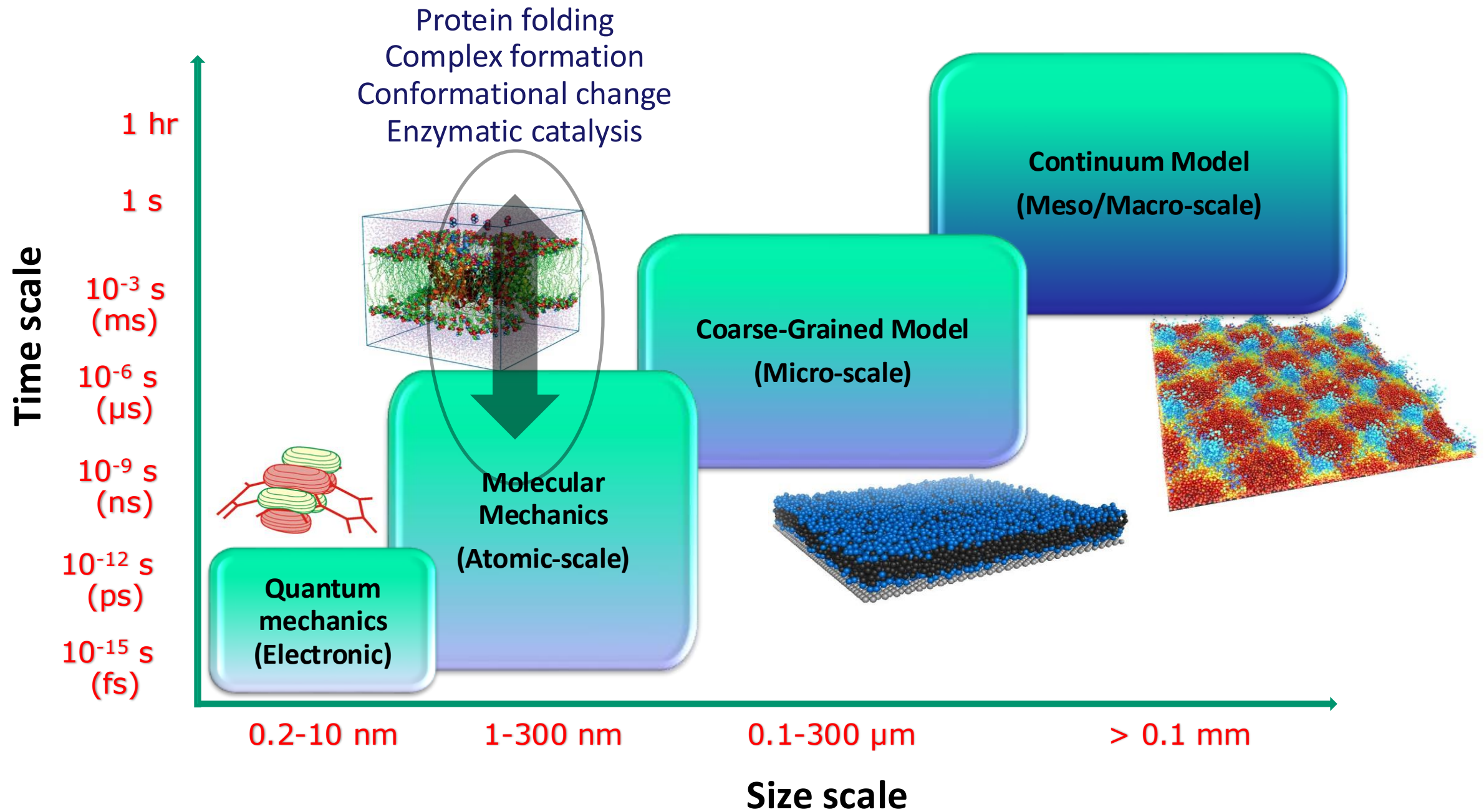


# **QM/MM Simulation (QM/MM Fundamentals)**

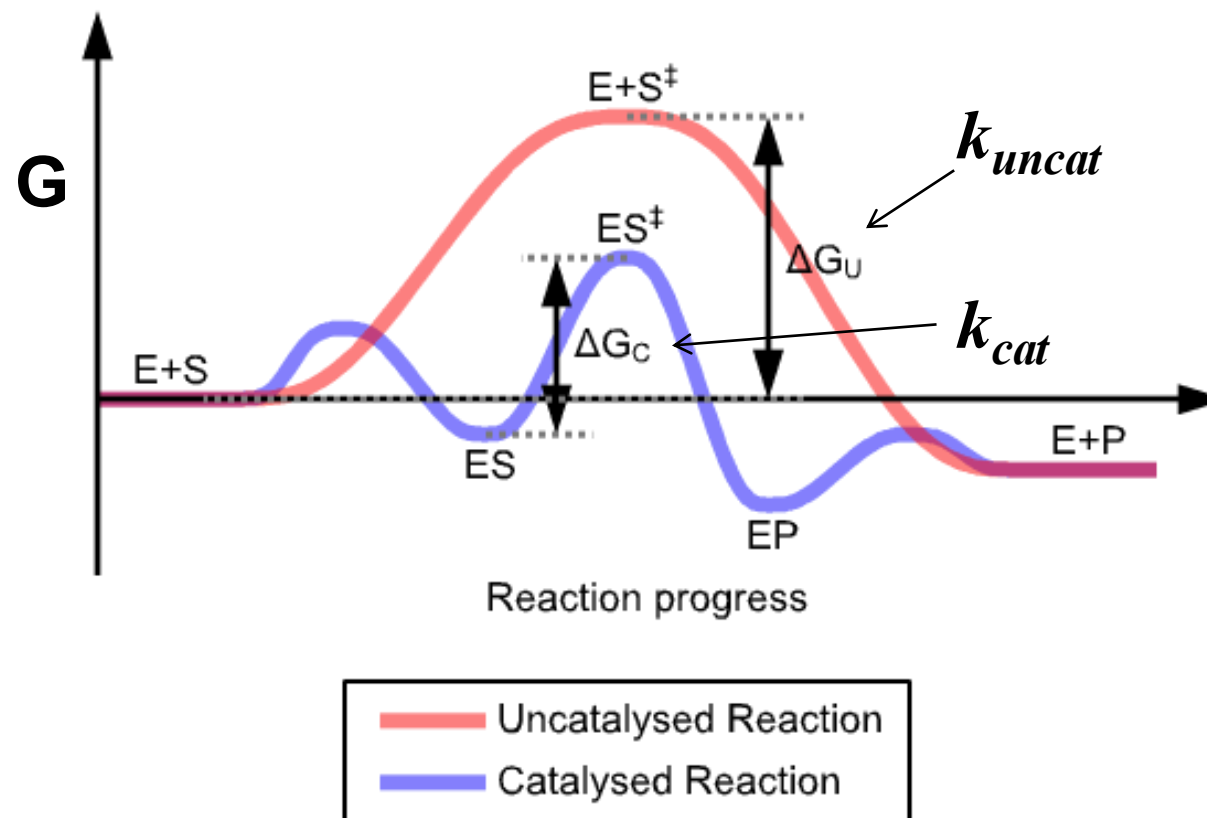
Kwangho Nam  
Dept. Chemistry and Biochemistry  
University of Texas at Arlington

kwangho.nam@uta.edu

# Multiscale challenges



Let's start from here...



Catalytic efficiency  
 $= k_{cat}/k_{uncat}$

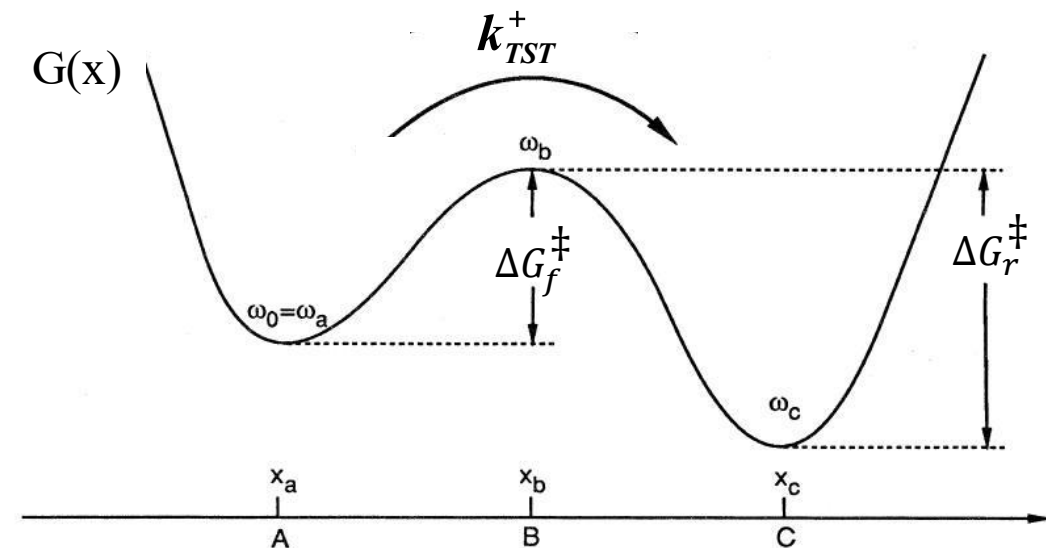
- OMP decarboxylase ( $\sim 10^{17}$ )
- $\beta$ -Amylase ( $\sim 10^{17}$ )
- Fumarase ( $\sim 10^{15}$ )
- Carbonic anhydrase ( $\sim 10^7$ )
- ...

# Transition state theory

Assumptions:

- Reactant states are in **local equilibrium** along a progress coordinate.
  - Trajectories that cross the transition state hypersurface **do not recross** it before becoming thermalized on the reactant or product side.
- => The forward reaction rate equals the **equilibrium flux in the product direction** through the transition state.

$$k_{TST}^+(T) = \frac{k_B T}{h} \frac{Q^\ddagger(T)}{Q^{RS}(T)} e^{\Delta U_f^\ddagger(T)/k_B T}$$
$$= \frac{k_B T}{h} e^{\Delta G_f^\ddagger(T)/k_B T}$$



$Q^\ddagger$  : transition state partition function  
 $\Delta G^\ddagger$  : free energy barrier

## Potential Energy Function

- \*Molecular Mechanical Method
  - CHARMM, AMBER, MM2, ...
- \*Coarse Grained Method
  - Go-potential, Martini, ...
- \*Quantum Mechanical Method
  - *Ab initio* quantum mechanical method  
⇒ HF, MP<sub>n</sub>, CCSD(T), and FCI
  - Density functional theory(DFT)
  - Semi-empirical method: AM1, PM3, PM6, DFTB, xTB, EVB, ...
- \*Hybrid Method
  - **QM/MM**, MM/CG, QM/MM/CG, ...
- \*Cost: Full QM >>> QM/MM >> MM > CG

## Searching Potential Energy Surface

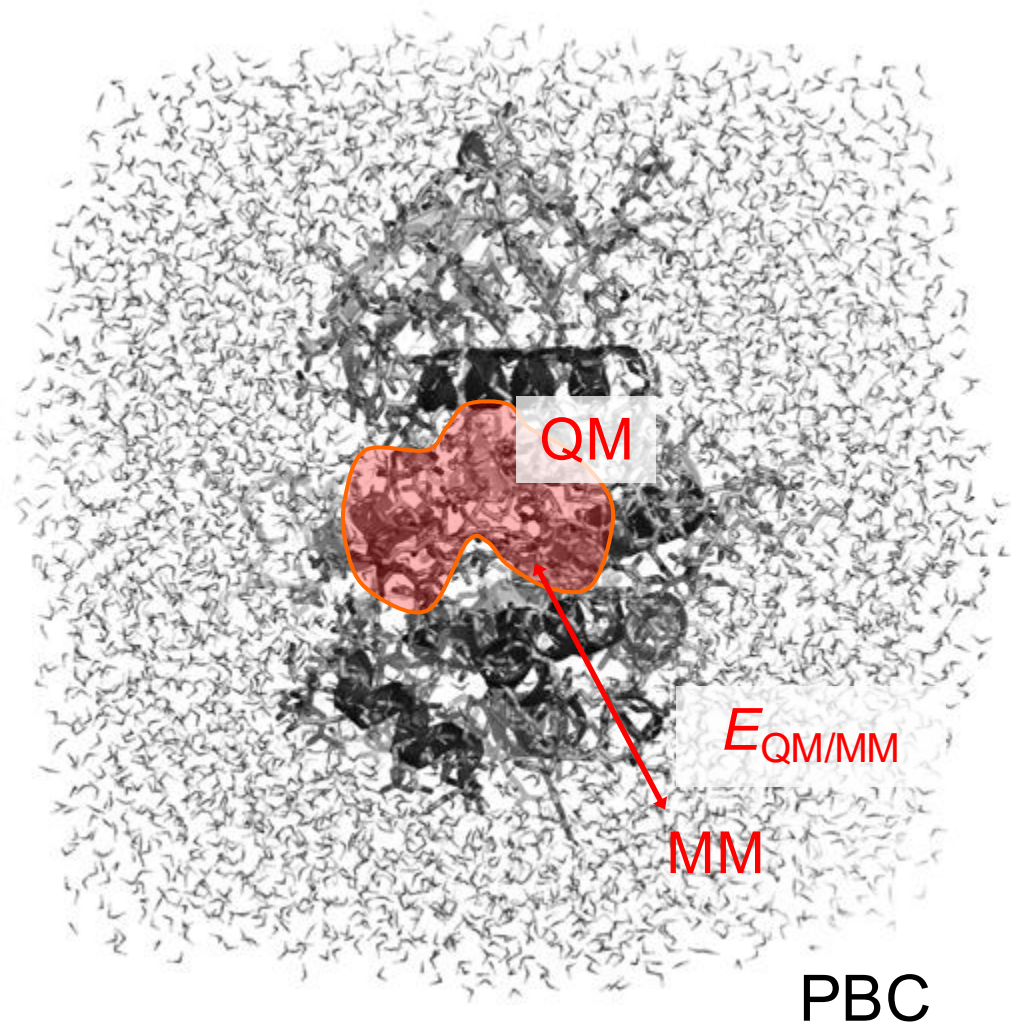
- \*Energy Minimizations
  - Geometry optimizations: SD, CG, NR, ...
  - Identification of reactant, transition and product states
  - Conformational search

## Statistical Simulations

- \*Molecular Dynamics
  - Integrators: Leapfrog Verlet, Velocity Verlet
  - Thermostats and barostats: NPT, NVT
  - Car-Parrinello method
- \*Monte Carlo Simulations
  - ⇒ Metropolis method, Gibbs ensemble MC, ...

# Hybrid QM/MM potential

- QM where it is needed; the rest by MM
- Low cost: small QM region for the treatment of reacting part
- Enable bond formation/breaking



$$H_{eff} = H_{QM} + H_{QM/MM} + H_{MM}$$

where

$H_{QM}$  : Hamiltonian for QM region

$$H_{QM/MM} = H_{QM/MM}^{elec} + H_{QM/MM}^{vdW} + H_{QM/MM}^{boundary}$$

$H_{MM}$  : Hamiltonian for MM region;  
i.e., MM force field energy



# Solving QM Hamiltonian

1. Schrödinger equation:

$$\hat{H}_{QM}^{tot}(r^N, R^M) \Psi_{QM}^{tot} = E_{QM}^{tot}(R^M) \Psi_{QM}^{tot}$$

where

$$\hat{H}_{QM}^{tot} = -\frac{1}{2} \sum_{A=1}^M \frac{1}{m_A} \nabla_A^2 - \frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^M \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}$$

$E_{QM}^{tot}(R^M)$ : Potential energy function

2. Born-Oppenheimer (BO) approximation to separate the motion of nuclei from electrons

$$\Psi_{QM}^{tot}(r^N, R^M) = \Psi_{QM}^{nu}(R^M) \Psi_{QM}^{elec}(r^N; R^M)$$

$$E_{QM}^{tot}(R^M) = \mathbf{E}_{QM}^{elec}(R^M) + E_{QM}^{nu}(R^M)$$

Typically,

$$E_{QM}^{nu}(R^M) = \sum_{A=1}^M \frac{P_A^2}{2m_A} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}, \text{ determined classically}$$

# Solving QM Hamiltonian

1. Schrödinger equation for the electronic degrees of freedom:

$$\hat{H}_{QM}^{elec}(r^N; R^M) \Psi_{QM}^{elec} = E_{QM}^{elec}(R^M) \Psi_{QM}^{elec}$$

where

$$\hat{H}_{QM}^{elec} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^M \frac{1}{r_{ij}}$$

2. Slater determinant of LCAO MO

$$\Psi_{QM}^{elec} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_N(2) \\ \dots & \dots & \dots & \dots \\ \phi_1(N) & \phi_2(N) & \dots & \phi_N(N) \end{vmatrix}, \quad \phi_i = \sum_{\mu}^N C_{i\mu} \chi_{\mu}$$

MO as a linear combination of atomic orbitals (LCAO-MO)

3.  $E_{QM}$

$$E_{QM}(R^M) = \langle \Psi_{QM}^{elec} | \hat{H}_{QM}^{elec} | \Psi_{QM}^{elec} \rangle + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}$$



# Solving QM Hamiltonian

E.g., Hartree-Fock (HF) method

- Variational principle:  $E_{QM}^{trial} \geq E_{QM}^{exact}$

-i.e., 
$$\frac{\partial E_{QM}}{\partial C_{\mu i}} = 0$$

-then...

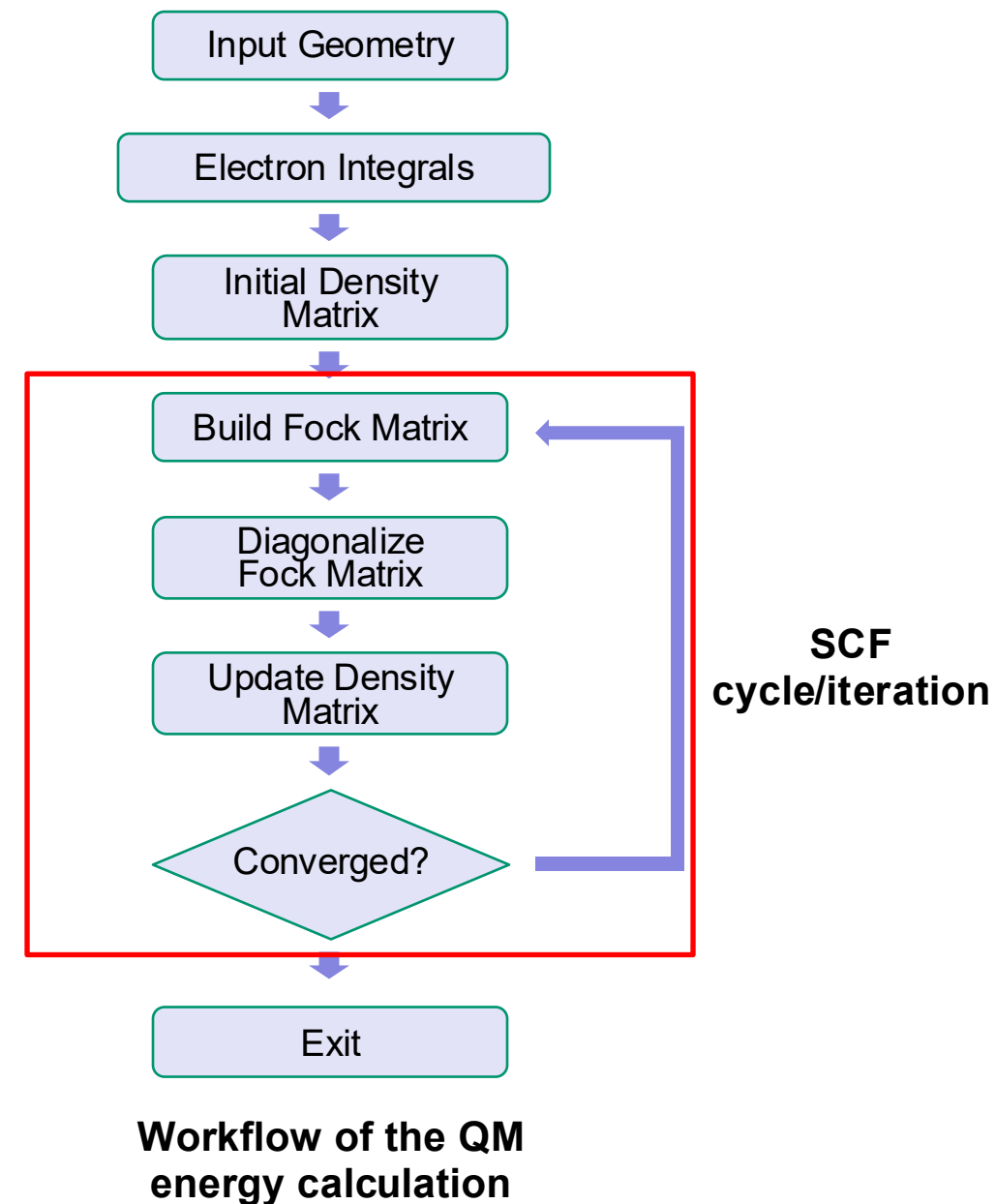
$$\mathbf{F}_{QM}^{elec} \mathbf{C} = \mathbf{S} \mathbf{C} \mathbf{E}_{QM}^{elec}$$

Diagram illustrating the equation  $\mathbf{F}_{QM}^{elec} \mathbf{C} = \mathbf{S} \mathbf{C} \mathbf{E}_{QM}^{elec}$  with labels:

- $\mathbf{F}_{QM}^{elec}$ : Fock matrix
- $\mathbf{C}$ : MO coefficients matrix
- $\mathbf{S}$ : Overlap matrix
- $\mathbf{E}_{QM}^{elec}$ : MO energy

Computational cost/scaling:

- MM with PME:  $N \log(N)$
- HF/DFT:  $O(N^3) \sim O(N^4)$
- MP2:  $O(N^5)$
- CCSD(T):  $O(N^7)$



# Depending on how QM/MM coupling is treated

## ❑ Mechanical embedding

- MM charges do not polarize QM density: i.e., QM calculation in vacuum.
- QM/MM interactions at the MM level: MM Coulomb and van der Waals interactions
- Produce spurious effects, e.g., hydrogen transfer.

## ❑ Electrostatic embedding

- MM charges polarize QM density.
- Most QM/MM implementation
- Need to parameterize van der Waals interactions, which are evaluated at the MM level.

## ❑ Polarization embedding

- MM polarization is included, i.e., MM and QM charges are mutually polarized.
- Micro-iteration (i.e, double SCF): MM dipoles & QM density

# Solving QM Hamiltonian

1. Schrödinger equation for the electronic degrees of freedom:

$$\hat{H}_{QM}^{elec}(r^N; R^M) \Psi_{QM}^{elec} = E_{QM}^{elec}(R^M) \Psi_{QM}^{elec}$$

where

$$\hat{H}_{QM}^{elec} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^M \frac{1}{r_{ij}}$$

2. Slater determinant of LCAO MO

$$\Psi_{QM}^{elec} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_N(2) \\ \dots & \dots & \dots & \dots \\ \phi_1(N) & \phi_2(N) & \dots & \phi_N(N) \end{vmatrix}, \quad \phi_i = \sum_{\mu}^N C_{i\mu} \chi_{\mu}$$

MO as a linear combination of atomic orbitals (LCAO-MO)

3.  $E_{QM}$

$$E_{QM}(R^M) = \langle \Psi_{QM}^{elec} | \hat{H}_{QM}^{elec} | \Psi_{QM}^{elec} \rangle + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}$$

# Solving QM + QM/MM Hamiltonians: electrostatic embedding

1. Schrödinger equation for the electronic degrees of freedom:

$$\hat{H}_{QM+QM/MM}^{elec}(r^N; R^M, R_{MM}^{N_{MM}}) \Psi_{QM}^{elec} = E_{QM+QM/MM}^{elec}(R^M, R_{MM}^{N_{MM}}) \Psi_{QM}^{elec}$$

where

$$\hat{H}_{QM+QM/MM}^{elec} = \hat{H}_{QM}^{elec} - \sum_{i=1}^N \sum_{M=1}^{N_{MM}} \frac{q_M}{r_{iM}}$$

2. Slater determinant of LCAO MO

$$\Psi_{QM}^{elec} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_N(2) \\ \dots & \dots & \dots & \dots \\ \phi_1(N) & \phi_2(N) & \dots & \phi_N(N) \end{vmatrix}, \quad \phi_i = \sum_{\mu}^N C_{i\mu} \chi_{\mu}$$

$$3. \quad E_{eff} = \langle \Psi_{QM}^{elec} | \hat{H}_{QM+QM/MM}^{elec} | \Psi_{QM}^{elec} \rangle + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} + \sum_{A=1}^M \sum_{M=1}^{N_{MM}} \frac{Z_A q_M}{R_{AM}} + E_{QM/MM}^{vdW} + E_{QM/MM}^{boundary} + E_{MM}$$

# Solving QM + QM/MM Hamiltonians: electrostatic embedding

E.g., Hartree-Fock (HF) method

- Variational principle:  $E_{QM}^{trial} \geq E_{QM}^{exact}$

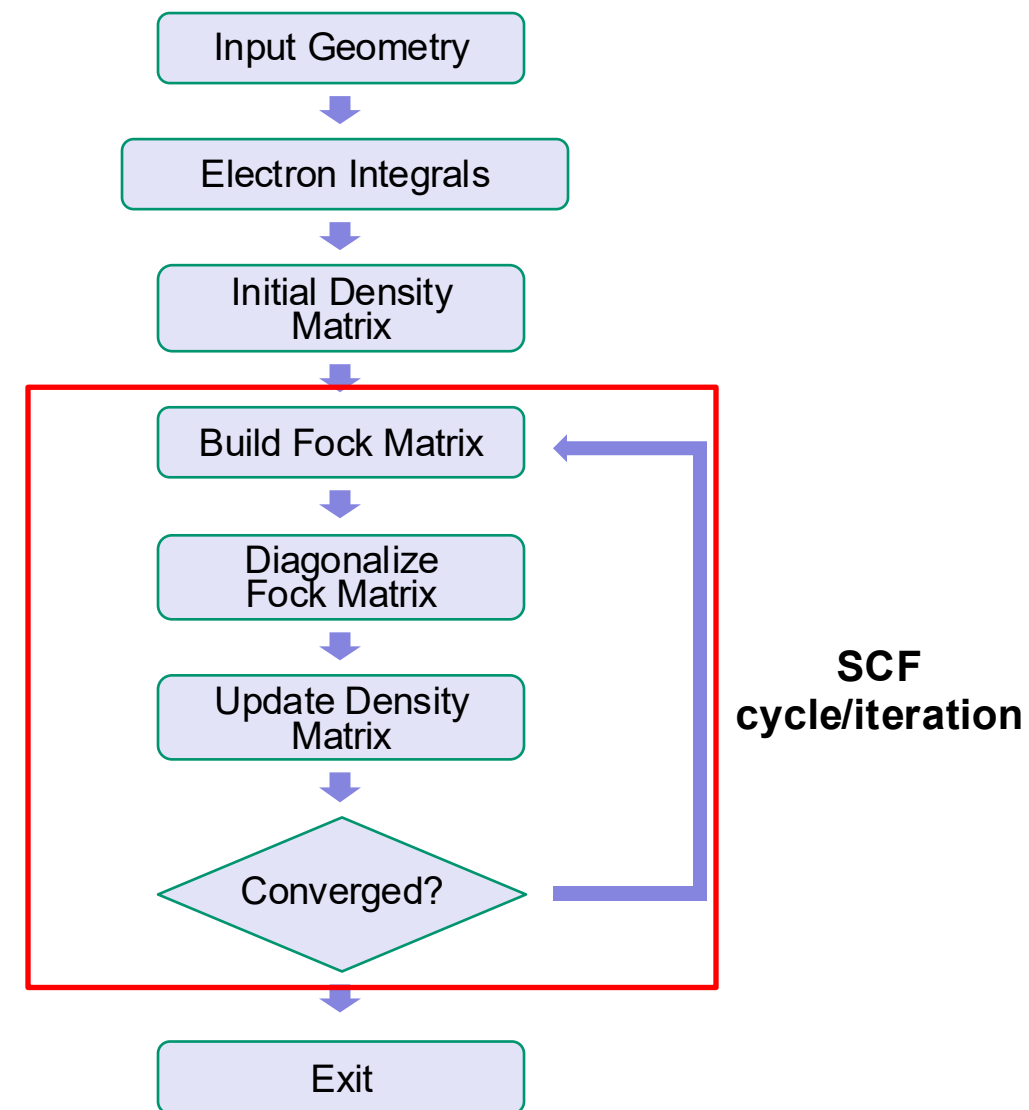
-i.e., 
$$\frac{\partial E_{QM}}{\partial C_{\mu i}} = 0$$

-then...

$$\mathbf{F}_{QM}^{elec} \mathbf{C} = \mathbf{S} \mathbf{C} \mathbf{E}_{QM}^{elec}$$

Diagram illustrating the equation  $\mathbf{F}_{QM}^{elec} \mathbf{C} = \mathbf{S} \mathbf{C} \mathbf{E}_{QM}^{elec}$  with labels for the components:

- $\mathbf{F}_{QM}^{elec}$ : Fock matrix
- $\mathbf{C}$ : MO coefficients matrix
- $\mathbf{S}$ : Overlap matrix
- $\mathbf{E}_{QM}^{elec}$ : MO energy



Workflow of the QM energy calculation

# Practical Issue: Details matter and know your system!

## ❑ QM region/method selection

- Know your chemistry; not a black-box method and use chemical intuition and knowledge of the reaction you are studying. Transition metals are particularly challenging.
- Depending on the required accuracy: DFT vs semi-empirical QM methods
- Expects some parameterizations of van der Waals parameters and SE-QM methods.
- Large QM region is better? QM theory/basis set limitation?
- Most X-ray structures do not have natural substrates: modeling and/or docking.

## ❑ QM-MM boundary treatment

- When cutting bond, cut along C-C single bond at least several bonds away from the region of chemical reaction.
- Polarization artifact around the cutting bond, especially, in the H-link approach. There exist many different approaches!

## ❑ Periodic boundary condition

- If possible, use PBC with PME scheme. Not all QM packages and QM theories support this.
- If not, use no-cutoff for the QM/MM and MM interactions. Do not use cut-off for MM and no-cutoff for QM/MM interactions.

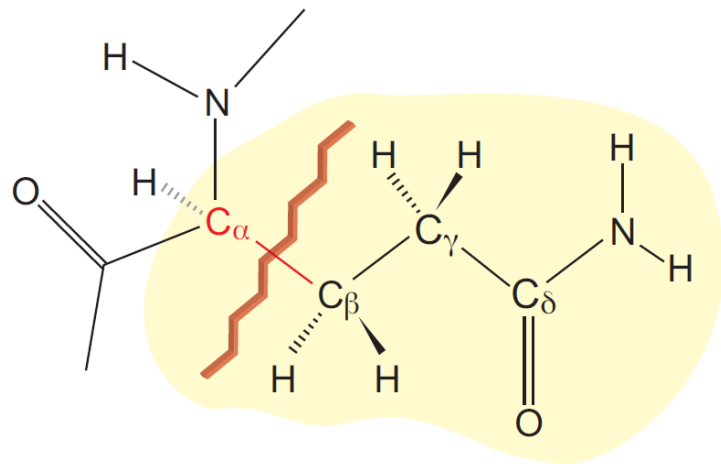
# Practical issue I: QM region/method selection

- ❑ Which atoms should be included in the QM region?
  - At least, substrates and key/catalytic protein residues
  - In some cases, metals and 1<sup>st</sup> solvation layer (i.e., coordinated groups)
  - < 200 atoms
- ❑ Which QM method (and basis set) should be used?
  - Ab initio MO theory: HF, MP2, etc
  - DFT: B3LYP, PBE, M06-2X, etc
  - Semi-empirical QM methods: AM1, PM3, DFTB, etc
  - For AI/DFT, use at least 6-31G(d) or larger basis sets. Also, check basis set convergence.
- ❑ Usually, the answers depending on the total cost of the calculations. Many QM/MM simulations have used semi-empirical QM methods; only recently employed AI-QM/MM methods.

|                     | Ab initio MO                                      |                       | DFT                     | SE QM methods  |
|---------------------|---|-----------------------|-------------------------|--|
|                     | HF  | MP2                   |                         |  |
| Accuracy            | Bad   | Good                  | Reasonable              | System dependent<br>Can be improved by reparameterization. |
| Speed (scaling)     | Slow<br>$O(N^4)$                                  | Very slow<br>$O(N^5)$ | Slow<br>$O(N^{3\sim4})$ | Very fast ( $\sim 10^3$ order faster)<br>$O(N^3)$          |
| Parallel efficiency | Reasonable ~ Good<br>(e-e integrals calculations) |                       |                         | Bad<br>(Implementation dependent)                          |



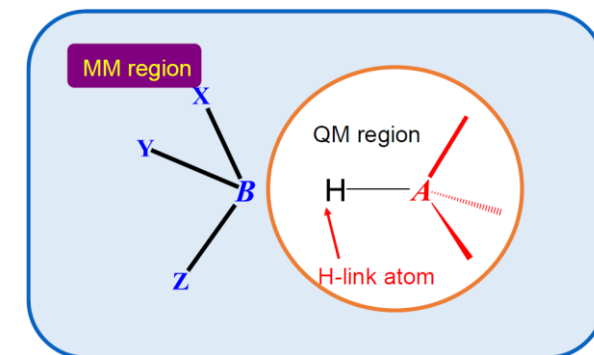
# Practical issue II: QM/MM boundary treatment



Along C-C single bond

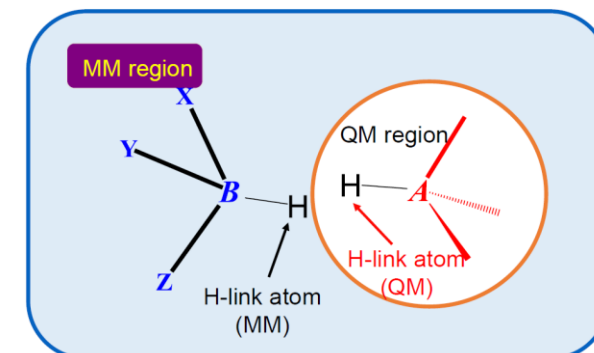
- **H-link atom approach**

- Cap the dangling bond of the QM region
- Simple to implement



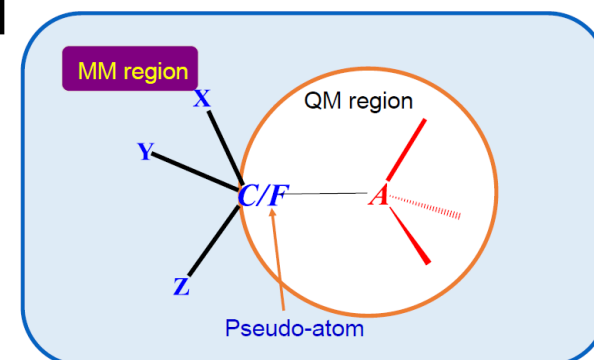
- **Double-link atom approach**

- Brooks and coworkers
- Also, use Gaussian blur for MM atoms.



- **Pseudo-bond method**

- Yang and co-workers: DFT method
- Thiel and co-workers: connection bond for SE-QM method

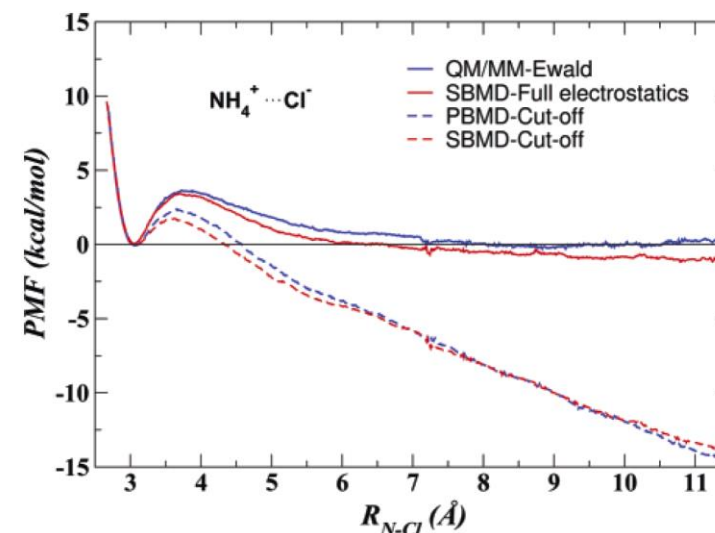
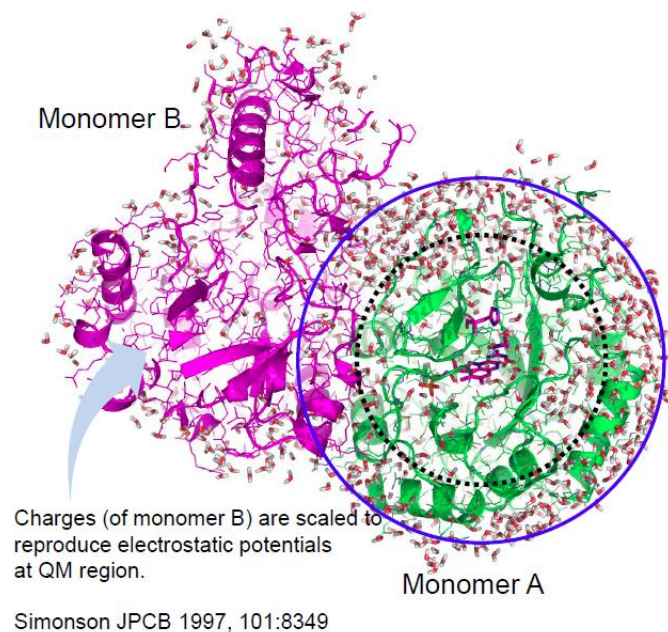




# Practical issue III: Periodic vs. non-periodic boundary

## Long-range QM/MM interactions for periodic boundary MD simulations?

Most QM/MM calculations are carried out with “cutoff” or “no-cutoff” under stochastic boundary conditions.



Nam, JCTC (2005)

$$E_{tot}^{PME} = E_{QM}^{RS}[\rho] + E_{QM/MM}^{RS}[\rho] + \Delta E_{QM}^{PME}[Q] + E_{MM}^{PME}[q]$$

- Allows to use PBC with PME scheme
  - Balanced QM, QM/MM and MM (long-range electrostatic) interactions
  - Avoid cut-off artifacts and produce stable MD trajectory
  - Available in many QM and QM/MM packages, CHARMM, AMBER, Q-Chem, etc
- \*Alternative models: Ambient-potential composite Ewald (York et al.), Multipole moments (Rothlisberger et al.), ESP/ChEIPG charges (Herbert et al.), Gen-Ew (Thiel et al.), Augmented charges (Shao et al.)

# Options, QM methods/packages, etc

|        | NDDO  |        | DFTB | AI packages   | Ewald/PME   | QM-MM Boundary               | FE methods                |
|--------|-------|--------|------|---|---|------------------------------|---------------------------|
|        | MOPAC | MNDO97 |      |   |   |                              |                           |
| CHARMM | ✓     | ✓      | ✓    | Gaussian<br>Q-Chem<br>GAMESS-US<br>Turbomole<br>...     | Support:<br>NDDO/DFTB<br>and some AI<br>QM packages | H-link<br>GHO<br>(NDDO/DFTB) | US<br>String<br>Method    |
| AMBER  | ✓     | ✓      | ✓    | Gaussian<br>Q-Chem<br>GAMESS-US<br>ORCA<br>QUICK<br>... | Support:<br>NDDO/DFTB<br>and some AI<br>QM packages | H-link                       | US<br>PLUMED<br>interface |