QM/MM in MOLCAS

Nicolas Ferré and RL

Université de Provence, Marseille, France

and

Lund University, Lund
Introduction: Luciferin chemistry
QM/MM principles
Recent QM/MM developments
The ESPF method
Current Model developments
Luciferin chemistry: results
Conclusions
Introduction

QM/MM methods is today used as a tool to study chemistry at action in large systems. I will give you here a brief example of what can be done.

This workshop will not include any exercises using QM/MM, however, a workshop addressing the practical parts of the technique is scheduled for the spring 2010 in Lund.
Luciferin chemistry by MOLCAS

Luciferin + O₂ → O=C–C–R₁

Spontaneous

CO₂ + O=C–C–R₁

Oxyluciferin

Light
Neutral or anion form of keto-L?

Neutral: $r$(CO) = 1.324 Å, $E_a = 23.4$ kcal/mol, concerted dissociation. $T_v = 3.32$ eV

Anion: $r$(CO) = 1.335 Å, $E_a = 7.3$ kcal/mol, biradical dissociation. $T_v = 2.54$ eV
Neutral or anion form of keto-L?

Oxy-LH2(-1): r(CO)=1.335 Å, $E_a = 7.3$ kcal/mol, biradical dissociation. $T_v = 2.54$ eV

- Activation energy consistent with a biochemical process.
- A long r(CO) bond closer to that of excited formaldehyde (1.362 Å)
- Dissociation process has the expected character
- Emission in the right energy range
- Note the $sp^2$ vs. $sp^3$ hybridization of the oxygen anchor carbon at the TS for the concerted and biradical mechanism, respectively.
Polarization in the micro environment

TD-DFT calculated $T_v$ values in eV.

keto-trans \hspace{1cm} \text{keto-trans} + \text{CH}_2\text{Cl}_2 \hspace{1cm} \text{keto-trans} + \text{H}_2\text{O} \hspace{1cm} \text{keto-trans}(-1)

\begin{align*}
3.04 & (2.99, 2.87) \\
3.19 & \\
3.27 & \\
3.32 & (3.35)
\end{align*}
Luciferin-Luciferase Complex: CASPT2/Tinker QM/MM
What’s the origin of the variation of the bioluminescence colour?

OxyLuciferin

Natural Firefly luciferase

Mutant

Wild-type (closed form)

S286N (open form)
What’s the origin of the variation of the bioluminescence colour?

• The tight pocket of the wild-type luciferase should not allow too much structural relaxation of the oxyluciferin before it emits light and decade to its ground state.

• Most of the chemical energy is then transformed into light with a short wavelength.
What’s the origin of the variation of the bioluminescence colour?

• The loose pocket of the mutant, on the other hand, should let the oxyluciferin structure to relax a bit before the decay on the ground state.

• Part of the chemical energy is “wasted”, and the emitted light wavelength is longer.
The QM/MM principles (1)

\[
E = \langle \psi | \hat{H} | \psi \rangle = \langle \psi | \hat{H}_{QM} + \hat{H}_{MM} + \hat{H}_{QM/MM} | \psi \rangle
\]

- \[ \langle \psi | \hat{H}_{QM} | \psi \rangle = E_{QM} \]: total energy of the isolated QM subsystem
- First approximation: the major part of the interactions are independent of the electronic coordinates
  \[ \langle \psi | \hat{H}_{MM} | \psi \rangle = E_{MM} \langle \psi | \psi \rangle = E_{MM} \]: molecular mechanics force-field energy of the isolated MM part
- Second approximation 'a la MM': only the QM/MM electrostatic interactions depend on the electronic coordinates
  \[ \langle \psi | \hat{H}_{QM/MM} | \psi \rangle = \langle \psi | \hat{H}_{QM/MM}^{\text{elect}} | \psi \rangle + E_{other}^{QM/MM} \]
The QM/MM principles (2)

\[ E = \langle \psi | \hat{H}_{QM} + \hat{H}_{QM/MM}^{\text{elect}} | \psi \rangle + E_{QM}^{\text{nuc}} + E_{QM/MM}^{\text{other}} + E_{MM} \]

- The wavefunction is (almost always) polarized by its electrostatic surroundings.
- Usually, the one-electron effective hamiltonian is modified

\[ h_{\mu \nu}^{QM/MM} = \langle \chi_{\mu} | \sum_a \frac{q_a}{r_a} | \chi_{\nu} \rangle \]

- The QM/MM and MM 'classical' interactions usually share the same functional form

\[ E_{QM/MM}^{\text{other}} = E_{\text{bonded}} + E_{\text{non-bonded}} \]
\[ = E(\text{stretch}) + E(\text{bend}) + E(\text{torsion}) + \cdots \]
\[ + E(\text{van der Waals}) \]
\[ + E(\text{QM nuclei} - \text{MM electrostatic potential}) \]
The QM/MM principles (3)

What about the MM part?

- Any MM forcefield is highly parametrized, for reproducing:
  - experimental data
  - QM results

Caution for a QM/MM hybrid scheme! One must be careful when using the standard parameters.
The QM/MM principles (4)

QM/MM frontier?

- Nothing if no bonds
- Link atom or link group
- Effective atomic or group pseudo-potentials
- Frozen orbitals

In principle, this requires (re-)parametrization of the MM force-field
Recent QM/MM developments

1. The ESPF scheme for sophisticated QM/MM electrostatics
2. The integrated ESPF QM/MM MD algorithm
3. The ESPF QM/MM scheme including a polarizable force-field

All these developments are now part of the Molcas package (version 7.0 and above), including the coupling to a modified Tinker program
The ESPF method

- Direct method:
  \[ V_{\mu \nu}^{MM} = \left\langle \mu \left| \frac{q_{MM}^{MM}}{r_{MM}} \right| \nu \right\rangle \propto N_{MM} \]
  
  no multipoles (MM2 ...)
  
  \( E^{\text{elec}} (A = \text{QM}; \ B = \text{MM}) \neq E^{\text{elec}} (A = \text{MM}; \ B = \text{QM}) \)
  
  all the MM multipoles usually polarize the wavefunction, the closest ones may overpolarize !

- Approximate method: multipolar atomic operators \( Q^A \) fitted to the electrostatic potential (ESPF)

\[
\Delta H_{\mu \nu} = \sum_A Q_{\mu \nu}^A \phi_{MM}^A = \sum_A \sum_K \left( (T^\dagger T)^{-1} T^\dagger \right)^{AK} V_{\mu \nu}^K \phi_{MM}^A
\]

\[ \phi_{MM}^A = \sum_i^{MM} \frac{q_i}{R_{iA}} + \cdots + \text{PBC} + \cdots \]

\[ V_{\mu \nu}^K = \left\langle \mu \left| \frac{1}{r_K} \right| \nu \right\rangle \propto N_K \ll N_{MM} \]

---

The ESPF method features

- Geometries / relative energies in good agreement with the 'direct' method
  \[ \Delta \Delta E(S_0 \rightarrow S_1 \text{ or } S_2) < 1 \text{ kcal mol}^{-1} \]
- Atomic multipoles: \[ q^A = \sum_{\mu \nu} P_{\mu \nu} Q_{\mu \nu}^A \]
- QM/MM electrostatic interaction energy can be partitioned: \[ E_{QM/MM}^A = q^A \times V^A \Rightarrow \text{unicity of the QM/MM electrostatic energy} \]
- Easy implementation of a QM/MM MD algorithm
Problems → improvements

- Usual fitting problems (overdetermination) → SVD decomposition. But complicates the gradient formulation.
- The fit depends heavily on the grid around the QM subsystem → constraints. But complicates the gradient formulation.
- Which density matrix must be used when SA-CASSCF wavefunctions are computed?
- A lot of other problems: still to be analyzed.
Current model developments

- On-the-fly CASPT2 scaling of the CASSCF gradient for QM/MM MD trajectories
- Polarizable MM: $\phi^A_{\text{MM}}$ depends on the QM electronic state $\rightarrow$ induced dipoles on the MM side
- (On-the-fly) QM-parametrized MM/MD trajectories $\rightarrow$ statistical sampling
Luciferin-Luciferase Complex: CASPT2/Tinker QM/MM

QM/MM model: Solvated protein, 10329 atoms (626 water molecules), QM system: OxyLH2(-1)
Results

- The same substrate in native and mutated enzyme exhibits the same emission spectra. That is the observed red-shift is not due to a structural difference of the active site/substrate upon mutation of the enzyme.
Conclusions

QM/MM for photochemistry available at

- http://www.teokem.lu.se/molcas