# QM/MM Methods, Energy Minimization

Lubomír Rulíšek, Martin Srnec

*Institute of Organic Chemistry and Biochemistry AS CR*

*J. Heyrovský Institute of Physical Chemistry AS CR, Prague, Czech Republic*

# Outline
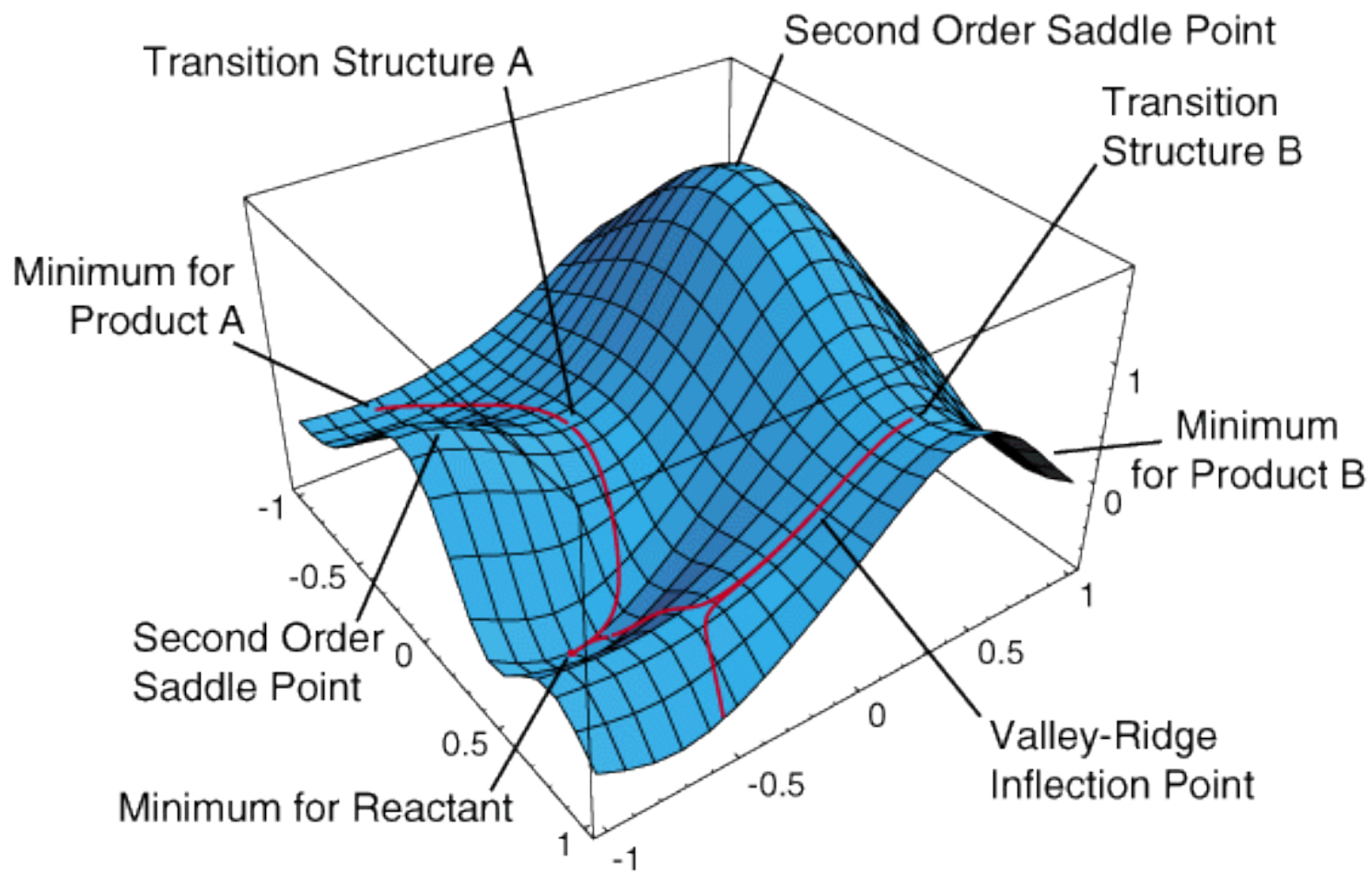
**Energy Minimization** ( ≈ *geometry optimization*)
- Steepest descent
- Conjugate gradient
- "Pseudo"-second-order methods

**QM/MM Methods**
- Motivation, Basic Concept
- Subtractive, Additive Schemes
- Mechanical, Electronic, Polarized Embedding
- Link Atoms
- Implementation (Example)

# Potential energy surface (hypersurface)

# Energy minimum

$f(x_1, x_2, \ldots, x_n)$

At a minimum $\dfrac{\partial f}{\partial x_i} = 0; \quad \dfrac{\partial^2 f}{\partial x_i^2} > 0$

Choice of variables:     MM – almost exclusively cartesians
                         QM – (redundant) internals
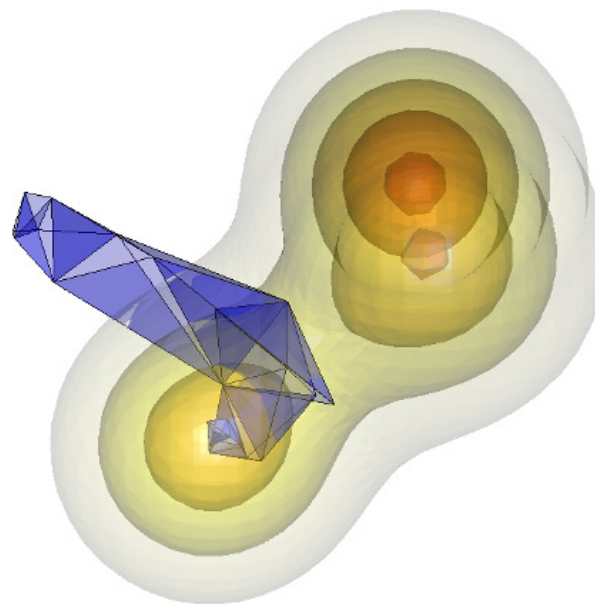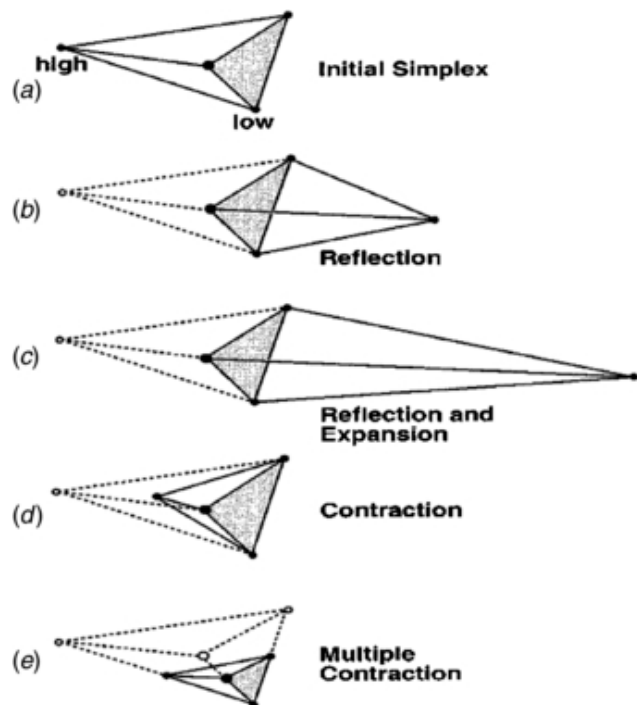
**Derivatives:** analytical, numerical

# Non-derivative Minimization Methods
(seldom used, not very efficient)

The Simplex Method

Simplex… geometrical figure with (M+1) connected vertices; "amoeba"

# The Sequential Univariate Method
(seldom used)

For each new coordinate, two new structures are generated ($x_i + \delta x_i$, $x_i + 2\delta x_i$)

Parabola is fitted through the three points

Minimum found and the coordinate changed to the position of the minimum

Next coordinate is treated

# Derivative Minimization Methods

Gradient: $\mathbf{grad}\, V(x_1, \ldots xn) = \boldsymbol{g} = \left(\dfrac{\partial V}{\partial x_1}, \ldots, \dfrac{\partial V}{\partial x_n}\right)$

Hessian (force constant matrix): *N* x *N* matrix

elements $\dfrac{\partial^2 V}{\partial x_i \partial x_j}$

# First-order Minimisation methods

**Steepest Descent**

"walking down" in the direction of the unit vector

$$s_k = -\frac{g_k}{|g_k|} \; ; \text{ line search or } x_{k+1} = x_k + \lambda_k x_k$$

**Conjugate Gradient**
(Fletcher-Reeves)

$$v_k = -g_k + \gamma_k v_{k-1}$$

$$\gamma_k = \frac{g_k \cdot g_k}{g_{k-1} \cdot g_{k-1}}$$

ÚOCHB AV ČR  Computer Modeling of Chemical Reactions and Enzyme Catalysis

# Second Derivative Methods: The Newton-Raphson

For purely quadratic multivariate function
$$x_{min} = x_k - g(x_k)H^{-1}(x_k)$$

Hessian (**H**) must be positive definite

## Quasi Newton-Raphson

$$x_{k+1} = x_k - A_k g(x_k) \quad \text{with} \quad \lim_{k \to \infty} A_k = H^{-1}$$

Hessian update: Davidon-Fletcher-Powell (DFP), Broyden-Fletcher-Goldfarb-Shanno (BFGS), …

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{\mathbf{q}_k \mathbf{q}_k{}^T}{\mathbf{q}_k{}^T \mathbf{p}_k} - \frac{\mathbf{H}_k \mathbf{p}_k \mathbf{p}_k{}^T \mathbf{H}_k}{\mathbf{p}_k{}^T \mathbf{H}_k \mathbf{p}_k}$$

# Quantum AND Classical World

ÚOCHB AV ČR
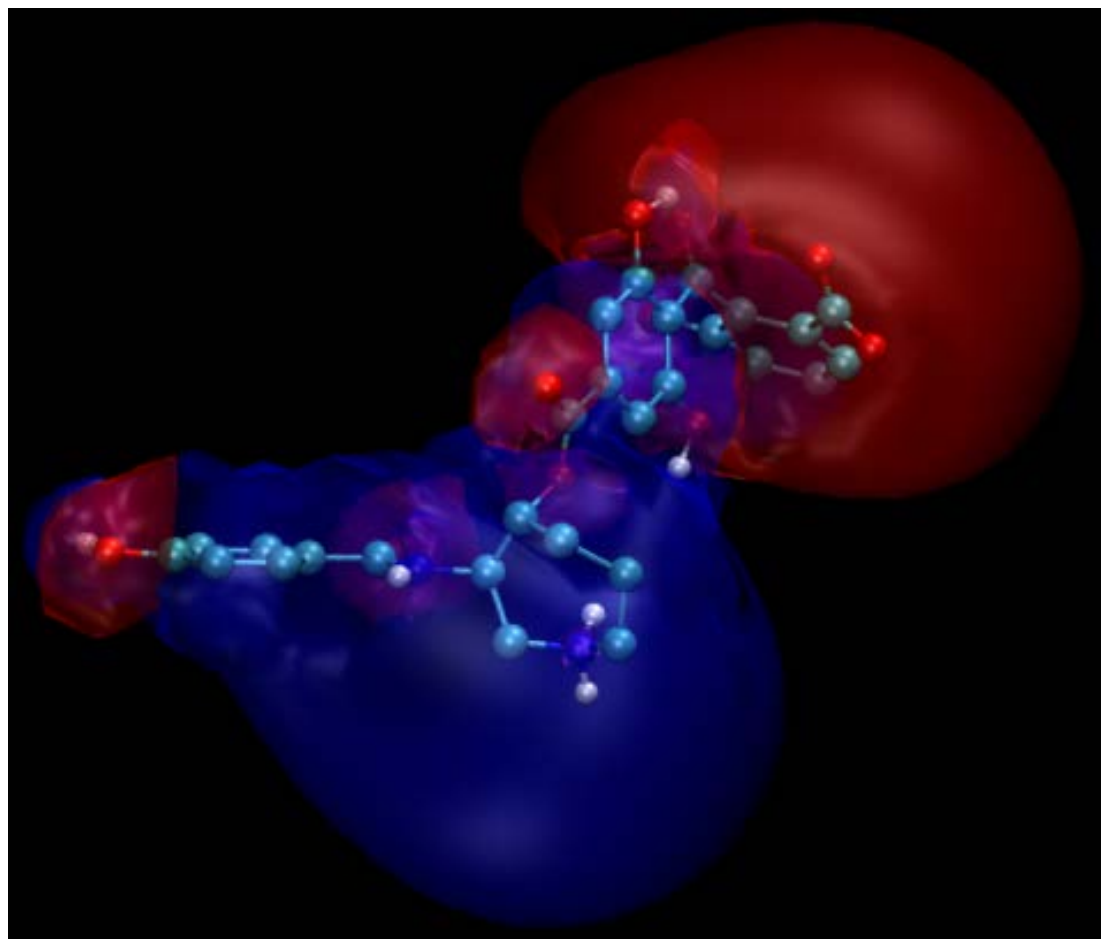
# QM/MM: Why not to couple these two?

**History:** Warshel, A.; Levitt, M.: Theoretical Studies of Enzymatic Reactions: Dielectric Electrostatic and Steric stabilisation of the carbonium ion in the reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227.

Warshel, A.; Karplus, M. Calculation of ground and excited state potential surfaces of conjugated molecules I: formulation and parametrization. *J. Am. Chem. Soc.* **1972**, *94*, 5612.



The energy of the total system is given as :

$$E_{tot} = E_{QM} + E_{MM} + E_{coupling}$$
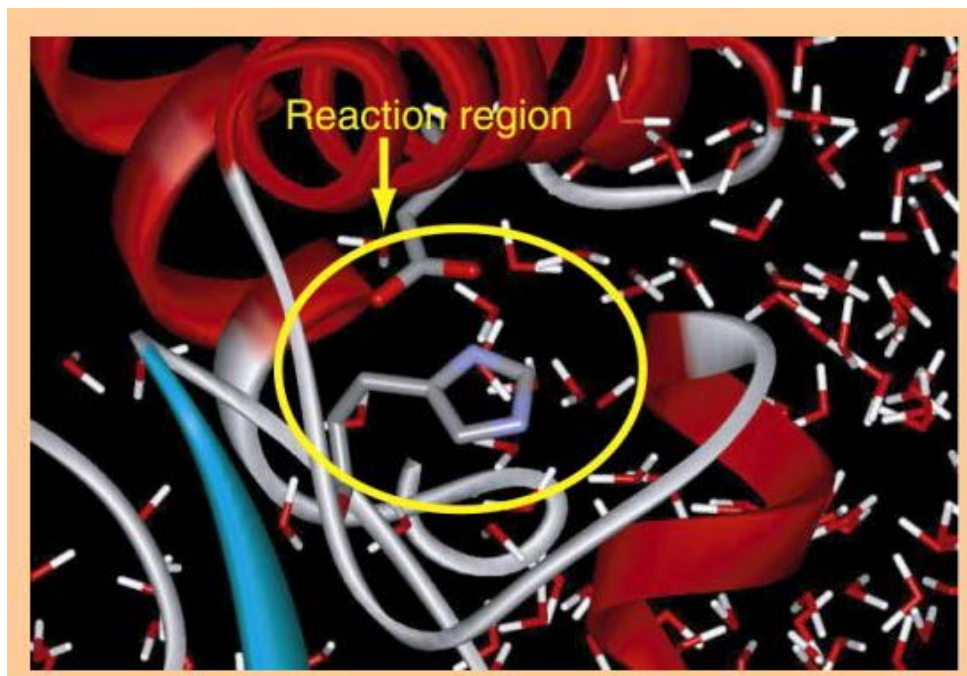
System is divided into 2 parts:

1/ Reaction region

described by QM methods

2/ The rest (spectator, bulk)

described by MM methods

**"Modern QM/MM":**

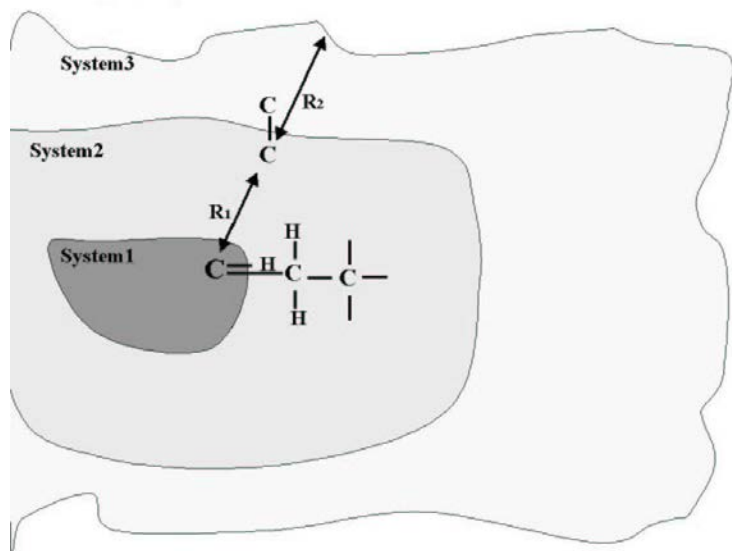Morokuma, Friesner, Thiel, Ryde, Harvey, Mulholland, Gao, Merz, Field, …

# QM/MM (Example: ONIOM-like approach)



**System 1 (S1)** is treated at the QM level. It is truncated using link atoms. The positions of the link atoms are linearly dependent on the cooresponding heayy atoms (hydrogen link approach)

**System 2+3** Treated with MM methods.

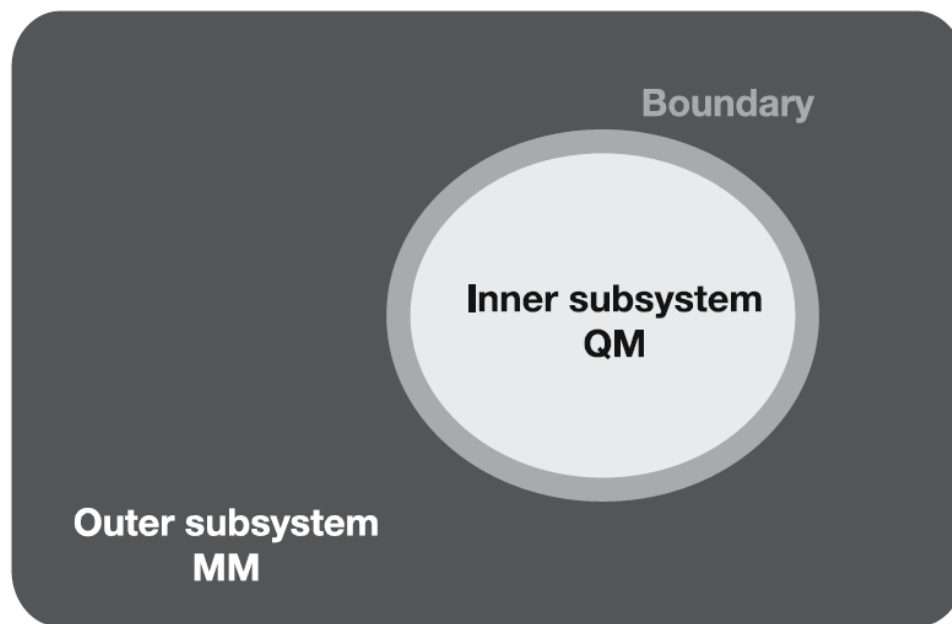The energy of the total system is given as :

$$E(QM/MM)=E(QM,S1) + E(MM,S3) - E(MM,S1),$$

where

$E(QM,S1)$ … QM energy of S1 in the field of point charges
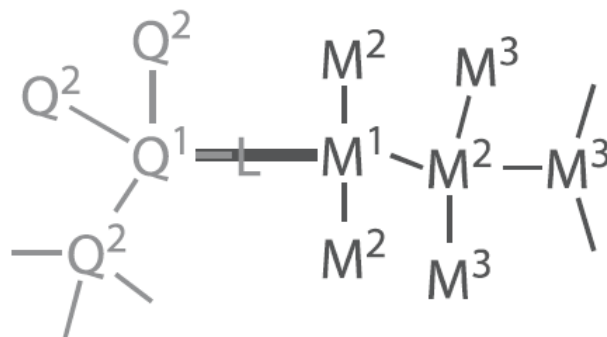
$E(MM,S3)$ … MM energy of S3 with charges of S1 set to zero

$E(MM,S1)$ … MM energy of S1 with charges of S1 set to zero

**Fig. 1** Partitioning of the entire system $\mathbb{S}$ into inner ($\mathbb{I}$) and outer ($\mathbb{O}$) subsystems

## MM energy

$$E_{\text{MM}} = \sum_{\text{bonds}} k_{\text{b}} \left(d - d_0\right)^2 + \sum_{\text{angles}} k_{\theta} \left(\theta - \theta_0\right)^2 + \sum_{\text{dihedrals}} k_{\phi} \left[1 + \cos\left(n\phi + \delta\right)\right]$$

$$+ \sum_{\substack{\text{non-bonded} \\ \text{pairs AB}}} \left\{ \varepsilon_{\text{AB}} \left[ \left(\frac{\sigma_{\text{AB}}}{r_{\text{AB}}}\right)^{12} - \left(\frac{\sigma_{\text{AB}}}{r_{\text{AB}}}\right)^{6} \right] + \frac{1}{4\pi\varepsilon_0} \frac{q_{\text{A}} q_{\text{B}}}{r_{\text{AB}}} \right\} , \qquad (1)$$

## Subtractive Scheme

$$E_{\text{QM/MM}}(\mathbb{S}) = E_{\text{MM}}(\mathbb{S}) + E_{\text{QM}}(\mathbb{I} + \mathbb{L}) - E_{\text{MM}}(\mathbb{I} + \mathbb{L})$$

## Additive Scheme

$$E_{\text{QM/MM}}(\mathbb{S}) = E_{\text{MM}}(\mathbb{O}) + E_{\text{QM}}(\mathbb{I} + \mathbb{L}) + E_{\text{QM-MM}}(\mathbb{I}, \mathbb{O}) .$$

*Mechanical, electronic embedding;*
Morokuma's **IMOMM, IMOMO, ONIOM**

ÚOCHB AV ČR

# Mechanical Embedding

QM–MM electrostatic interaction is treated MM level

Charge model of the MM method used – typically rigid atomic point charges, but other approaches, e.g., bond dipoles, are also possible – applied to the QM region

**major disadvantages and limitations:**
(i)     the charges in the outer region do not interact with the QM density

(ii) as the charge distribution in the QM region changes, for instance during a reaction, the charge model needs to be updated, may lead to discontinuities in the potential-energy surface

(iii) The derivation of, e.g., MM point charges for the inner region is often not trivial

# Electrostatic Embedding

$$\hat{H}^{\text{el}}_{\text{QM-MM}} = -\sum_{i}^{\text{electrons}} \sum_{M \in \mathbb{O}} \frac{q_M}{|r_i - R_M|} + \sum_{\alpha \in \mathbb{I}+\mathbb{L}} \sum_{M \in \mathbb{O}} \frac{q_M Z_\alpha}{|R_\alpha - R_M|}$$

*Special care is required at the QM–MM boundary, where the MM charges are placed in immediate proximity to the QMelectron density, which can lead to overpolarization*
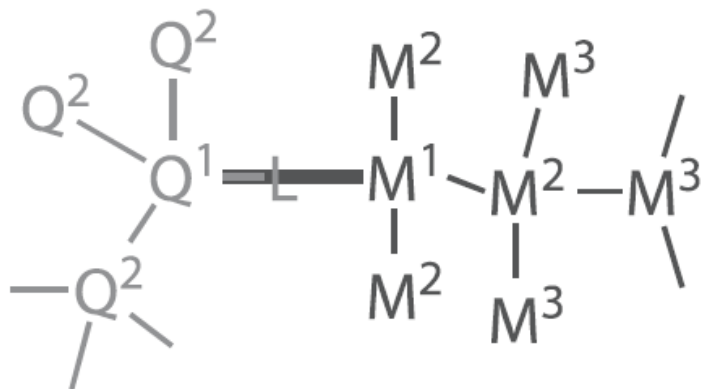
# QM/MM simulation protocol
## (general, protein setup)

➢ Start with a crystal (NMR) structure

➢ Add hydrogens, missing atoms (His protonation, Cys-Cys, Asn/Gln)

➢ parameter files for unknown parts of the protein

➢ Run MM minimization for the system *in vacuo* with atoms resolved in c.s. as fixed or restrained

➢ Add *solvation shell, sphere* or solvation box (PBC) to the whole system

➢ Run minimization of the whole solvated system (c.s. atoms fixed)

➢ Run 300-ps of simulated annealing protocol (c.s. atoms fixed) (equilibration)

➢ Run the final MM minimization

➢ Define QM region, surrounding region (S2) that is relaxed in MM

➢ Start QM/MM production calculations

# QM/MM simulation protocol
## (technical details)

- Evaluate the wave function of System 1 in the field of the pointcharges of System 2 and 3 (**QM**).

- Evaluate the forces of system1 including the electrostatics of system2 and 3 (**QM**).

- Evaluate the forces of system1 and 2 with any electrostatic interactions (**MM**).

- Add the QM and MM forces to obtain the QM/MM forces on the atoms of System 1 (FixForce).

- Relax the atoms of System 1 using the QM/MM forces (**QM**).

- Use the relaxed coordinates of System 1 to construct the new coordinate representation of System 1 (FixCoord1).

- Insert the charges (Mulliken, ESP,...) of the atoms of System 1 obtained from the QM calculation into the MM representation (FixCharge).

  - Relax the atoms of System 2 with System 1 fixed (**MM**).

  - Insert the new coordinates of the pointcharges of the atoms in System 2 and 3 to be used in the next QM calculation. (FixCoord2).

  - Calculate the QM energy of System 1 (**QM**).

  - Calculate the MM energy of Systems 1 and 2 (**MM**).

  - Add the energies appropriately (FixEnergy)

  - Check for convergence. If not converged, then go to step 2 above.

# QM/MM (ComQum) +/-

+ QM/MM method is the natural way to include the environment (electrostatics, protein restraints on the active site of interest) into the 'accurate' QM calculations (at minimum cost) => 'good' structures

+ QM (~200 atoms) *vs* MM (~1000 atoms) system size: full MM minimization of S2 for each QM step

- Poor representation of electrostatics in the boundary region (polarizable force fields?, charge-fitting procedures for S2 (< 5 Å) ?, multipole expansions)

- QM/MM sampling might be necessaryto get a satisfactory accurate energetics? => QTCP method

(T. H. Rod & U. Ryde (2005) Phys. Rev. Lett., 94, 138302)

# Other QM/MM coupling schemes

**Quantum Refinement** (QM/MM with structure refinement)

improving locally crystal structures

$$E_{QM/MM/X\text{-ray}} = E_{QM/MM} + w_a E_{X\text{-ray}}$$

## QM/MM-EXAFS

more accurate information about the local structure

$$E_{QM/MM/EXAFS} = w_{QM/MM} E_{QM/MM} + w_{EXAFS} E_{EXAFS}$$

## QM/MM-NMR

(group of Prof. Ryde, Lund University)

ÚOCHB AV ČR   Computer Modeling of Chemical Reactions and Enzyme Catalysis

# Reductive Cleavage of the O-O Bond in Multicopper Oxidases: A QM/MM and QM Study
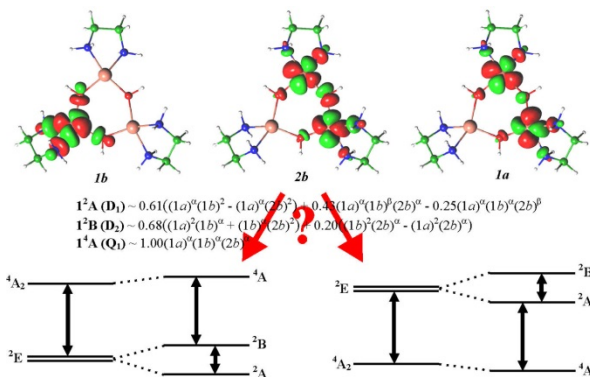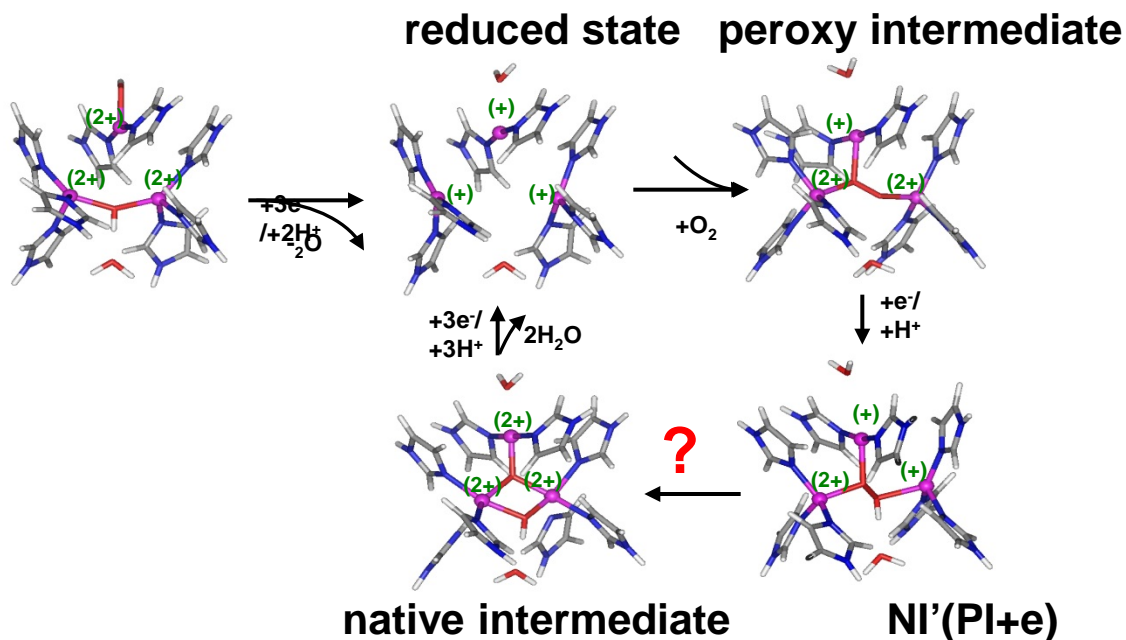
**Martin Srnec, Ulf Ryde, Lubomír Rulíšek**

*Institute of Organic Chemistry and Biochemistry and Gilead Sciences Research Center & IOCB, Prague, Czech Republic*
*Department of Theoretical Chemistry, Lund University, Sweden*

ÚOCHB AV ČR

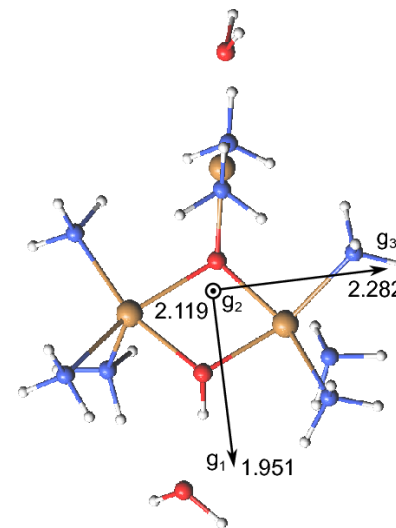Computer Modeling of Chemical Reactions and Enzyme Catalysis

# **Multicopper Oxidases** couple four 1$e^-$ oxidations of substrates with 4$e^-$ reduction of $O_2$ to $H_2O$ (plant and fungal laccases, ascorbate oxidase, bilirubin oxidase, phenoxazinone synthase, CotA protein, and highly specific enzymatic oxidants of metal ions, such as ceruloplasmin, Fet3p, CueO and MnxG…)
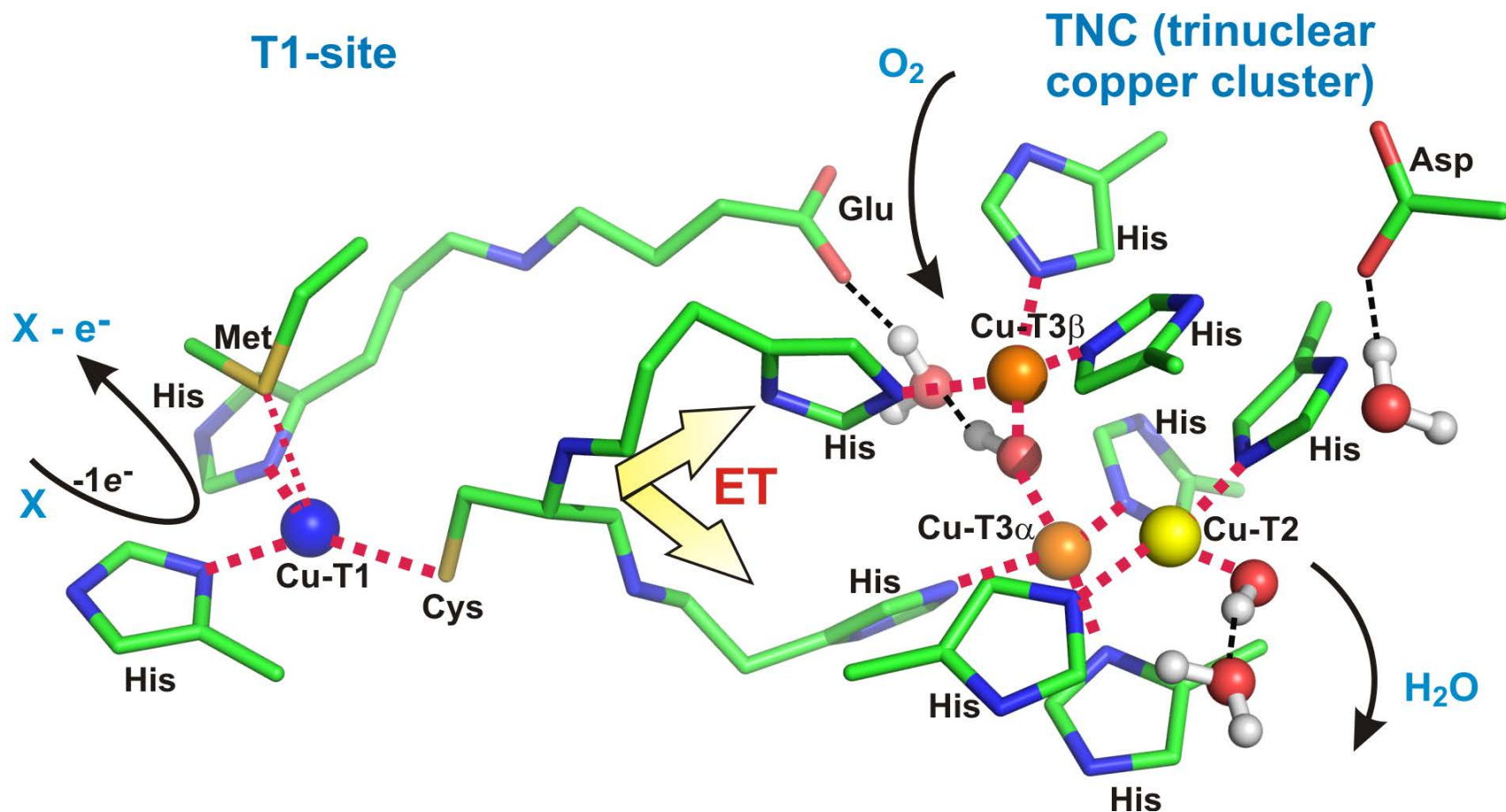
# MCO Catalytic Cycle



reduced state    peroxy intermediate

native intermediate    NI'(PI+e)

**Selected references:**
Ed Solomon and coworkers
*Chem. Rev.* 1996, *96*, 2563
*J. Am. Chem. Soc.* 1996, *118*, 3202;
2002, *124*, 6180; 2005, *127*, 13680;
2007, *129*, 13127; 2007, *129*, 13118;
*Dalton Trans.* 2008, 3921

Rulíšek, Ryde, Solomon *et al.*
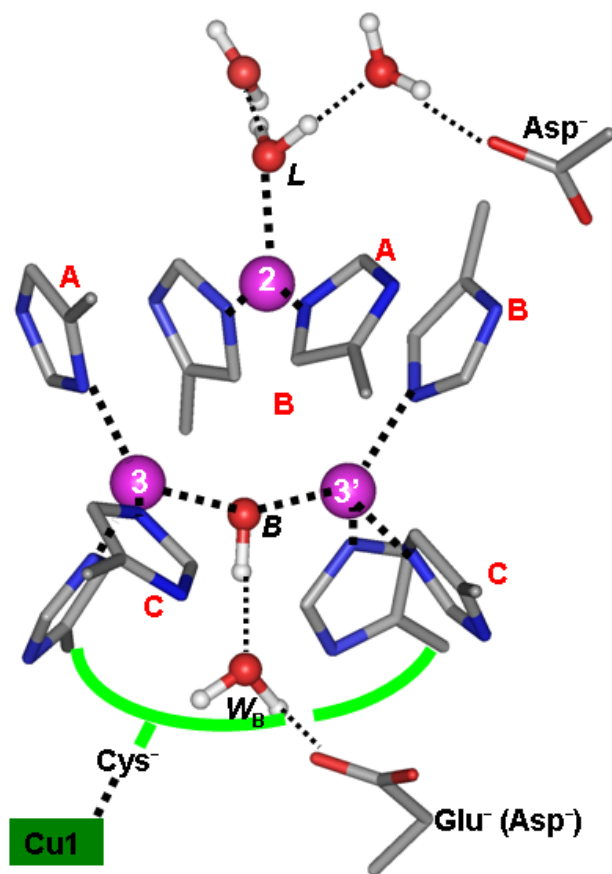*Inorg. Chem.* 2005, *44*, 5612; 2006,
*45*, 11051; *J. Am. Chem. Soc.* 2007,
*129*, 726; *J. Phys. Chem. B*
2010, *114*, 7692.

**Correlating quantum chemistry with spectroscopy (absorption, EXAFS, EPR)**
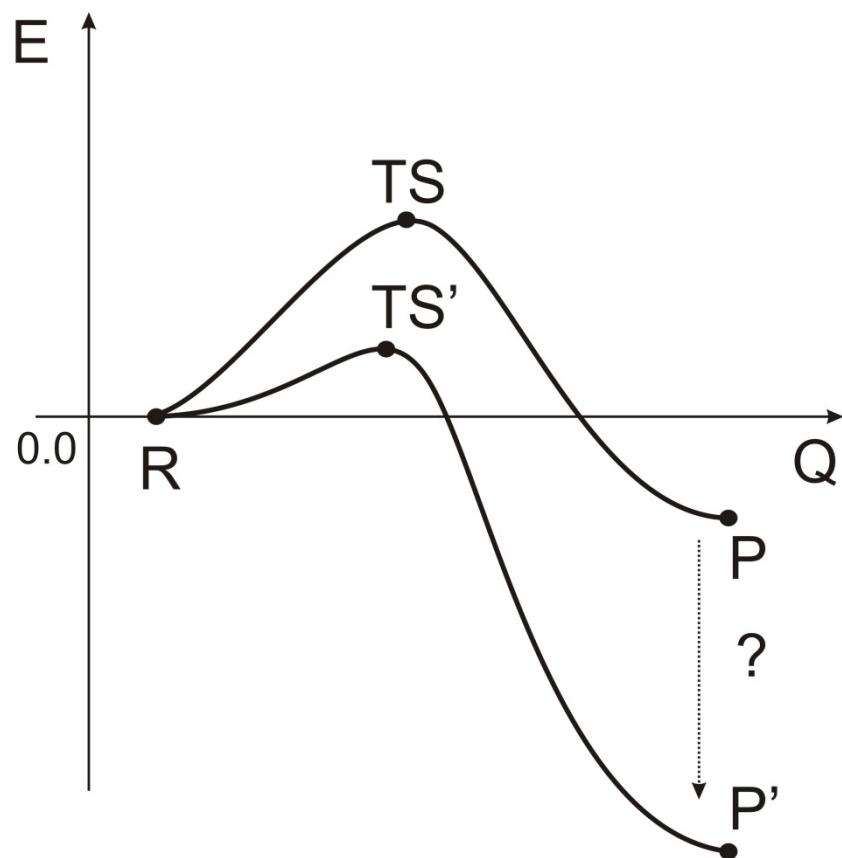
# Quantum System



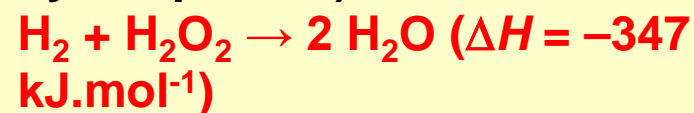**CuT2T3 (TNC) structure: partially reduced copper ions**

# Multi-copper oxidases – challenging systems for quantum (theoretical) chemistry
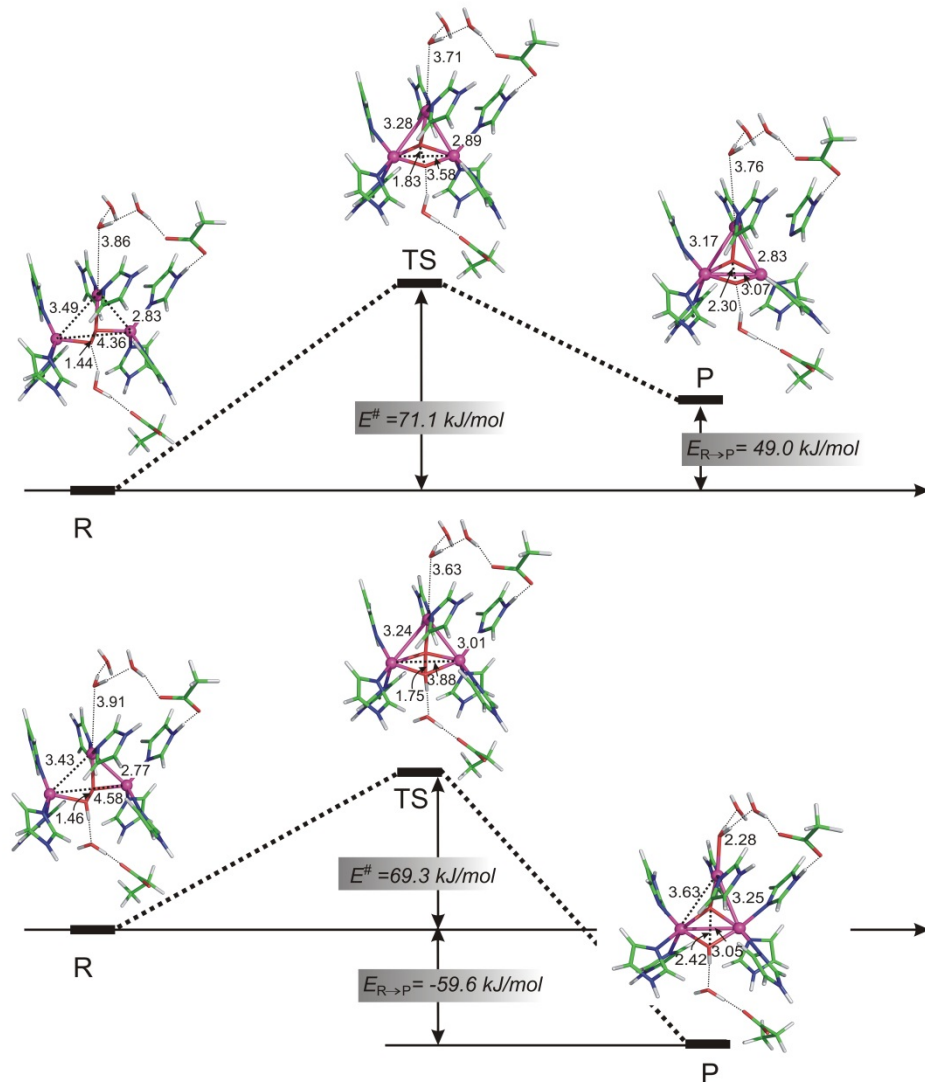


1/ **Large exothermicity of the 'parent reaction' (must be somehow buffered by the protein)**

$H_2 + H_2O_2 \rightarrow 2\ H_2O$ ($\Delta H = -347$ kJ.mol$^{-1}$)

2/ **Spin states (spin frustration)**

3/ *In vacuo* **instability of many states (formal charge of the 'small model' cluster may vary between +2 and +5)**

# Reaction pathway for O₂ cleavage



| Cu-T2 ligand | Central ligand | Charge of System 1 | $\Delta G^{\neq}_{QM/MM}$ |
|---|---|---|---|
| $H_2O$ | $O_2H^-$ | +1 | **66.8** |
| | | | **67.0** |
| $OH^-$ | $O_2H^-$ | 0 | – |
| | | | – |
| – | $O_2H^-$ | +1 | 79.4 |
| | | | 73.7 |
| $H_2O$ | $O_2^{2-}$ | 0 | 72.4 |
| | | | 71.3 |
| $OH^-$ | $O_2^{2-}$ | -1 | **67.0** |
| | | | **63.5** |
| – | $O_2^{2-}$ | 0 | 78.4 |
| | | | 77.5 |

ÚOCHB AV ČR    Computer Modeling of Chemical Reactions and Enzyme Catalysis