UNIVERSITY OF NORTH TEXAS
CASCaM
CENTER FOR ADVANCED SCIENTIFIC COMPUTING AND MODELING

UNT
EST. 1890

QM/MM Calculations

Alice R. Walker
Erik Vázquez Montelongo

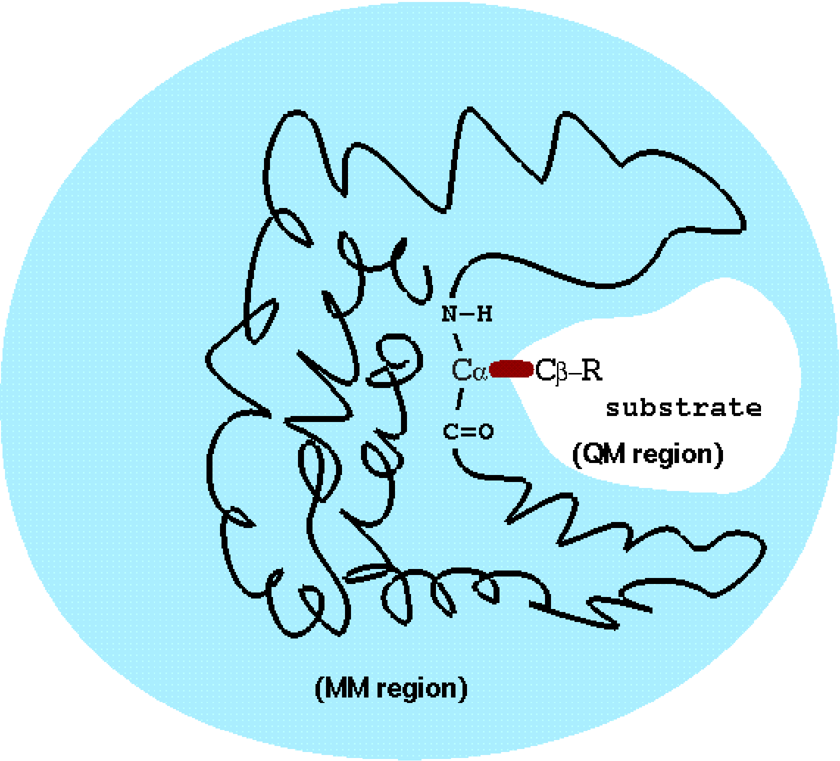
07/19/2017

Outline

- Introduction
- QM/MM energy function
- Boundary approximations
- Long range electrostatics
- QM/MM optimization
- Reaction path calculations
- Chain–of–replica methods
- FEP

Introduction

- Molecular Mechanics (MM): thousands of atoms, no reactions.
- Quantum Mechanics (QM): electronic structure, 1000 atoms at most.
- Solution: Combine both (QM/MM, ONIOM).
- Investigate reaction mechanism, analyze sources of catalysis, compare with experiments.



QM/MM energy function

- The total energy of a QM/MM system can be written as:

$$E_{Total} = E_{QM}(QM) + E_{MM}(MM) + E_{QM/MM}(QM/MM)$$

- where the first and second terms of Eq. 1 correspond to the QM and MM energies of the respective QM and MM subsystems, and the last term represents the interaction according to:

$$E_{QM/MM}(QM/MM) = E_{Coulomb}(QM/MM) + E_{vdW}(QM/MM) \\ + E_{MM-bonded}(QM/MM)$$

- the second and third terms of Eq. 2 are calculated with the force field (FF).

Additive QM/MM

- The first term is calculated together with EQM(QM) from an effective Hamiltonian to polarize the wavefunction,

$$E_{QM}(QM) + E_{Coulomb}(QM/MM) = \left\langle \psi \,|\, H_{eff} \,|\, \psi \right\rangle$$

$$\begin{aligned} H_{eff} = \frac{1}{2}\sum_{i}^{N}\nabla_{i}^{2} + \sum_{i \neq j}^{N}\frac{1}{r_{ij}} - \sum_{i}^{N}\sum_{\alpha \in QM}\frac{Z_{\alpha}}{r_{\alpha i}} + \sum_{\alpha_1 \neq \alpha_2 \in QM}\frac{Z_{\alpha}Z_{\beta}}{r_{\alpha 1}r_{\beta 2}} \\ - \sum_{i}^{N}\sum_{\beta \in MM}\frac{q_{B}}{r_{iB}} + \sum_{\alpha_1 \in QM, B \in MM}\frac{Z_{\alpha}q_{B}}{r_{\alpha}r_{B}} \end{aligned}$$

- Hamiltonian can be semiempirical or *ab initio*

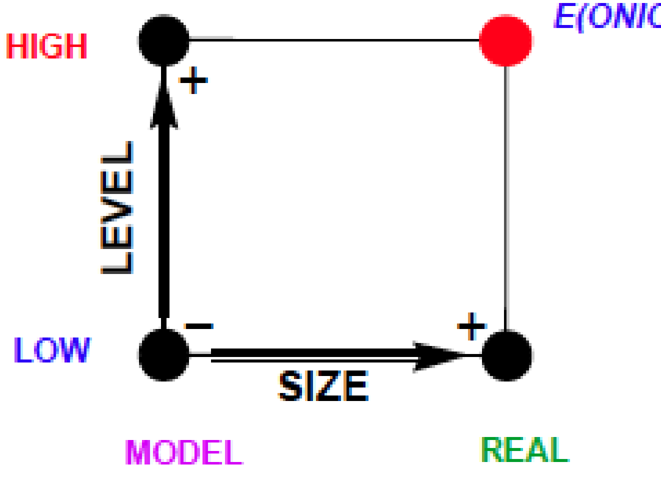
Zhang, Liu, Yang, JCP, **112**, 3483, (2000).

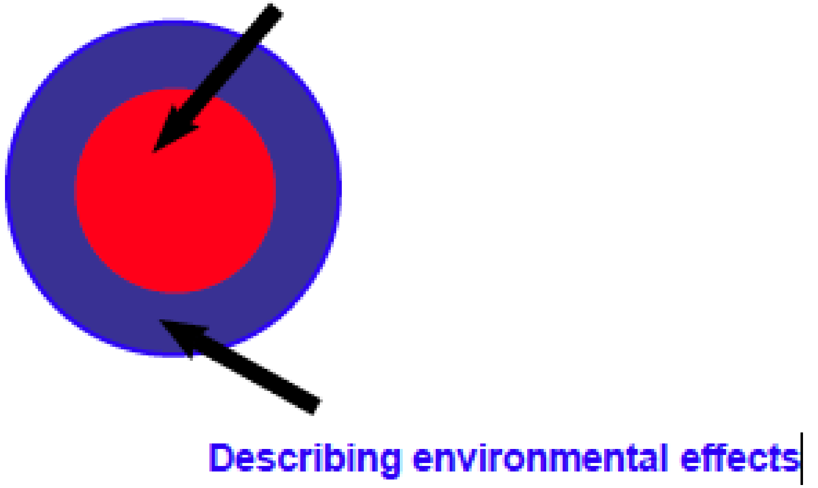
Subtractive QM/MM

ONIOM method

E(HIGH,REAL) ≈ E(ONIOM)

= E(LOW,MODEL) + [E(LOW,REAL)-E(LOW,MODEL)] + [E(HIGH,MODEL)-E(LOW,MODEL)]



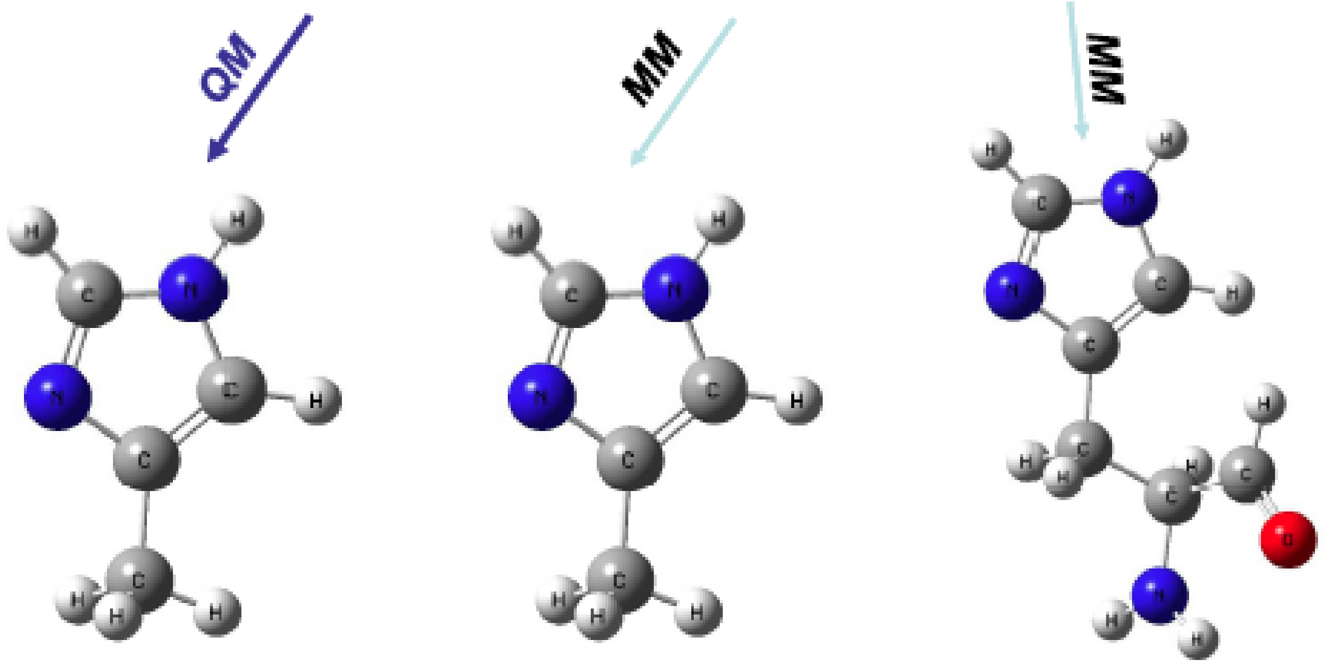


Describing chemical process

Describing environmental effects

Subtractive QM/MM (cont.)

Energy(ONIOM) = E(high, model) – E(low, model) + E(low, real)



Boundary Approximations

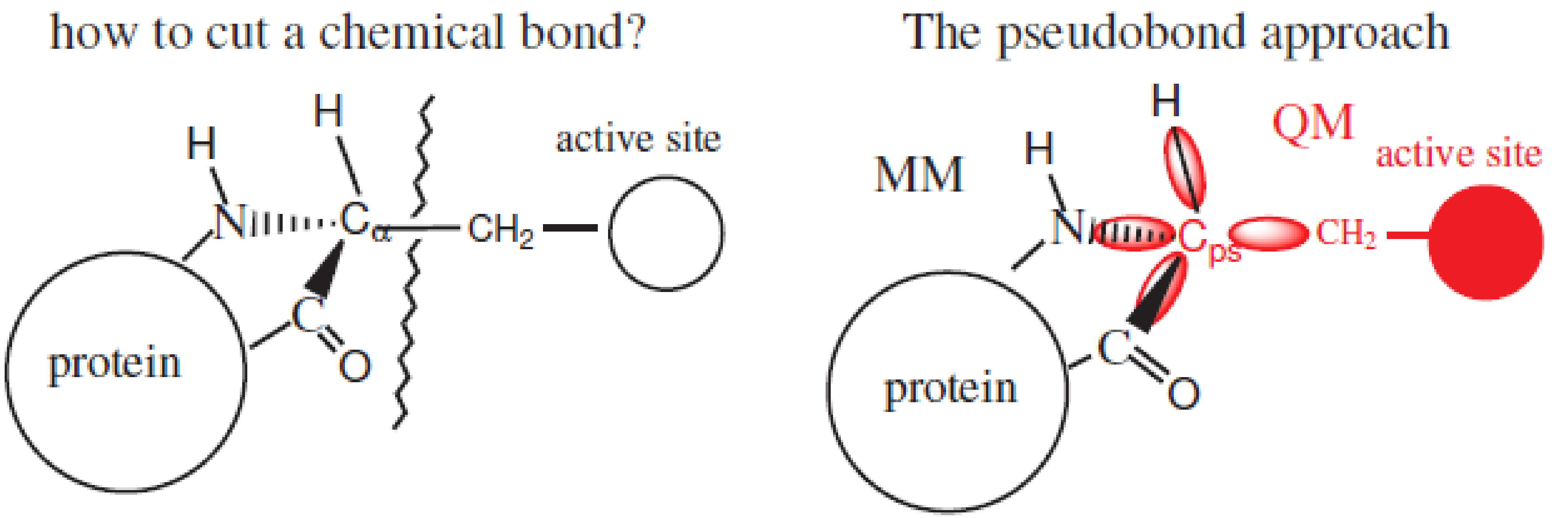
Boundary problem in QM/MM: cutting chemical bonds

- Link atom (single and double)
- Localized self-consistent field (LSCF)
- Generalized hybrid orbital (GHO)
- Pseudobond

Advantages and disadvantages to each, e.g., excluded partial charges, distributed gaussian, etc.

König *et al.*, JPCB, **109**, 9082, (2005).

Pseudobond Method



- *Cps:* pseudoatom with one free valence, composed of a minimal basis set and pseudopotential.
- The pseudopotential is "tailored" to reproduced C-C or C-N bond length, angle, force.

Zhang, JCP, 122, 024114, (2005), Parks, et al., JCP, 129, 154106, (2008).

Long range interactions

In general QM/MM calculations are performed with stochastic boundary conditions and the FF are non–polarizable. Recently:

- Generalized solvent boundary potential (PB charges outside cutoff range). Schaefer *et al.*, JCP, **123**, 014905, (2005).
- QM/MM–Ewald (approximate QM ESP by Mulliken charges for reciprocal space). Nam, Gao and York, JCTC, **1**, 2, (2005).
- QM/MM-Ewald improved (represent QM charge density for reciprocal space). Giese and York, JCTC, **12**, 2611, (2016).

Long range interactions

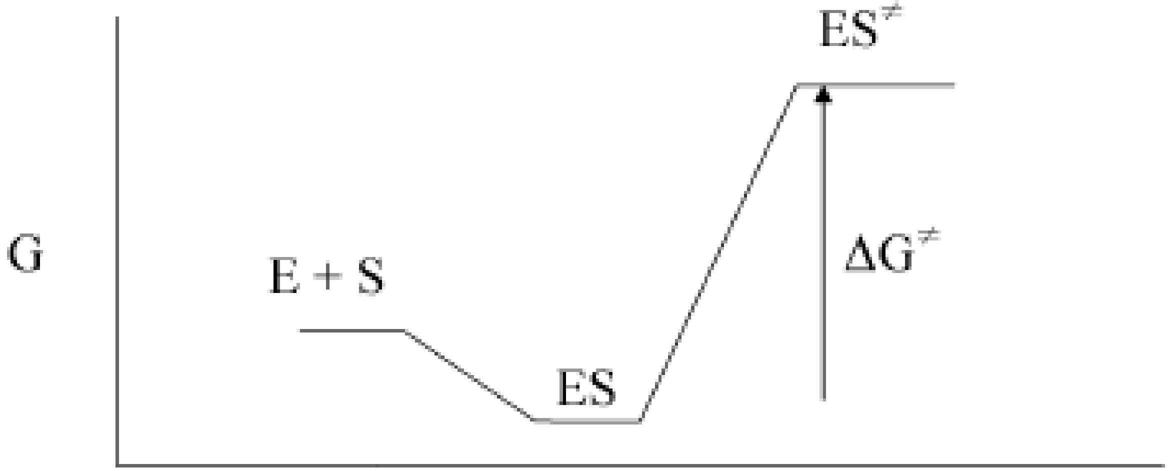
- QM/MM lattice summation (use Ladd lattice summation). Dehez *et al.*, JCP, **122**, 234503, (2005).
- Potential (or its force) displacement or reduction (use functions to chance the potential in the boundary). McCann *et al.*, JCTC, **9**, 944, (2013).
- Smoothing functions. Fang *et al.*, JCP, **143**, 044103, (2015). Kratz *et al.*, TCA, 135:166 (2016).

QM/MM optimization

How to optimize a QM/MM system?

- Start by optimizing QM subsystem in field of MM subsystem keeping MM fixed (quasi–Newton, redundant internal coordinates).
- Once QM converges, optimize MM subsystem with QM fitted charges (MK) keeping QM fixed (truncated Newton, cartesian coordinates).
- Iterate until neither subsystem changes

Modeling enzyme reactions



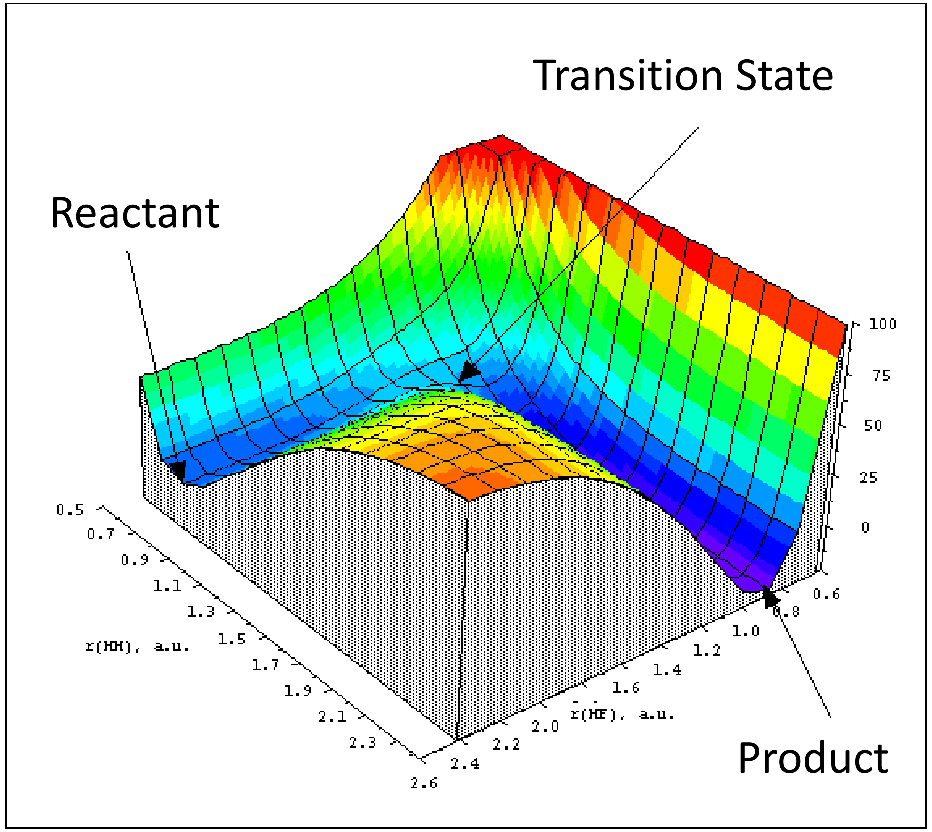
- Classical reaction rate (Eyring's Transition State Theory):

$$k = \kappa \frac{1}{\beta h} \exp(-\beta \Delta G^{\ddagger})$$

$$\Delta G^{\ddagger} = \text{free energy activation barrier}$$

$$\kappa = \text{transmission coefficient}$$

Reaction Paths



- Minimum energy path (MEP): steepest descent path connecting reactants and products, or reaction path at "0K".
- Free energy can further be calculated by including the fluctuations around the MEP.

$$\frac{dX_i}{ds} = c\frac{\partial E}{\partial X_i}$$

$$s = Reaction\,coordinate$$

$$\frac{dX_i}{ds} = Tangent\,to\,path$$

$$\frac{\partial E}{\partial X_i} = Energy\,gradient$$

Calculation of the MEP

- One–dimensional case:

$$A - B + C \underset{\leftarrow}{\overset{\rightarrow}{}} A + B - C$$

Reaction coordinate driving method: use restraint potential to drive system from reactant to product

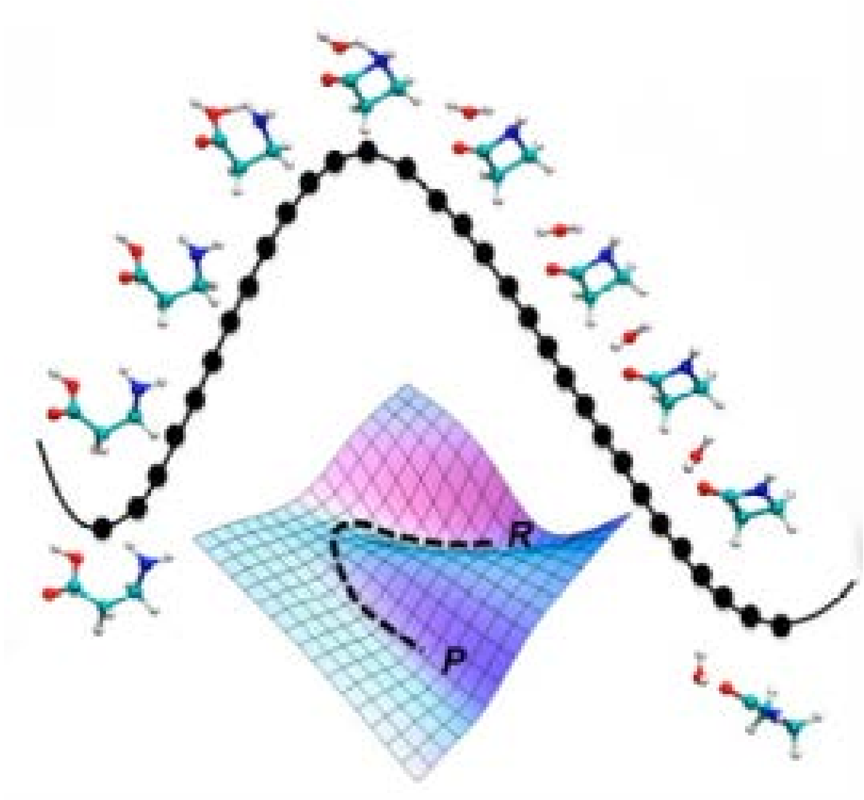
$$V_{restraint} = k(s - s_0)^2$$

$$s = R_{BC} - R_{AB}$$

- Multi–dimensional case:

s may be hard to identify, multidimensional mapping is computationally expensive.

Chain-of-replicas



- Given the information of the reactant, (intermediate) and product, a discretized chain of points can be obtained along the path.
- All the points can be optimized to the MEP simultaneously by using the force perpendicular to the path tangent.
- Points are kept spaced by imposing an equidistant constraint.

Tao *et. al*, JCTC, 10, 10, (2014)

Nudged Elastic Band (NEB)

1st order methods use only energies and gradients to optimize MEP.

- “Spring” between atoms.

- Projection of spring forces (parallel to MEP).

- Add perpendicular force to spring force.

- Minimize in direction of force (velocity Verlet).

Henkelman and Jónsson, JCP, **113**, 9978

2^{nd} order methods

Include approximate Hessians to optimize MEP.

- QSM. Burger and Yang, JCP, **124**, 054109
    - Adaptive step–size integration (Runge–Kutta).
    - Use quadratic surface approx. by damped BFGS
    - Updated Hessian.
    - Reaction path approximated by 3rd order B–spline.

- 2nd order path optimizer. Ayala and Schlegel, JCP, **107**, 375
    - Explicit critical point optimizer.
    - Hessian update with BFGS/Powell mixture.
    - Reaction path approximated by arc of circle to calculate tangents.

Enzyme reaction paths

For QM/MM calculations it is possible to do a combined iterative procedure to reduce time.

1) Start optimization with NEB.
2) At each NEB step, do MM optimization with restraints.
3) Once NEB converged, finish with 2nd order method.

4OT example: 348 (765); 306(618) for first and second steps with combined procedure (2nd order only).

Cisneros, Lu, Liu, Yang, JCP, **122**, 114502

QM/MM-FEP

- Average fluctuations of the enzyme can be included:

$$\Delta G = \Delta E_{QM}^{\ddagger} + \Delta G_{QM/MM}^{\ddagger}$$

$$\begin{aligned}\Delta G_{QM/MM}^{\ddagger} &= \sum \Delta G^{i \to i+1} \\ &= \sum \left\{ -\frac{1}{\beta} \left\langle \exp\left(-\beta \Delta E_{QM/MM}^{i \to i+1} \right) \right\rangle_i \right\}\end{aligned}$$

- MEP obtained in enzyme environment.
- FEP due to QM fluctuations can be included.

Zhang, Liu, Yang, JCP, **112**, 3483

Other methods for Free Energy calculations

Recently other methods have been proposed to determine FE of reaction with QM/MM:

- Minimum free energy path (QM/MM–MFEP)
  - Thermodynamics of the reaction described by a PMF
  - MM is sampled and averaged, followed by optimization of QM on average MM field (PMF)

- Born-Oppenheimer MD (QM/MM–BOMD)
  - On–the–fly BOMD with QM/MM
  - Umbrella sampling used for high energy regions of PES

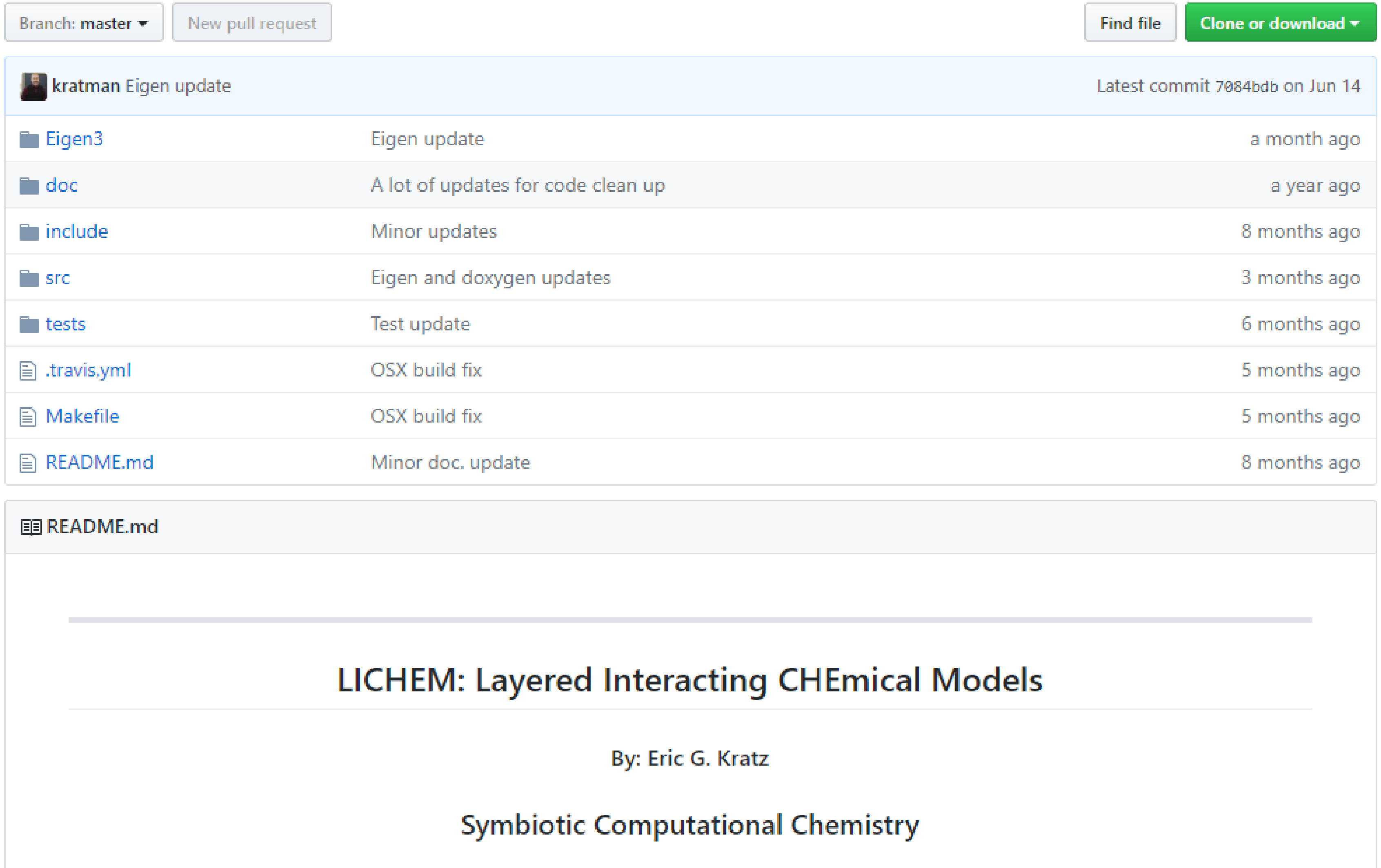
Applications and calculations using LICHEM

Electrostatic models

- Point-charges
- Point-multipoles
- Polarizable sites
- Diffuse charge density

- AMBER/CHARMM: point-charge based models
- Core-Shell/COS: point-charge and polarization
- AMOEBA: point-multipoles and polarization
- Gaussian electrostatic model (GEM): charge density and polarization

LICHEM: Layered Interacting CHEmical Models



kratman Eigen update — Latest commit `7084bdb` on Jun 14

Eigen3	Eigen update	a month ago
doc	A lot of updates for code clean up	a year ago
include	Minor updates	8 months ago
src	Eigen and doxygen updates	3 months ago
tests	Test update	6 months ago
.travis.yml	OSX build fix	5 months ago
Makefile	OSX build fix	5 months ago
README.md	Minor doc. update	8 months ago

README.md

LICHEM: Layered Interacting CHEmical Models

By: Eric G. Kratz

Symbiotic Computational Chemistry

LICHEM

- Design principles:
  - Primarily an interface
  - A single executable
  - Open source

Mode	Potential	QM wrapper	MM wrapper	Calc. type
File conv.	Point-charge	Gaussian	TINKER	Energies
QM only	Mtp/pol	NWCHem	TINKER-HP	Optimization
MM only	Pseudo-bond	PSI4	LAMMPS	PIMC
QM/MM	Diffuse charge		AMBER	Monte Carlo
	GEM			Reaction Path
				Dynamics
				Frequencies

Multipole moments

- Few QM packages can include atomic multipole moments
- Most QM packages can include point-charges
- Why not convert the atomic multipole moments to point-charges?

Simplifying multipoles

- Spherical harmonic multipoles can be converted to six point-charges (Devereux et al., JCTC, 10, 4229, 2014).
- Now nearly any QM package can be used for calculations with multipole moments
- Of course, this is only needed for QM packages which do not support external multiple moments

$$q_{\pm x} = \frac{Q_{00}}{6} \pm \frac{Q_{11c}}{2d} - \frac{Q_{20}}{6d^2} + \frac{Q_{22c}}{2\sqrt{3}d^2}$$

$$q_{\pm y} = \frac{Q_{00}}{6} \pm \frac{Q_{11s}}{2d} - \frac{Q_{20}}{6d^2} + \frac{Q_{22c}}{2\sqrt{3}d^2}$$

$$q_{\pm z} = \frac{Q_{00}}{6} \pm \frac{Q_{10}}{2d} + \frac{Q_{20}}{6d^2} + \frac{Q_{22c}}{2\sqrt{3}d^2}$$

Polarization

- Polarizable QM/MM requires two coupled SCF (wfn/μ_{ind})

(Lipparini et al. J. Chem. Phys, 141, 184108, 2014)

- The polarization energy can be expressed as:

$$E_{pol} = E_{pol}^{QM} + E_{pol}^{QM/MM} + E_{pol}^{MM}$$

- ~80% of the polarization energy comes from interactions with the static field (Wang et al., J. Phys. Chem. B, 117, 9956, 2013)

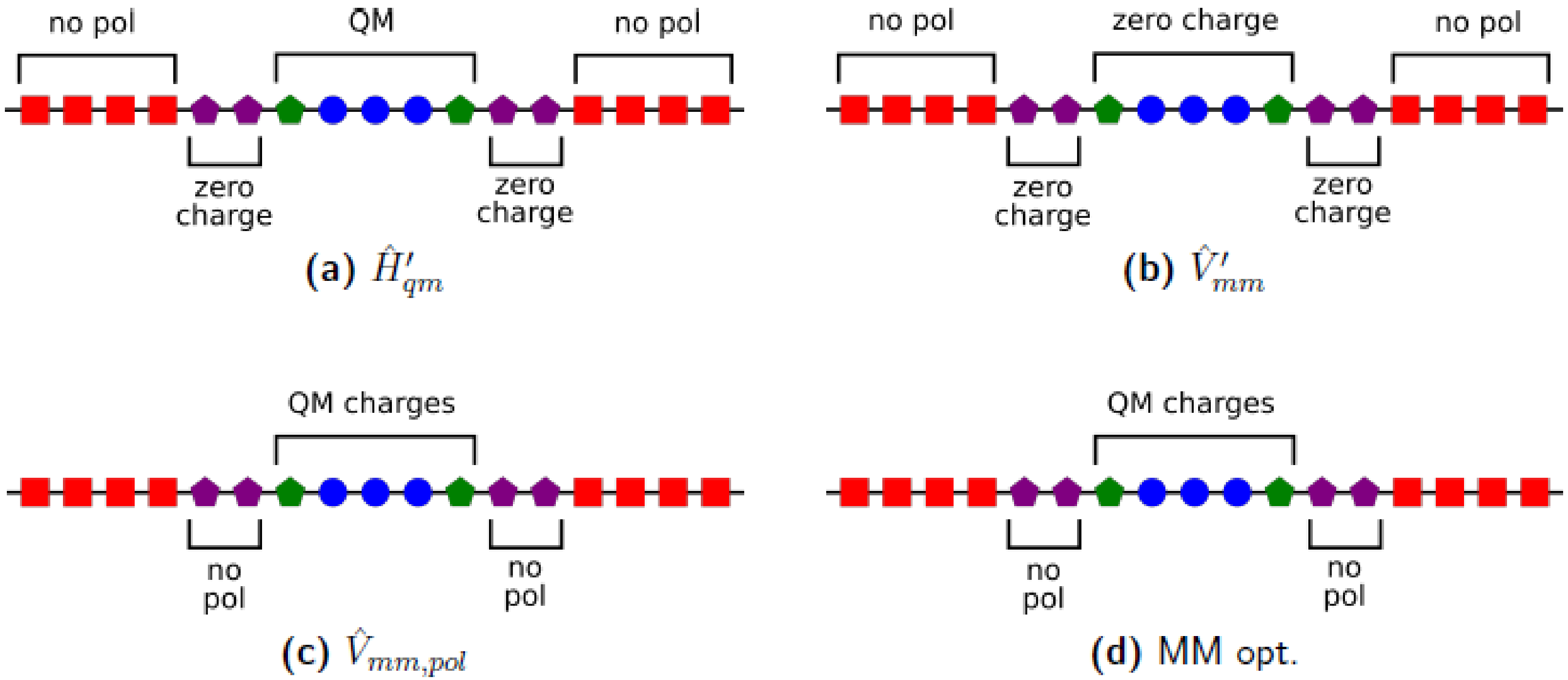
- The polarization energy is approximated by neglecting μ_{ind}^{MM} in H_{eff} during the QM evaluation

Polarization

- QM charges effectively include the induced atomic dipoles
- E_{pol} includes full polarization from the MM and QM terms and the QM/MM interaction includes polarization from static MM field
- In essence this approximation captures $\geq$80% of the mutual QM/MM polarization energy

Polarizable pseudo-bond method

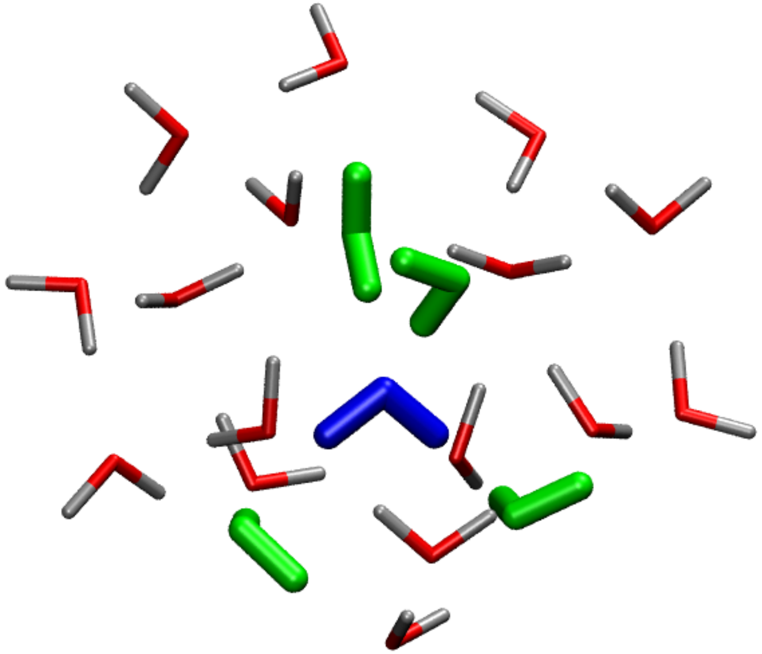
- Regions: QM, MM, pseudo-bond atoms (PB), boundary atoms (BA)



Kratz et al. DOI:10.1002/jcc.24295
Zhang et al., DOI:10.1063/1.478083
Parks et al., DOI:10.1063/1.2994288

Importance of Polarization

- Testing QMMM polarization requires properties that can easily be calculated with full QM



$$(H_2O)_5(H_2O)_{16} \rightarrow (H_2O) + (H_2O)_4(H_2O)_{16}$$

$$E_{solv} = E_{5,16} - E_{4,16} - E_{1,0}$$

Importance of Polarization

QM^*	MM	$E_{Solv}(eV)$
MP2	-	-1.5730
B97D	-	-1.5239
B97D	AMOEBA	-1.4663
B97D	TIP3	-1.2817
B97D	AMOEBA**	-1.2387
PBE0	-	-1.5166
PBE0	AMOEBA	-1.5814
PBE0	TIP3	-1.3999
PBE0	AMOEBA**	-1.3575
-	AMOEBA	-1.5419
	TIP3	-1.1948
	AMOEBA**	-0.5091

*= 6-311++G(d,p)
**= Without polarization

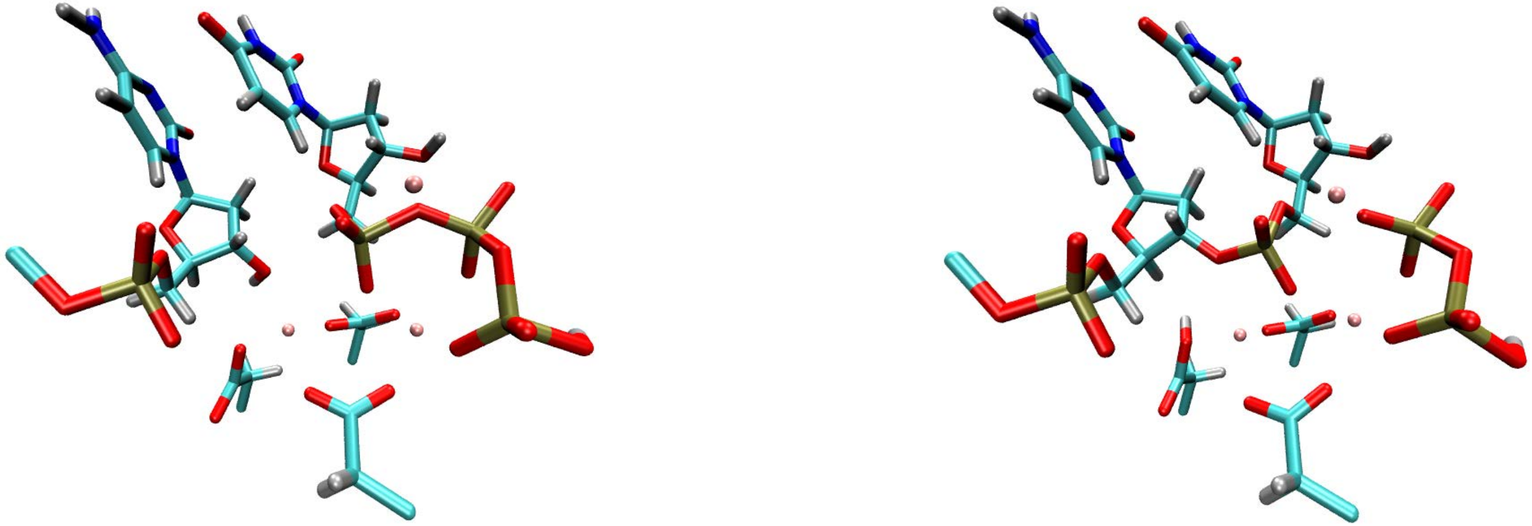
DNA polymerase $\boldsymbol{\lambda}$ $(\boldsymbol{Pol\ \lambda})$

Brief synopsis of $(Pol\ \lambda)$:

- Human DNA polymerase λ is an enzyme which replicates DNA and is one of several enzymes involved in the DNA repair mechanisms
- The ions in the active site are transient and play a large role in the efficiency of the reaction
- Pol β and Pol μ have been shown to have 3 Mg^{2+} ions in the active site
- Only 2 Mg^{2+} ions have been observed in the Pol λ active site

$\boldsymbol{Pol\ \lambda}$

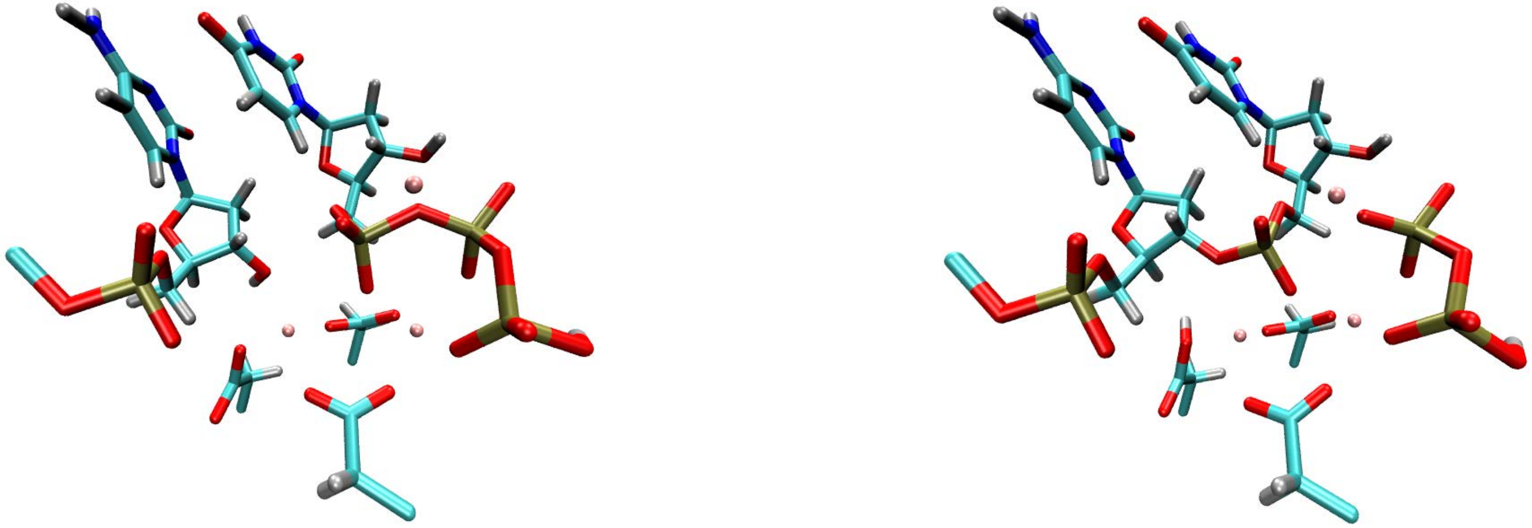
- Addition of a dUTP to the nascent DNA chain



- The pseudo-bond method can be used to describe bonds between the QM and MM regions
  - Method: B97D/6-31G(d)/AMBER99
  - 102 QM atoms, 4 pseudo-bonds, 18 boundary atoms, 18401 MM atoms

$\boldsymbol{Pol\ \lambda}$

- Preliminary QMMM results predict that 3 Mg^{2+} ions are vital to the Pol λ reaction
- Our previous results (without the 3rd ion) predicted that the reaction was endothermic



- Calculations with LICHEM suggest that the reaction is only exothermic when the third Mg^{2+} ion is present

Acknowledgements

- Prof. G. Andres Cisneros
- Dr. Robert E. Duke
- Dr. Hatice Gökcan
- Dr. Hedieh Torabifard
- Alice R. Walker
- Pavel Silvestrov
- Yi-Jung Tu
- Mark Alan Hix
- Phuong (Penny) Nguyen
- Neha Arora
- Michelle Fang
- Sarah Maier (REU Summer Student)
- Edgar Blanco (Visitor Student)

National Institutes of Health
NIGMS GM108583