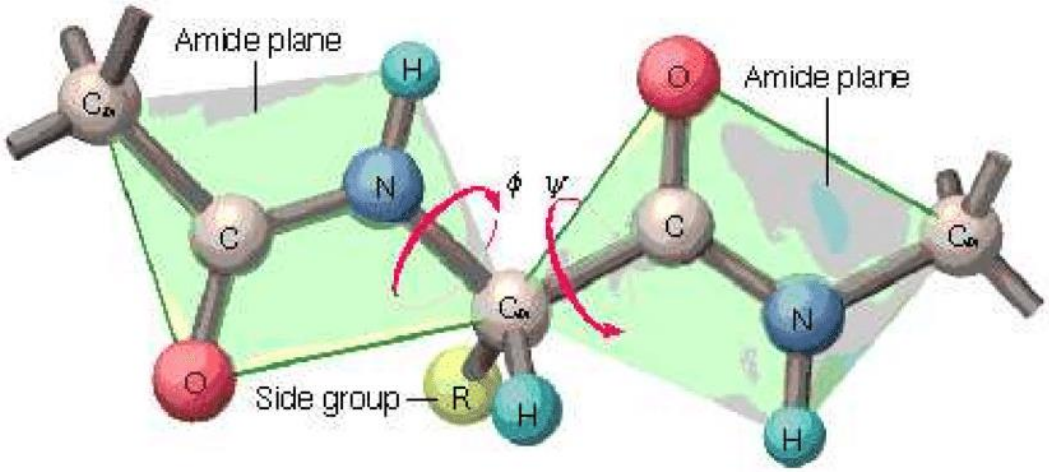
oh: O in hydroxyl group

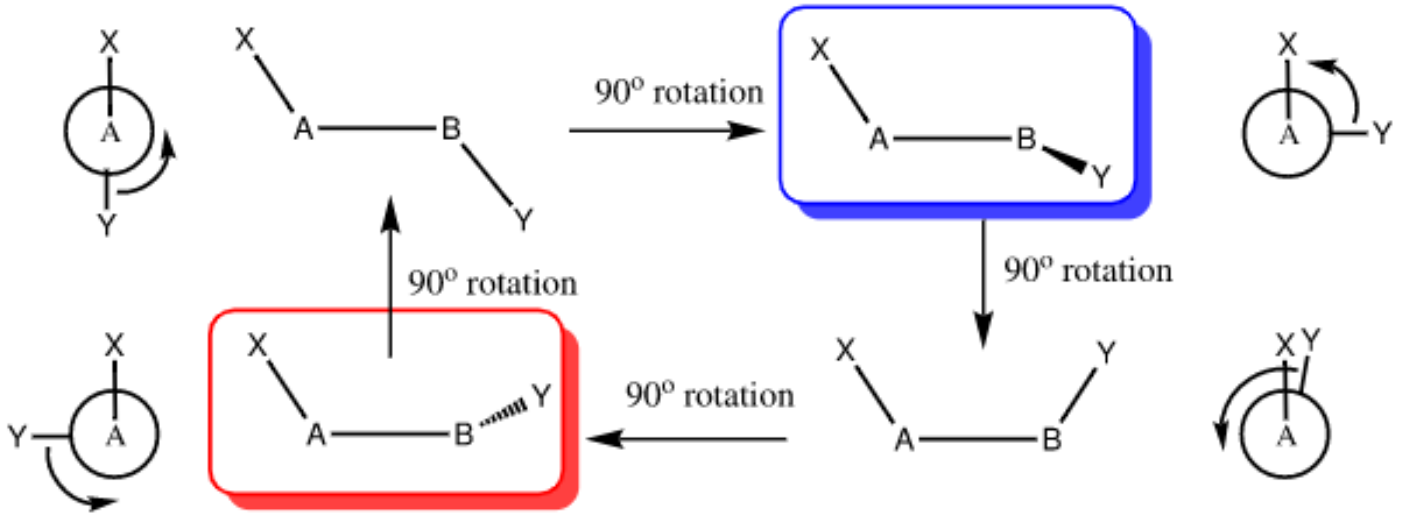
ho: H in hydroxyl group

General Amber Force Field (GAFF)
Wang et al. J Comput. Chem. 2004, 15, 1157.

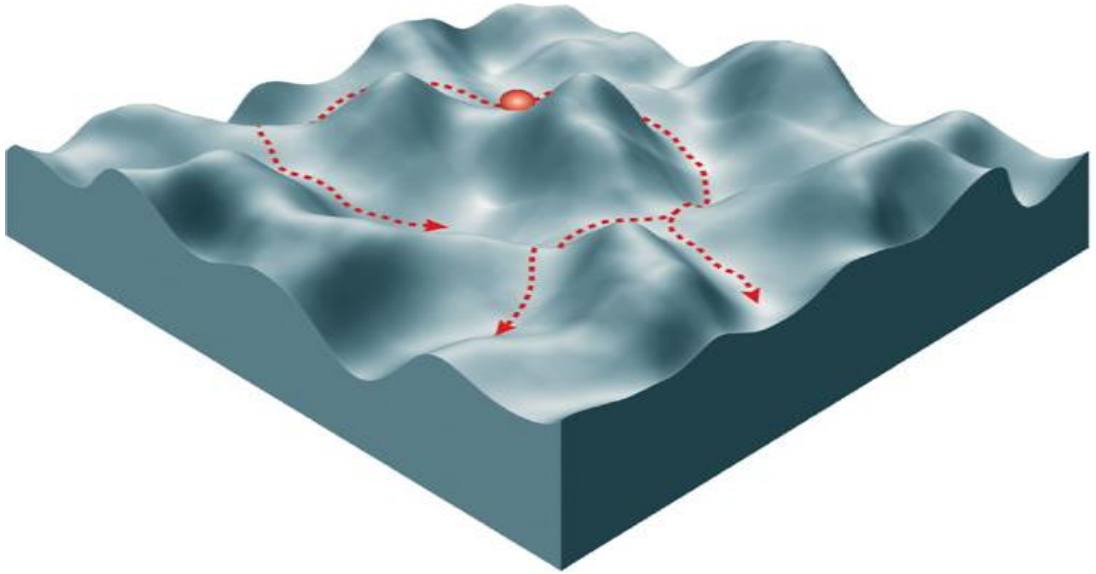
Concept of Torsion rotations

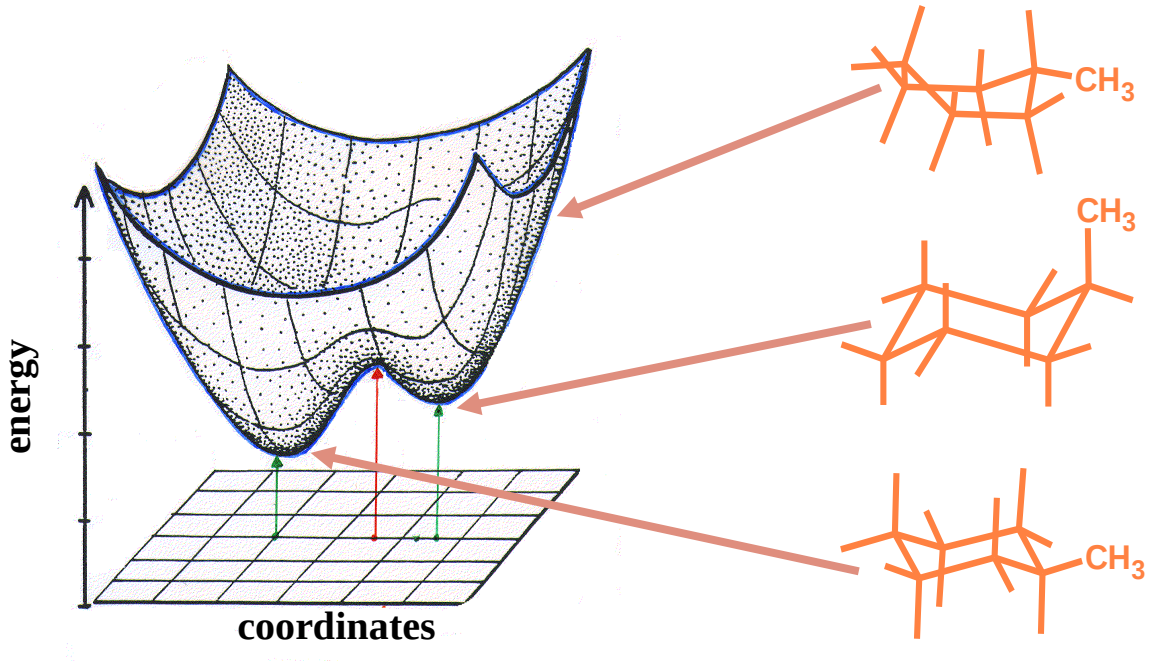




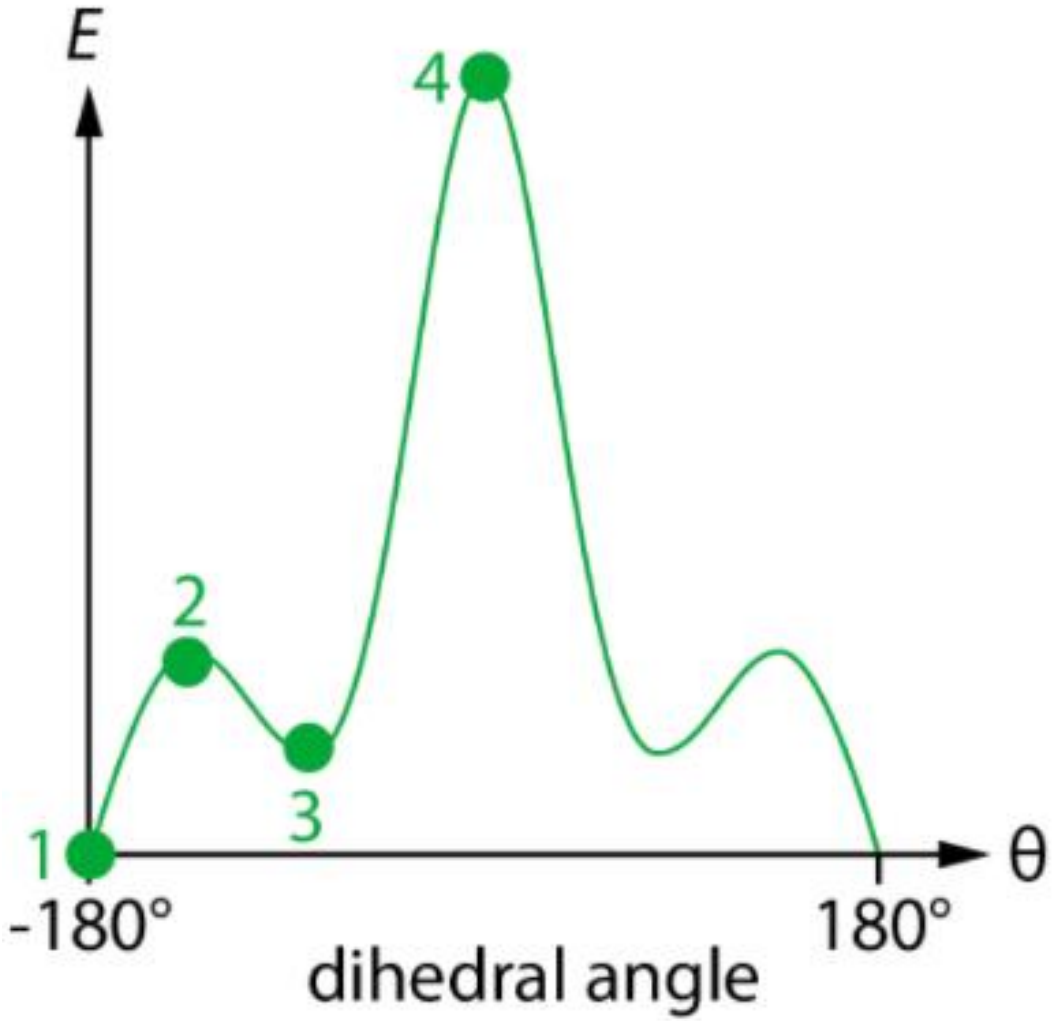


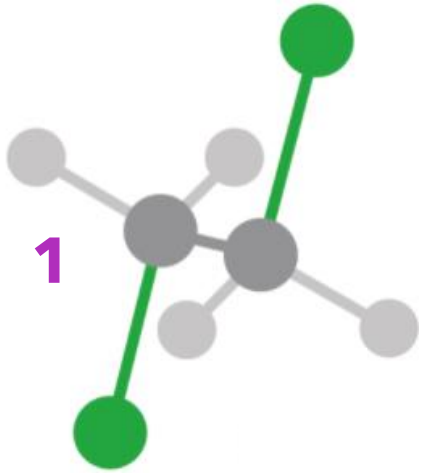
The Potential Energy Surface





The Potential Energy Surface

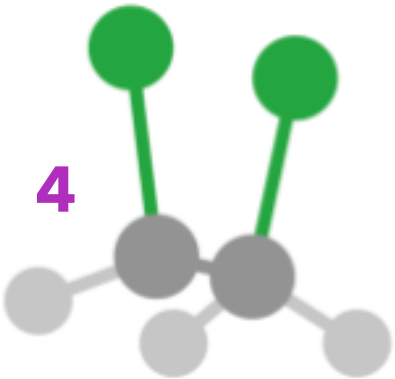




Global Minimum

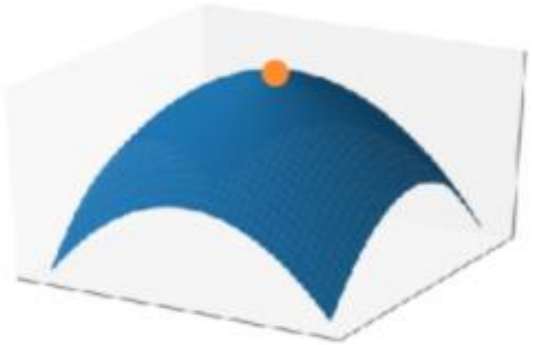
Local Maximum

Local Minimum

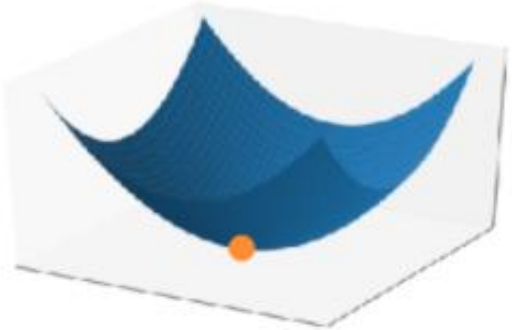


Global Maximum

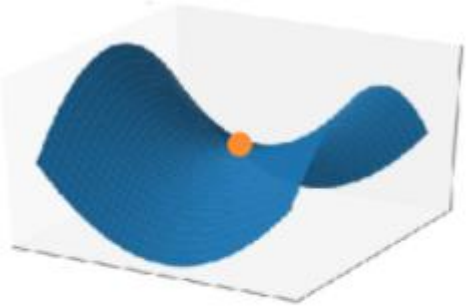
Classification of stationary points



local maximum

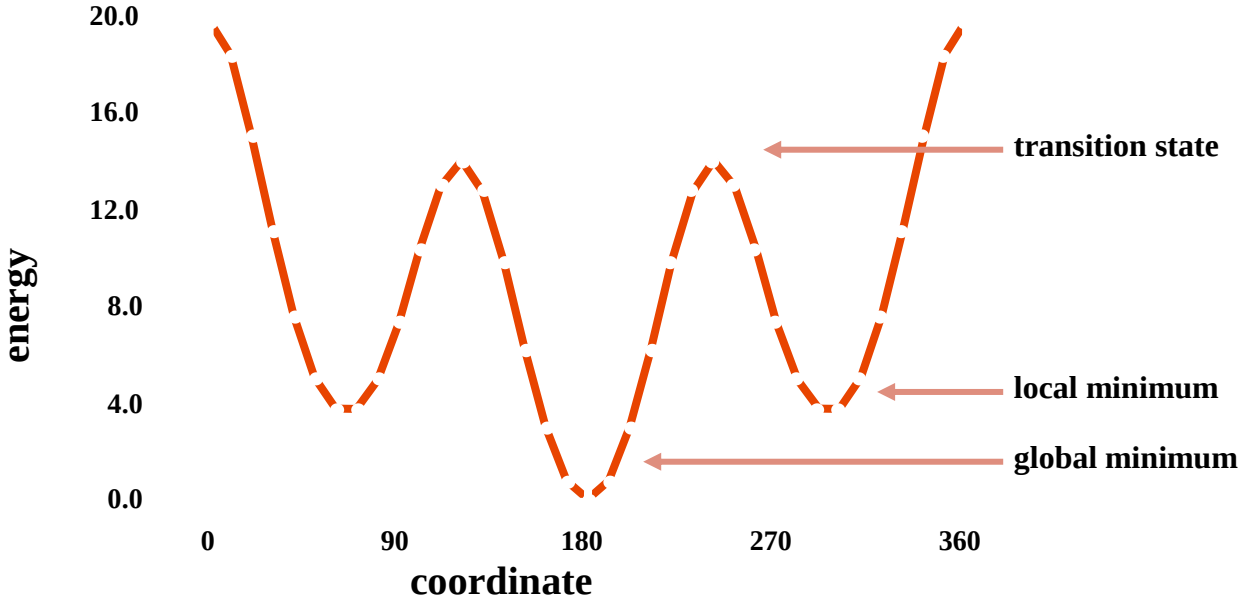


local minimum



saddle point
(transition state)

$$\frac{\partial E}{\partial q_1} = \frac{\partial E}{\partial q_2} = \cdots = 0$$



Classification of stationary points

$$\frac{\mathrm{d}E}{\mathrm{d}x} = 0.$$

$$\left(\frac{\partial^2 E}{\partial x^2}\right) > 0, \left(\frac{\partial^2 E}{\partial y^2}\right) > .0$$

$$\left(\frac{\partial^2 E}{\partial x^2}\right) < 0, \left(\frac{\partial^2 E}{\partial y^2}\right) < 0.$$

$$\left(\frac{\partial^2 E}{\partial x_{\parallel}^2}\right) > 0, \left(\frac{\partial^2 E}{\partial x_{\perp}^2}\right) < 0.$$

1st Derivative

2nd Derivative

Type	1st Derivative	2nd Derivative*
Minimum	0	positive
Maximum	0	negative
Saddle point	0	negative

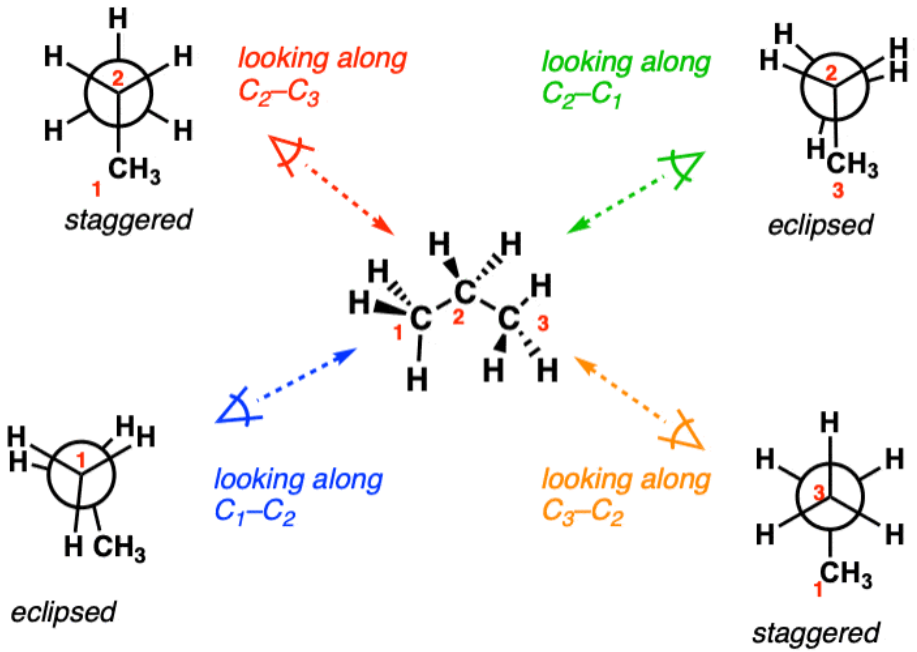
Methods

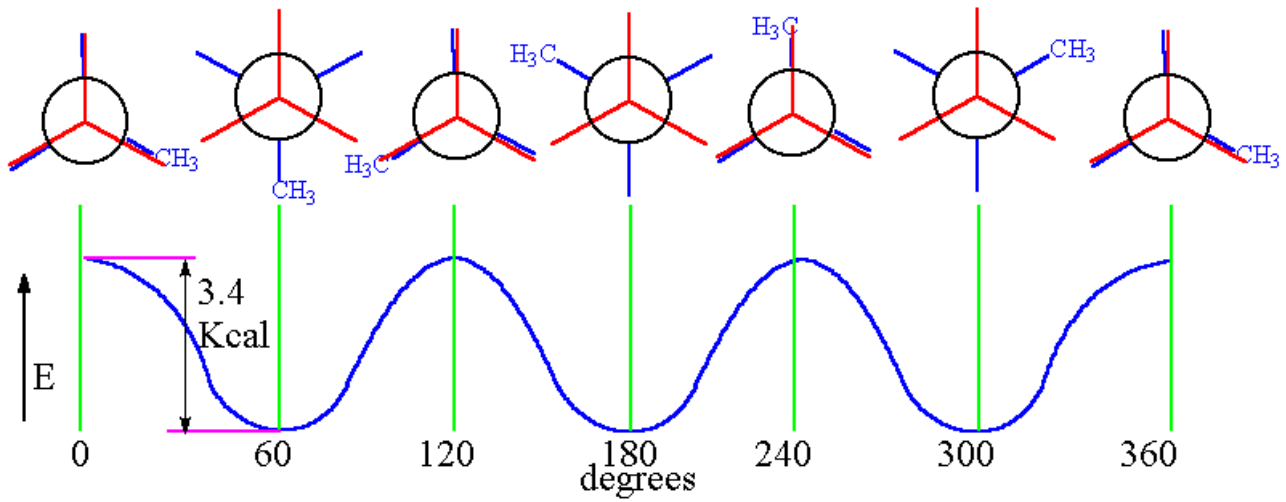
- Systematic search
- Model-building approach
- Random search
- Molecular dynamics
- Monte- Carlo
- Simulated annealing
- Genetic algorithms

Systematic Search

Systematic search

- **Propane conformations**



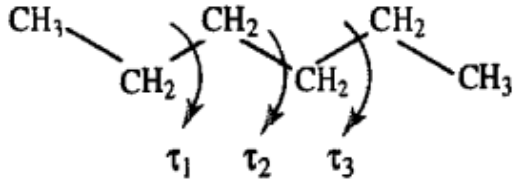


Systematic search

- **Combinatorial explosion**

$$\text{Number of conformations} = \prod_{i=1}^{N} \frac{360}{\theta_i}$$

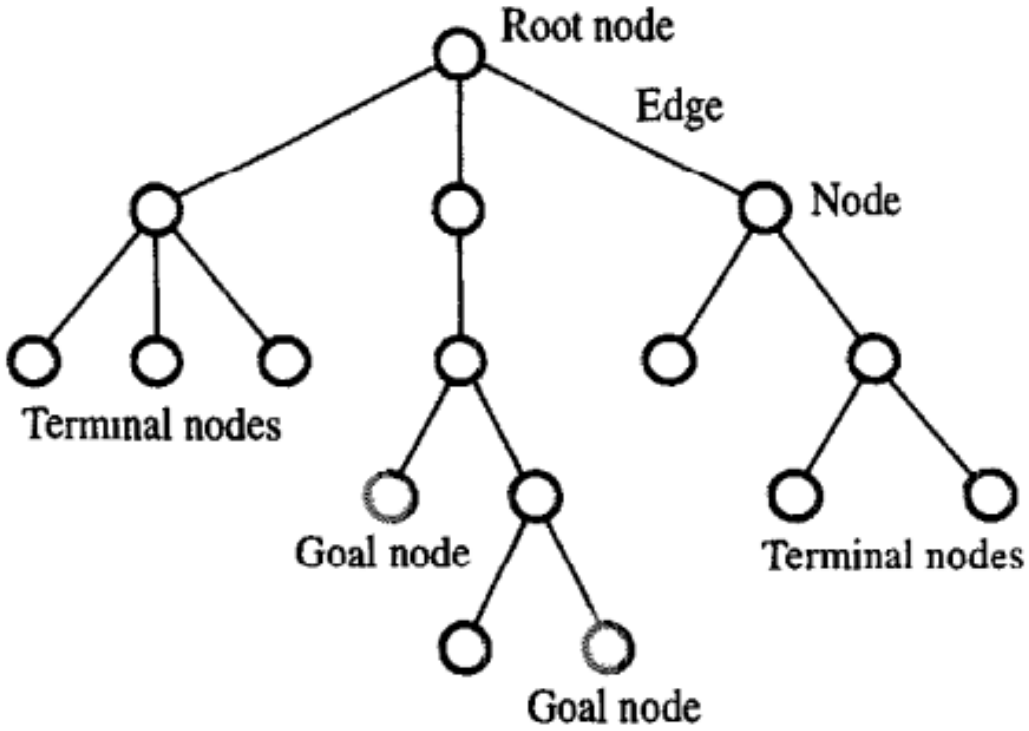
Dihedral over bond i

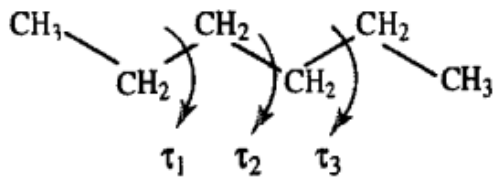


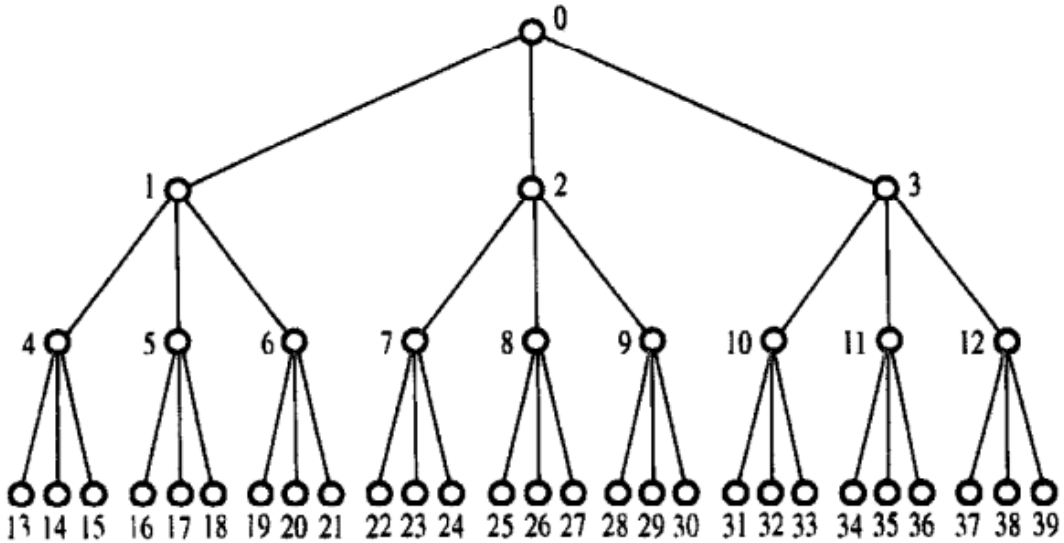
- For **5 bonds** and 30 increments, the total resultant structures is **248832** conformers→ (if 1s per structure = 69 hours)
- For **7 bonds** and 30 increments, the total resultant structures is 36 million (**36, 000,000**) conformers → (if 1s per structure = 415 days)

Systematic search

- **Search Trees**

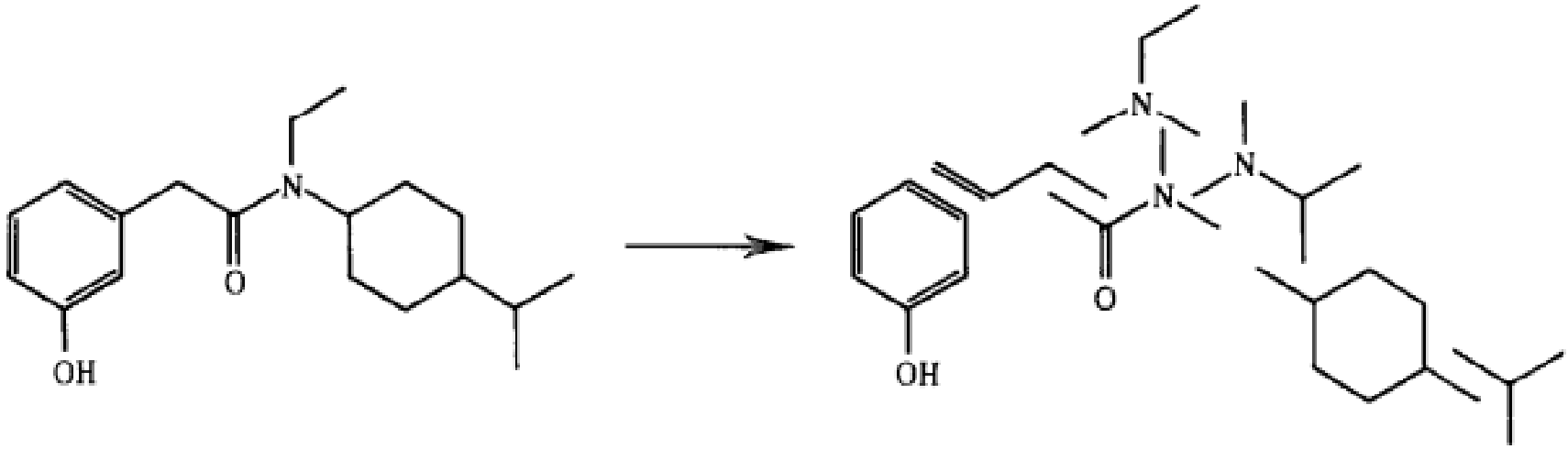






Model-building Approach

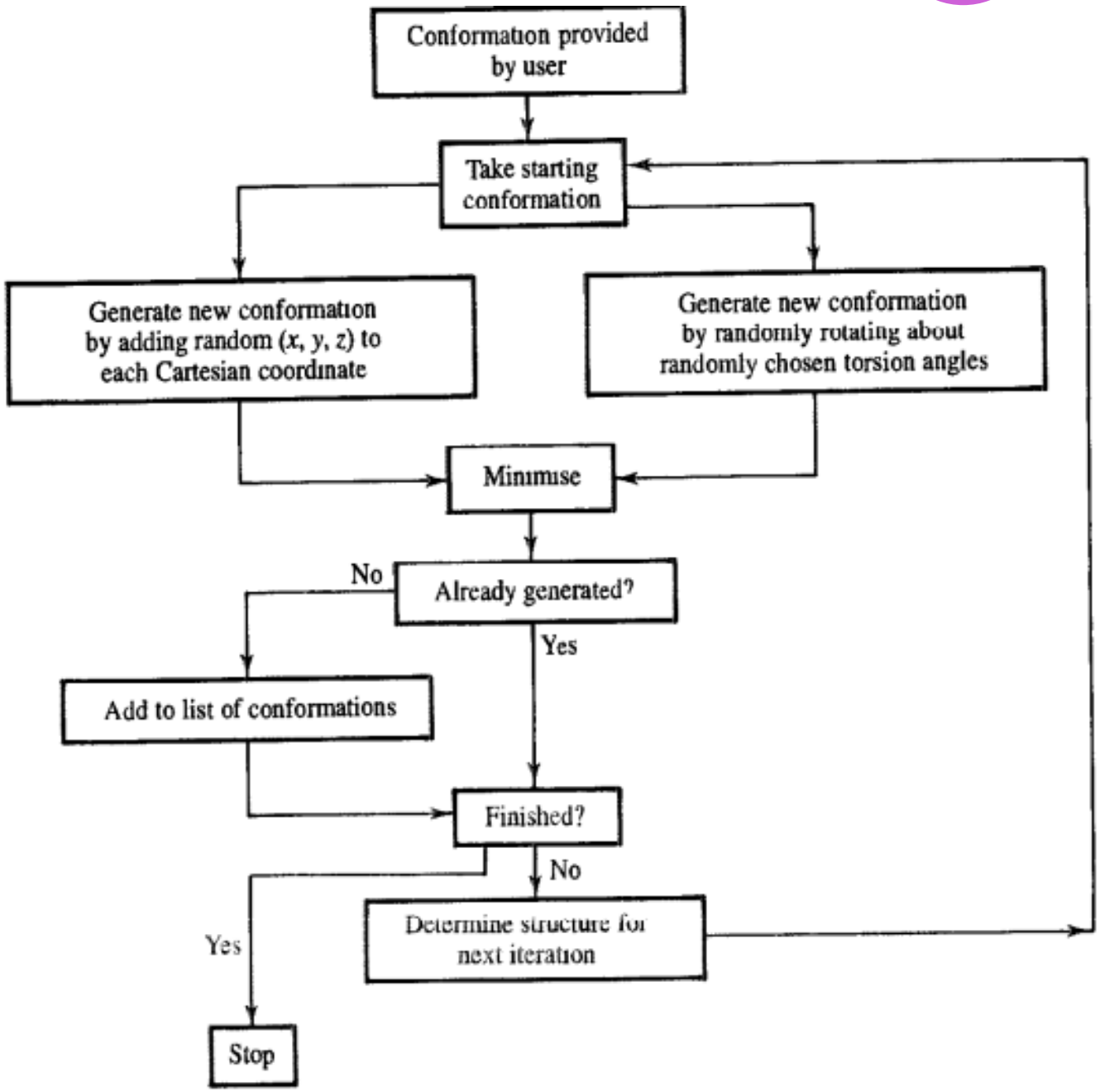
Model-building approach



- **Decide on the fragments → "substructure search"**
- **Generate the conformations**

Random Search

Random Search



Molecular Dynamics

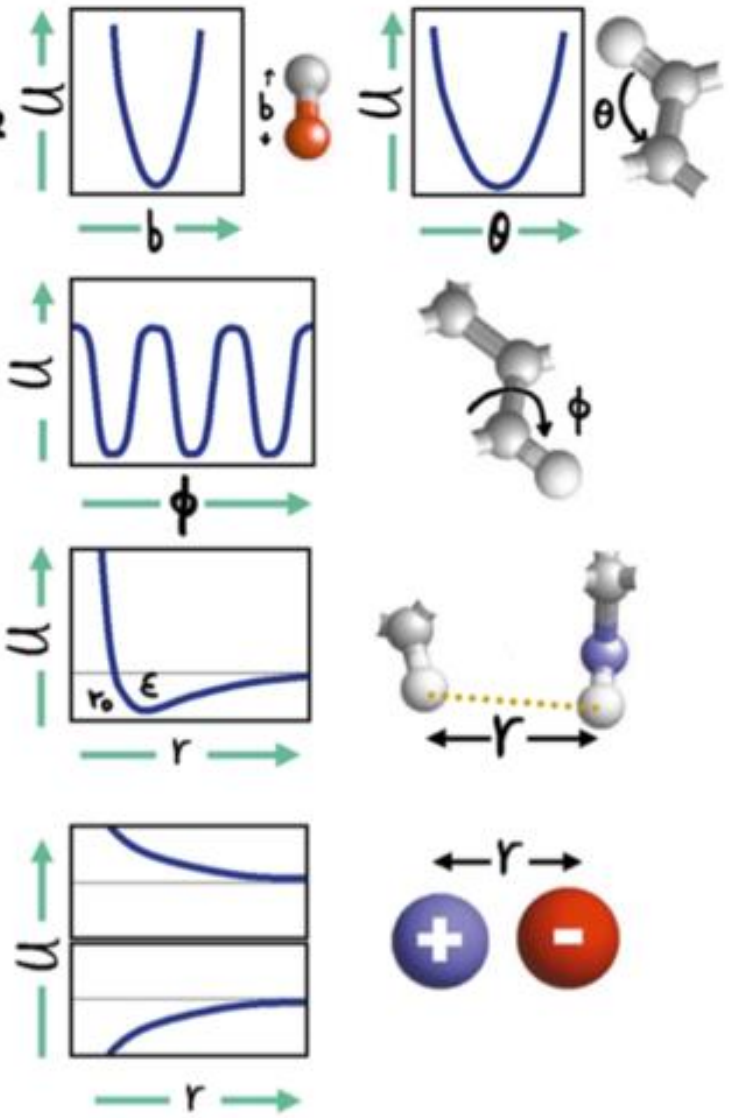
Force Fields

$$U = \sum_{\text{All Bonds}} \tfrac{1}{2} K_b (b - b_o)^2 + \sum_{\text{All Angles}} \tfrac{1}{2} K_\theta (\theta - \theta_o)^2$$

$$+ \sum_{\text{All Torsion Angles}} K_\phi [1 - \cos(n\phi + \delta)]$$

$$+ \sum_{\text{All nonbonded pairs}} \varepsilon \left[(r_o/r)^{12} - 2(r_o/r)^6 \right]$$

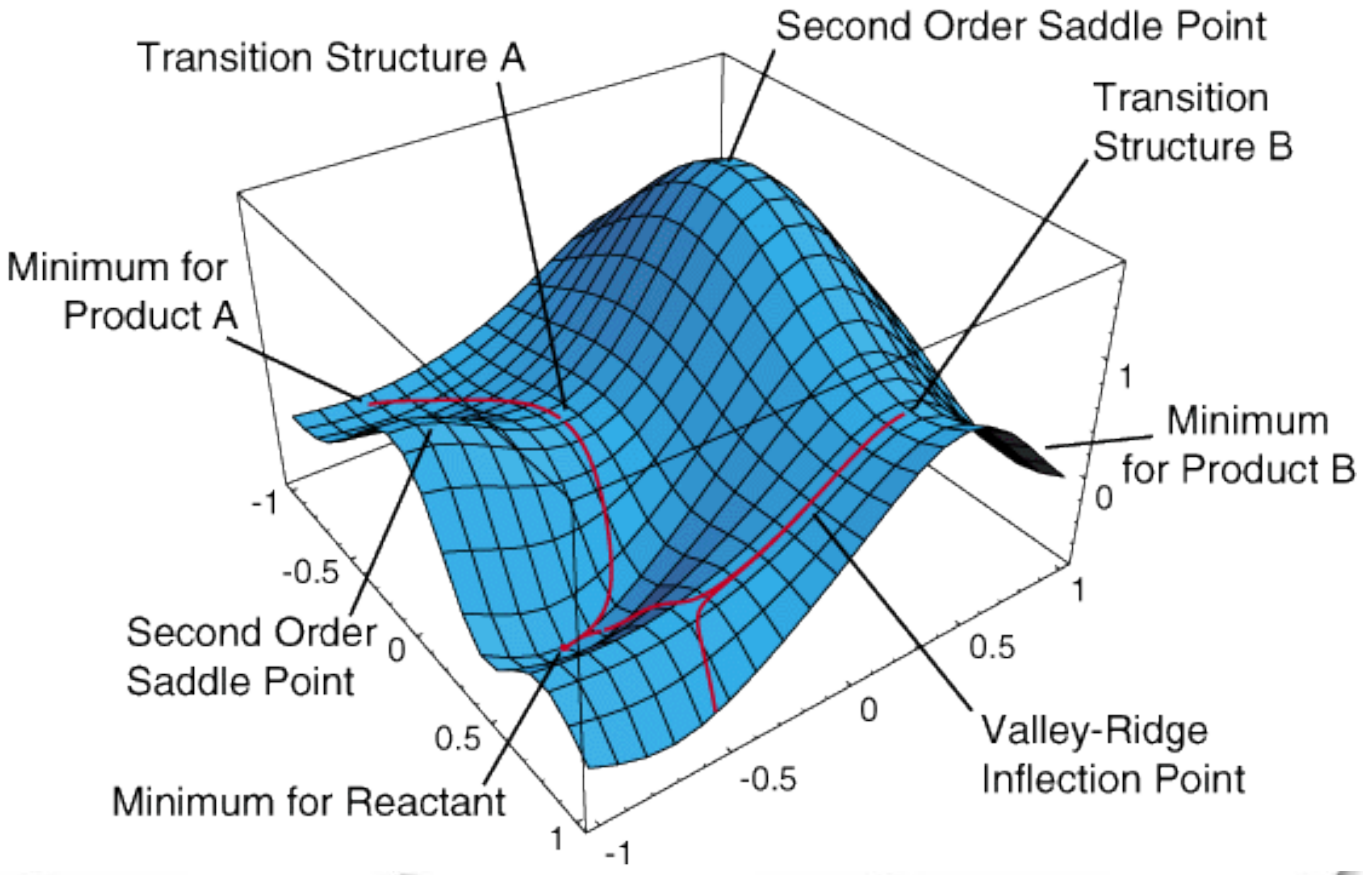
$$+ \sum_{\text{All partial charges}} 332 q_i q_j / r$$

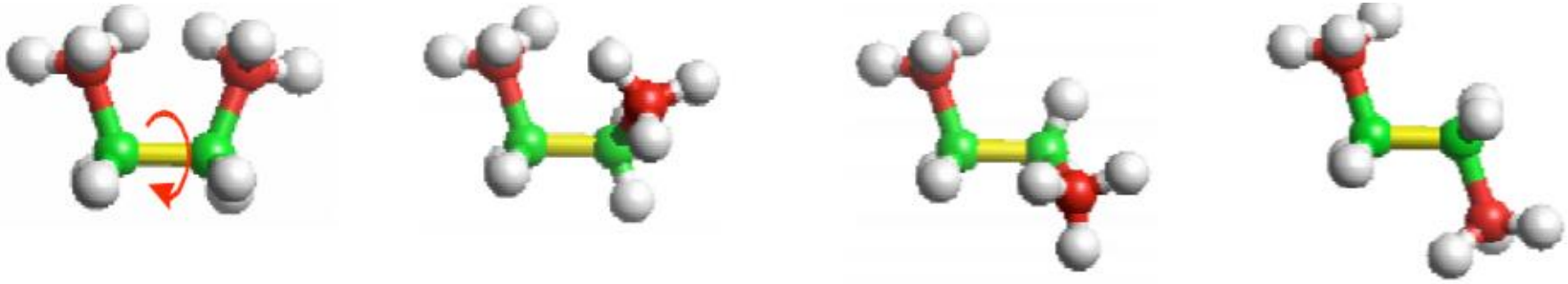


Popular molecular mechanics force fields

- AMBER (Assisted Model Building and Energy Refinement) - widely used for proteins and DNA
- CHARMM - originally developed at Harvard, widely used for both small molecules and macromolecules
- CVFF - also broadly used for small molecules and macromolecules
- GROMOS - A force field that comes as part of the GROMACS (GROningen MOlecular Simulation package
- OPLS-aa, OPLS-ua, OPLS-2001, OPLS-2005 - Members of the OPLS family of force fields developed by William L. Jorgensen at Yale Department of Chemistry.
- ECEPP/2 - free energy force field

Potential Energy Surface

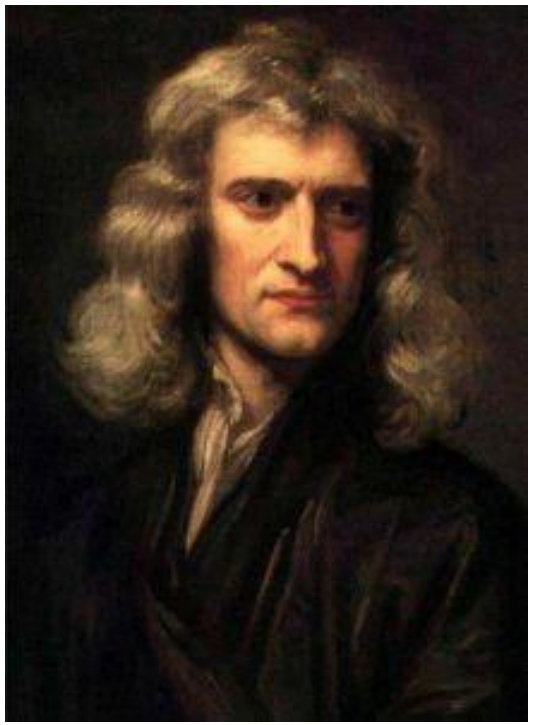




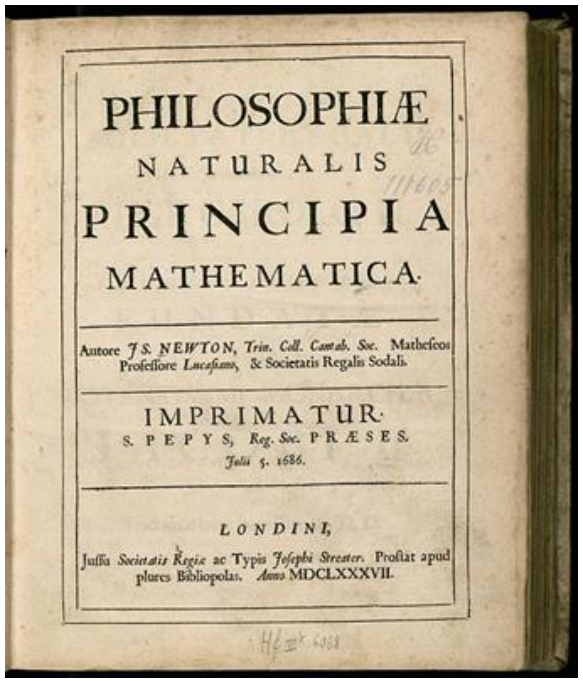
Why MD simulations?

- Link physics, chemistry and biology
- Model phenomena that cannot be observed experimentally
- Understand protein folding…
- Access to thermodynamics quantities (free energies, binding energies,…)

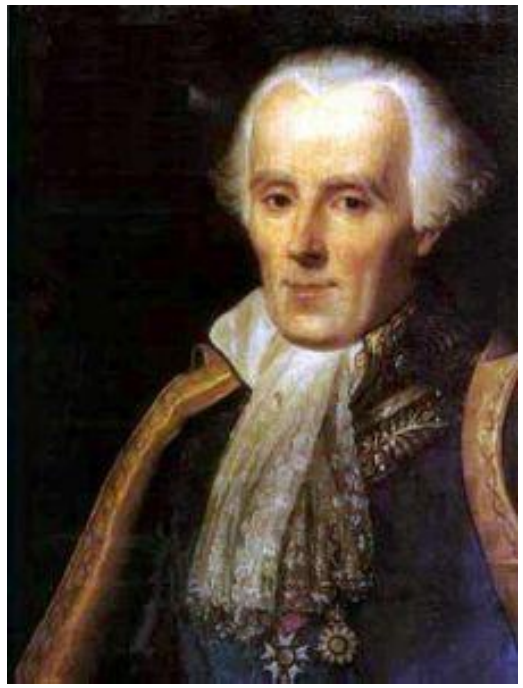
Molecular Dynamics



Sir Isaac Newton



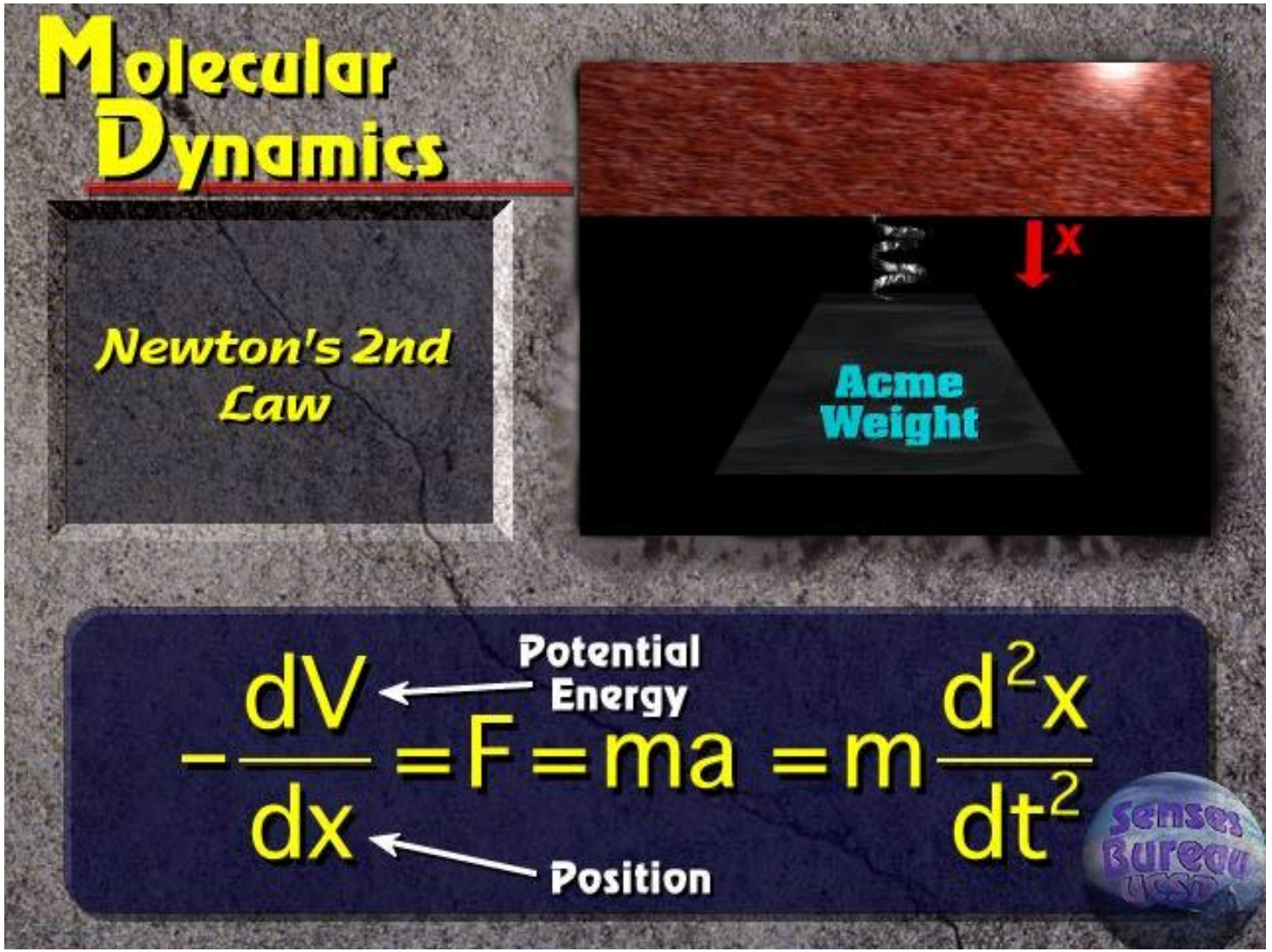
$$F = m_i a_i$$



Pierre-Simon Laplace

Deterministic nature of MD

Molecular Dynamics



Representing an Atom

- Classical mechanics: a point particle
- Defined by its position (x,y,z) and its mass
- May carry an electric charge (positive or negative), usually partial (less than an electron)

Molecular Dynamics

- Time dependent integration of classical equations of motion

$$F = -\frac{\partial U}{\partial x}$$

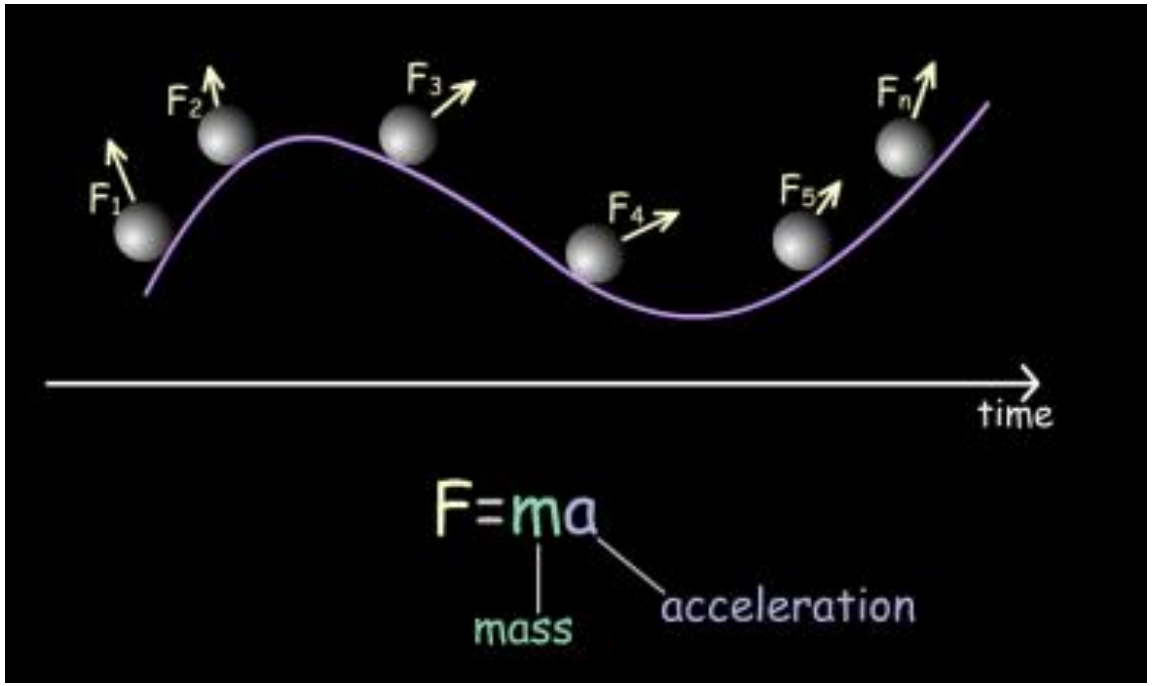
$$F = ma$$

$$a = \frac{v_2 - v_1}{\partial t}$$

$$v = \frac{x_2 - x_1}{\partial t}$$

$$E = U + K$$

$$\partial t = 2 \text{ fs}$$



Molecular Dynamics: Mathematically

- Time dependent integration

$$F = -\frac{\partial U}{\partial x}$$

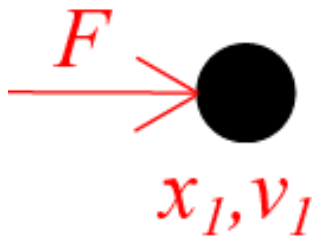
$$F = ma$$

$$a = \frac{v_2 - v_1}{\partial t}$$

$$v = \frac{x_2 - x_1}{\partial t}$$

$$E = U + K$$

$$\partial t = 2 \text{ fs}$$



Molecular Dynamics: Mathematically

- Time dependent integration

$$F = -\frac{\partial U}{\partial x}$$

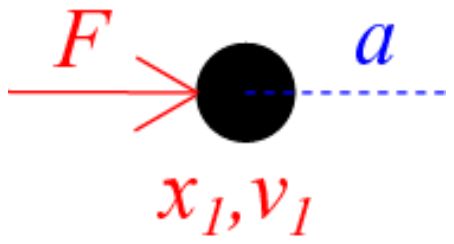
$$F = ma$$

$$a = \frac{v_2 - v_1}{\partial t}$$

$$v = \frac{x_2 - x_1}{\partial t}$$

$$E = U + K$$

$$\partial t = 2 \text{ fs}$$



Molecular Dynamics: Mathematically

- Time dependent integration

$$F = -\frac{\partial U}{\partial x}$$

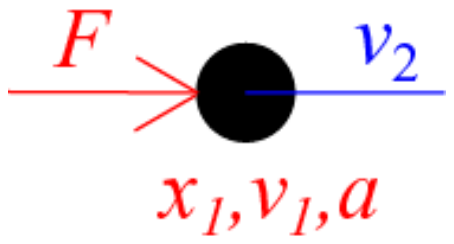
$$F = ma$$

$$a = \frac{v_2 - v_1}{\partial t}$$

$$v = \frac{x_2 - x_1}{\partial t}$$

$$E = U + K$$

$$\partial t = 2 \text{ fs}$$



Molecular Dynamics: Mathematically

- Time dependent integration

$$F = -\frac{\partial U}{\partial x}$$

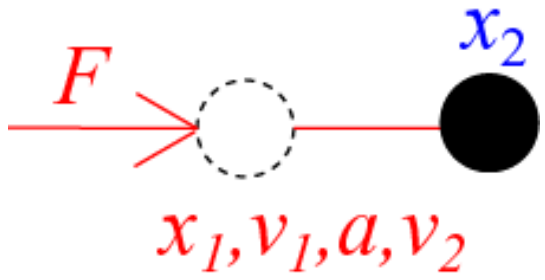
$$F = ma$$

$$a = \frac{v_2 - v_1}{\partial t}$$

$$v = \frac{x_2 - x_1}{\partial t}$$

$$E = U + K$$

$$\partial t = 2 \text{ fs}$$



Molecular Dynamics: Mathematically

- Time dependent integration

$$F = -\frac{\partial U}{\partial x}$$

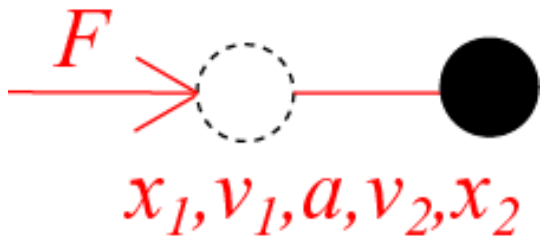
$$F = ma$$

$$a = \frac{v_2 - v_1}{\partial t}$$

$$v = \frac{x_2 - x_1}{\partial t}$$

$$E = U + K$$

$$\partial t = 2 \text{ fs}$$



Molecular Dynamics: Mathematically

- Time dependent integration

$$F = -\frac{\partial U}{\partial x}$$

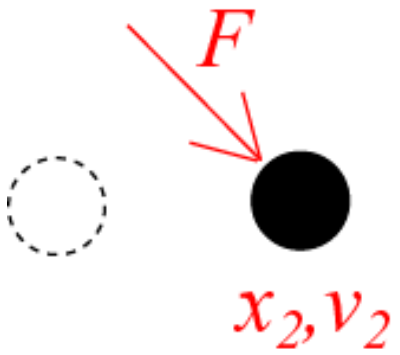
$$F = ma$$

$$a = \frac{v_3 - v_2}{\partial t}$$

$$v = \frac{x_3 - x_2}{\partial t}$$

$$E = U + K$$

$$\partial t = 2 \text{ fs}$$



Molecular Dynamics: Mathematically

- Time dependent integration

$$F = -\frac{\partial U}{\partial x}$$

$$F = ma$$

$$a = \frac{v_3 - v_2}{\partial t}$$

$$v = \frac{x_3 - x_2}{\partial t}$$

$$E = U + K$$

$$\partial t = 2 \text{ fs}$$

- Evaluate forces and perform integration for every atom
- Each picosecond of simulation time requires 500 iterations of cycle
- E.g. w/ 50,000 atoms, each ps (10^{-12} s) involves 25,000,000 evaluations

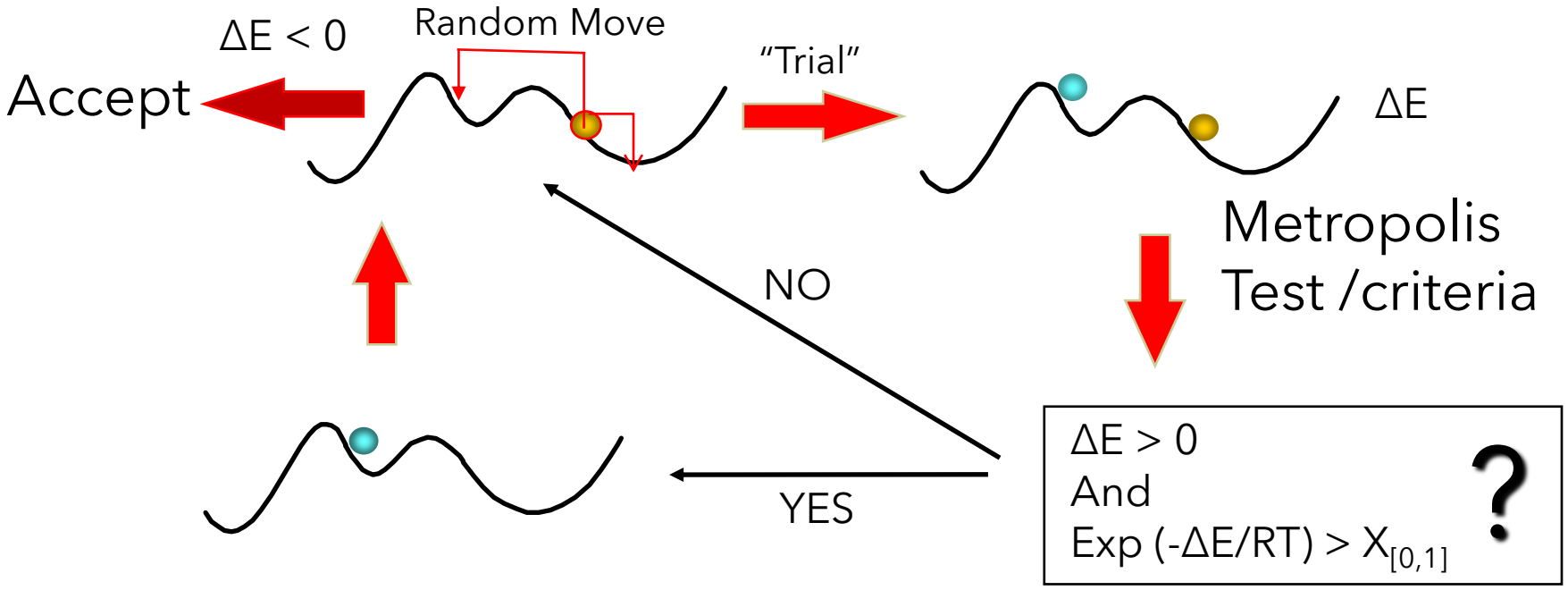
Stochastic search- Monte Carlo Simulations

Basics

- **System: The physical process of interest**
- **Model: Mathematical representation of the system**
  - Models are a fundamental tool of science, engineering, business, etc.
  - Abstraction of reality
  - Models always have limits of credibility
- **Simulation: A type of model where the computer is used to imitate the behavior of the system**
- **Monte Carlo simulation: Simulation that makes use of internally generated (pseudo) random numbers**

Monte Carlo Simulation

- A method for sampling the PES.
- Allows for uphill moves (but have difficulty in climbing over barriers)



- Can get stuck in a conformational region.

$X_{[0,1]}$ is a random number in the range 0 to 1, *i.e.* PROABABLITY

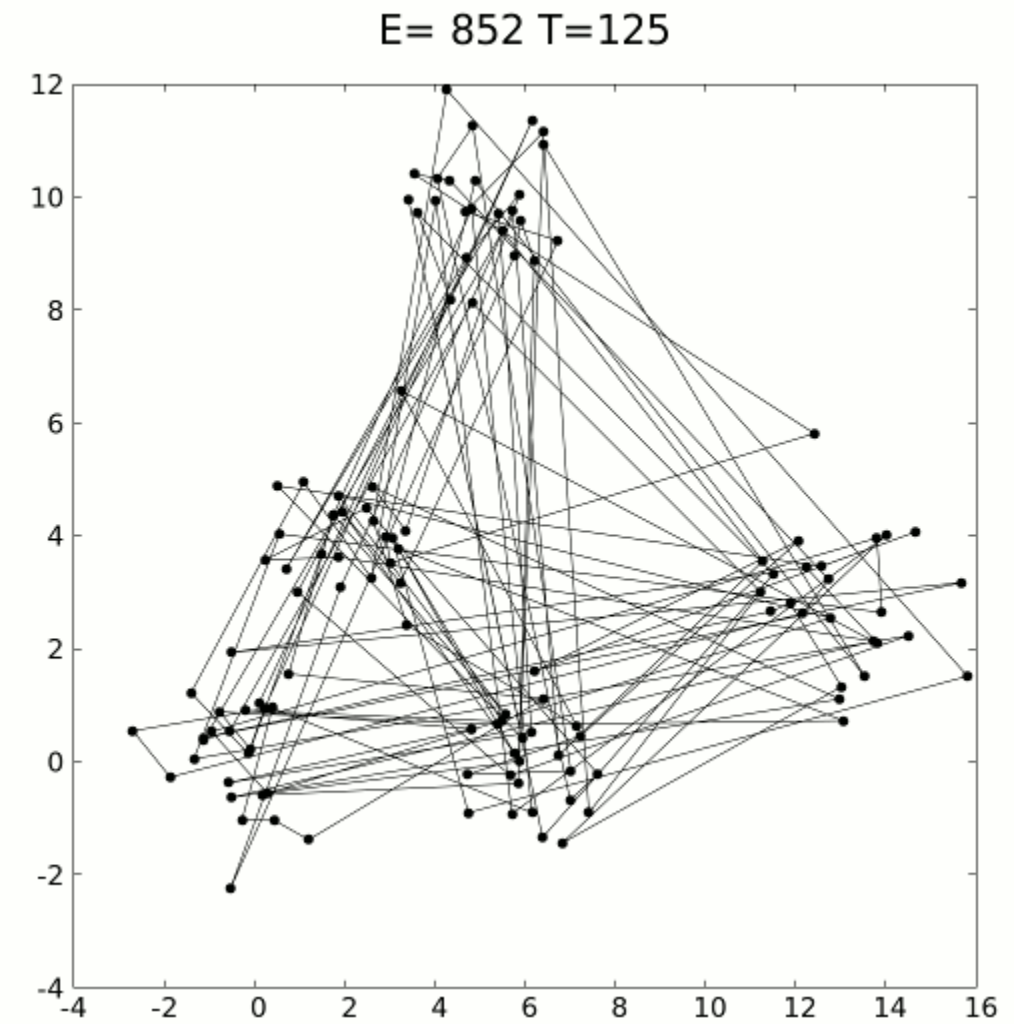
42

Theory of Monte Carlo: Basic Principles

- A computing method that relies on probabilities
- exact solutions to statistical mechanical problems
- relies on transition probabilities between different states of the simulated system
- These transitions are traced according to the following scheme:
  - generation of an initial configuration
  - trial of a randomly generated system configuration, and
  - evaluation of an “acceptance criterion” for the trial configuration
- The acceptance criterion is usually formulated in terms of the potential energy change between trial (new) and existing (old) states and some other properties of the new and old configuration

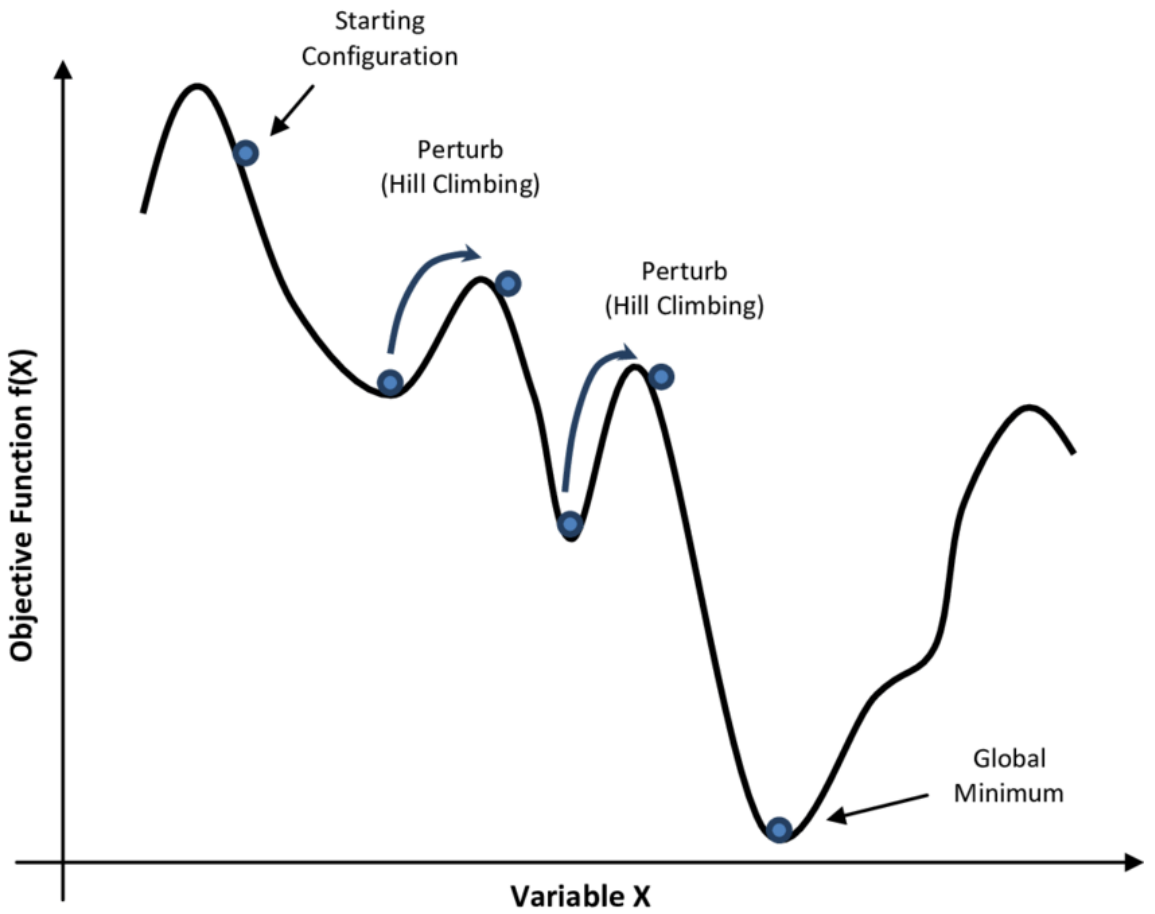
Simulated Annealing

Simulated Annealing



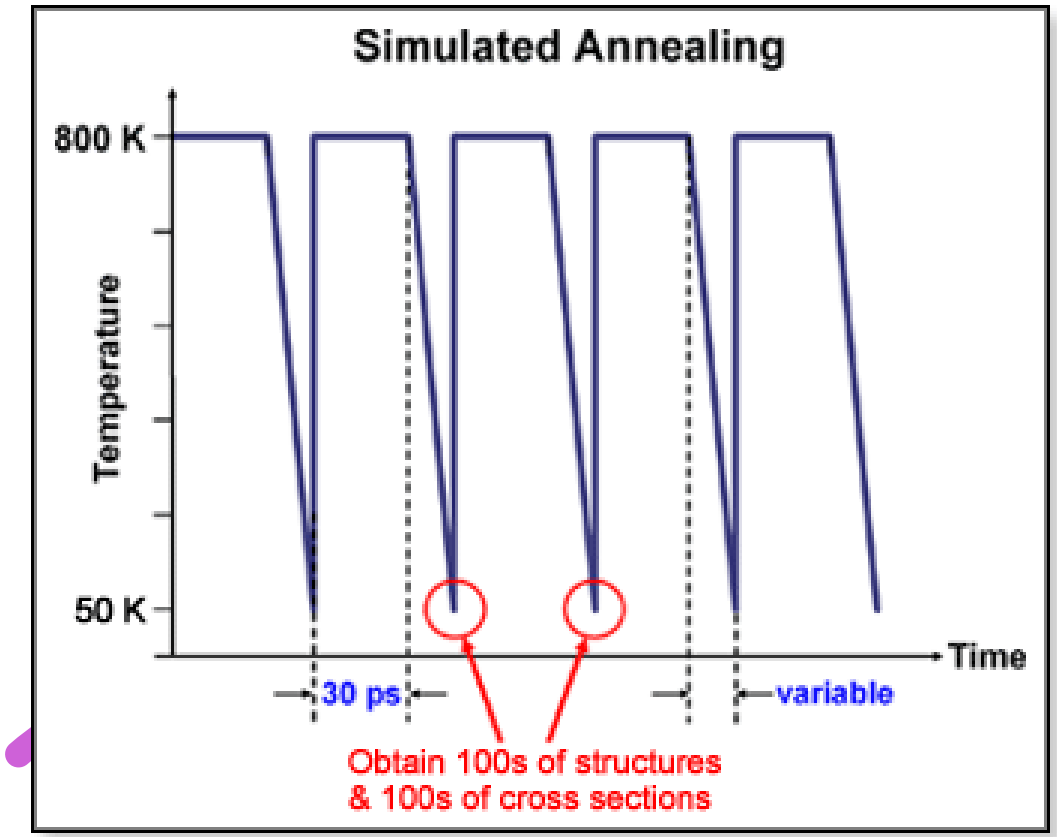
Simulated Annealing

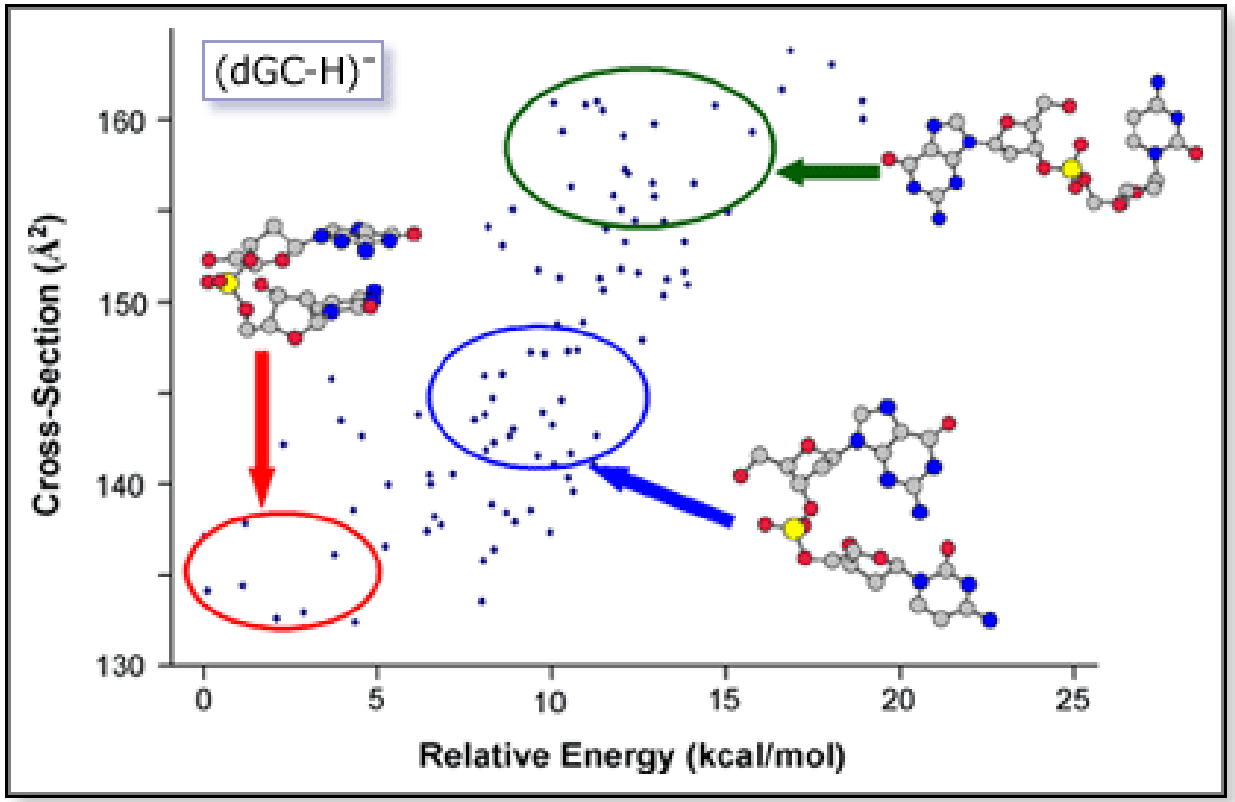
- **Hill Climbing Problem**



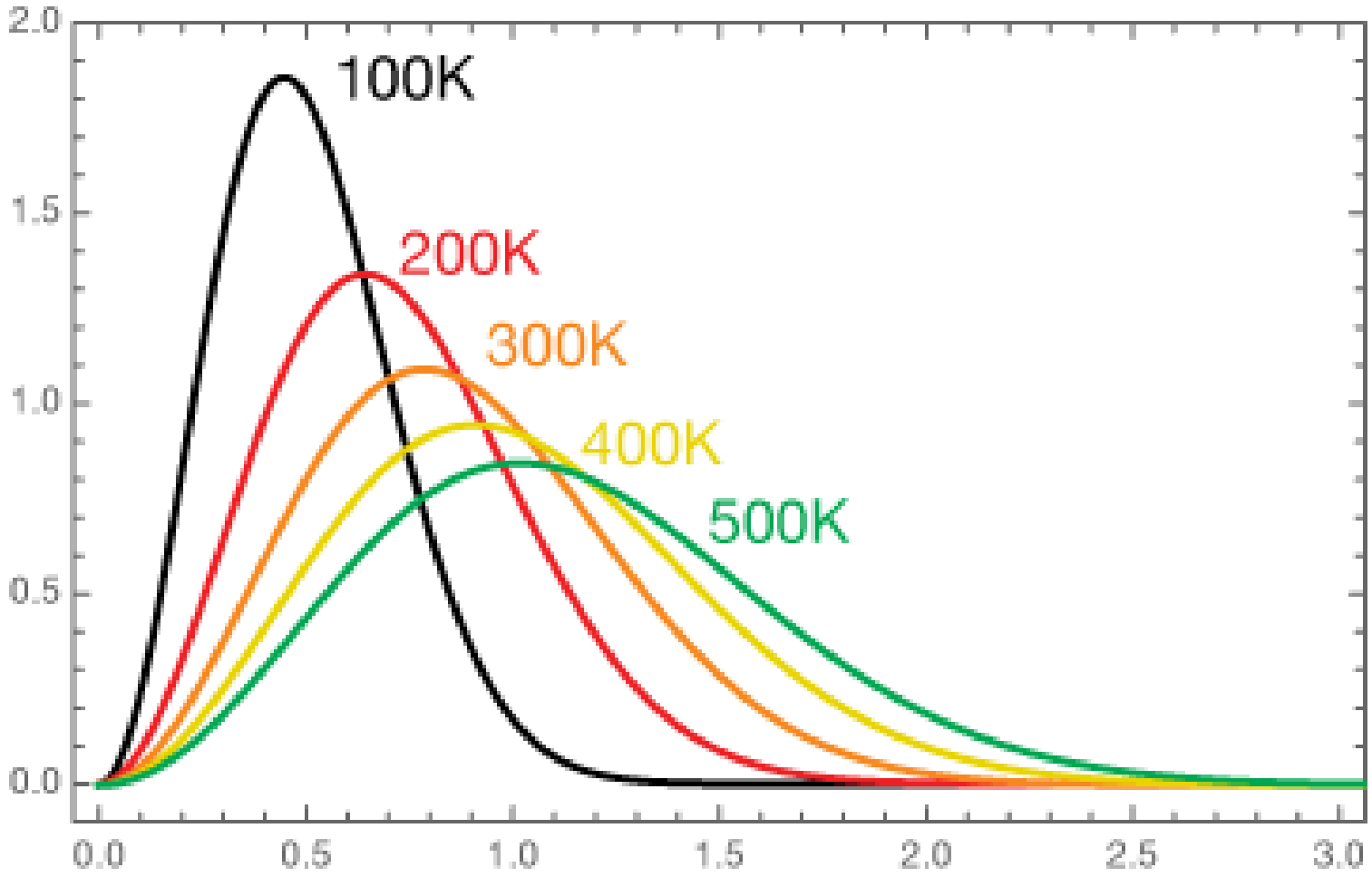
Simulated Annealing

- **Conformational Analysis**



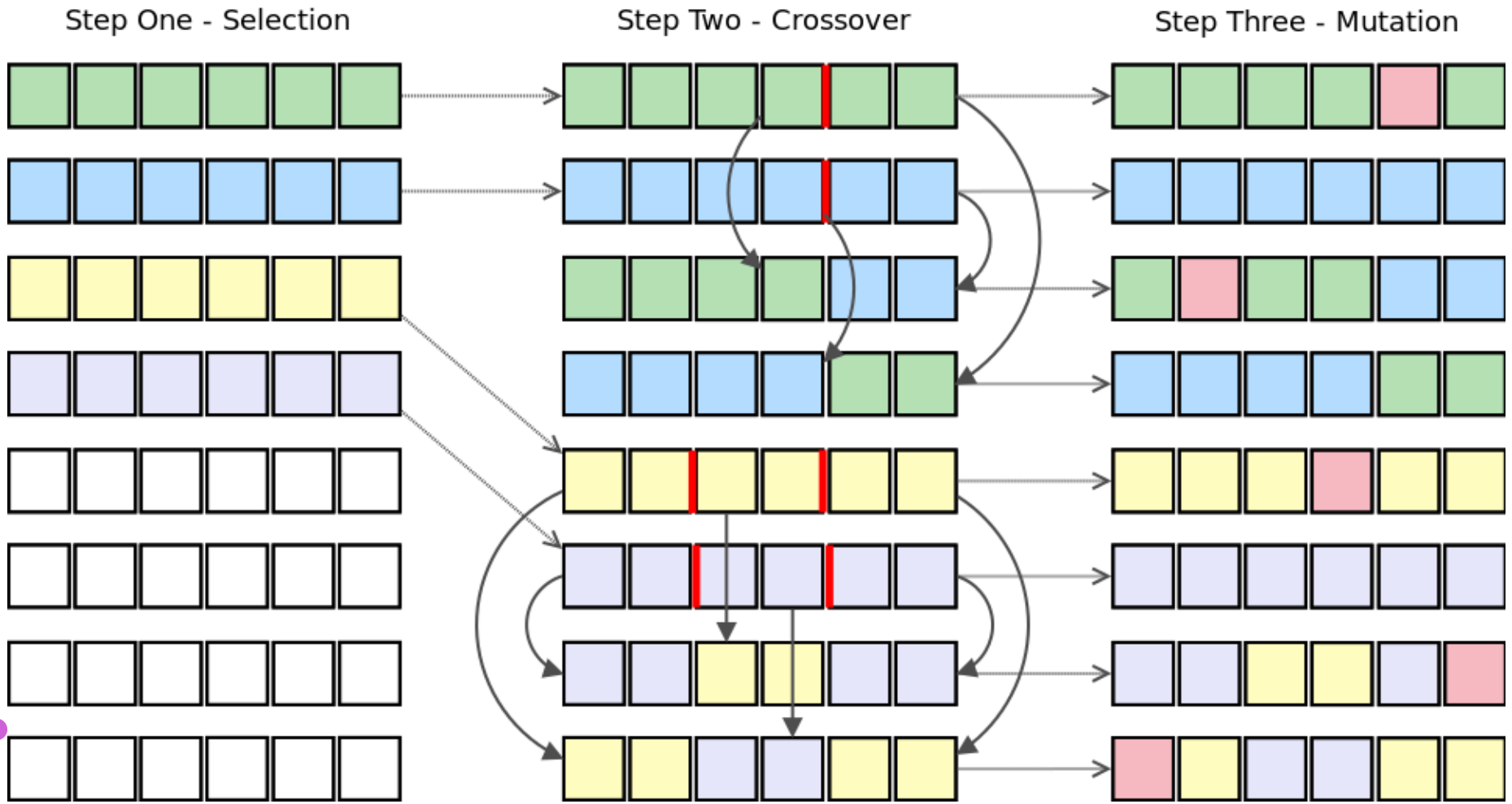


Boltzmann Distribution

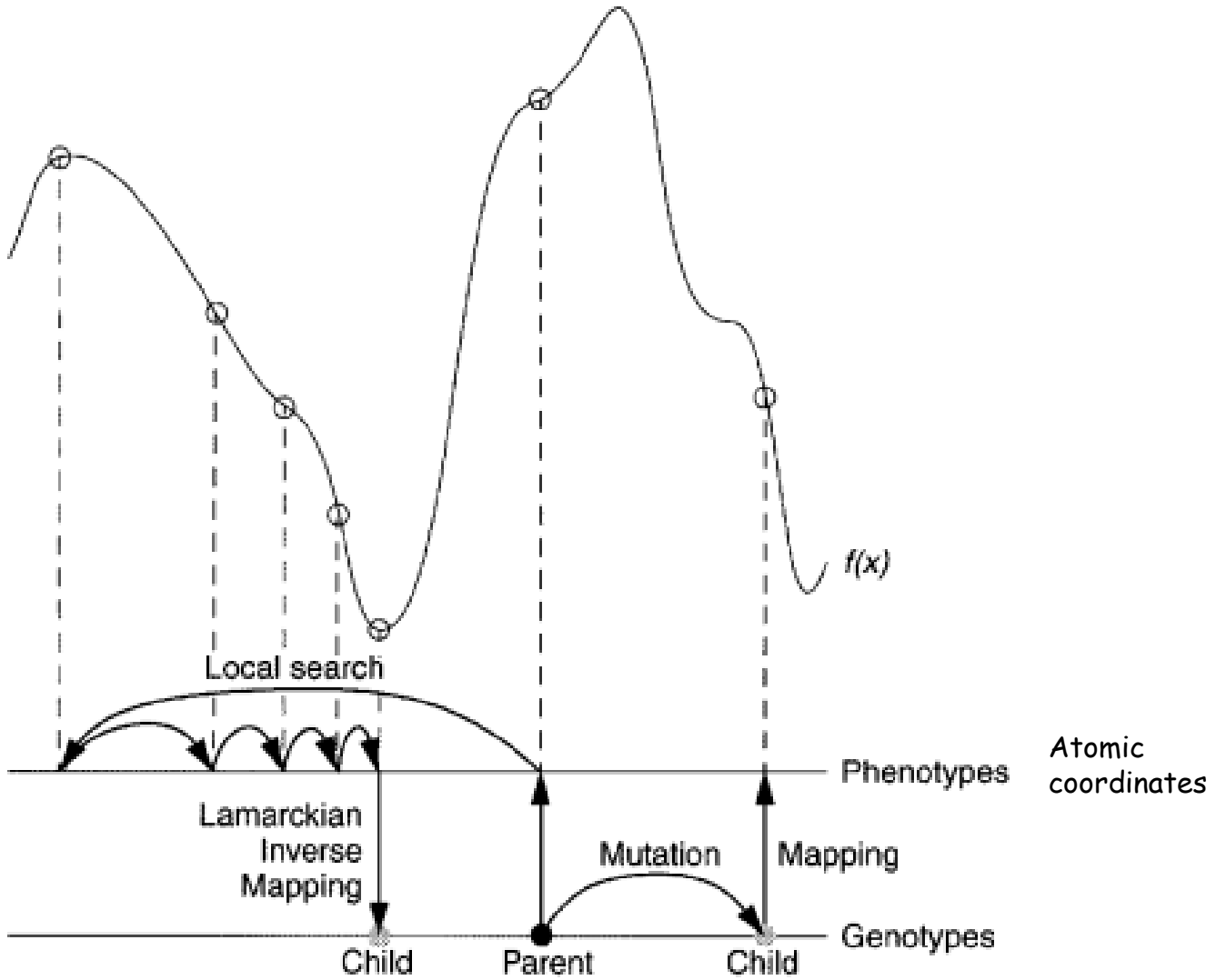


Genetic Algorithms

Genetic Algorithms: Darwinian

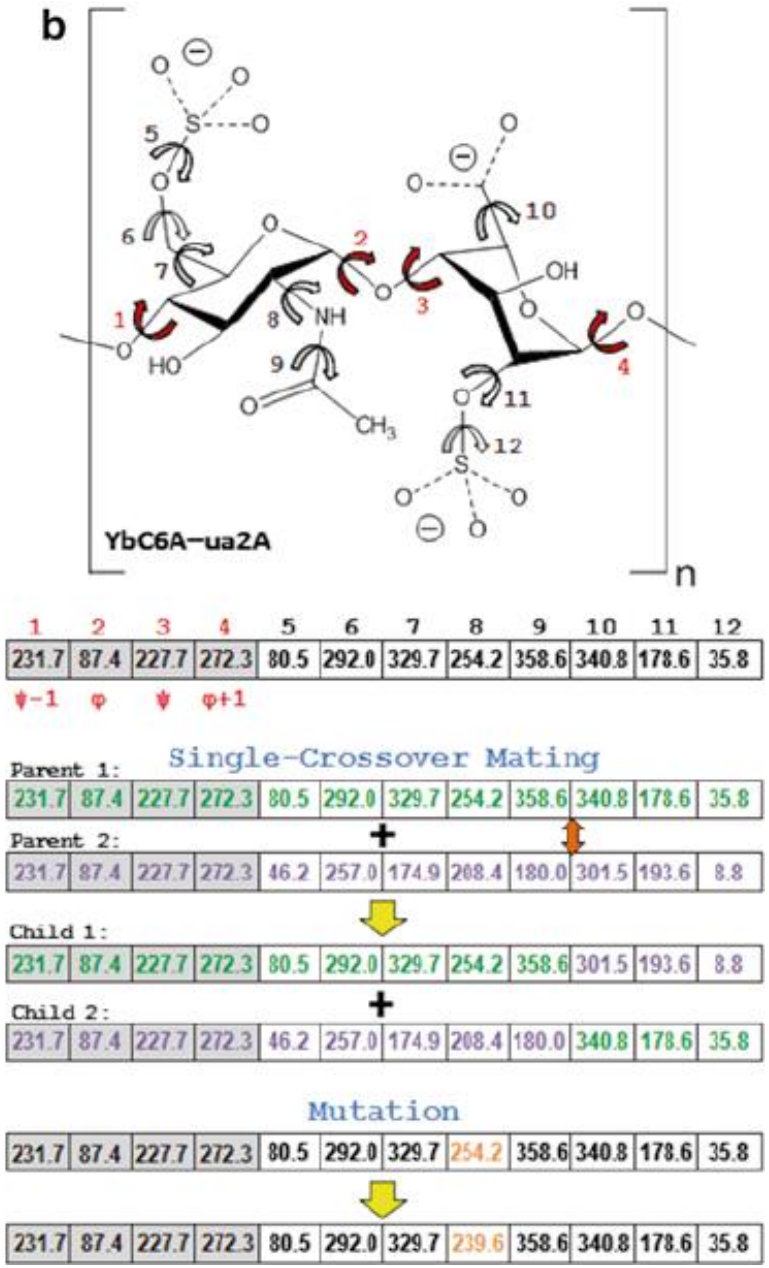


Genetic Algorithms: Lamarckian



Atomic coordinates


a
GA
Create Population of Random Conformations
Fitness Function Evaluation
Reproduction: Mating & Mutation
Fitness Function Evaluation
Replace Worst-Scoring Conformations in Population
At Iteration Limit?
NO
YES
Final Docked Solutions


b
YbC6A-ua2A
n
1 2 3 4 5 6 7 8 9 10 11 12
231.7 87.4 227.7 272.3 80.5 292.0 329.7 254.2 358.6 340.8 178.6 35.8
ψ-1 φ ψ φ+1
Single-Crossover Mating
Parent 1:
231.7 87.4 227.7 272.3 80.5 292.0 329.7 254.2 358.6 340.8 178.6 35.8
Parent 2:
231.7 87.4 227.7 272.3 46.2 257.0 174.9 208.4 180.0 301.5 193.6 8.8
Child 1:
231.7 87.4 227.7 272.3 80.5 292.0 329.7 254.2 358.6 301.5 193.6 8.8
Child 2:
231.7 87.4 227.7 272.3 46.2 257.0 174.9 208.4 180.0 340.8 178.6 35.8
Mutation
231.7 87.4 227.7 272.3 80.5 292.0 329.7 254.2 358.6 340.8 178.6 35.8
231.7 87.4 227.7 272.3 80.5 292.0 329.7 239.6 358.6 340.8 178.6 35.8

THANK YOU