

# Molecular Dynamics Simulations

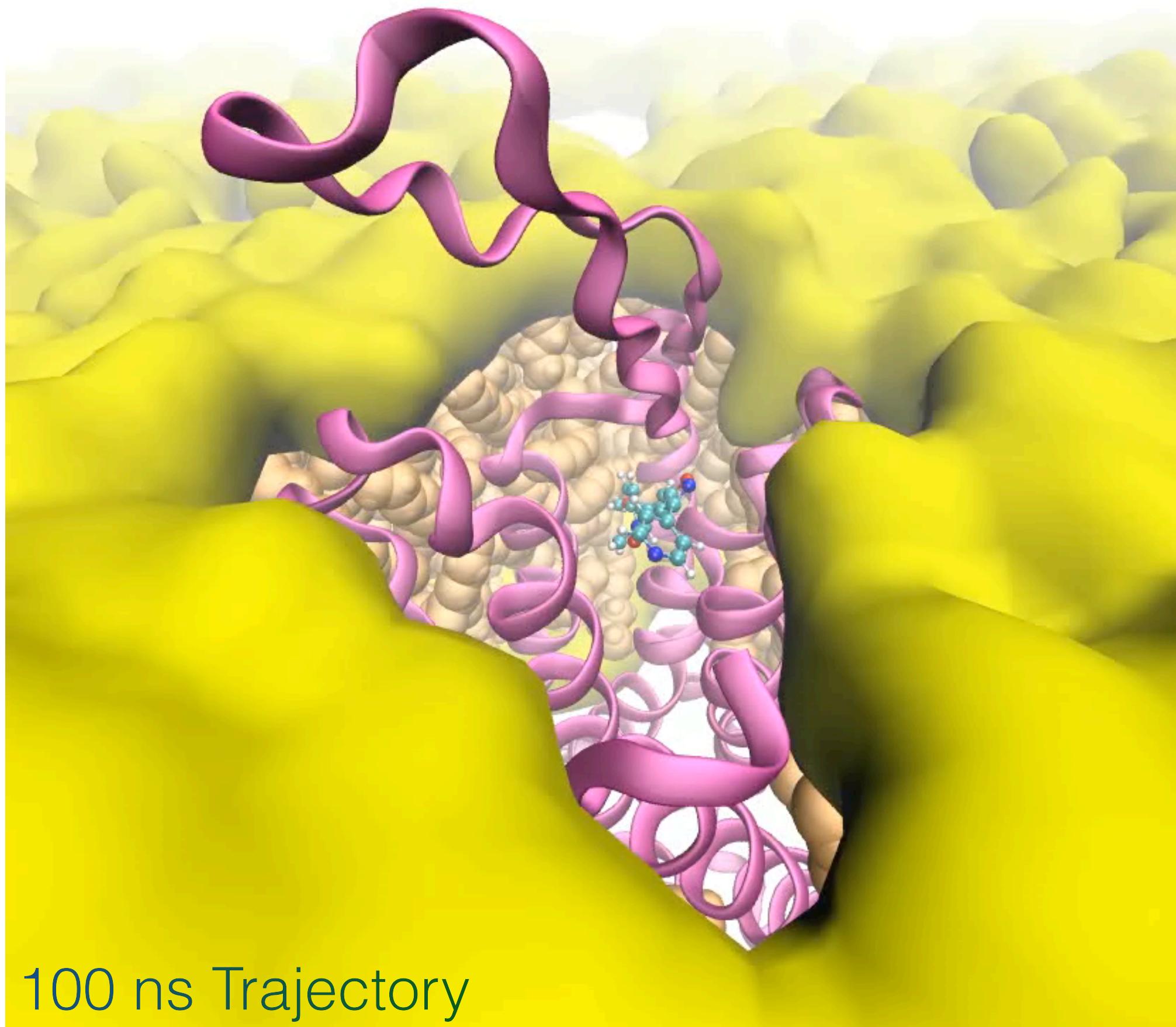
---

**CCATS Group**

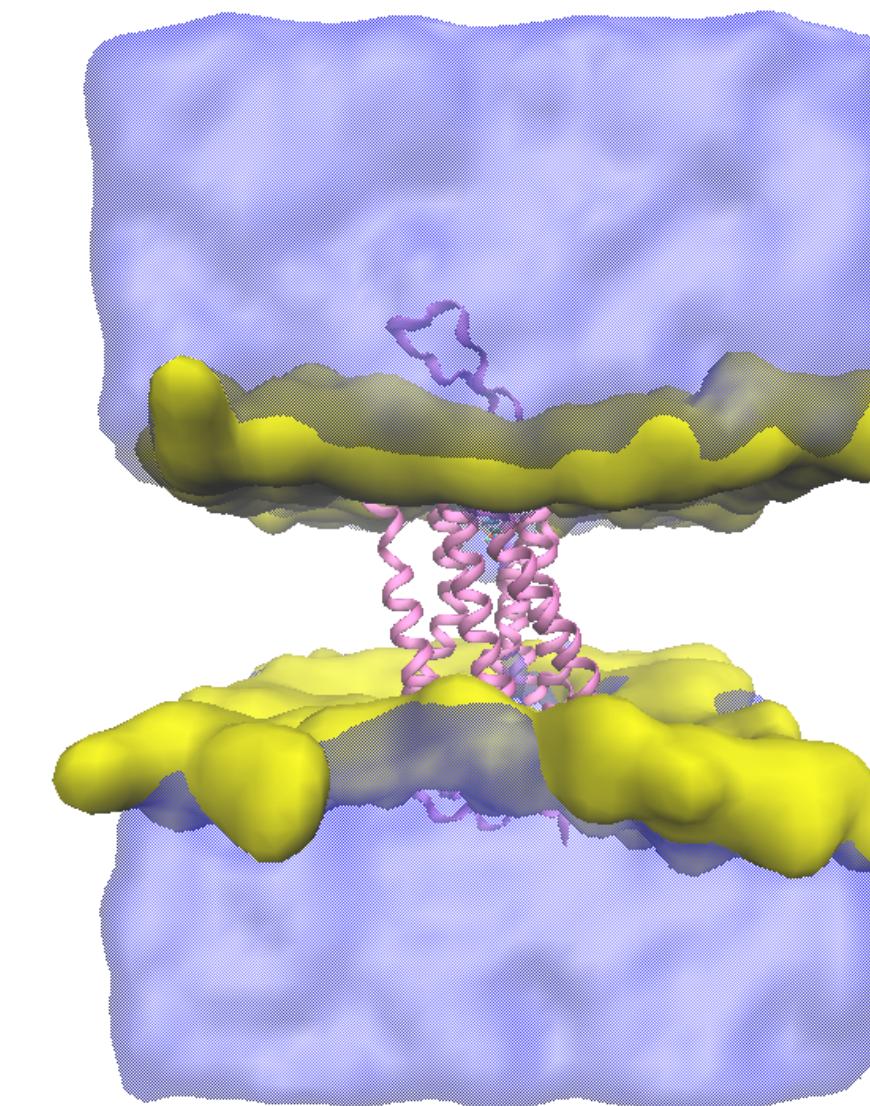


# What Do We Do in Molecular Dynamics Simulations?

- Use Newtonian mechanics to capture the real-time motion of chemical or biological systems



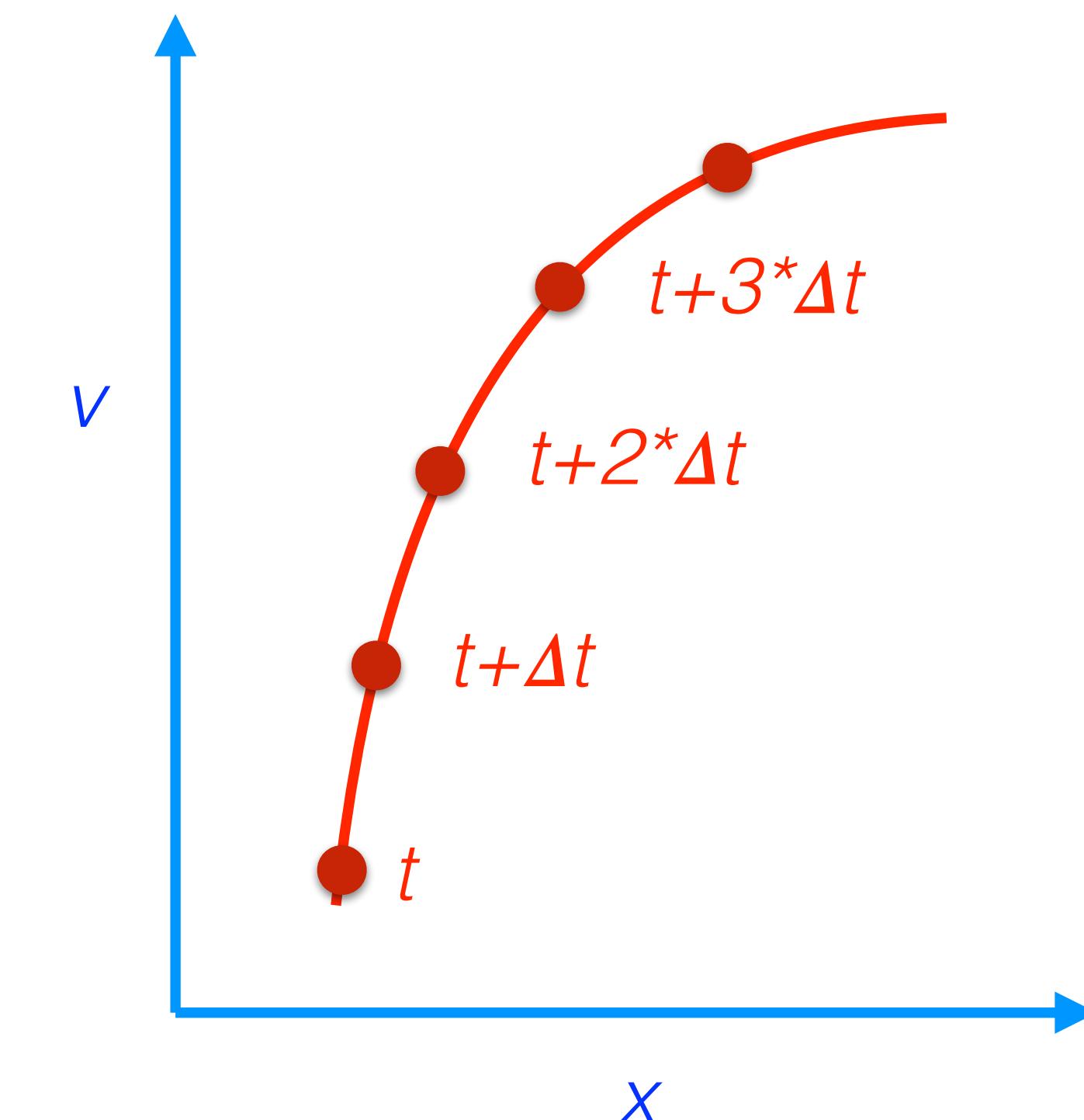
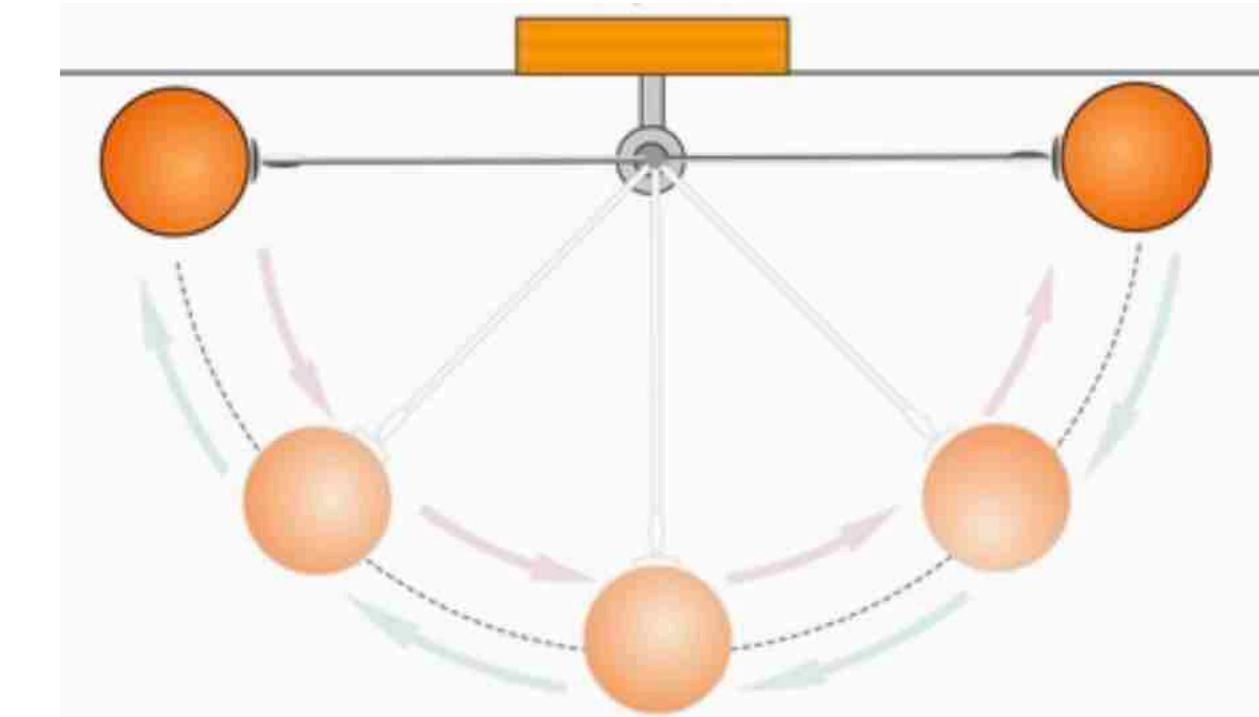
Truncated mGluR7 (PDB 7EPC)  
MMPIP (ligand; AutoDock Vina)  
POPC (lipid membrane)  
Water



\*Lipid tail group not shown

# What Do We Do in Molecular Dynamics Simulations?

- Use Newtonian mechanics to capture the real-time motion of chemical or biological systems
- Coordinates:  $\mathbf{x} = [x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N, z_N]$
- Velocities:  $\mathbf{v} = [v_{x1}, v_{y1}, v_{z1}, v_{x2}, v_{y2}, v_{z2}, \dots, v_{xN}, v_{yN}, v_{zN}]$
- Classical trajectory:  $\mathbf{x}(t_0), \mathbf{v}(t_0) \longrightarrow \mathbf{x}(t_0 + \Delta t), \mathbf{v}(t_0 + \Delta t)$ 
  - coordinate propagation:  $\mathbf{x}(t_0 + \Delta t) = \mathbf{x}(t_0) + \mathbf{v}(t_0)\Delta t$
  - velocity propagator:  $\mathbf{v}(t_0 + \Delta t) = \mathbf{v}(t_0) - \frac{1}{m} \frac{\partial U(\mathbf{x})}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\mathbf{x}(t_0)} \Delta t$



# What Do We Need to Perform a MD Simulation?

---

- A good computer (like schooner)
- MD simulation software
  - AMBER (Peter Kollman, David Case, Carols Simmingling, Adrian Roitberg, Ken Merz, Junmei Wang, Tom Cheatham)
  - CHARMM (Martin Karplus, Charles Brooks, Bernard Brooks, Alex McKerell, Benoit Roux, ....)
  - GROMACS (Erik Lindahl, Berend Smit)
  - NAMD (Klaus Schulten)
  - OpenMM (John Chodera, Tom Markland, Bernard Brooks)
- An initial structure (and atomic velocity):  $\mathbf{x}(t_0), \mathbf{v}(t_0)$ 
  - Adding missed residues, hydrogen atoms, solvent molecules, counter ions
- A trajectory propagation scheme:  $\mathbf{x}(t_0), \mathbf{v}(t_0) \longrightarrow \mathbf{x}(t_0 + \Delta t), \mathbf{v}(t_0 + \Delta t)$ 
  - Velocity Verlet, leapfrog

# Different Types of Basic MD Simulations

---

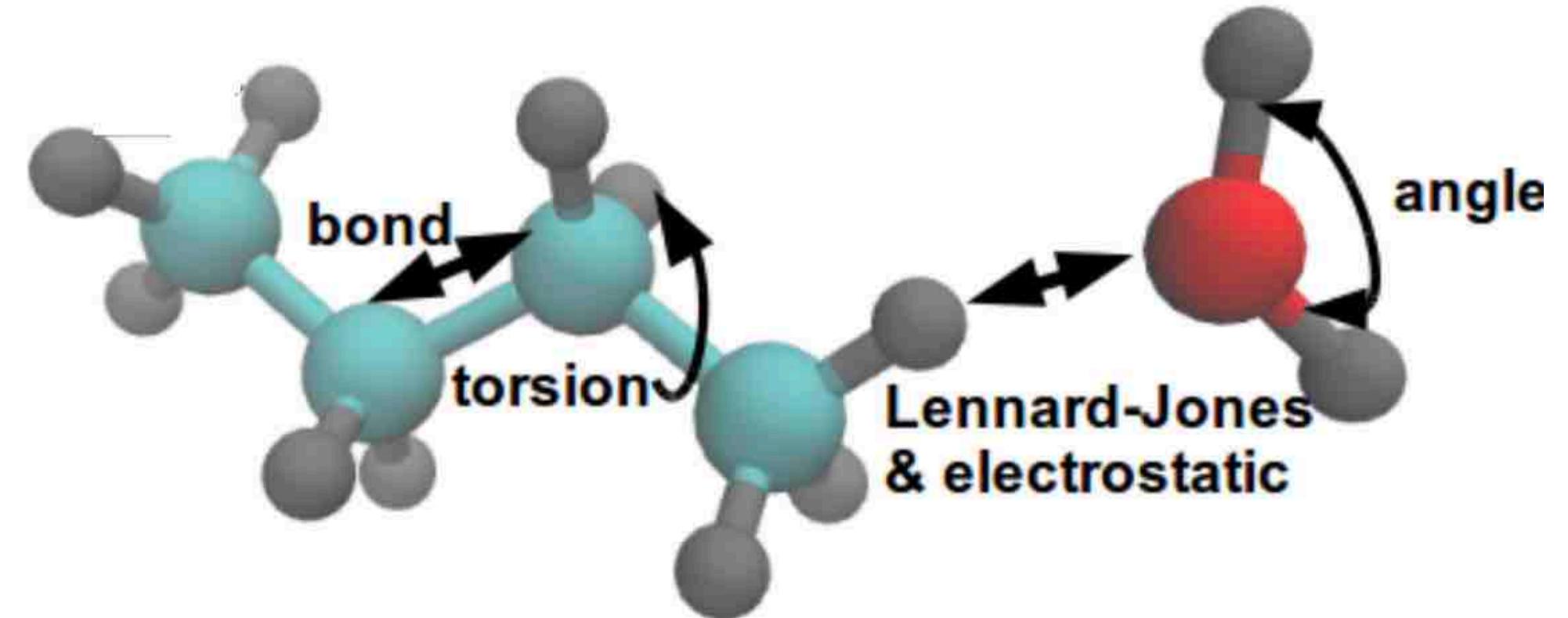
- NVE (microcanonical ensemble)
  - Constant number of particles, constant volume, and constant energy
  - The system does not have an uniform temperature
- NPT (canonical ensemble)
  - Constant number of particles, constant pressure, and constant temperature
- NVT (canonical ensemble)
  - Constant number of particles, constant volume, and constant temperature

# Potential Energy: Classical Force Fields

- $$U = U_{\text{bond}} + U_{\text{angle}} + U_{\text{torsion}}$$
  
$$+ U_{\text{van-der-Waals}} + U_{\text{electrostatic}}$$

bonded terms

non-bonded terms



$$U_{\text{bond}} = \sum_{b \in \text{bonds}} k_b (r_b - r_b^{\text{eq}})^2$$

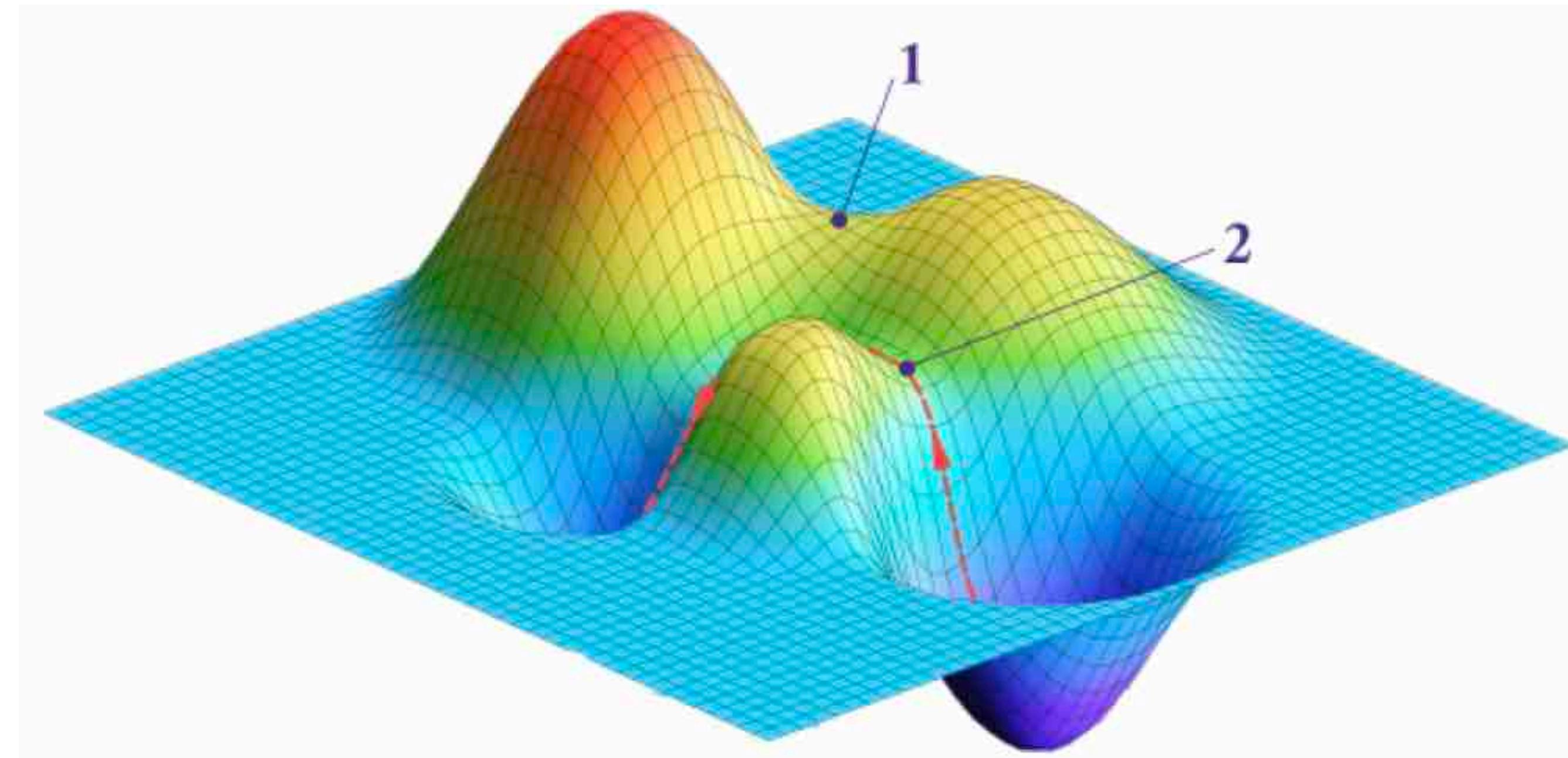
$$U_{\text{angle}} = \sum_{a \in \text{angles}} k_a (\theta_a - \theta_a^{\text{eq}})^2$$

$$U_{\text{torsion}} = \sum_{t \in \text{torsions}} \sum_{n=1}^3 k_t [1 + \cos(n\phi_t - \gamma_n)]$$

$$U_{\text{vdW}} = \sum_{i < j}^{\text{atoms}} \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - 2 \left( \frac{\sigma}{r_{ij}} \right)^6 \right]$$

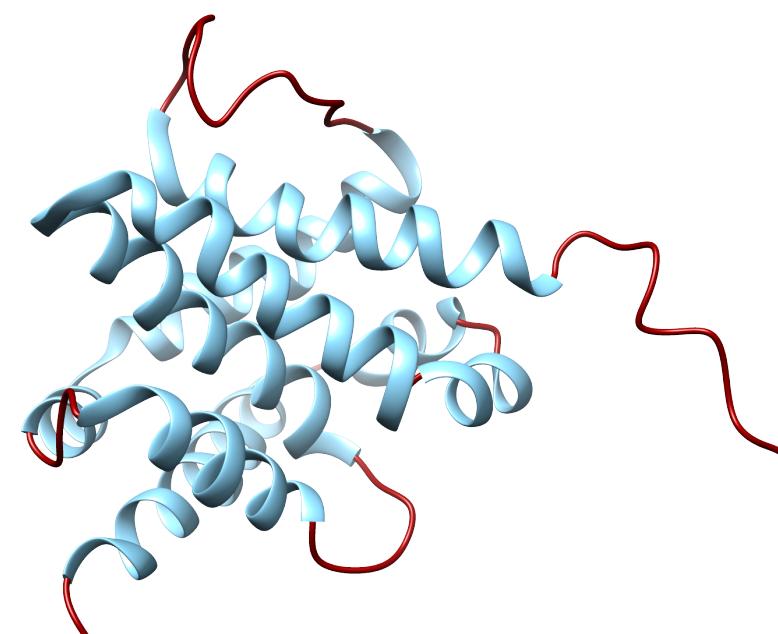
$$U_{\text{electrostatic}} = \sum_{i < j}^{\text{atoms}} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

# Potential Energy Surfaces



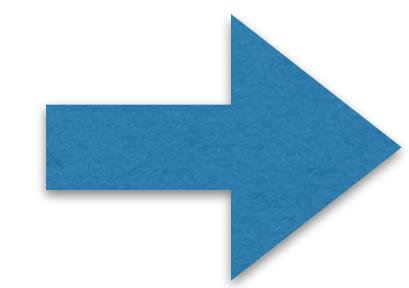
- Classical mechanical mechanics (MM) force fields
  - Macromolecular conformational changes; ligand-receptor binding
- Ground-state QM/MM energy functions: enzyme reactions
- Excited-state QM/MM energy functions: photochemistry and photobiology

# General Workflow



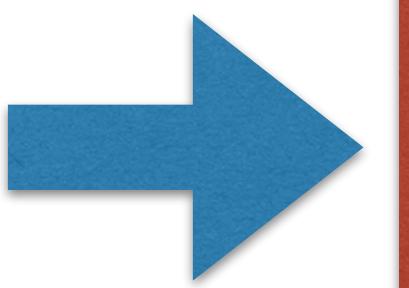
## Initial coordinates

- RCSB
- Molecular Docking
- Alpha-RoseTTA-  
Omega-Fold

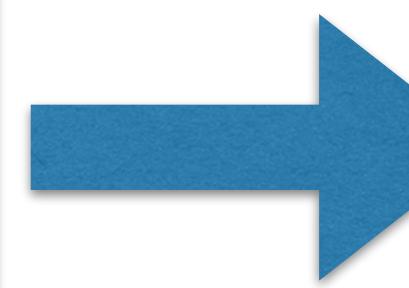


Parameter / Topology File  
• atom/bond/angles information

Coordinate File  
• Atom coordinates  
• Box information  
• Sometimes velocities



Equilibration (0.5 - 20 ns)  
Minimization  
• Solvent then solute  
Heating (NVT)  
• Heat to temp (300K)  
Pressure (NPT)  
• Density to 1 bar



Production (100 ns <)  
• NVT or NPT

# OpenMM with AMBER Force Field

---

- We will follow the Colab notebook prepared by Arantes and co-workers:

<https://github.com/pablo-arantes/making-it-rain>

- Copy the modified version to your Google Drive:

<https://colab.research.google.com/drive/1ZNvKJV2zIZg85H5Ja6QILhLUW9ozCxdM?usp=sharing>

- Changes I made are:

- Example is Bax (PDB 1F16)
- Input/Output names are kept consistent
- Units are kept consistent

# **MM-GBSA / MM-PBSA**

**&**

# **Trajectory Analysis**

---

**CCATS Group**



# **MM-GBSA / MM-PBSA**

---

- Molecular Mechanics - Generalized Born Surface Area (MM-GBSA)
- Molecular Mechanics - Poisson-Boltzmann Surface Area (MM-PBSA)
- Predict the binding free energy of complexes or transitions between conformations (State A →B)
  - More accurate than scoring functions (Molecular Docking)
  - Not computationally expensive as other methods (Alchemical, etc.)

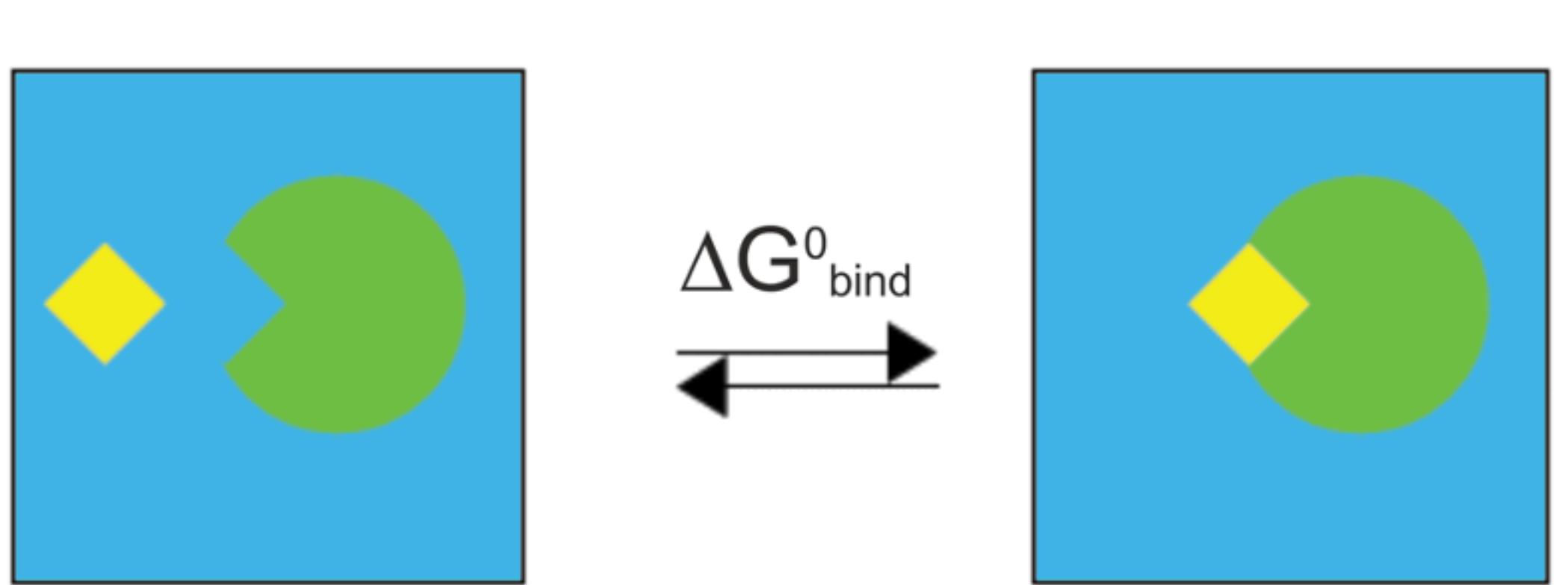
---

MM      Type of energy function used to compute potential energy of molecular structure (force fields).

---

GBSA/  
PBSA      Calculates the electrostatic potentials of a solute in solution.

# MM-GBSA / MM-PBSA



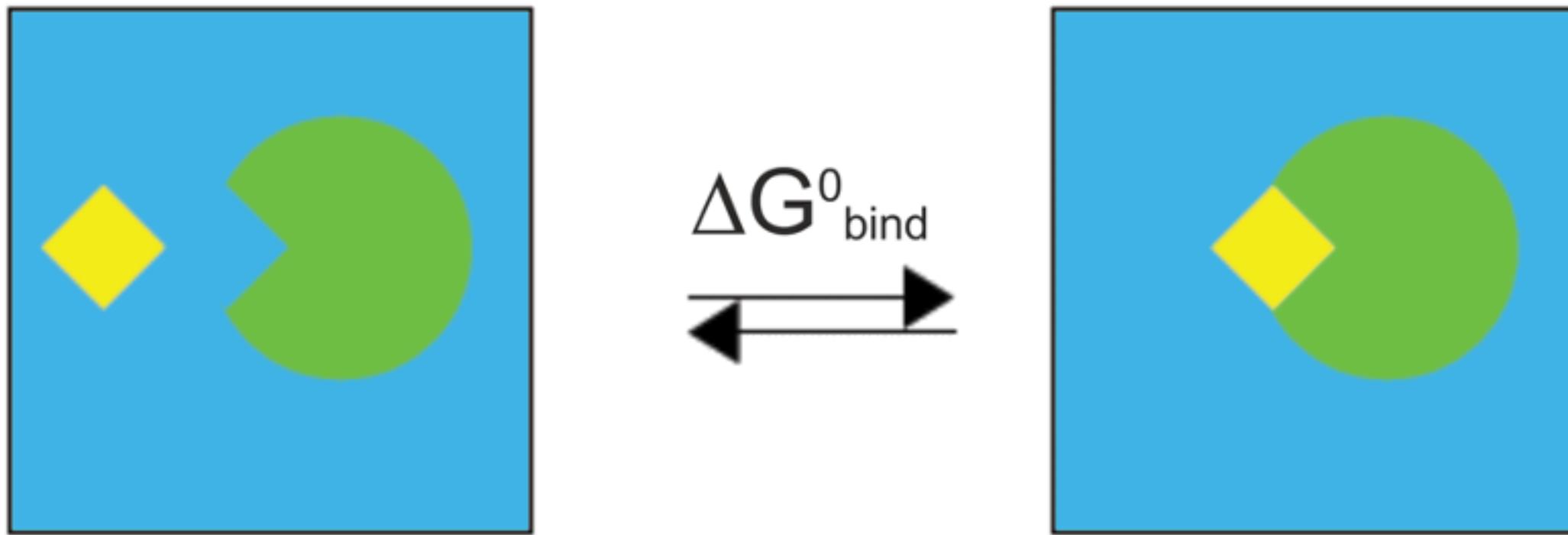
- Ligand Binding Example
- Goal is to calculate the free energy difference between two states (bound and unbound) of *solvated* molecules
- PB/GB = Implicit Solvent Model
- GB: Generalized Born
- PB: Poisson-Boltzmann

$$(1) \quad \Delta G_{bind} = \langle G_{Complex} \rangle - [\langle G_{Protein} \rangle + \langle G_{Ligand} \rangle]$$

$$(1a) \quad \underline{\Delta G_{bind} = E_{bonded} + E_{nonbonded} + G_{polar} + G_{nonpolar} - TS}_{\text{MM}}$$

GB/PB      SASA  
 $\Delta G_{solvation}$

# MM-GBSA



- Generalized Born (1/4)

$$\Delta G_{polar} \approx -\frac{1}{2} \sum_{ij} \frac{q_i q_j}{f_{GB}(r_{ij} R_i R_j)} \left( 1 - \frac{\exp[-\kappa f_{GB}]}{\epsilon} \right)$$

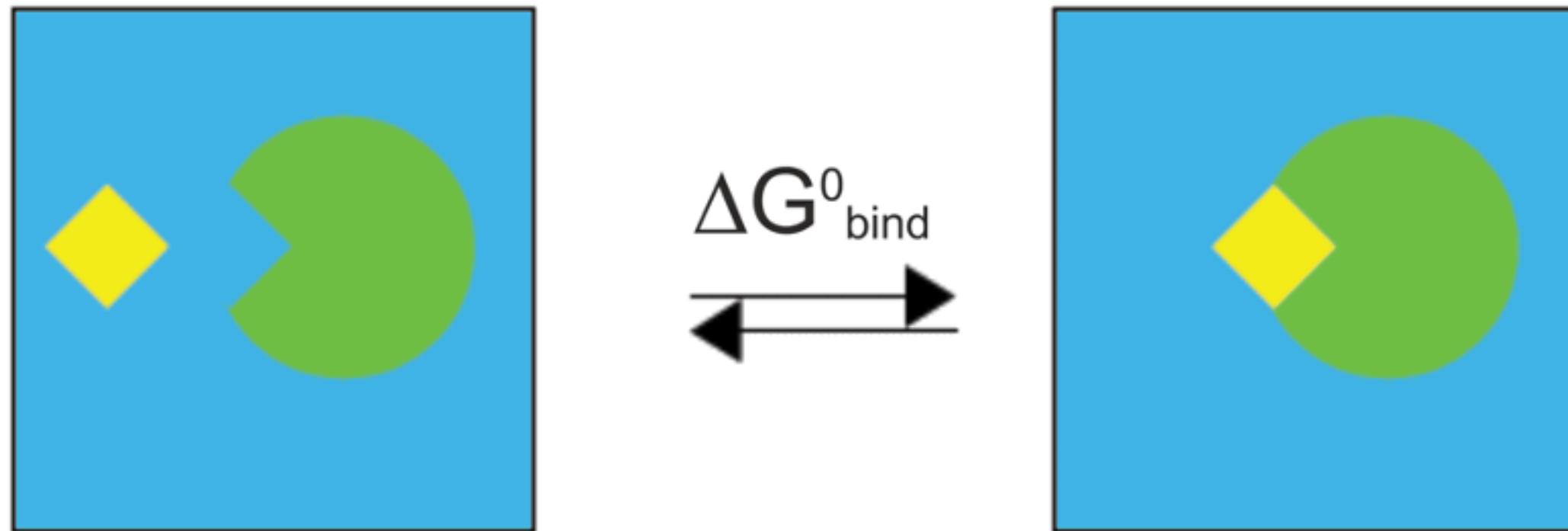
(1)  $\Delta G_{bind} = \langle G_{Complex} \rangle - [\langle G_{Protein} \rangle + \langle G_{Ligand} \rangle]$

(1a)  $\Delta G_{bind} = E_{bonded} + E_{nonbonded} + G_{polar} + G_{nonpolar} - TS$

MM

GB/PB      SASA  
 $\Delta G_{solvation}$

# MM-PBSA



- Generalized Born (1/4)

$$\Delta G_{polar} \approx -\frac{1}{2} \sum_{ij} \frac{q_i q_j}{f_{GB}(r_{ij} R_i R_j)} \left( 1 - \frac{\exp[-\kappa f_{GB}]}{\epsilon} \right)$$

- Poisson-Boltzmann (ALPB, 1/4)

$$\Delta G_{polar} \approx -\frac{1}{2} \left( \frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_{ex}} \right) \frac{1}{1 + \alpha\beta} \sum_{ij} q_i q_j \left( \frac{1}{f_{GB}} + \frac{\alpha\beta}{A} \right)$$

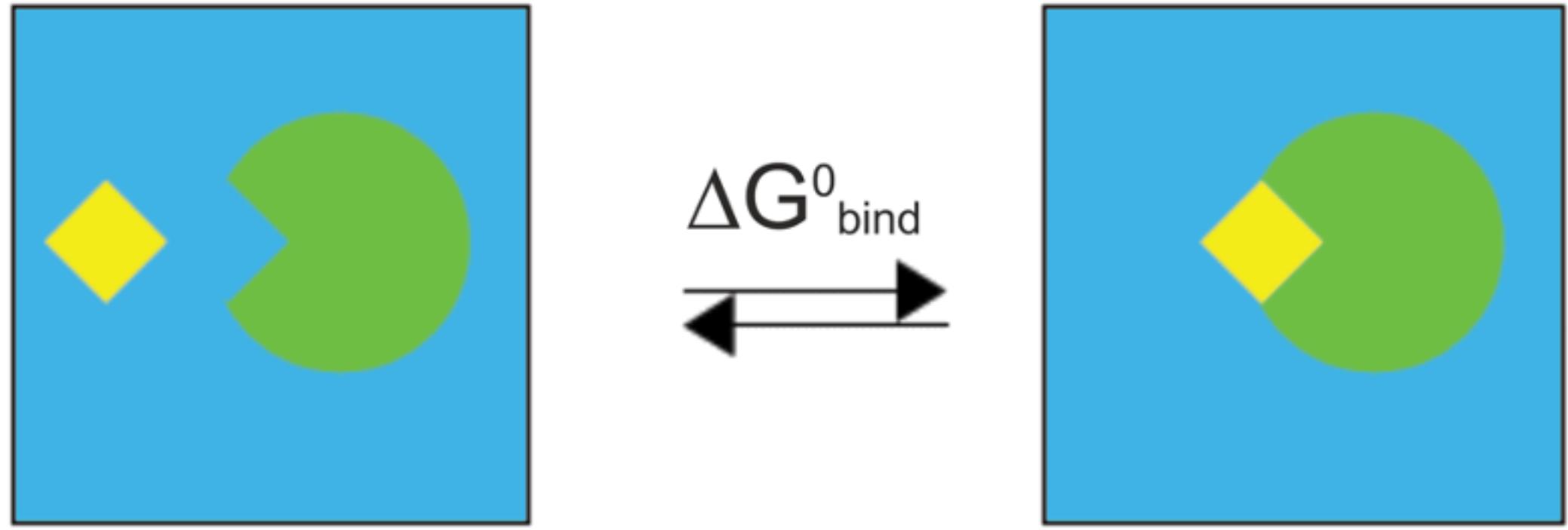
(1)  $\Delta G_{bind} = \langle G_{Complex} \rangle - [\langle G_{Protein} \rangle + \langle G_{Ligand} \rangle]$

(1a)  $\Delta G_{bind} = E_{bonded} + E_{nonbonded} + G_{polar} + G_{nonpolar} - TS$

MM

GB/PB      SASA  
 $\Delta G_{solvation}$

# SASA



- Generalized Born (1/4)

$$\Delta G_{polar} \approx -\frac{1}{2} \sum_{ij} \frac{q_i q_j}{f_{GB}(r_{ij} R_i R_j)} \left( 1 - \frac{\exp[-\kappa f_{GB}]}{\epsilon} \right)$$

- Poisson-Boltzmann (ALPB, 1/4)

$$\Delta G_{polar} \approx -\frac{1}{2} \left( \frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_{ex}} \right) \frac{1}{1 + \alpha\beta} \sum_{ij} q_i q_j \left( \frac{1}{f_{GB}} + \frac{\alpha\beta}{A} \right)$$

- Solvent Accessible Surface Area (SASA)

$$\Delta G_{nonpolar} = \gamma SASA$$

$$(1) \quad \Delta G_{bind} = \langle G_{Complex} \rangle - [\langle G_{Protein} \rangle + \langle G_{Ligand} \rangle]$$

- Linear Combination of Pairwise Overlaps (LCPO)

$$S_i = 4\pi r^2$$

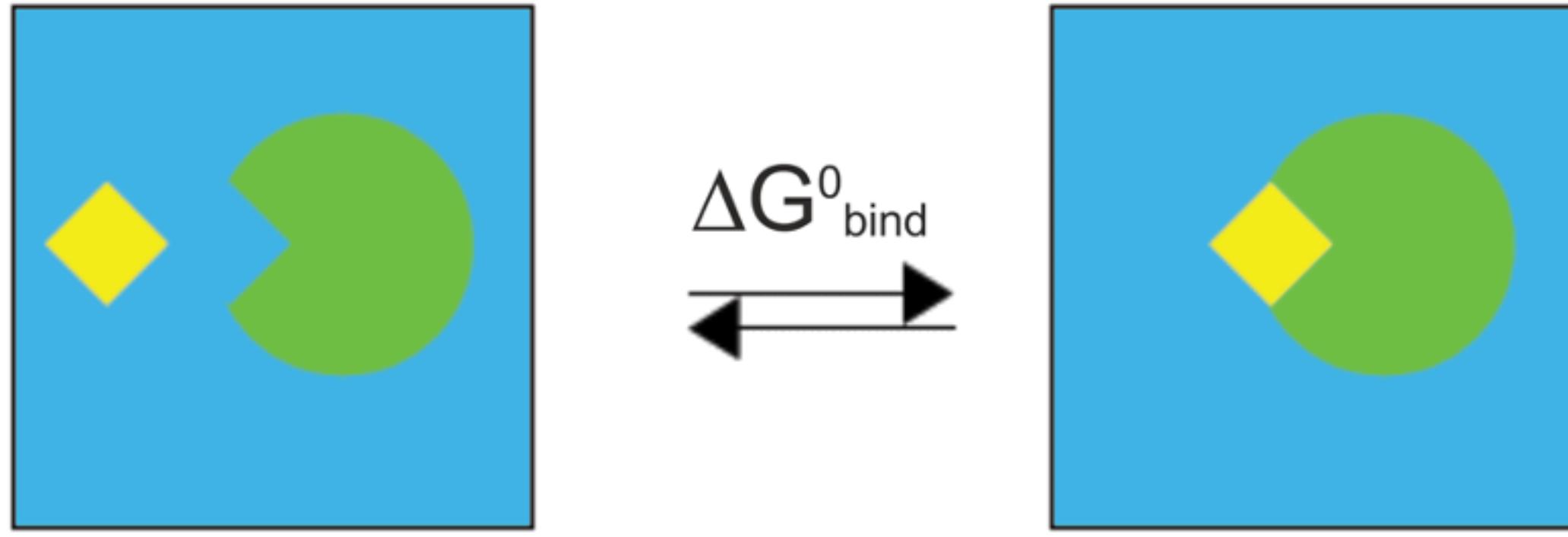
$$A_{ij} = 2\pi r_i \left( r_i - \frac{r_{ij}}{2} - \frac{r_i^2 - r_j^2}{2r_{ij}} \right)$$

$$SASA_i = \beta_1 S_i + \beta_2 \sum_{j \in N(i)} A_{ij} + \beta_3 \sum_{\substack{j, k \in N(i) \\ k \in N(j)}} A_{jk} + \beta_4 \sum_{j \in N(i)} \left( A_{ij} \sum_{\substack{k \in N(i) \\ k \neq j}} A_{jk} \right)$$

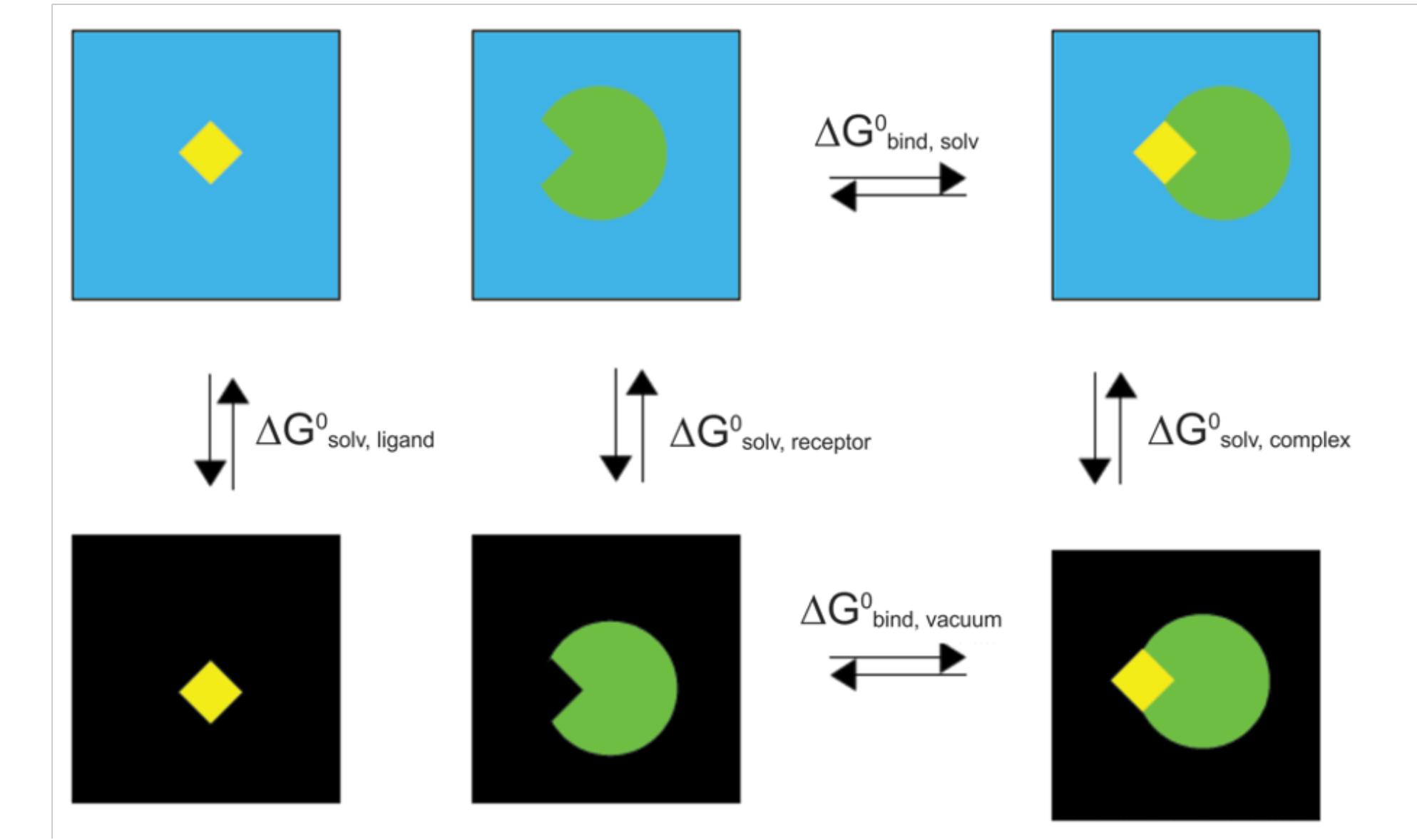
GB/PB      SASA  
 $\Delta G_{solvation}$

**MM**

# SASA



- Thermodynamic Cycle



$$\Delta G^0_{binding,complex} = \Delta G^0_{bind,vacuum} + \Delta G^0_{solv,complex} - (\Delta G^0_{solv,ligand} + \Delta G^0_{solv,protein}) \quad (2)$$

$$(1) \quad \Delta G_{bind} = \langle G_{Complex} \rangle - [\langle G_{Protein} \rangle + \langle G_{Ligand} \rangle]$$

$$(1a) \quad \underline{\Delta G_{bind} = E_{bonded} + E_{nonbonded} + G_{polar} + G_{nonpolar} - TS}$$

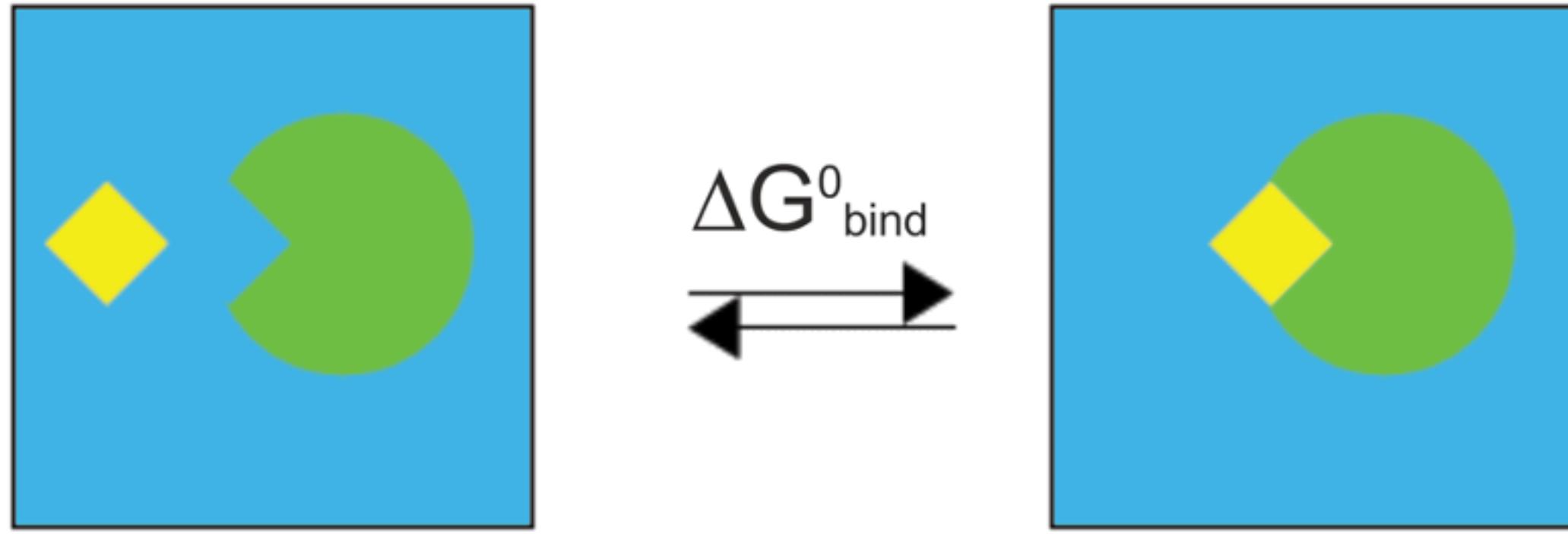
MM

GB/PB	SASA
$\Delta G_{solvation}$	

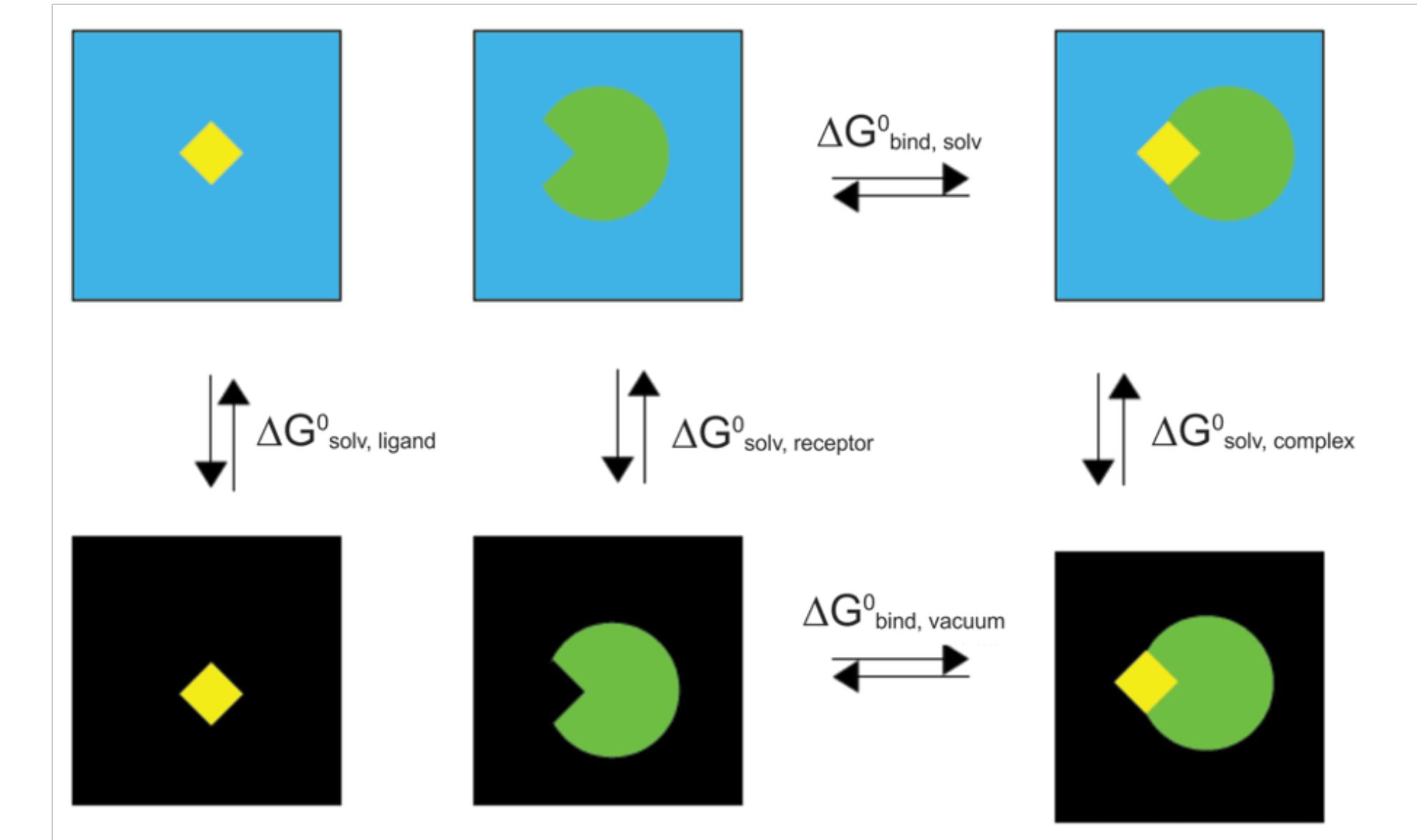
$$\Delta G^0_{solv} = \Delta G_{polar,\epsilon=80} - \Delta G_{polar,\epsilon=1} + \Delta G_{nonpolar} \quad (2a)$$

$$\Delta G_{polar,\epsilon=1} = \Delta G_{vacuum} = \Delta E_{MM} - T\Delta S_{NMA} \quad (2b)$$

# SASA



- Thermodynamic Cycle



$$\Delta G_{solv,complex}^0 = \Delta G_{bind,vacuum}^0 + \Delta G_{solv,complex}^0 - (\Delta G_{solv,ligand}^0 + \Delta G_{solv,protein}^0) \quad (2)$$

$$(1) \quad \Delta G_{bind} = \langle G_{Complex} \rangle - [\langle G_{Protein} \rangle + \langle G_{Ligand} \rangle]$$

$$(1a) \quad \underline{\Delta G_{bind} = E_{bonded} + E_{nonbonded} + G_{polar} + G_{nonpolar} - TS}$$

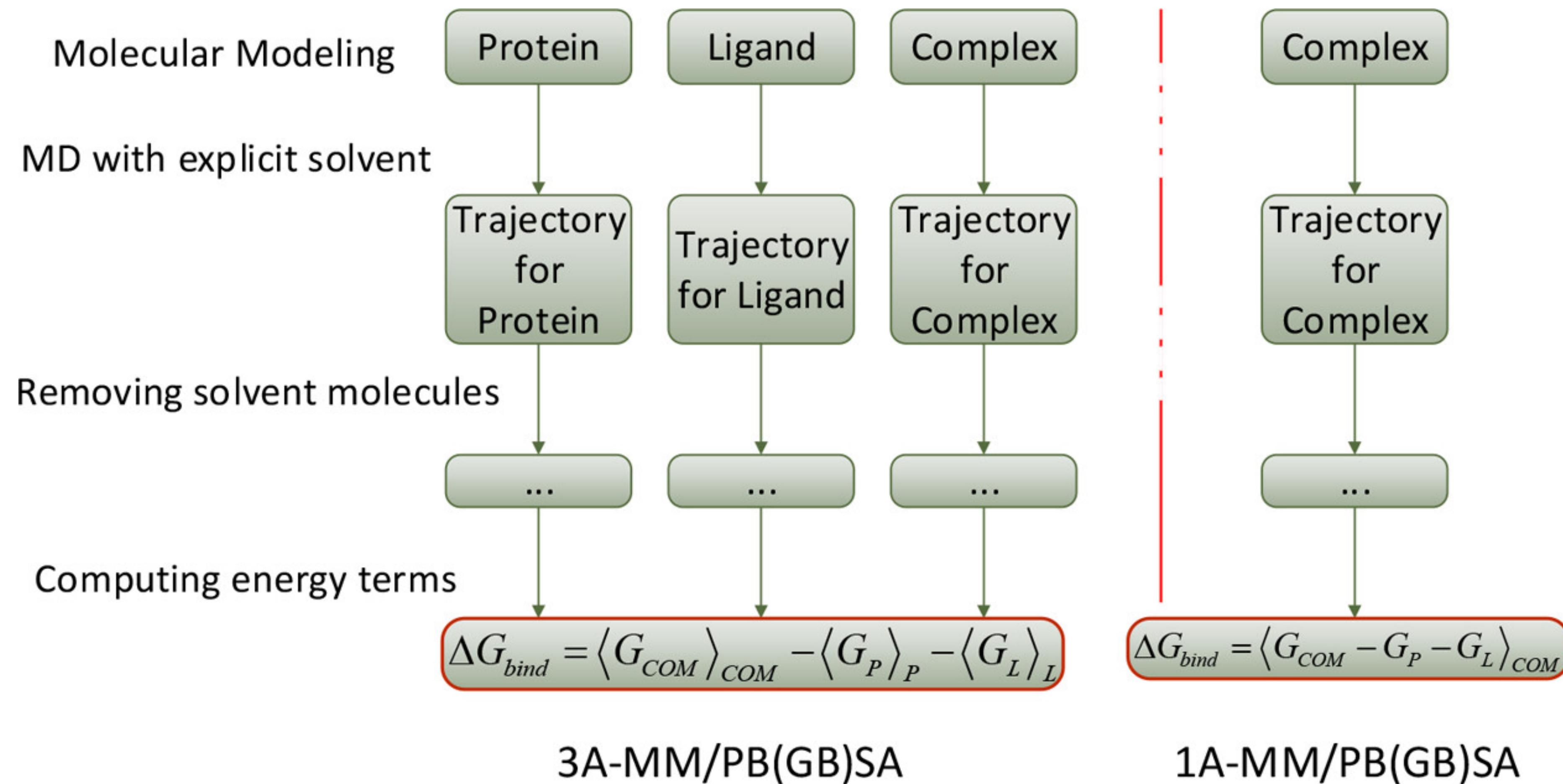
**MM**

GB/PB	SASA
$\Delta G_{solvation}$	

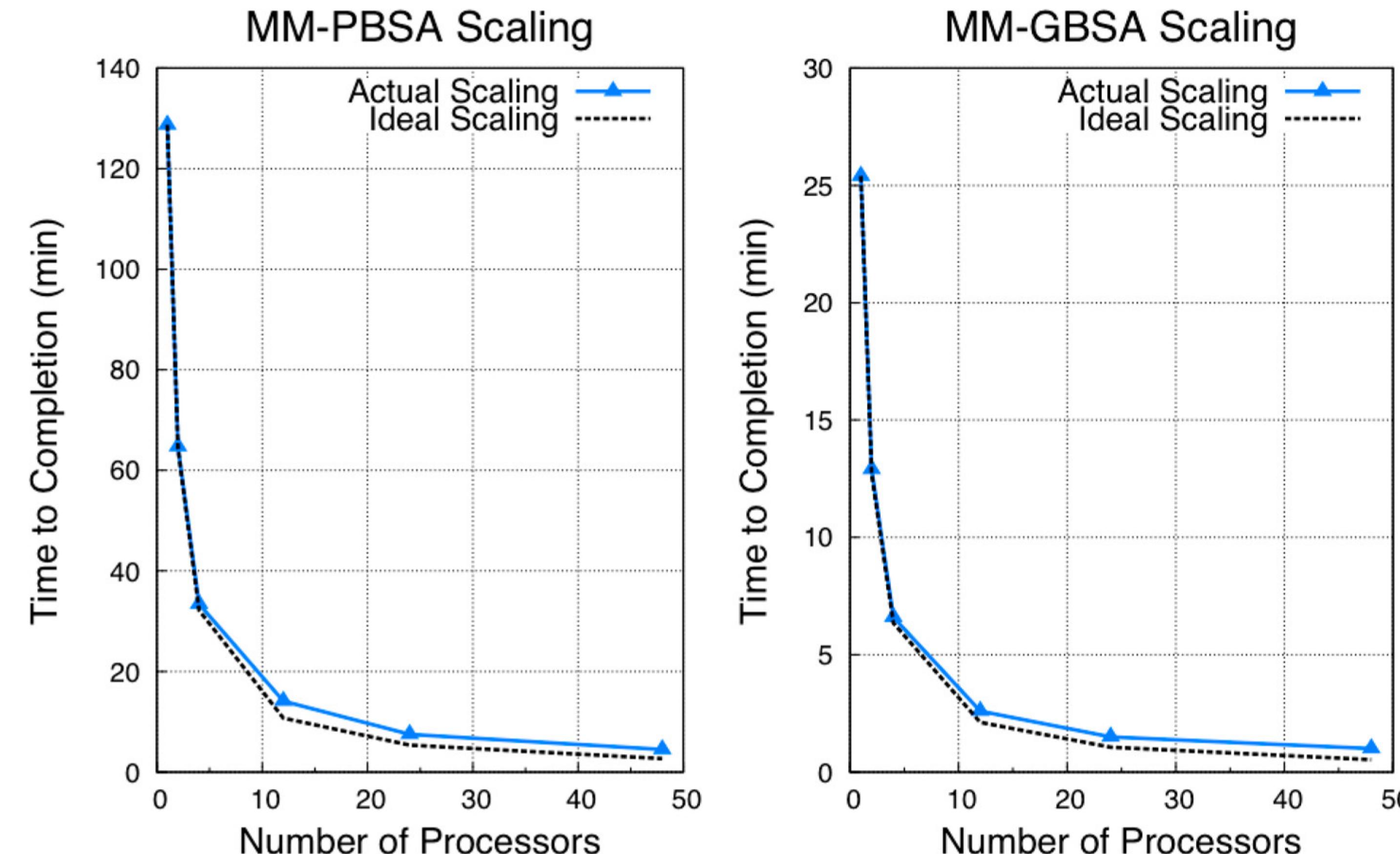
$$\Delta G_{solv}^0 = \Delta G_{polar,\epsilon=80} - \Delta G_{polar,\epsilon=1} + \Delta G_{nonpolar} \quad (2a)$$

$$\Delta G_{polar,\epsilon=1} = \Delta G_{vacuum} = \Delta E_{MM} - T\Delta S_{NMA} \quad (2b)$$

# Workflow



# GB verus PB



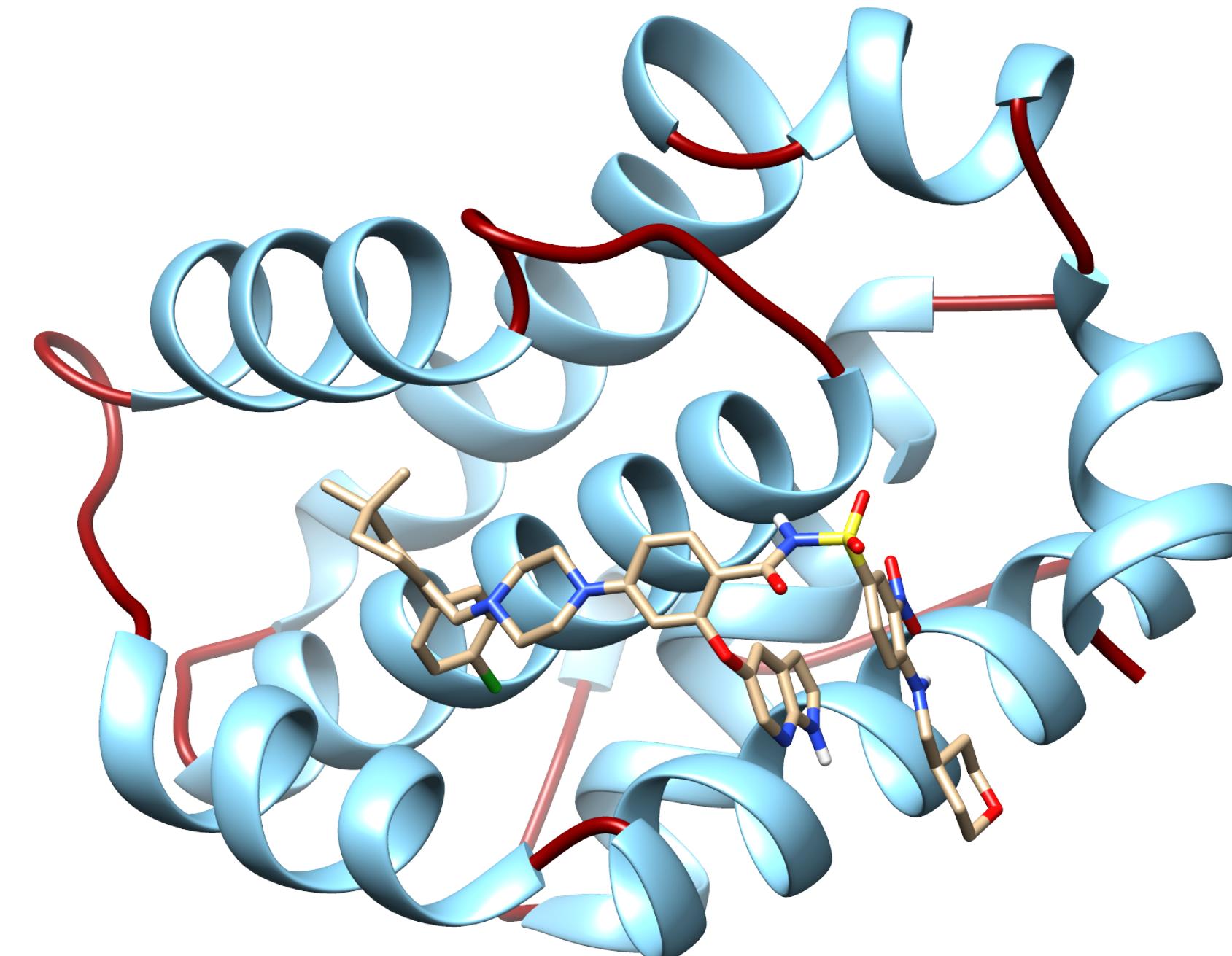
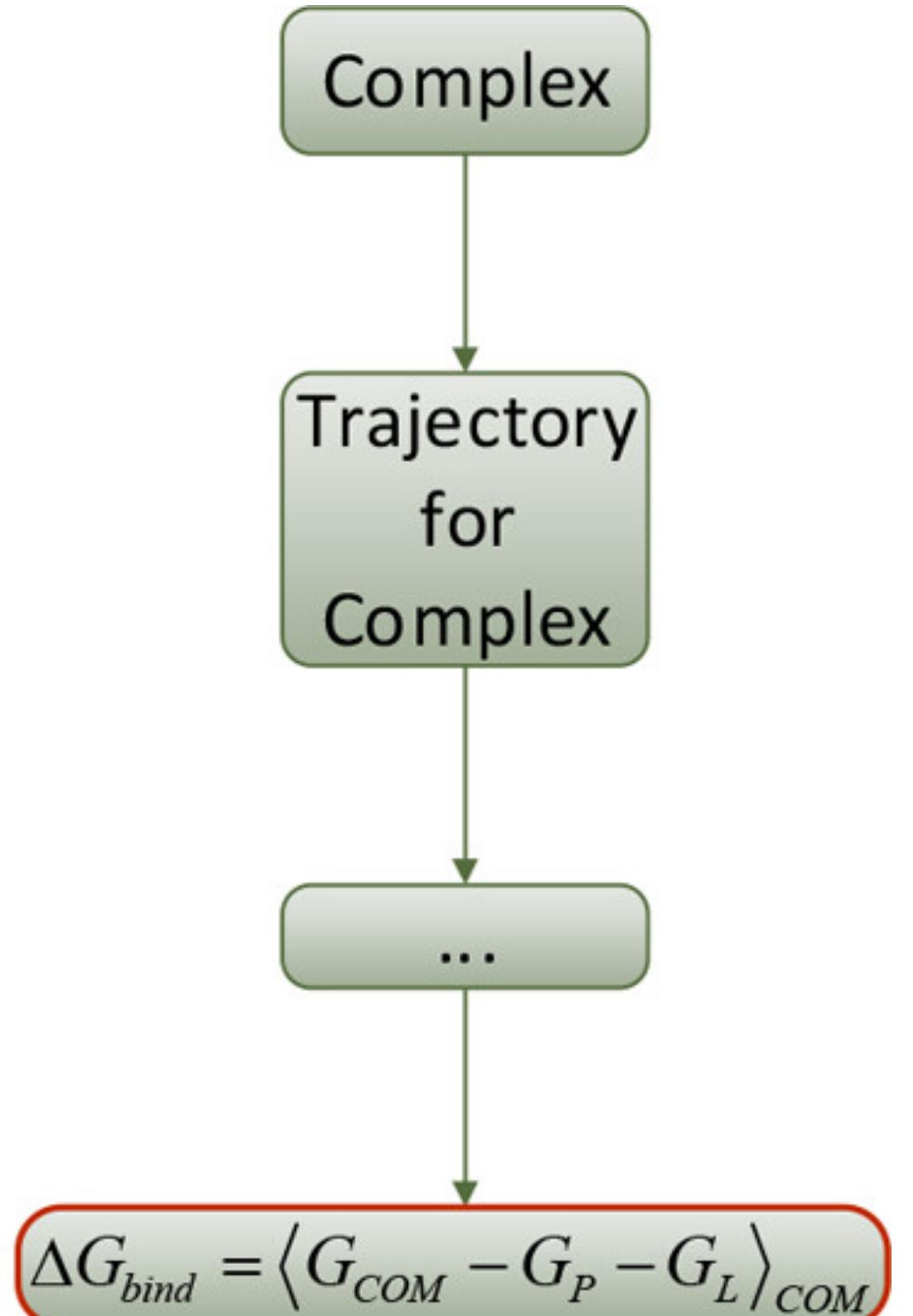
MMPBSA.py scaling comparison on 200 frames of a 5910-atom complex.  
All calculations were performed on NICS Keeneland (2 Intel Westmere 6-core CPUs per node, QDR infiniband interconnect).

# Caveats

---

- Method (CG or PB) accuracy depends on system
  - Error often  $> \sim 2$  kcal/mol
- Single trajectory assumes no significant conformational changes
  - Can lead to overstabilization of binding
- Multiple trajectories often have much higher standard error
- Entropy of water is ignored at binding site
  - Normal Mode Analysis can compute entropy
  - Significantly increases computational cost

# Bcl-2 & Venetoclax



1A-MM/PB(GB)SA