# Hybrid Quantum Mechanics/ Molecular Mechanics (QM/MM) Approach
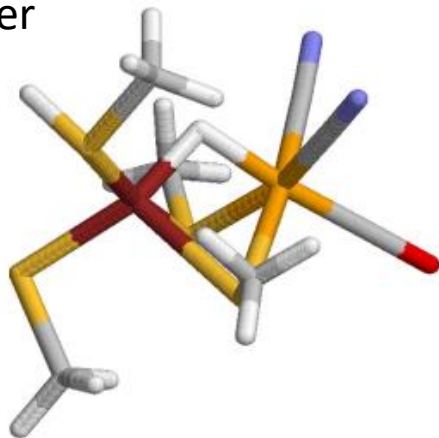
Sivanujan S (Sivanu)

# Introduction

- **Quantum-mechanical (QM)** methods are required for describing chemical reactions and other electronic processes, such as charge transfer or electronic excitation.
  - Restriction: QM methods are restricted to systems of up to a few hundred atoms
- Highly efficient, force-field-based **molecular mechanics (MM)** methods are capable of treating up to several 100 000 atoms and allowing for simulations over time scales of tens of nanoseconds.
- A logical model is built to use;
  - a QM method for the **chemically active region** (e.g., substrates and co-factors in an enzymatic reaction) and
  - an MM treatment for the **surroundings** (e.g., protein and solvent).
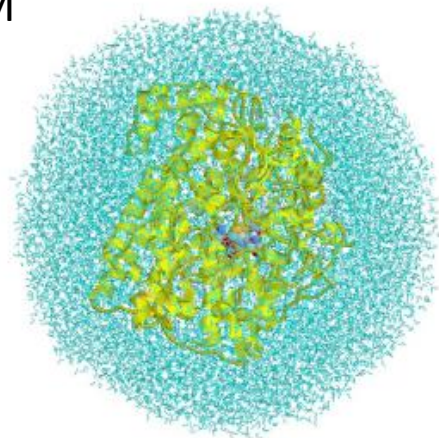
# Introduction (Contd.)

- The prime **challenges** with QM/MM calculations are;
  - bonds between the QM and MM systems,
  - the selection of the QM system, and
  - the local-minima problem.
- There are two approaches to treat proteins with QM methods.
  1. **QM-cluster approach**: In the QM-cluster approach, a small number of residues (typically 30–200 atoms) are cut out from the protein and are studied by QM in a continuum solvent.
     - continuum solvent - (sometimes known as Implicit solvation ) is a method to represent solvent as a continuous medium instead of individual "explicit" solvent molecules.
  2. The active site is treated by a similar QM model, but the surrounding protein and solvent are modeled by molecular mechanics (MM), giving the **QM/MM** approach.
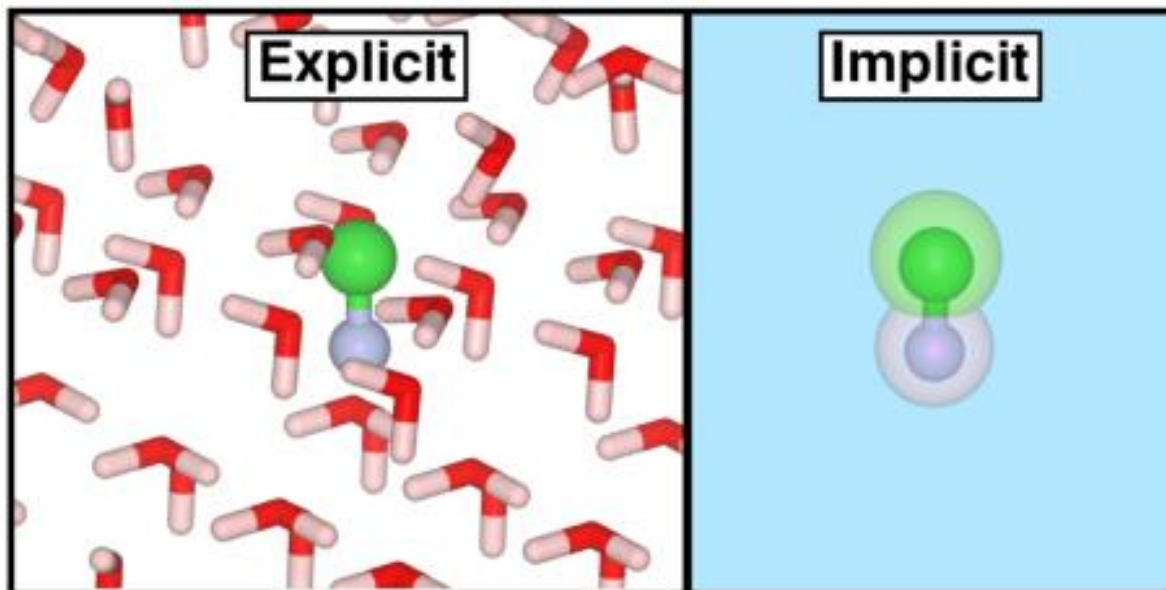
# Introduction (Contd.)

QM-Cluster

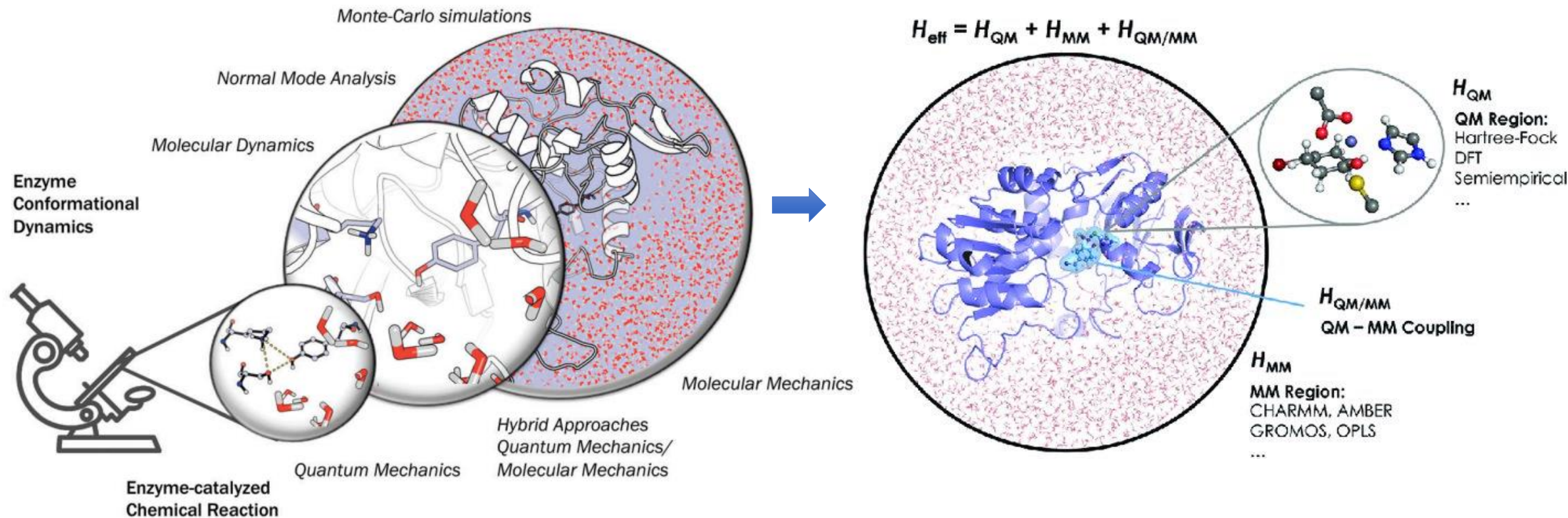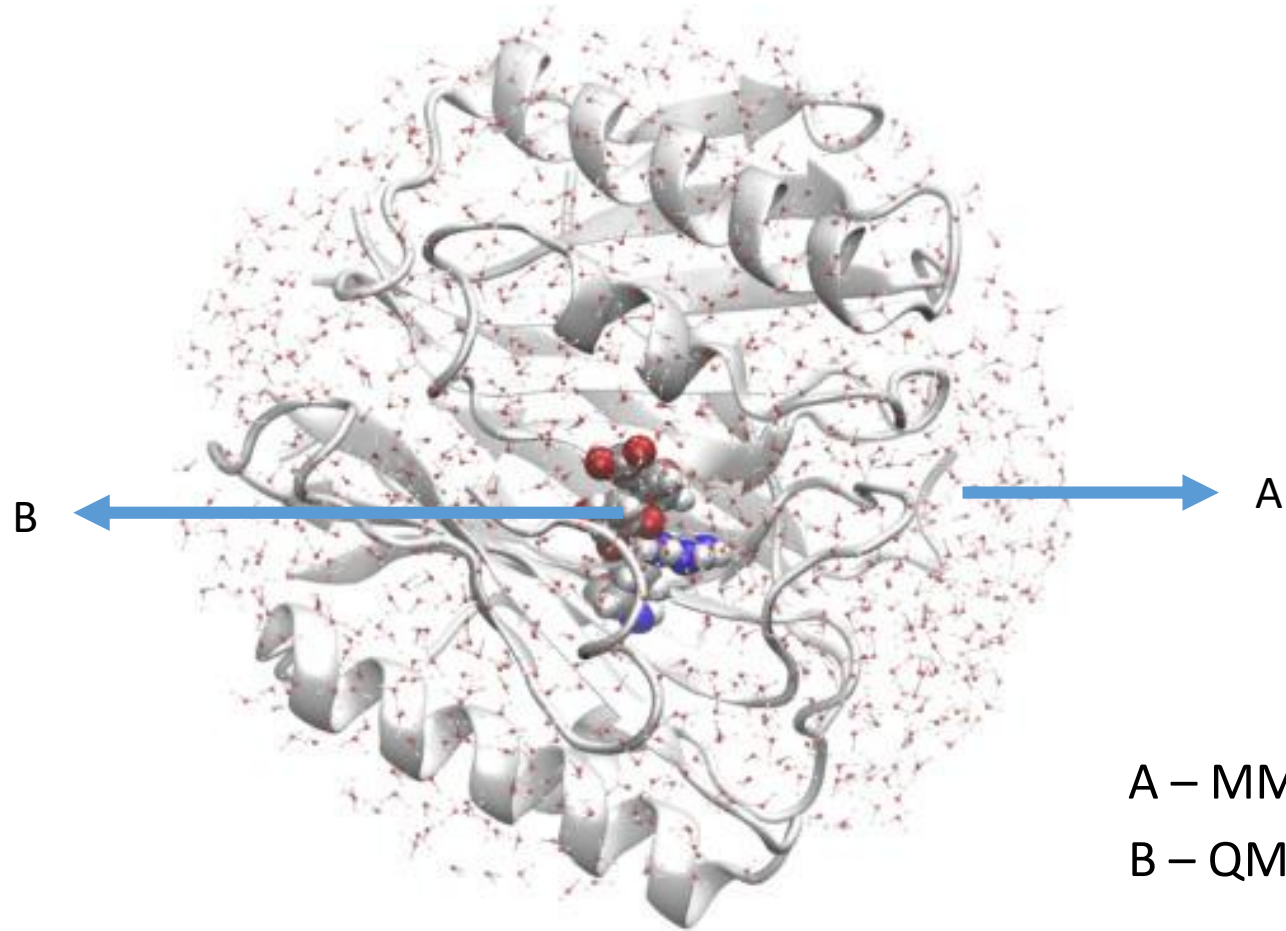Solvation Models

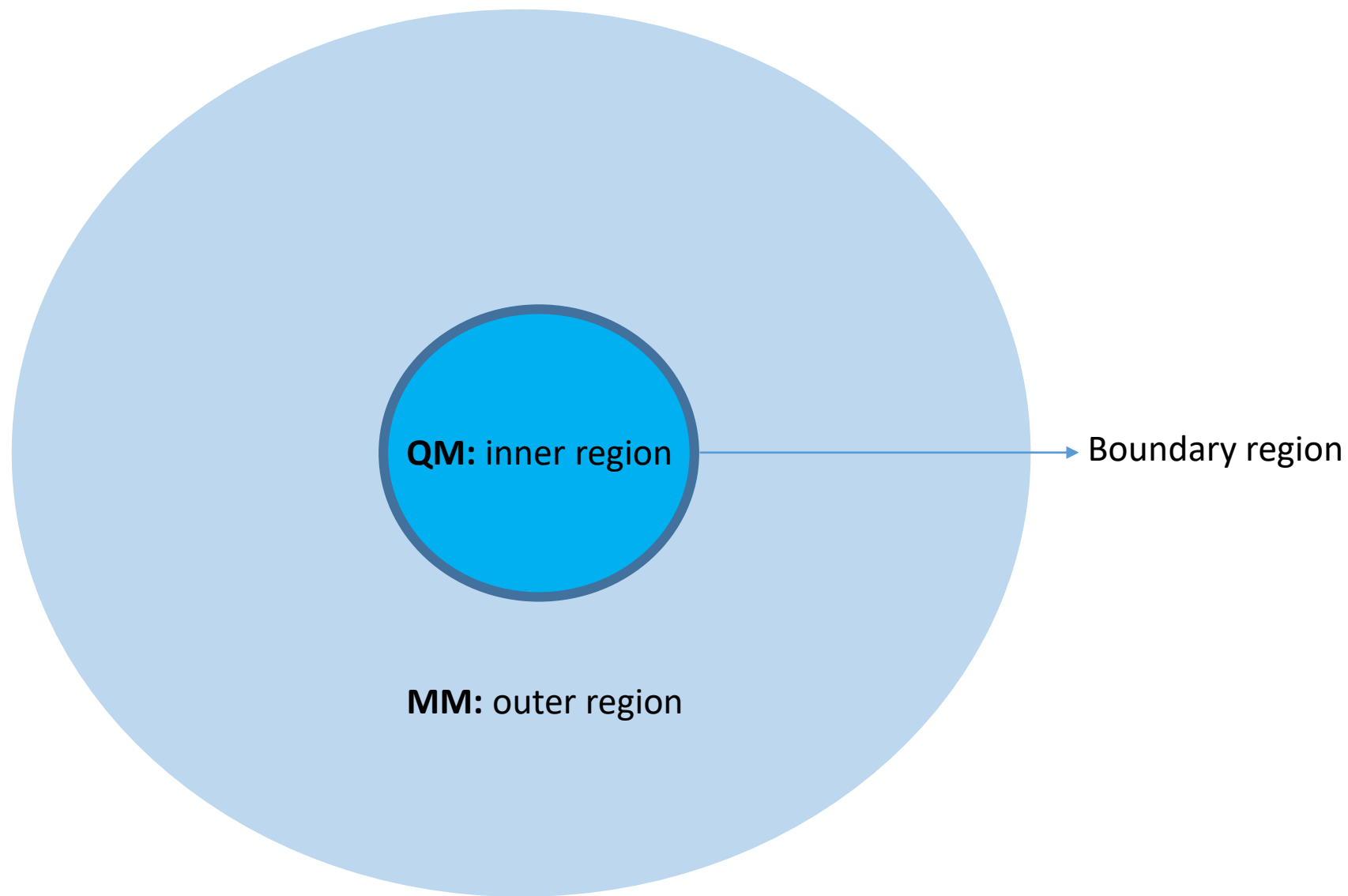

QM/MM

Ryde, U., 2016. QM/MM Calculations on Proteins. *Methods in Enzymology*, pp.119-158.

Iyemperumal, S. and Deskins, N., 2017. Evaluating Solvent Effects at the Aqueous/Pt(111) Interface. *ChemPhysChem*, 18(16), pp.2171-2190.

# QM/MM approach; a big picture



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

Monte-Carlo simulations

Normal Mode Analysis

Molecular Dynamics

Enzyme Conformational Dynamics

Molecular Mechanics

Hybrid Approaches Quantum Mechanics/ Molecular Mechanics

Quantum Mechanics

Enzyme-catalyzed Chemical Reaction

$H_{QM}$ QM Region: Hartree-Fock DFT Semiempirical ...

$H_{QM/MM}$ QM – MM Coupling

$H_{MM}$ MM Region: CHARMM, AMBER GROMOS, OPLS ...

Romero-Rivera, A., Garcia-Borràs, M. and Osuna, S., 2017. Computational tools for the evaluation of laboratory-engineered biocatalysts. *Chemical Communications*, 53(2), pp.284-297.

A – MM region

B – QM region

# QM–MM partitioning

**QM:** inner region

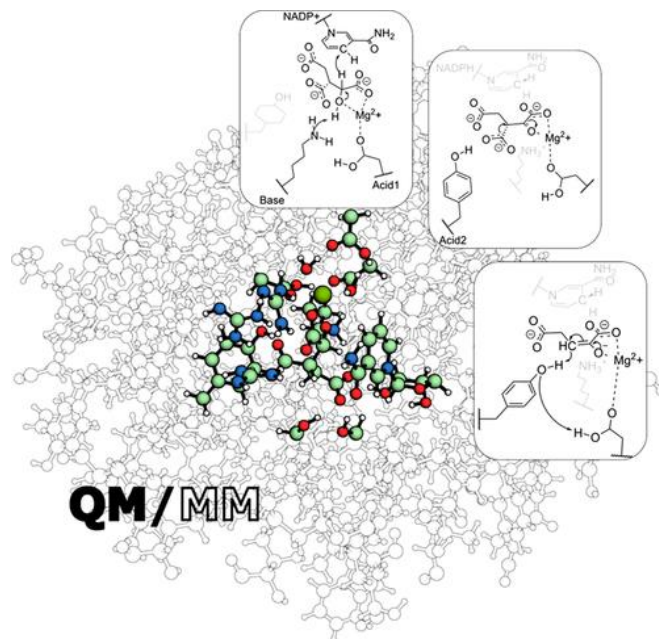Boundary region

**MM:** outer region

# QM/MM Boundary region

The term boundary region is used rather loosely to designate the region where the standard QM and MM procedures are modified or augmented in any way.

# Choice of the QM/MM boundary

- Ideally far away from reactive centre.
- Bond formation and bond-breaking processes?

Sousa, S., Ribeiro, A., Neves, R., Brás, N., Cerqueira, N., Fernandes, P. and Ramos, M., 2016. Application of quantum mechanics/molecular mechanics methods in the study of enzymatic reaction mechanisms. *WIREs Computational Molecular Science*, 7(2).

# Choice of the QM level

- In QM calculations, the Schrodinger equation is solved for a molecular system. The input is the <u>coordinates</u>, the <u>net charge</u>, and the total <u>spin</u>. The result is the total energy and the wave function, from which all measureable properties of the system can be calculated.

- Self-consistent-field (SCF) treatment: presence of the external point-charge field that represents the MM charge model in the case of electronic or polarized embedding.

- The most basic QM approach is the Hartree–Fock (HF) method. Unfortunately, it gives quite approximate results and it is therefore mainly used as a starting point of more accurate methods, based on perturbation theory or series expansions, eg, Møller–Plesset second-order perturbation theory (MP2).

# Choice of the QM level (Contd.)

- Currently, the great majority of QM calculations are performed with density functional theory (DFT).

- It is not based on the wave function, which is a function of the coordinates of all involved particles, but instead on the electron density, which is a function of only the <u>three coordinates of Cartesian space</u>.

- All mentioned QM methods expand the wave function in a set of known functions, the **basis set**.

- Generically, a basis set is a collection of vectors which spans (defines) a space in which a problem is solved.

- In QM, the "basis set" usually refers to the set of (non-orthogonal) one-particle functions used to **build molecular orbitals.**

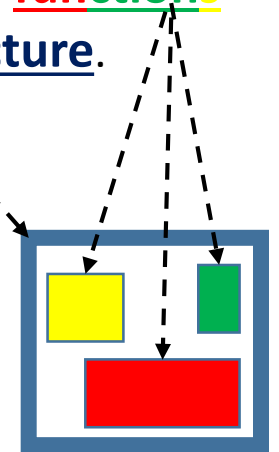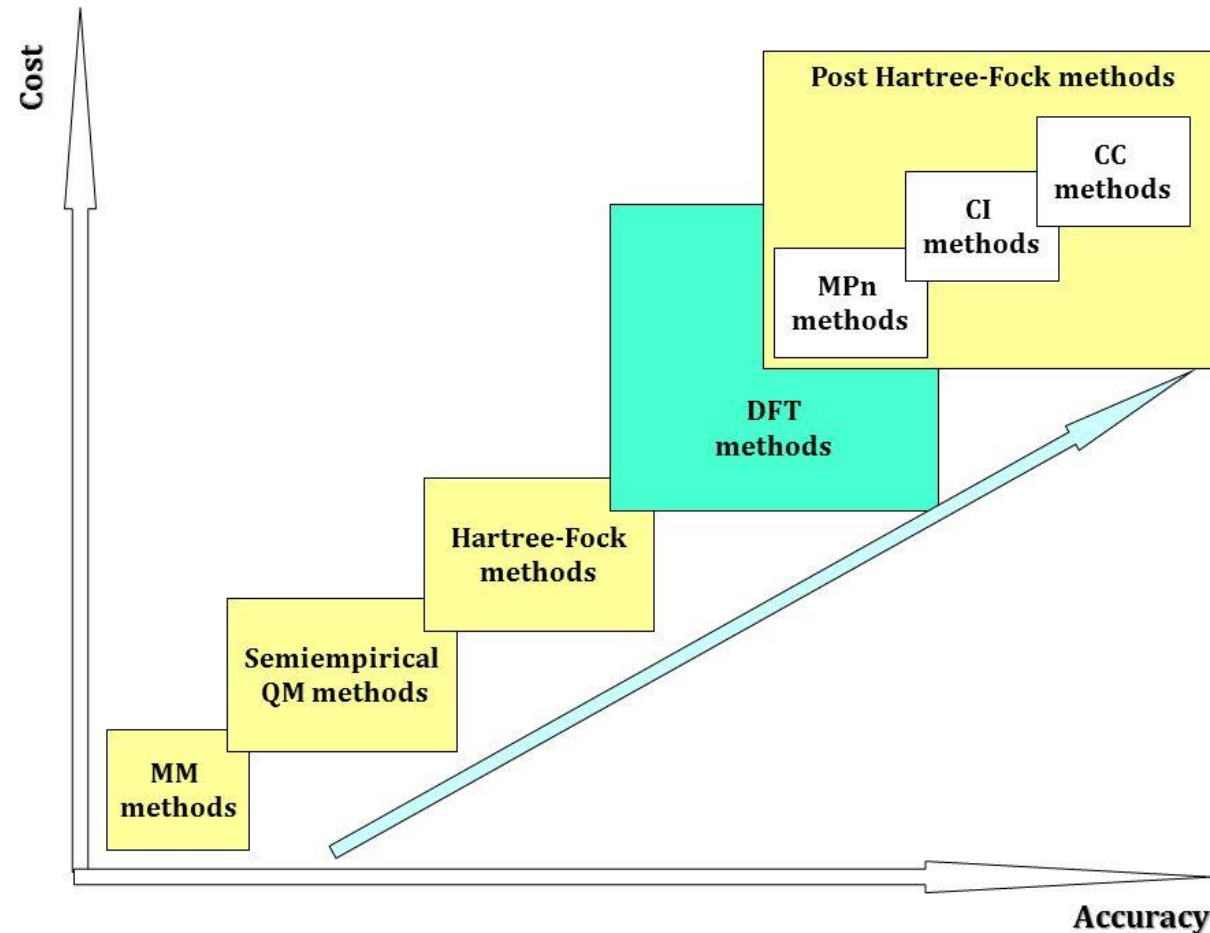# What is a Basis Set?
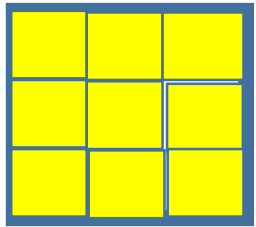
Basis set: 
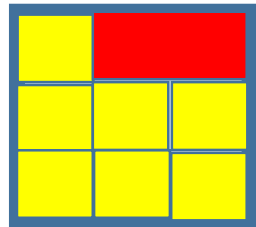
Basis set: 

Basis set: 

Basis sets determine the number of **functions** used to describe the **electronic structure**.

# QM methods; accuracy vs cost

The quest for a universal density functional: The accuracy of density functionals across a broad spectrum of databases in chemistry/ slideshare
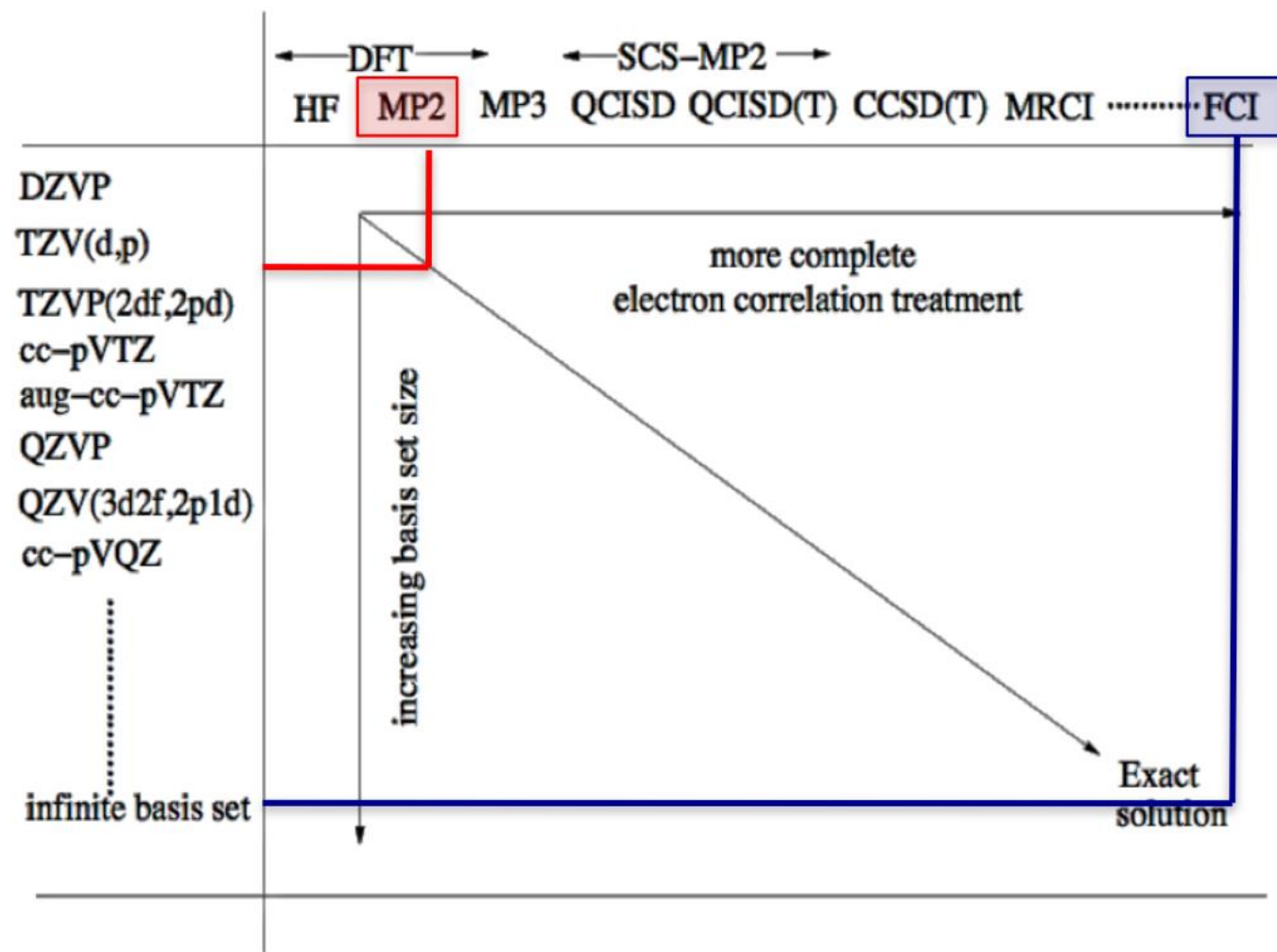
- How do you define basis sets?

A basis set is a mathematical description of orbitals of a system, which is used for approximate theoretical calculation or modeling.

# QM methods; basis sets and levels

# QM methods; solvent phases

- QM calculations are normally performed on isolated molecules in gas phase. However, most experiments are performed in condensed phases (aqueous solution).

- Methods have been developed to model a surrounding homogeneous solvent in QM calculations, continuum solvation methods.
    - Polarizable continuum model (PCM)
    - Conductor-like solvent model (COSMO)

- For protein-sized systems, methods based on the Poisson–Boltzmann (PB) equation or the generalized Born (GB) approach are more common.

# Potential Energy of QM region

**Born-Oppenheimer Approximation** : This is a well established mathematical approximation used in Quantum Mechanics (QM). This describes the electrons neglecting the motion of the nucleus. The underlying physical basis is that the mass of nucleus is much larger than that of an electron. **The density of electron $\rho_{elec}(r)$ in a molecule is obtained in the field of fixed nuclei.**

$$\hat{H}_{elec} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{|r_i - R_A|} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{|r_i - r_j|}$$
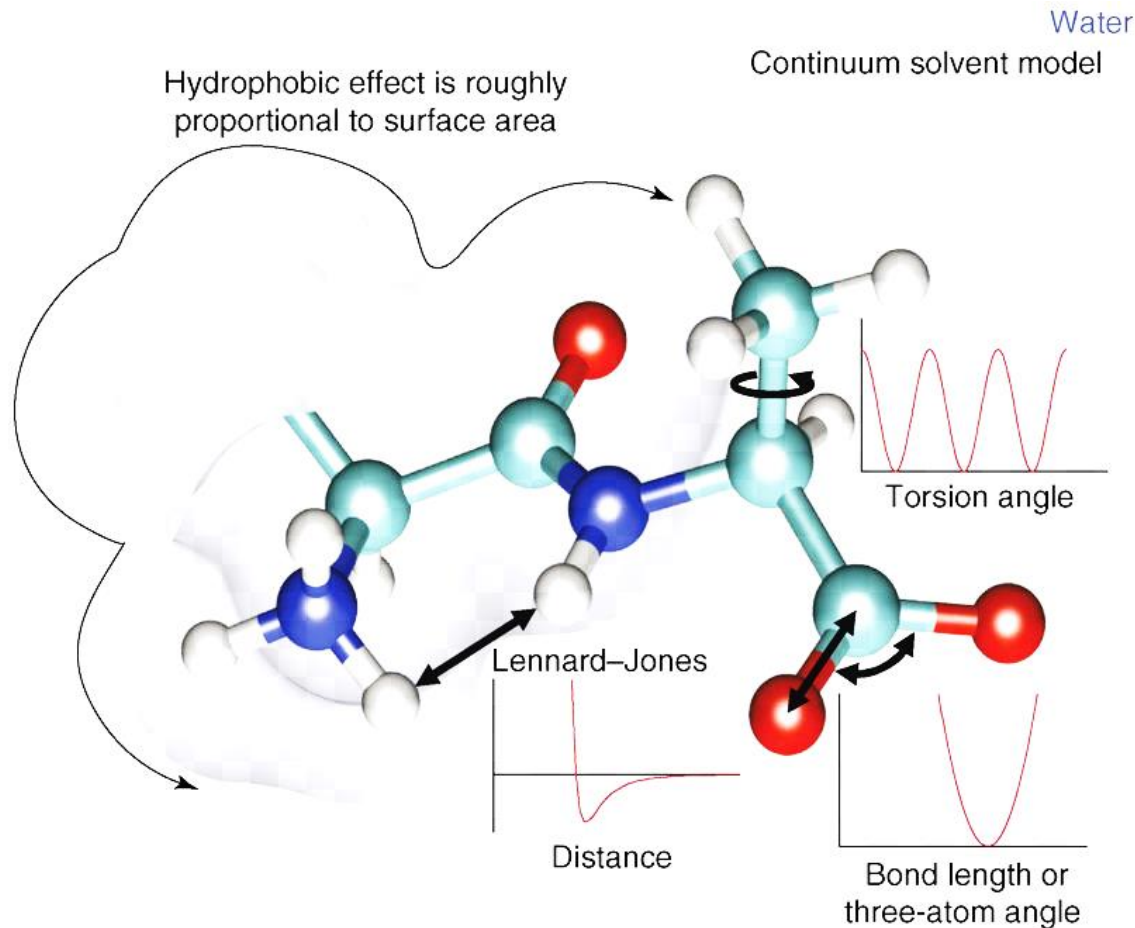
$$E_{elec} = \langle \Psi_{elec} | \hat{H}_{elec} | \Psi_{elec} \rangle$$

$$E_{QM} = E_{elec} + \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{1}{|R_A - R_B|}$$

# MM Methods

No attempt is made to solve the Schrodinger equation and electrons are ignored. Instead, molecules are considered as a collection of balls, connected by springs.

The energy of the system as a function of the coordinates is described by an empirical function, a **force field** (FF).

# Potential Energy of MM region



Hydrophobic effect is roughly proportional to surface area

Water
Continuum solvent model

Torsion angle

Lennard–Jones

Distance

Bond length or three-atom angle

Intra-molecular potential energy:

$$E_{\text{intra}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}}$$

$$E_{\text{bond}} = \frac{k}{2}(r - r_0)^2$$

Inter-molecular potential energy:

$$E_{\text{inter}} = E_{\text{LJ}} + E_{\text{Coul}}[+E_{\text{pol}}]$$

$$E_{\text{LJ}} = \sum_{A>B} 4\epsilon_{AB}\left[\left(\frac{\sigma_{AB}}{R_{AB}}\right)^{12} - \left(\frac{\sigma_{AB}}{R_{AB}}\right)^{6}\right],$$

$$E_{\text{Coul}} = \sum_{A>B} \frac{q_A\, q_B}{R_{AB}}$$

# MM (Contd.)

- Energy can be calculated for a whole protein in seconds, allowing for extensive sampling of the accessible phase space by molecular dynamics (MD) or Monte Carlo methods.

- The protein is typically solvated with several thousands of explicit water molecules and possibly counter ions to provide a more realistic account of the surroundings.

- To mimic infinite systems, periodic boundary conditions are employed.

- Choice of MM method: biomolecular force fields available.

  - Popular examples include AMBER, CHARMM, GROMOS and OPLS-AA.
  - Biomolecules - typically includes proteins and in most cases also nucleic acids, but less frequently carbohydrates or lipids.

Lopes PE, Guvench O, MacKerell AD Jr. Current status of protein force fields for molecular dynamics simulations. *Methods Mol Biol*. 2015;1215:47-71. doi:10.1007/978-1-4939-1465-4_3

# QM/MM methods for biomolecular systems

- Problem: There are covalent linkages between QM atoms and MM atoms.

- Consider that the QM region is connected to the MM region via one covalent bond: CB-CA.

- The 'CB' atom belongs to the QM region and the 'CA' atom to the MM region.

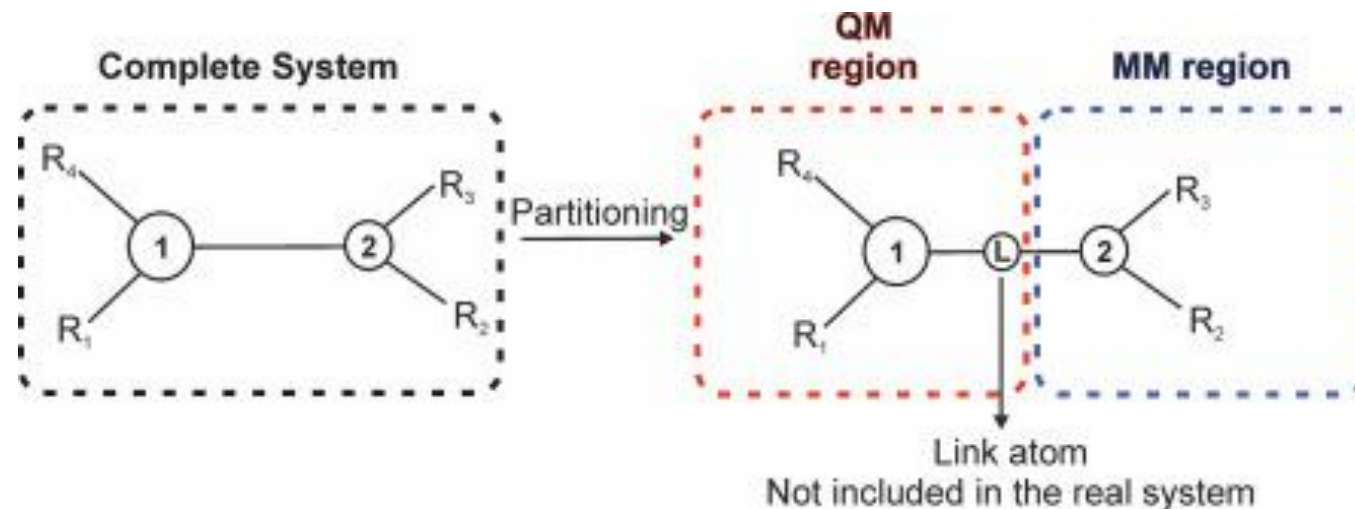- The link-atom (L) method: the free valency is capped by an additional atom (H):

  CB-CA –> CB-H.
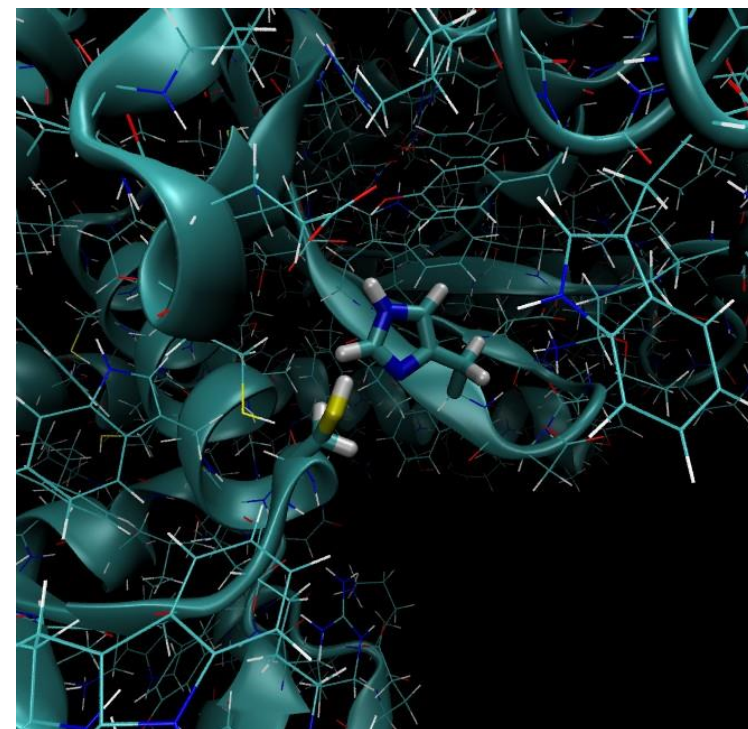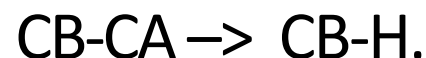
**Before adding link-atoms.**
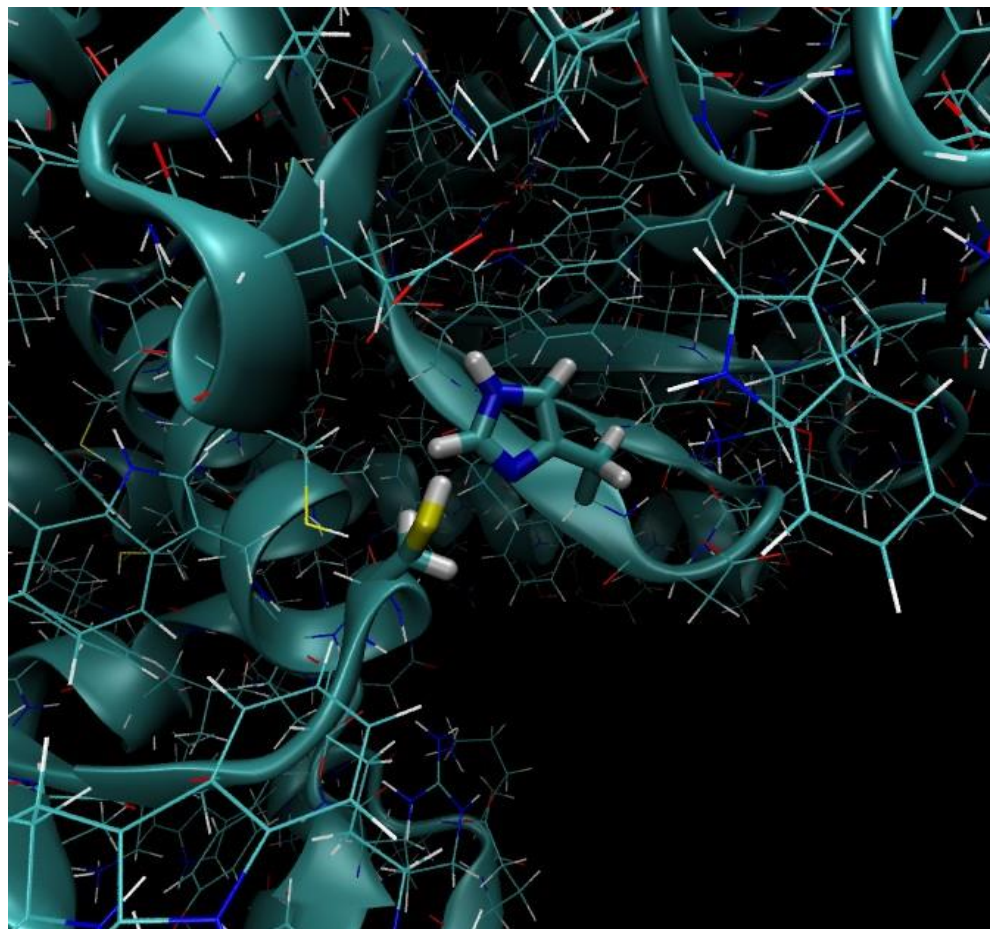
**After adding link-atoms.**
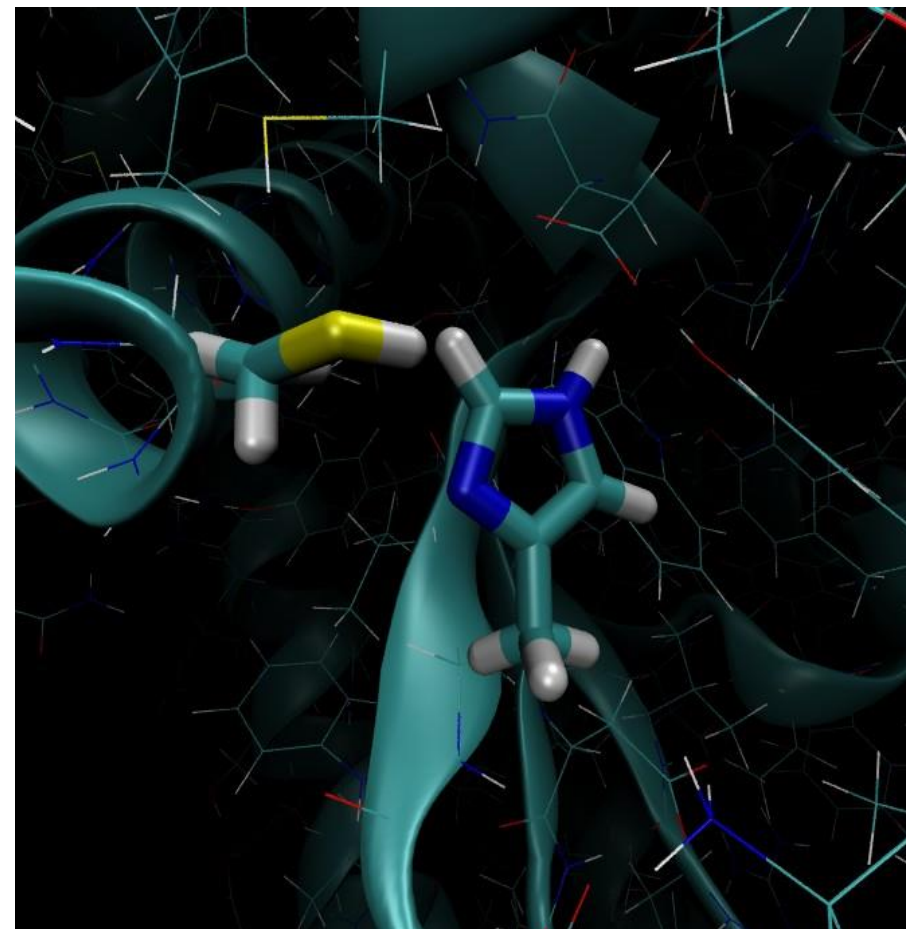
# The QM/MM Energy Expression;
## Subtractive and Additive QM/MM Schemes



Subtractive scheme:

Additive scheme:

Navizet, I., 2020. QM/MM Study of Bioluminescent Systems. *Challenges and Advances in Computational Chemistry and Physics*, pp.227-270.

# The QM/MM Energy Expression (Contd.)

- Subtractive QM/MM schemes require-

    (1) an MM calculation on the entire system;
    (2) a QM calculation on the inner subsystem; and
    (3) an MM calculation on the inner subsystem.

    ➢ The QM/MM energy of the entire system is then obtained by summing (1) and (2) and subtracting (3).

# The QM/MM Energy Expression (Contd.)

## Subtractive Scheme: ONIOM Method

- own N-layered Integrated molecular Orbital and molecular Mechanics

$$E_{\text{sys}} = E_{\text{MM}}(L + R) - E_{\text{MM}}(L) + E_{\text{QM}}(L)$$

- $L$ is a ligand in the QM region and $R$ is a receptor (protein) in the MM region

## Additive Scheme

-

$$E_{\text{sys}} = E_{\text{MM}}(R) + E_{\text{QM}}(L) + E_{\text{QM/MM}}$$

- There is an explicit QM-MM coupling term $E_{\text{QM/MM}}$.
- $E_{\text{QM/MM}}$ collects the interaction between QM and MM regions.

$$E_{\text{QM/MM}} = E_{\text{QM/MM}}^{\text{bonded}} + E_{\text{QM/MM}}^{\text{vdW}} + E_{\text{QM/MM}}^{\text{elec}}$$

https://github.com/swillow/modelingworkshop

# The Electrostatic QM–MM Interaction

- The electrostatic coupling between the QM charge density and the charge model used in the MM region can be handled at different levels of sophistication, characterized essentially by the extent of mutual polarization and classified accordingly as;

    **mechanical** embedding  **electrostatic** embedding, and **polarized** embedding.

1. mechanical embedding: use the MM-MM electrostatics.

$$E^{\text{elec}}_{\text{QM/MM}} = E^{\text{Coul}}_{\text{QM/MM}} = \sum_{A \in \text{QM}} \sum_{B \in \text{MM}} \frac{q_A q_B}{|\boldsymbol{R}_A - \boldsymbol{R}_B|}$$
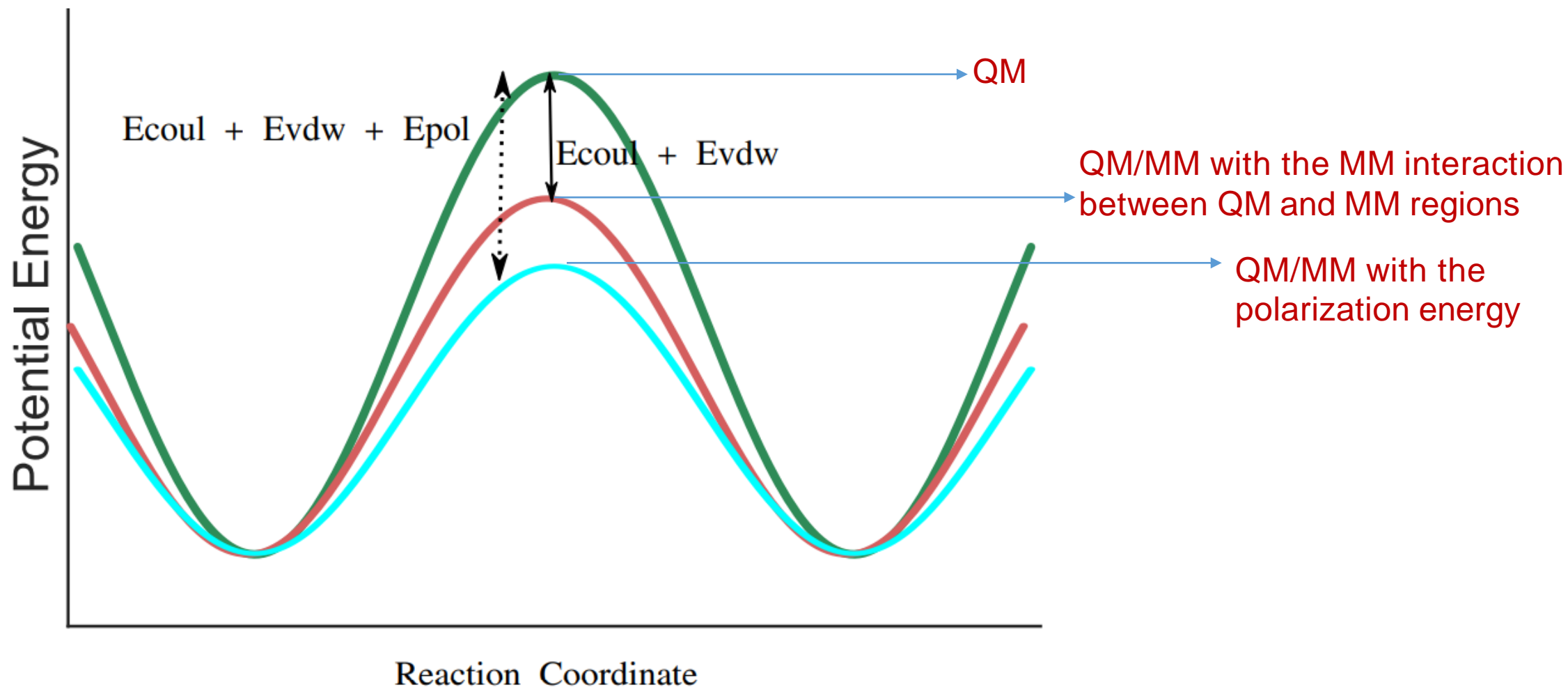
2. electrostatic embedding: perform the QM calculation as follows:

$$E^{\text{elec}}_{\text{QM/MM}} = \langle \Psi_{I:Q_I} | \hat{H}_{I:Q_I} | \Psi_{I:Q_I} \rangle - \langle \Psi_I | \hat{H}_I | \Psi_I \rangle$$
$$= E^{\text{Coul}}_{I:Q_I} + E^{\text{pol}}_{I:Q_I}$$

# Other **Nonbonded** and **Bonded** QM–MM Interactions

- There are also **van der Waals** and **bonded contributions** to the QM–MM coupling term.
- Their treatment is considerably simpler as they are handled purely at the MM level, irrespective of the class (subtractive or additive) of QM/MM scheme.
- The van der Waals interaction is typically described by a **Lennard–Jones potential**
- The formal reservations against using standard MM parameters to describe QM–MM interactions apply of course also to the bonded (**bond stretching, angle bending, torsional**, etc.) interactions.
- For electrostatic or polarized embedding, overpolarization of the QM density by the MM charges close to the cut has to be prevented, especially when using link atoms.

# Potential Energy Surface for a reaction

# Google Colab Session