C8855 Advanced Molecular Modelling Methods

Lesson 2

Empirical Valence Bond Theory

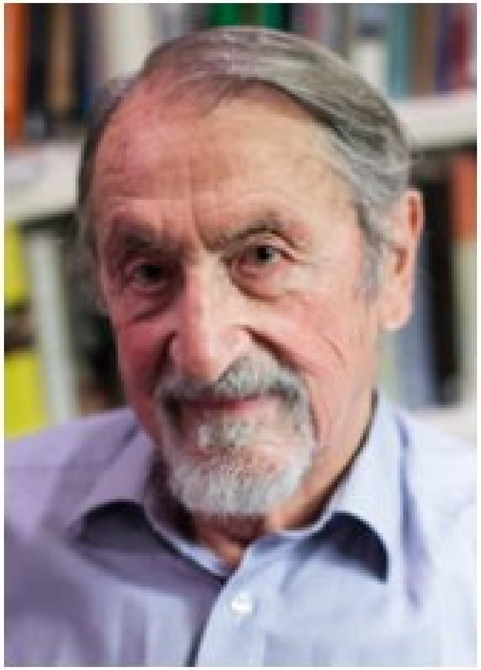
JS/2022 Present Form of Teaching: Rev1

Petr Kulhánek

kulhanek@chemi.muni.cz

National Centre for Biomolecular Research, Faculty of Science
Masaryk University, Kamenice 5, CZ-62500 Brno

Nobel Laureates in Chemistry 2013



© Harvard University

Martin Karplus

Université de Strasbourg, Strasbourg, France, Harvard University, Cambridge, MA, USA

Photo: © S. Fisch

Michael Levitt

Stanford University School of Medicine, Stanford, CA, USA

Photo: Wikimedia Commons

Arieh Warshel

University of Southern California, Los Angeles, CA, USA

http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/

Essential Works

Lifson, S.; Warshel, A. Consistent Force Field for Calculations of Conformations Vibrational Spectra and Enthalpies of Cycloalkane and N-Alkane Molecules. *J. Chem. Phys.* **1968**, *49*, 5116–&.

Levitt, M.; Lifson, S. Refinement of Protein Conformations Using a Macromolecular Energy Minimization Procedure. *J. Mol. Biol.* **1969**, *46*, 269–&.

Honig, B.; Karplus, M. Implications of Torsional Potential of Retinal Isomers for Visual Excitation. *Nature* **1971**, *229*, 558–&.

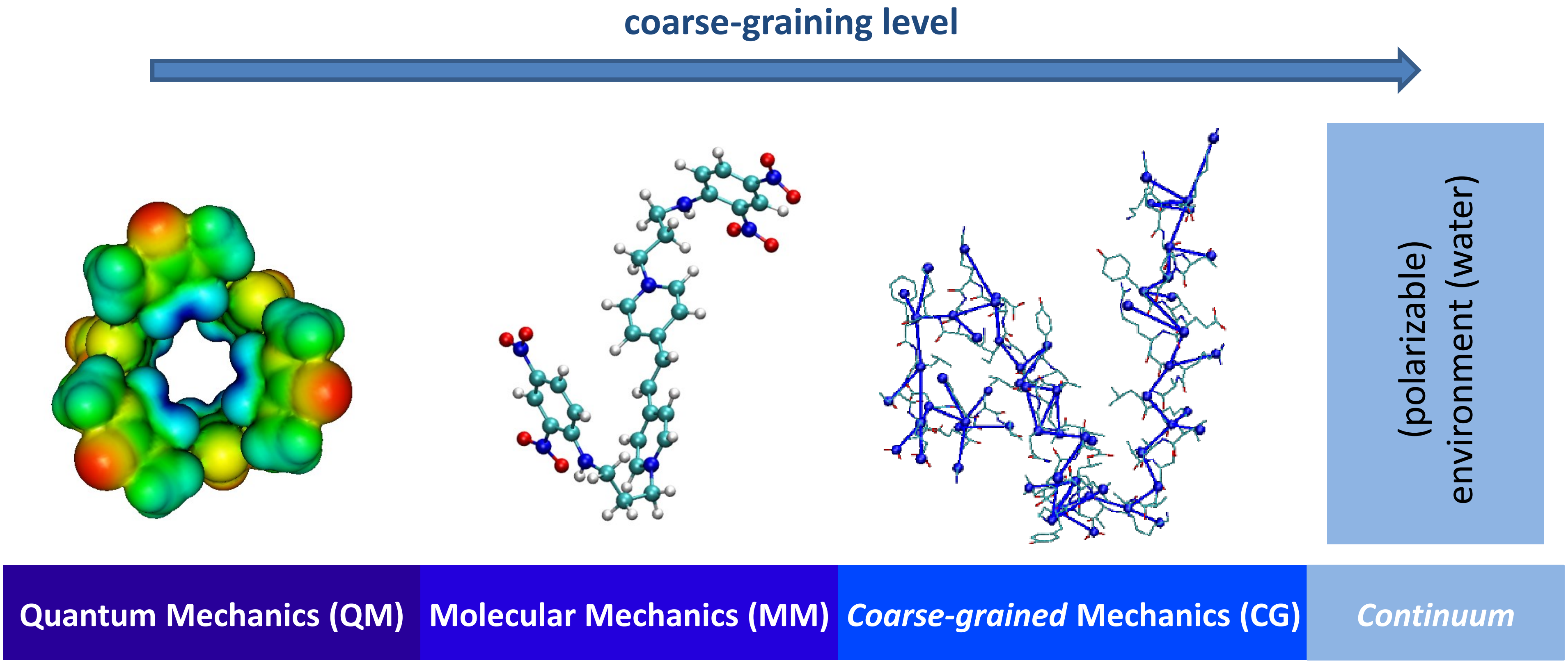
Warshel, A.; Karplus, M. Calculation of Ground and Excited-State Potential Surfaces of Conjugated Molecules .1. Formulation and Parametrization. *J. Am. Chem. Soc.* **1972**, *94*, 5612–&.

Levitt, M.; Warshel, A. Computer-Simulation of Protein Folding. *Nature* **1975**, *253*, 694–698.

Levitt, M. Simplified Representation of Protein Conformations for Rapid Simulation. *J. Mol. Biol.* **1976**, *104*, 59–107.

Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions - Dielectric, Electrostatic and Steric Stabilization of Carbonium-Ion in Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.

Foundation of Methods



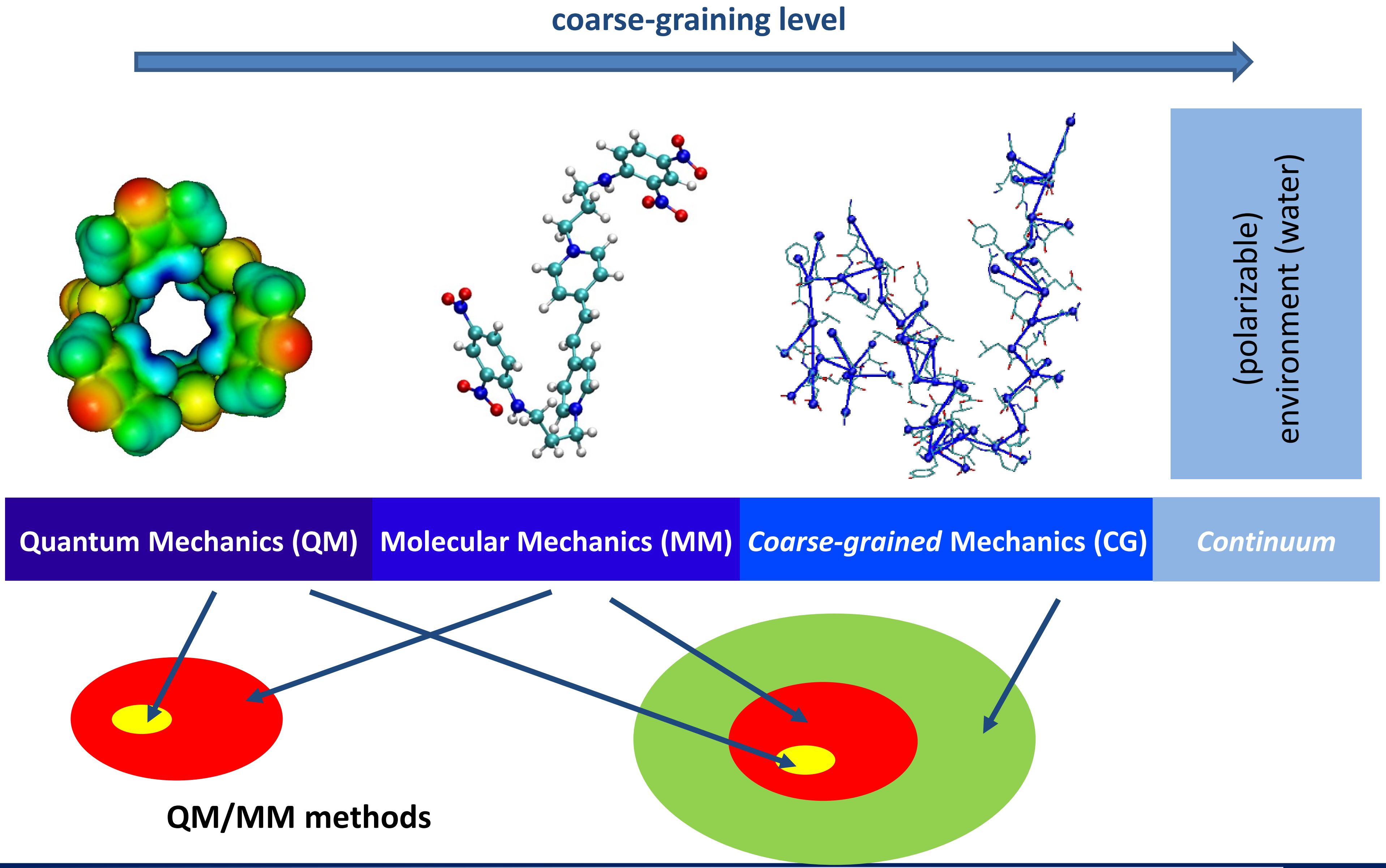
coarse-graining level

(polarizable)
environment (water)

Quantum Mechanics (QM) | **Molecular Mechanics (MM)** | ***Coarse-grained* Mechanics (CG)** | ***Continuum***

Newtonian (classical) mechanics

Multiscale Methods



coarse-graining level

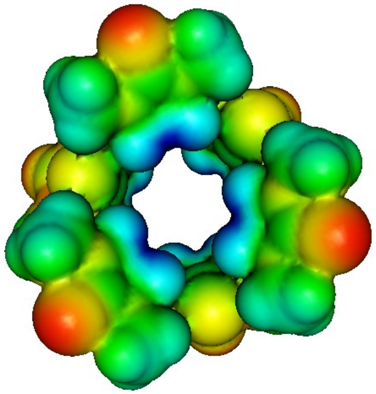
(polarizable) environment (water)

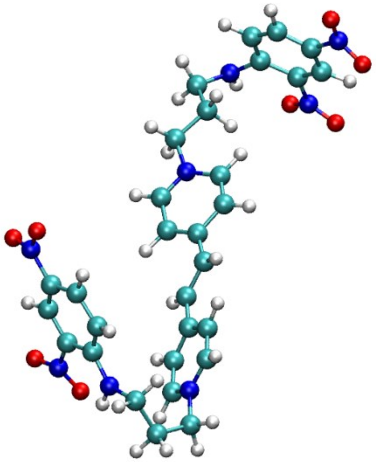
Quantum Mechanics (QM) | **Molecular Mechanics (MM)** | ***Coarse-grained*** **Mechanics (CG)** | ***Continuum***

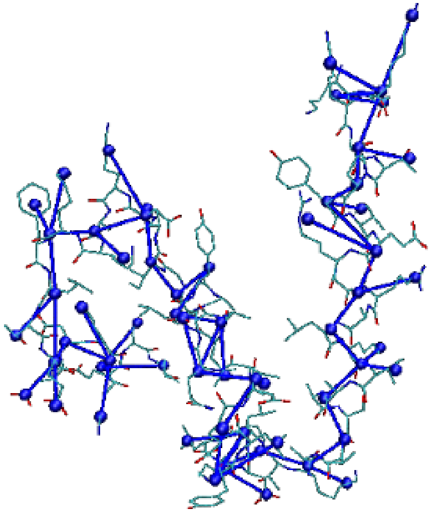
QM/MM methods

Multiscale Methods

coarse-graining level →

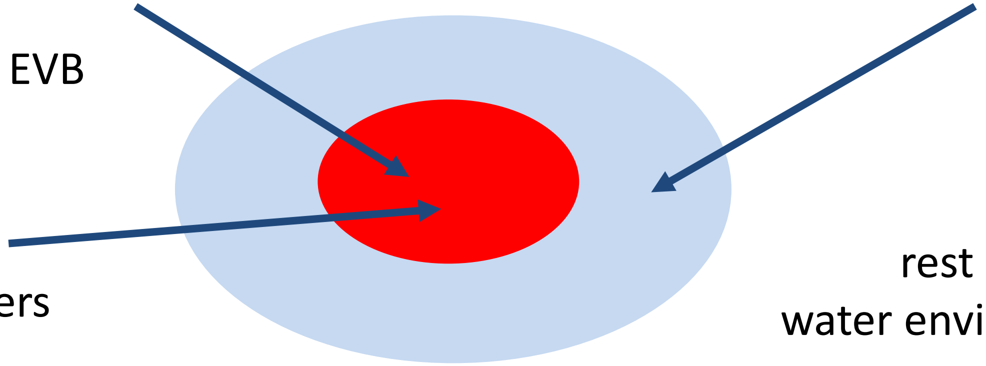






(polarizable) environment (water)

Quantum Mechanics (QM)	Molecular Mechanics (MM)	*Coarse-grained* Mechanics (CG)	*Continuum*

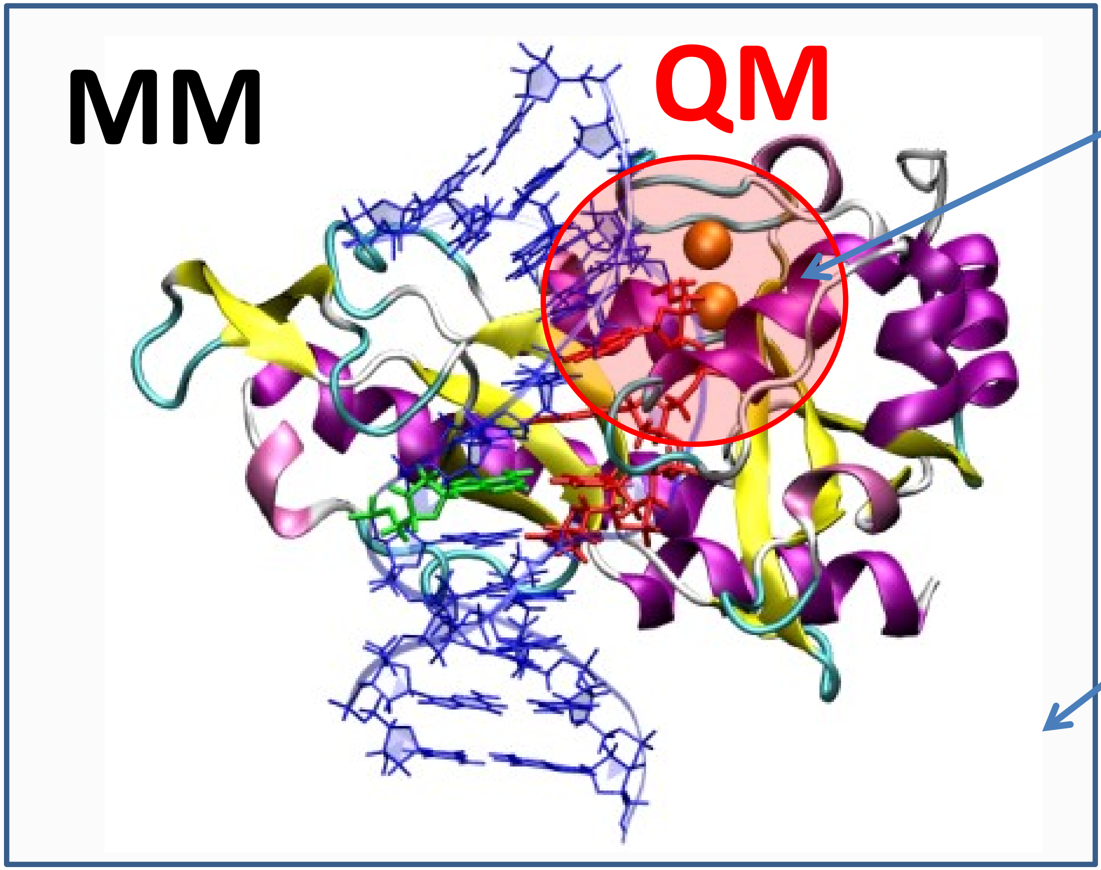


Valence state for EVB

protein or
reference system
and a few solvation layers

rest of
water environment

Enzymatic Reactions



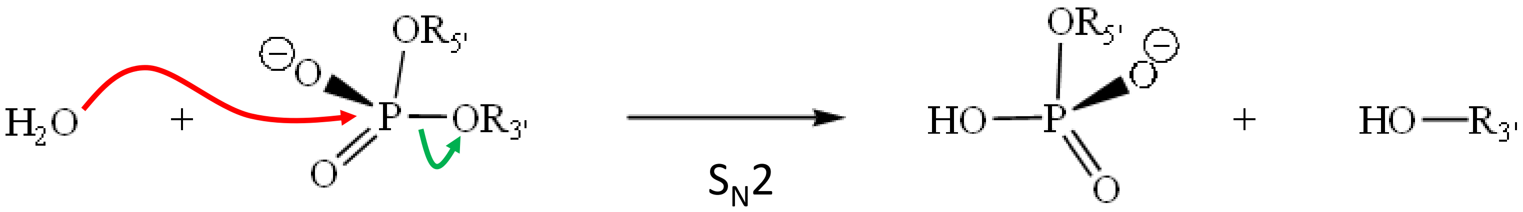
QM

active site only
tens to hundreds atoms

MM

rest of enzyme
solvent (water)

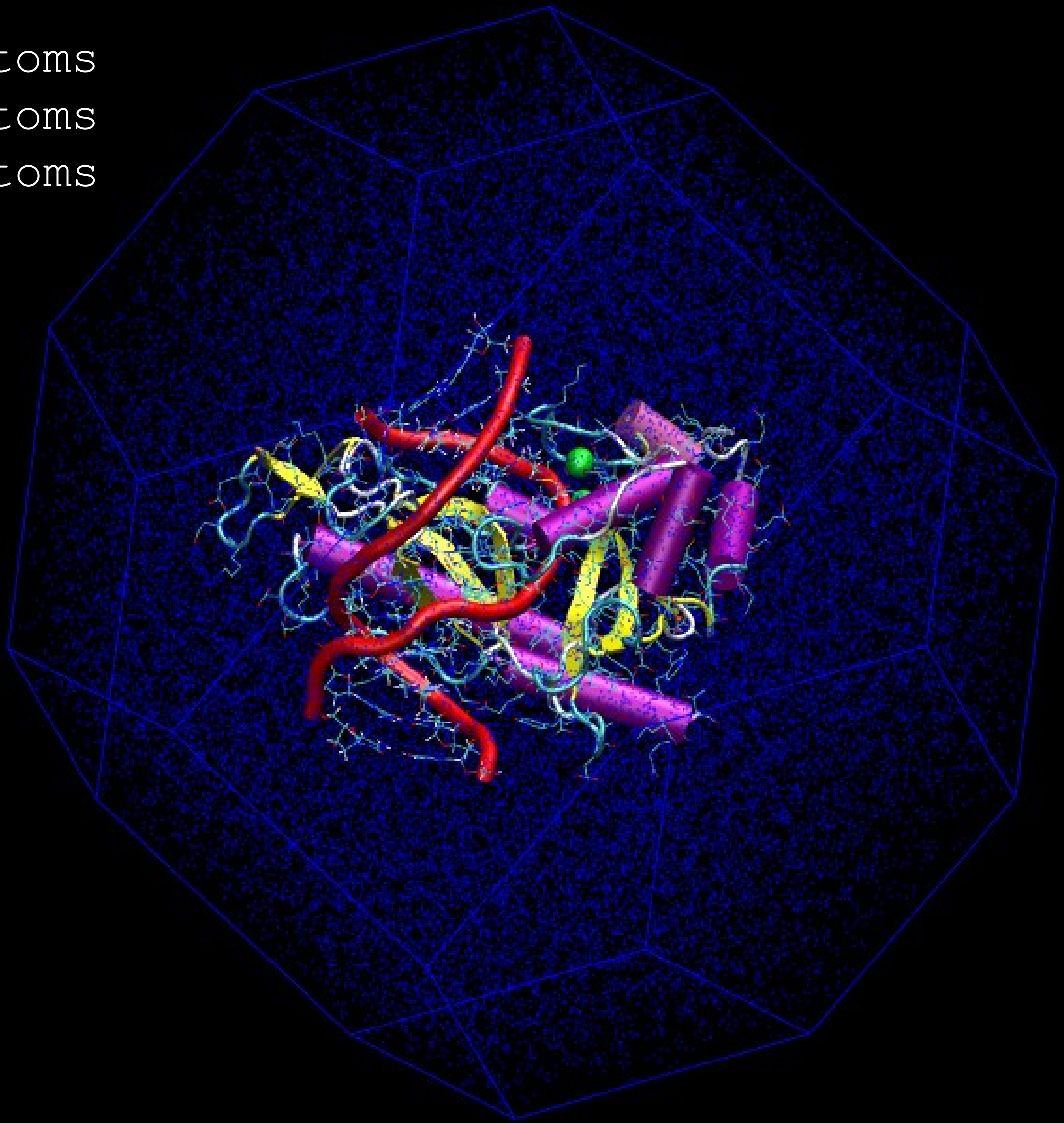
Catalyzed reaction – hydrolysis of phosphodiester bond



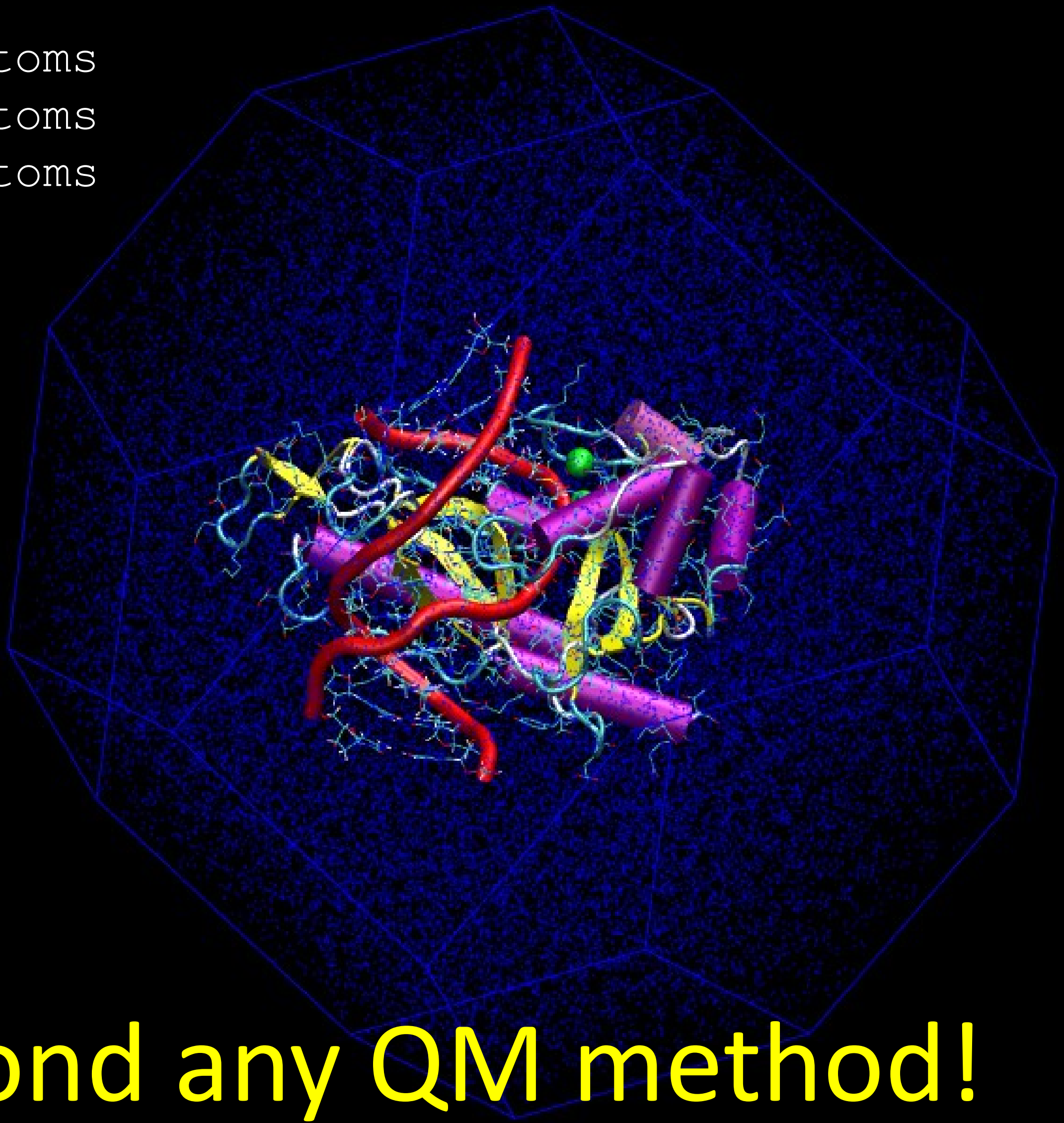
Enzymatic Reactions

```
Enzyme: ~4,300 atoms
Water: ~42,000 atoms
Total: ~46,500 atoms
```

MutH



Enzymatic Reactions

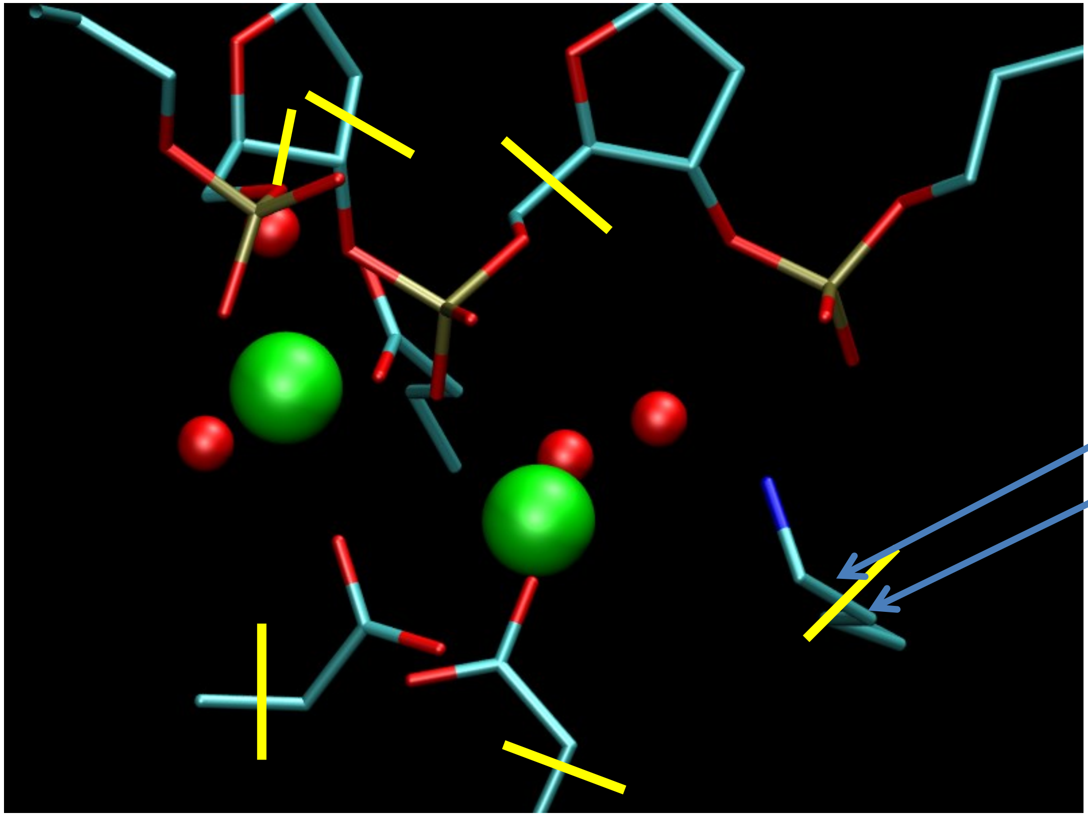


```
Enzyme: ~4,300 atoms
Water: ~42,000 atoms
Total: ~46,500 atoms
```

MutH

Beyond any QM method!

Boundary Problems



Active site of MutH

(hydrogen atoms are not shown for clarity)

e^-

e^-

Two electrons formally contribute to σ-bond (single bond).

- How to cut covalent bonds?
- There is an unphysical tension at boundary due to incompatible precision of QM and MM potentials.
- MM atoms cannot be polarized by QM zone
- QM atoms can be (over)polarized by MM atoms

Quantum Mechanics

$$\hat{H}_e \Psi_m(\mathbf{r},\mathbf{R}) = E_m(\mathbf{R}) \Psi_m(\mathbf{r},\mathbf{R})$$

How to describe wave function (WF) of many-body systems?

- Molecular Orbital (MO) Theory
- Valence Bond (VB) Theory

MO is now prevalent because MO was simpler to implement than VB in early stages of QM development.

Molecular Orbital Theory

Molecular Orbital (MO) theory is about construction of many-body system wavefunction (WF).

One-electron approximation:

$$\Psi(r_1, r_2, \ldots, r_n, \sigma_1, \sigma_2, \ldots, \sigma_n) = \sum_P sign(P)\{\varphi_1(r_1)\chi_1(\sigma_1)\varphi_2(r_2)\chi_2(\sigma_2)\ldots\varphi_n(r_n)\chi_n(\sigma_n)\}$$

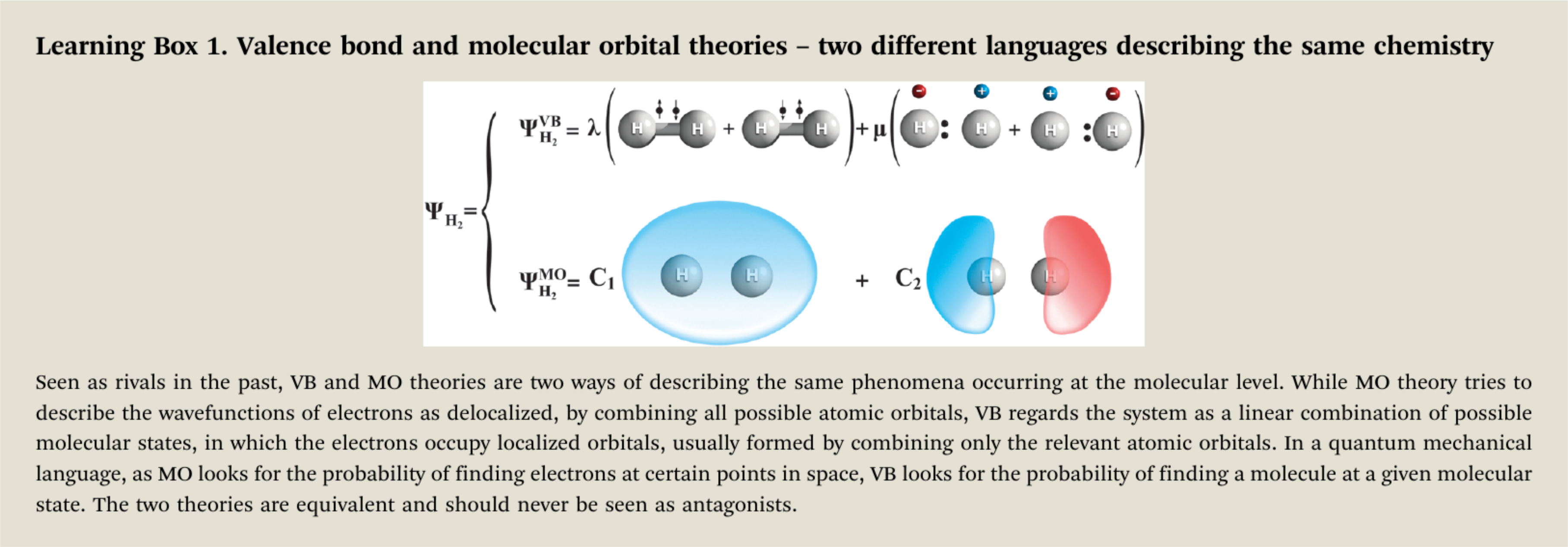
One-electron functions expressed using basis functions:

$$\varphi_i(\mathbf{r}_i) = \sum_{j=1}^{m} \boxed{c_{ij}} \chi_j(\mathbf{r}_i)$$

MO-LCAO - molecular orbital as linear combination of atomic orbitals

Valence Bond Theory

Valence Bond (VB) theory is about construction of many-body system wavefunction (WF).



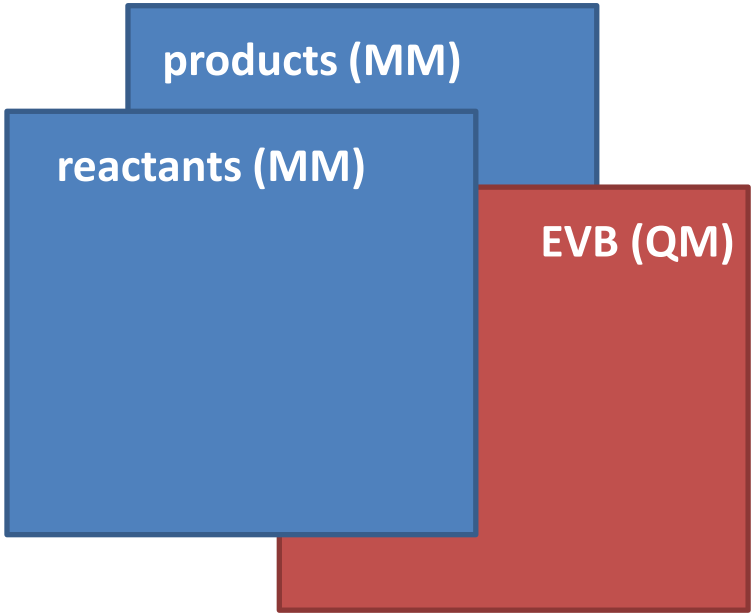
Learning Box 1. Valence bond and molecular orbital theories – two different languages describing the same chemistry

Seen as rivals in the past, VB and MO theories are two ways of describing the same phenomena occurring at the molecular level. While MO theory tries to describe the wavefunctions of electrons as delocalized, by combining all possible atomic orbitals, VB regards the system as a linear combination of possible molecular states, in which the electrons occupy localized orbitals, usually formed by combining only the relevant atomic orbitals. In a quantum mechanical language, as MO looks for the probability of finding electrons at certain points in space, VB looks for the probability of finding a molecule at a given molecular state. The two theories are equivalent and should never be seen as antagonists.

Shurki, A.; Derat, E.; Barrozo, A.; Kamerlin, S. C. L. **How Valence Bond Theory Can Help You Understand Your (Bio)Chemical Reaction.** *Chem. Soc. Rev.* **2015**, *44* (5), 1037–1052. https://doi.org/10.1039/C4CS00241E.

EVB *vs* hybrid QM/MM

Empirical Valence Bond (EVB)



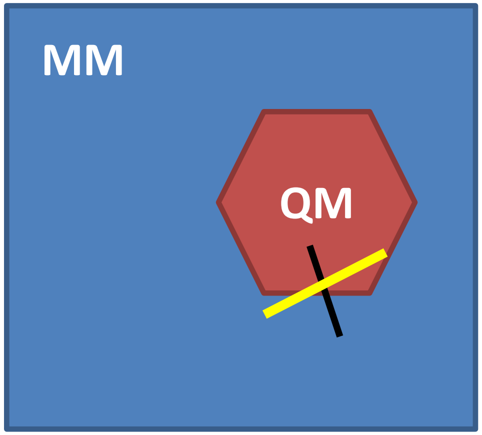
products (MM)

reactants (MM)

EVB (QM)

- no boundary problems
- simulation on mapping potentials
- reference reaction for calibration

Hybrid QM/MM



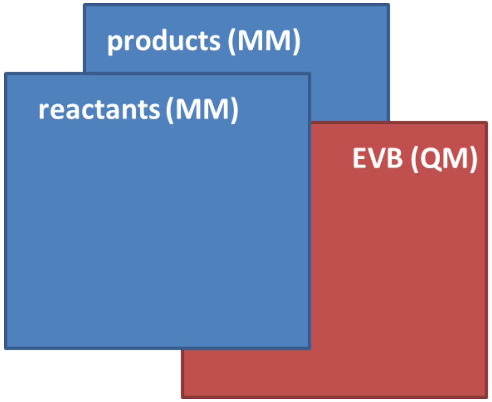
MM

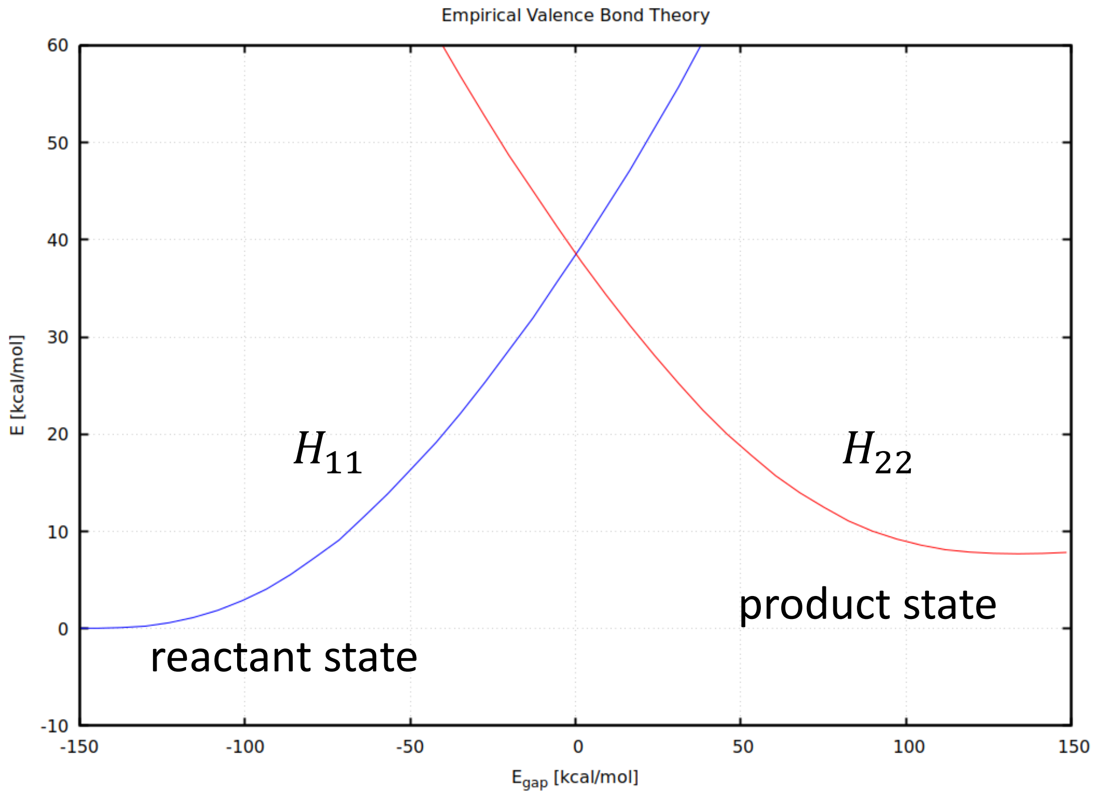
QM

- boundary problems
- any QM theory (DFT, semiempirics, etc.)

Warshel, A.; Weiss, R. **An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes**. *J. Am. Chem. Soc.* **1980**, *102* (20), 6218–6226. https://doi.org/10.1021/ja00540a008.

Reactants and products as VB states



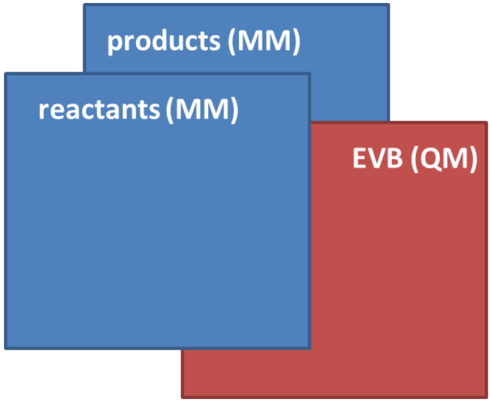


$$H_{11} = \varepsilon_1 \qquad H_{22} = \varepsilon_2 + \alpha$$

valence states described by MM

a constant, which is a correction to different MM reference states

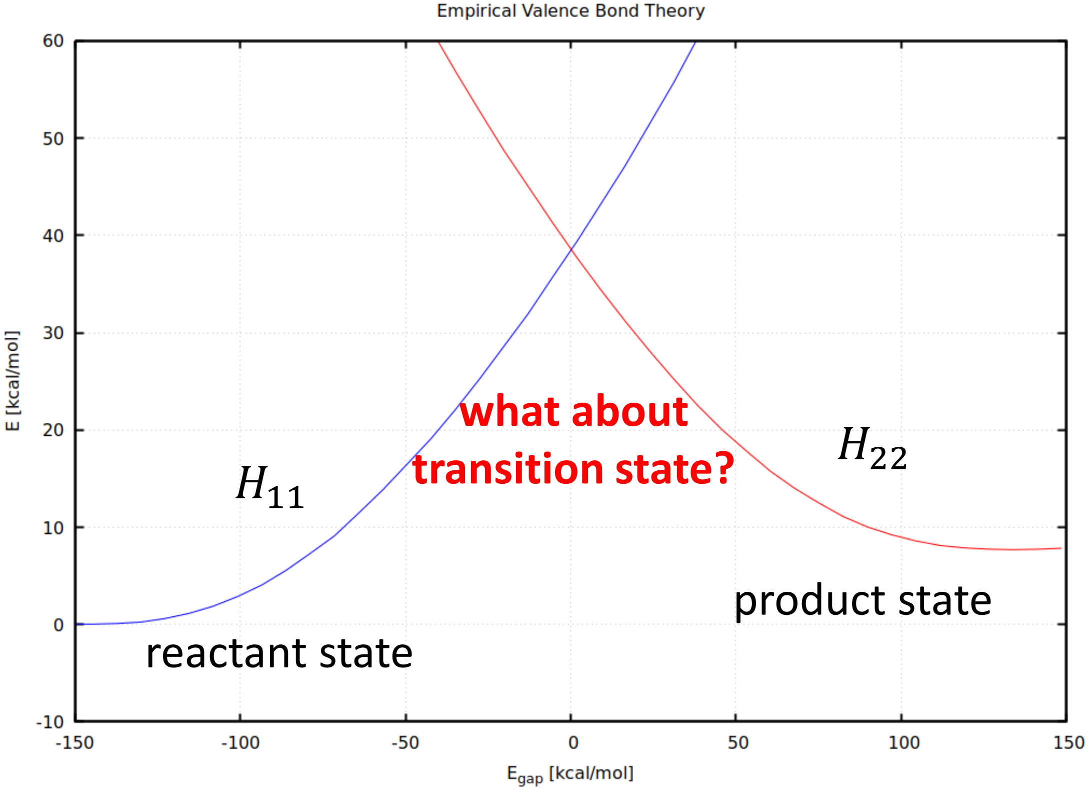
Reactants and products as VB states



products (MM)

reactants (MM)

EVB (QM)



H_{11}

H_{22}

what about transition state?

reactant state

product state

$$H_{11} = \varepsilon_1 \quad H_{22} = \varepsilon_2 + \alpha$$

valence states described by MM

a constant, which is a correction to different MM reference states

Empirical Valence Bond

valence states described by MM

$$H_{11} = \varepsilon_1 \qquad H_{22} = \varepsilon_2 + \alpha$$

off-diagonal mixing terms

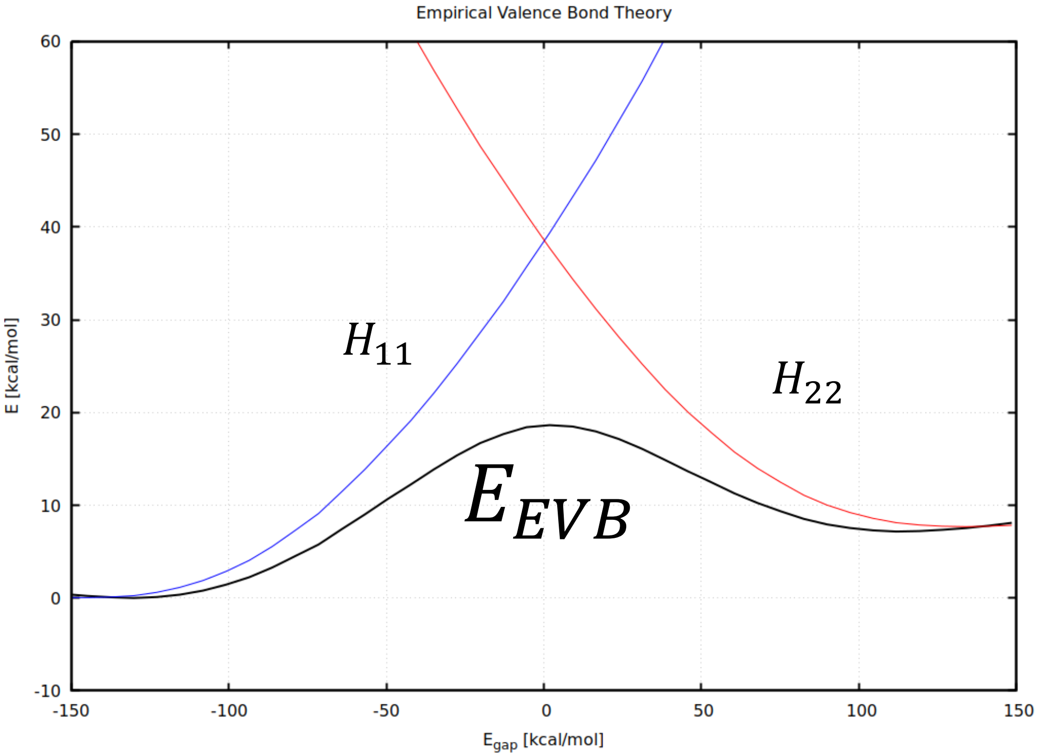
$$H_{12} = H_{21} = ?$$

Hamiltonian matrix

$$\mathbf{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$$

Secular problem

$$\mathbf{H}\psi_k = E_k \psi_k$$



Secular problem has **two solutions**:

$$E_2 > E_1 = E_{EVB}$$

ground state energy = EVB potential

Empirical Valence Bond

valence states described by MM

$$H_{11} = \varepsilon_1 \qquad H_{22} = \varepsilon_2 + \alpha$$

off-diagonal mixing terms

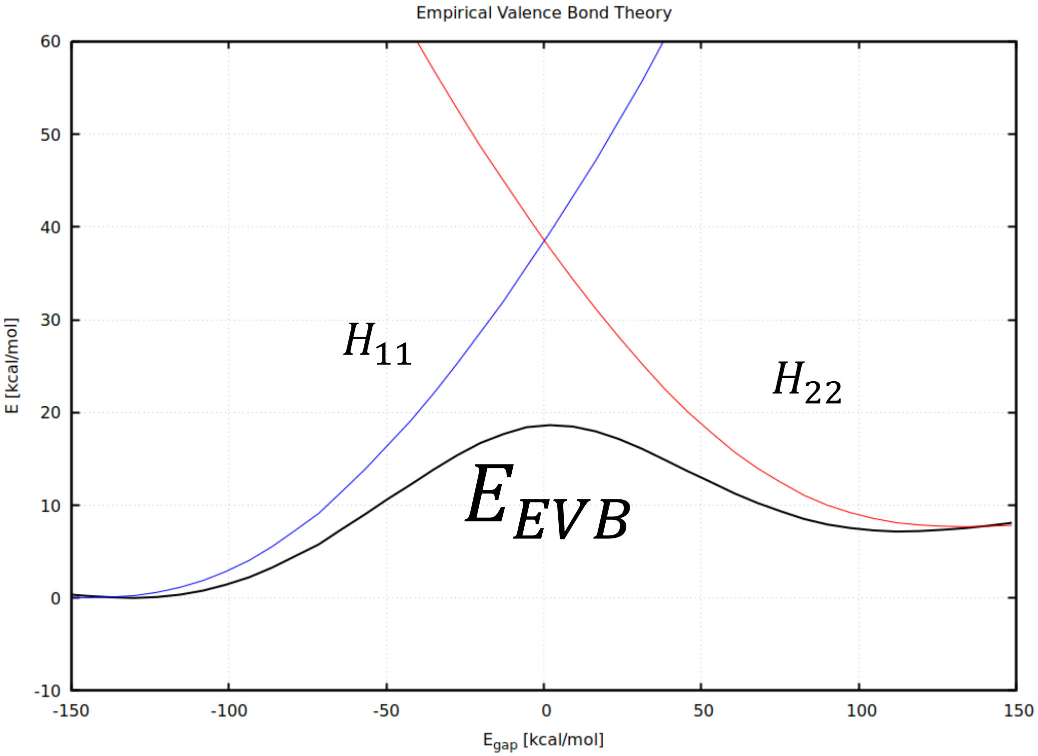
$$H_{12} = H_{21} = ?$$

Hamiltonian matrix

$$\mathbf{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$$

Secular problem

$$\mathbf{H}\psi_k = E_k \psi_k$$



WF describes contributions of individual valence states to the total system description.

EVB Assumption

Hong, G.; Rosta, E.; Warshel, A. **Using the Constrained DFT Approach in Generating Diabatic Surfaces and Off Diagonal Empirical Valence Bond Terms for Modeling Reactions in Condensed Phases.** *The Journal of Physical Chemistry B* **2006**, *110* (39), 19570–19574. https://doi.org/10.1021/jp0625199.

The empirical valence bond (EVB) model provides an extremely powerful way for modeling and analyzing chemical reactions in solutions and proteins. However, **this model is based on the unverified assumption that the off diagonal elements of the EVB Hamiltonian do not change significantly upon transfer of the reacting system from one phase to another**. This ad hoc assumption has been rationalized by its consistency with empirically observed linear free energy relationships, as well as by other qualitative considerations. Nevertheless, this assumption has not been rigorously established. The present work explores the validity of the above EVB key assumption by a rigorous numerical approach. This is done by exploiting the ability of the frozen density functional theory (FDFT) and the constrained density functional theory (CDFT) models to generate convenient diabatic states for QM/MM treatments, and thus to examine the relationship between the diabatic and adiabatic surfaces, as well as the corresponding effective off diagonal elements. It is found that, at least for the test case of SN2 reactions, the off diagonal element does not change significantly upon moving from the gas phase to solutions and thus the EVB assumption is valid and extremely useful.

EVB Assumption

Warshel, A.; Weiss, R. **An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes**. *J. Am. Chem. Soc.* **1980**, *102* (20), 6218–6226. https://doi.org/10.1021/ja00540a008.

Hong, G.; Rosta, E.; Warshel, A. **Using the Constrained DFT Approach in Generating Diabatic Surfaces and Off Diagonal Empirical Valence Bond Terms for Modeling Reactions in Condensed Phases**. *The Journal of Physical Chemistry B* **2006**, *110* (39), 19570–19574. https://doi.org/10.1021/jp0625199.

off-diagonal mixing terms

$$H_{12} = H_{21} = Ae^{-\mu(r_{ij}-r_o)}$$

26 years !!!!!

some characteristic geometry parameter (broaken/formed bond)

off-diagonal mixing terms - typical form

$$H_{12} = H_{21} = A$$

🟥 **empirical parameters**

EVB Parameters

valence states described by MM

$$H_{11} = \varepsilon_1 \qquad H_{22} = \varepsilon_2 + \alpha$$

off-diagonal mixing terms

$$H_{12} = H_{21} = A$$

EVB in its simplest form has only two parameters, which are characteristic for given reaction but independent on environment (water, protein interior).

🟥 **empirical parameters**

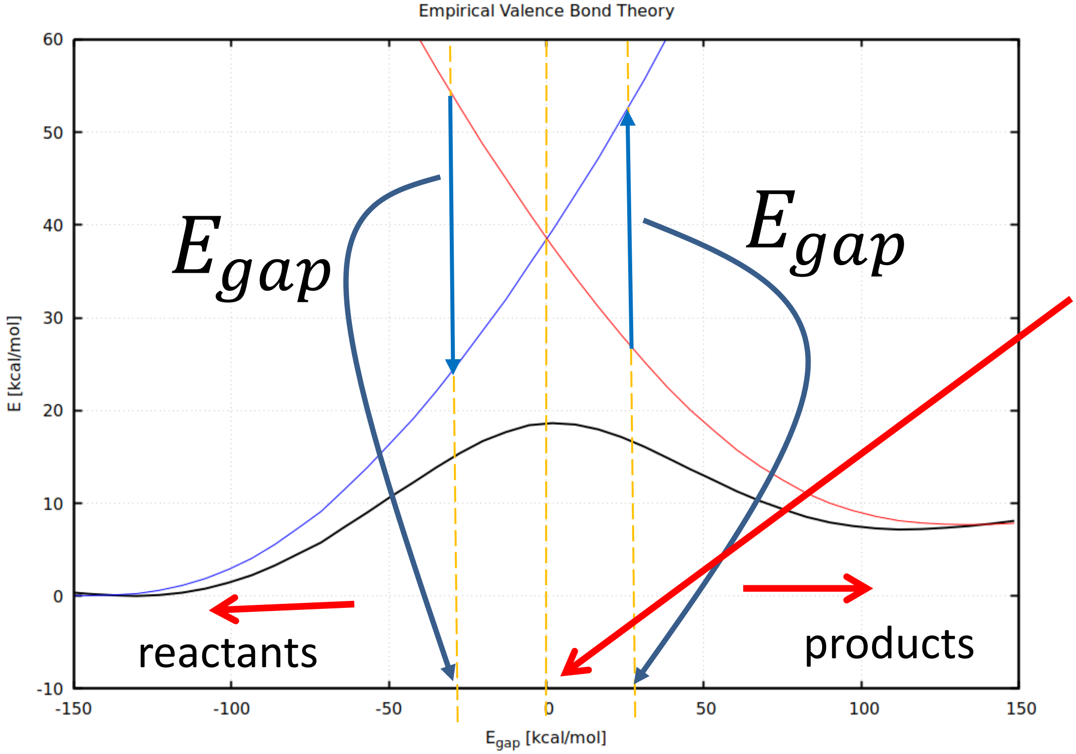
Reaction coordinate

How to promote the reaction, i.e., the transition between reactant and product states?

Suitable reaction coordinate is the **energy gap** between two valence states:

$$\xi = E_{gap} = H_{22} - H_{11} = (\varepsilon_2 + \alpha) - \varepsilon_1$$

it can be omitted



transition state at $E_{gap} = 0$

$$\xi = E_{gap}$$

Simulations on EVB potential

- It is possible to run MD simulations on EVB potential.
- However, the EVB parameters need to be provided before the simulation.
- Free energies (**reaction and activation free energies**) can be calculated using any biasing techniques:
  - Metadynamics
  - Umbrella Sampling
  - Adaptive Biasing Force method
  - etc.

kinetics

standard activation Gibbs energy

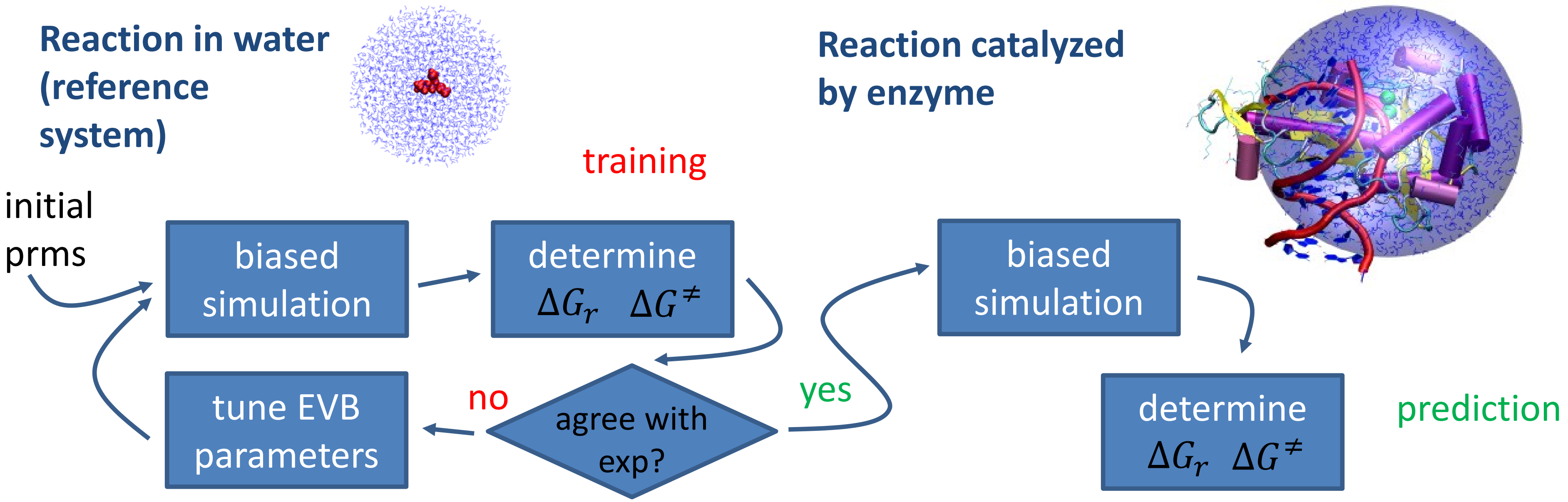
rate constant
(experiment)

$$k = \kappa \frac{k_B T}{h} e^{-\frac{\Delta G^{\neq}}{RT}}$$

Eyring equation (theoretical model)

Simulations on EVB potential

1. training/parametrization
   model the reaction in water environment (reference system) until the reaction and activation free energies are the same as obtained by experiments
2. prediction
   use the optimized EVB parameters to predict impact of the enzyme on the reaction and activation free energies (determine the catalytic effect of the enzyme)



any new EVB parameters require to run a new simulation

→ **the EVB parametrization is thus tedious and computationally demanding** ....

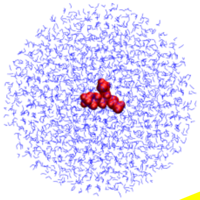
Simulations on EVB potential

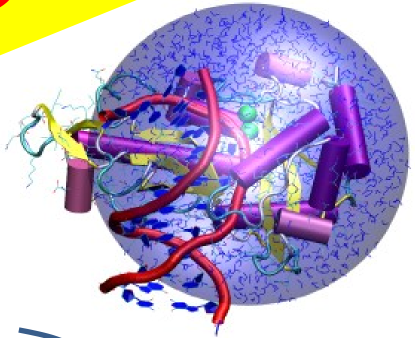
1. training/parametrization

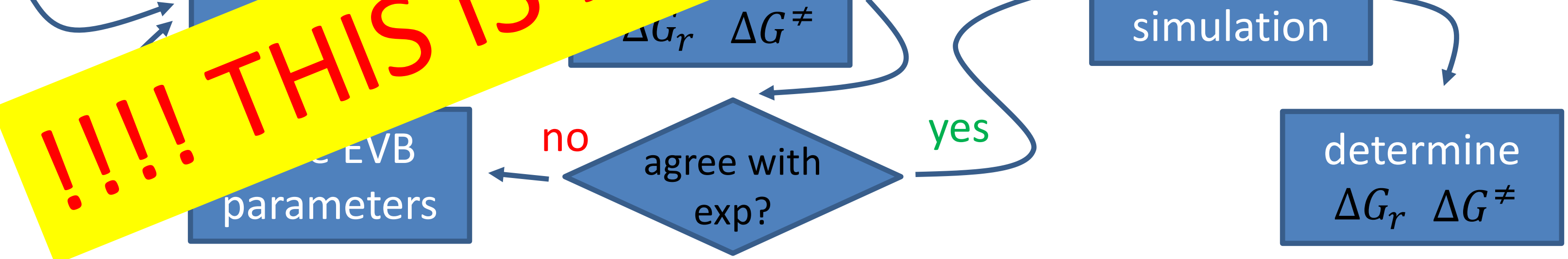
   model the reaction in water environment (reference system) until the reaction and activation free energies are the same as obtained by experiments
2. prediction

   use the optimized EVB parameters to predict impact of the enzym
   and activation free energies (determine the catalytic effect

Reaction in water (reference system)







initial prms

bia

ne
$\Delta G_r \quad \Delta G^{\neq}$

e EVB parameters

no

agree with exp?

yes

biased simulation

determine
$\Delta G_r \quad \Delta G^{\neq}$

prediction

!!!! THIS IS NOT USE IN PRACTICE !!!!

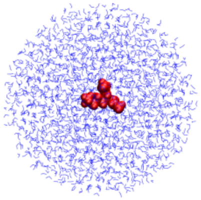
any new EVB parameters require to run a new simulation

→ **the EVB parametrization is thus tedious and computationally demanding** ….

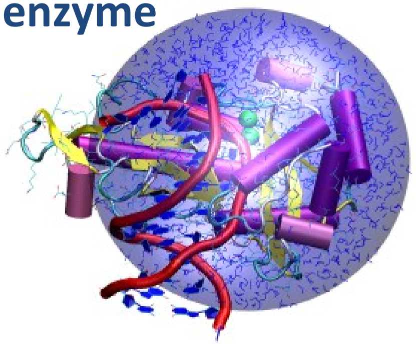
Mapping Potential and Reference Reaction

Reaction in water (reference system)

Reaction catalyzed by enzyme



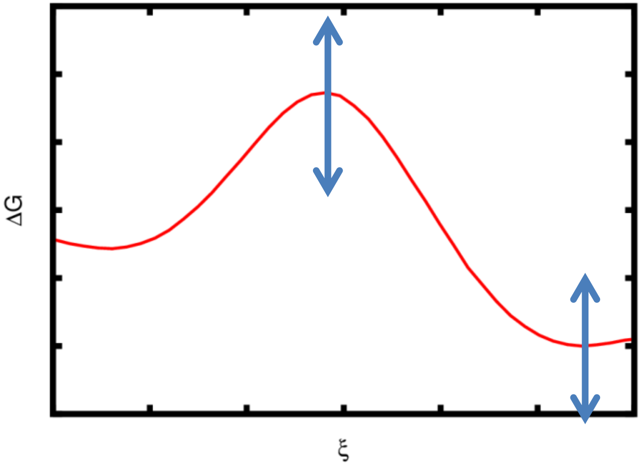
$$E_{map} = (1-\lambda)\varepsilon_1 + \lambda\varepsilon_2$$



Simulation on **mapping potential**

Simulation on **mapping potential**

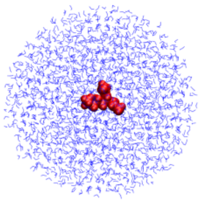
no EVB parameters are required to run simulations on the mapping potential
both simulations are thus independent; they can be run in a parallel



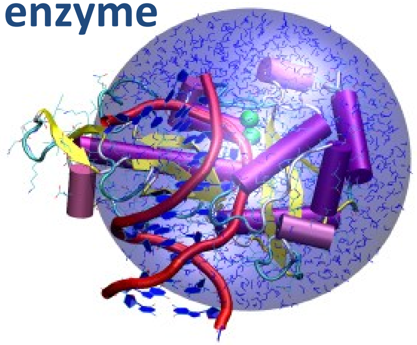
computationally demanding
but no EVB parameters are needed
virtually describe all possible reactions

Mapping Potential and Reference Reaction

Reaction in water (reference system)



Reaction catalyzed by enzyme



$$E_{map} = (1-\lambda)\varepsilon_1 + \lambda\varepsilon_2$$

Simulation on **mapping potential**

Simulation on **mapping potential**

computationally demanding
but no parameters are needed
virtually describe all possible reactions

training

EVB reconstruction via FEP/US (repeat)

computationally cheap

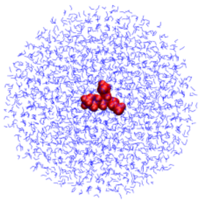
EVB parameters that reproduces
experimental reaction and activation
Gibbs free energies.

FEP/US – Free Energy Perturbation/Umbrella Sampling

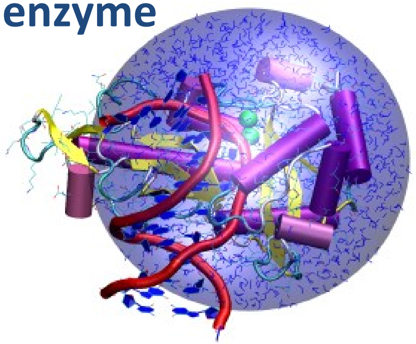
Mapping Potential and Reference Reaction

Reaction in water (reference system)

Reaction catalyzed by enzyme



$$E_{map} = (1-\lambda)\varepsilon_1 + \lambda\varepsilon_2$$



Simulation on **mapping potential**

Simulation on **mapping potential**

computationally demanding
but no parameters are needed
virtually describe all possible reactions

EVB reconstruction via FEP/US (repeat)

computationally cheap

EVB parameters that reproduces **experimental** reaction and activation Gibbs free energies.

prediction

EVB reconstruction via FEP/US

Predicted reaction and activation Gibbs free energies for **catalyzed reaction**.

FEP/US – Free Energy Perturbation/Umbrella Sampling

Results

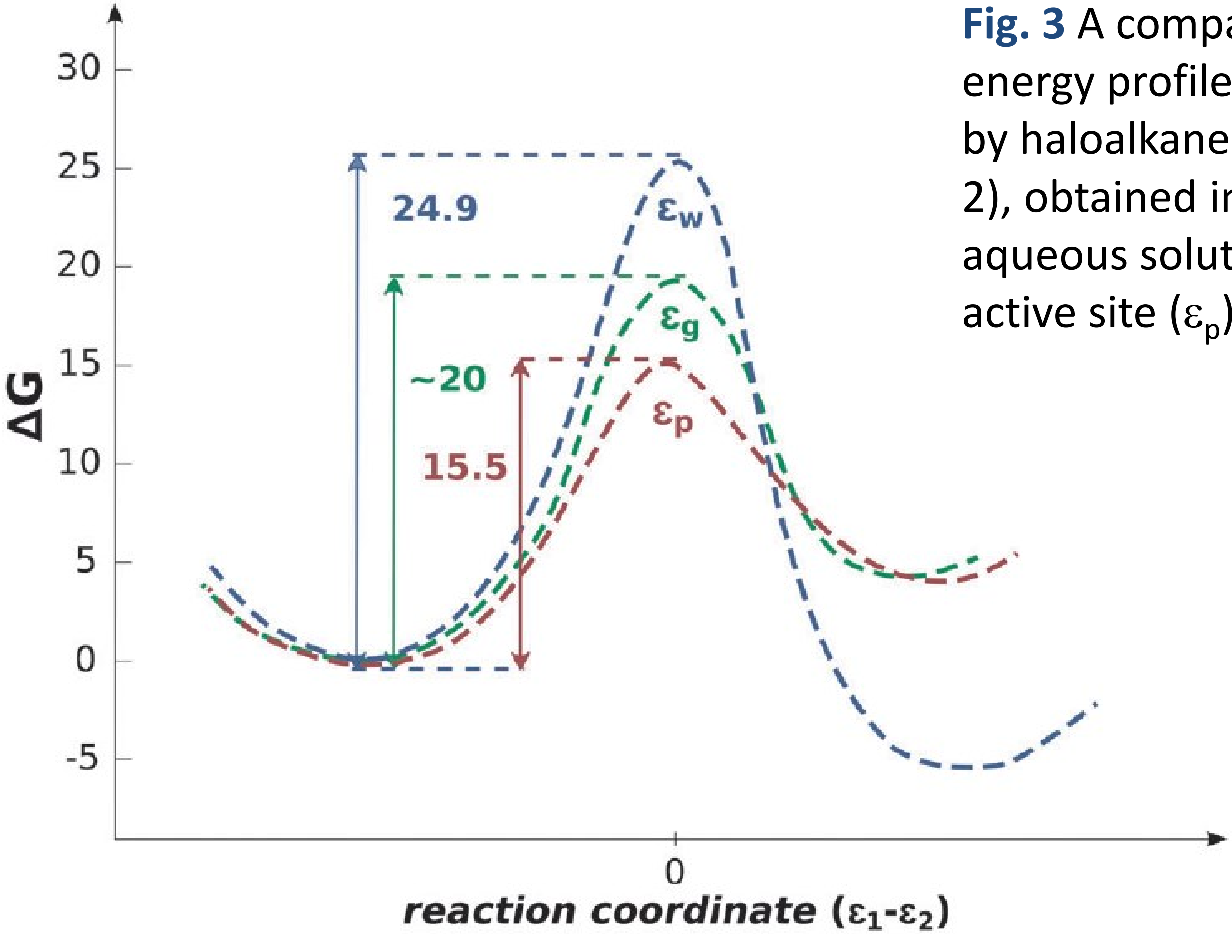
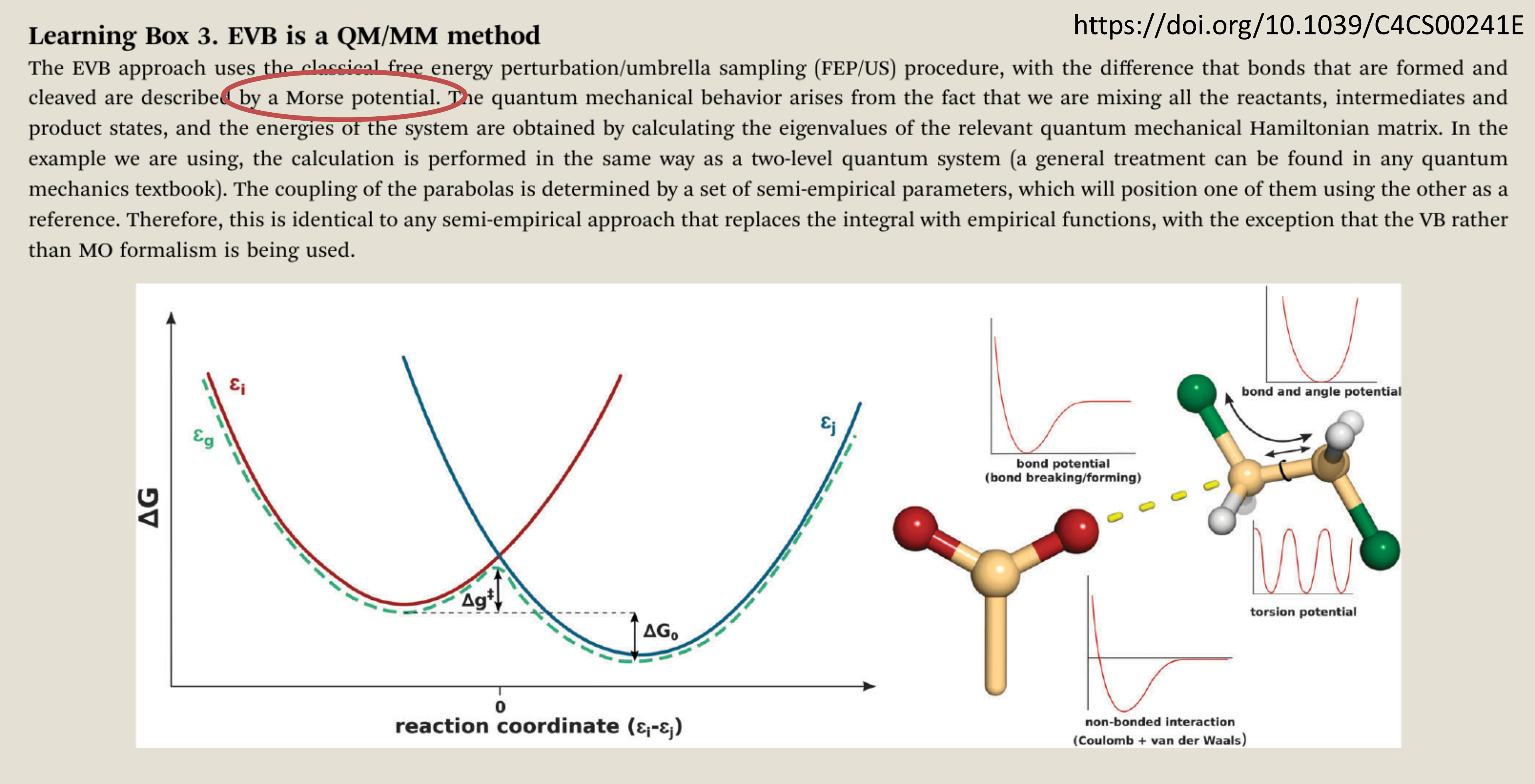


Fig. 3 A comparison of schematic EVB free energy profiles for the reaction catalysed by haloalkane dehalogenase (DhlA, Fig. 2), obtained in the gas-phase (ε_g), aqueous solution (ε_w) and in the enzyme active site (ε_p).

https://doi.org/10.1039/C4CS00241E

Technical Notes



https://doi.org/10.1039/C4CS00241E

Learning Box 3. EVB is a QM/MM method

The EVB approach uses the classical free energy perturbation/umbrella sampling (FEP/US) procedure, with the difference that bonds that are formed and cleaved are described by a Morse potential. The quantum mechanical behavior arises from the fact that we are mixing all the reactants, intermediates and product states, and the energies of the system are obtained by calculating the eigenvalues of the relevant quantum mechanical Hamiltonian matrix. In the example we are using, the calculation is performed in the same way as a two-level quantum system (a general treatment can be found in any quantum mechanics textbook). The coupling of the parabolas is determined by a set of semi-empirical parameters, which will position one of them using the other as a reference. Therefore, this is identical to any semi-empirical approach that replaces the integral with empirical functions, with the exception that the VB rather than MO formalism is being used.

- ➢ Breaking and forming bonds must be described by a Morse potential? Why?
- ➢ LJ interaction for atoms involved in breaking and forming bonds must be softened in the repulsive region.

EVB Pros & Cons

Pros:

- it is computationally cheap in comparison to hybrid QM/MM method (no expensive QM calculations are preformed)
- it can be very accurate (all errors related to models used can be suppressed by tuned empirical parameters)
- Egap as reaction coordinate (non-local but simple to calculate)
- Simulations on mapping potential simplify the EVB parametrization and prediction.

Cons:

- it cannot be used if mechanism of reaction in reference system and enzyme differs (empirical parameters are not then transferable)
- it cannot go beyond valence state setup (it cannot reveal new reaction pathways)

Questions

- Is EVB still used in research?

Further Readings

Mones, L.; Kulhanek, P.; Florian, J.; Simon, I.; Fuxreiter, M. **Probing the Two-Metal Ion Mechanism in the Restriction Endonuclease BamHI**. *Biochemistry* **2007**, *46* (50), 14514–14523. https://doi.org/10.1021/bi701630s.

Mones, L.; Kulhanek, P.; Simon, I.; Laio, A.; Fuxreiter, M. **The Energy Gap as a Universal Reaction Coordinate for the Simulation of Chemical Reactions**. *J. Phys. Chem. B* **2009**, *113* (22), 7867–7873. https://doi.org/10.1021/jp9000576.

Shurki, A.; Derat, E.; Barrozo, A.; Kamerlin, S. C. L. **How Valence Bond Theory Can Help You Understand Your (Bio)Chemical Reaction.** *Chem. Soc. Rev.* **2015**, *44* (5), 1037–1052. https://doi.org/10.1039/C4CS00241E.