KU LEUVEN

SCIENCE
PASSION
TECHNOLOGY

TU Graz

QM/MM models for solvated and embedded systems (ground state energies, properties and reactivity)

Jeremy Harvey, KU Leuven, Belgium

Lecture 2: Applications to Biochemistry

Workshop on Theoretical Chemistry: Modelling of Environment Effects in Quantum Chemistry
Mariapfarr, 20-23 February 2018
Scientific Organizer: Benedetta Mennucci

Case 1. The Birth of QM/MM, by Warshel and Levitt: Lysozyme

The first QM/MM study is widely considered to be: *Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme*, A. Warshel and M. Levitt, *J. Mol. Biol. 1976*, **103**, 227-249.

A general method for detailed study of enzymic reactions is presented. The method considers the complete enzyme–substrate complex together with the surrounding solvent and evaluates all the different quantum mechanical and classical energy factors that can affect the reaction pathway. These factors include the quantum mechanical energies associated with bond cleavage and charge redistribution of the substrate and the classical energies of steric and electrostatic interactions between the substrate and the enzyme. The electrostatic

The whole paper is worth reading!

The conformation of the enzyme–substrate complex can be adequately studied using empirical energy functions based on the "classical" contributions of bond stretching, bond angle bending, bond twisting, non-bonded interactions, etc. (Levitt, 1974*a*; Platzer *et al.*, 1972). The mechanism and energetics of the enzymic reaction can only be studied using a quantum mechanical approach. Previous quantum mechanical calculations on enzymic reactions have been limited in several respects. In the first place, they deal with an over-simplified model system consisting of only small fractions of the atoms involved in the real enzyme–substrate interaction (Loew & Thomas, 1972; Umeyama *et al.*, 1973; Scheiner *et al.*, 1975). Second, all available quantum mechanical methods treat the reaction as if in a vacuum and are, therefore, unable to include the dielectric, which has a very important effect on the energy and charge distribution of the system. This limitation does not depend on the actual quantum mechanical scheme, being equally valid for the simplest Hückel treatment and most extensive *ab initio* calculation.

Lysozyme

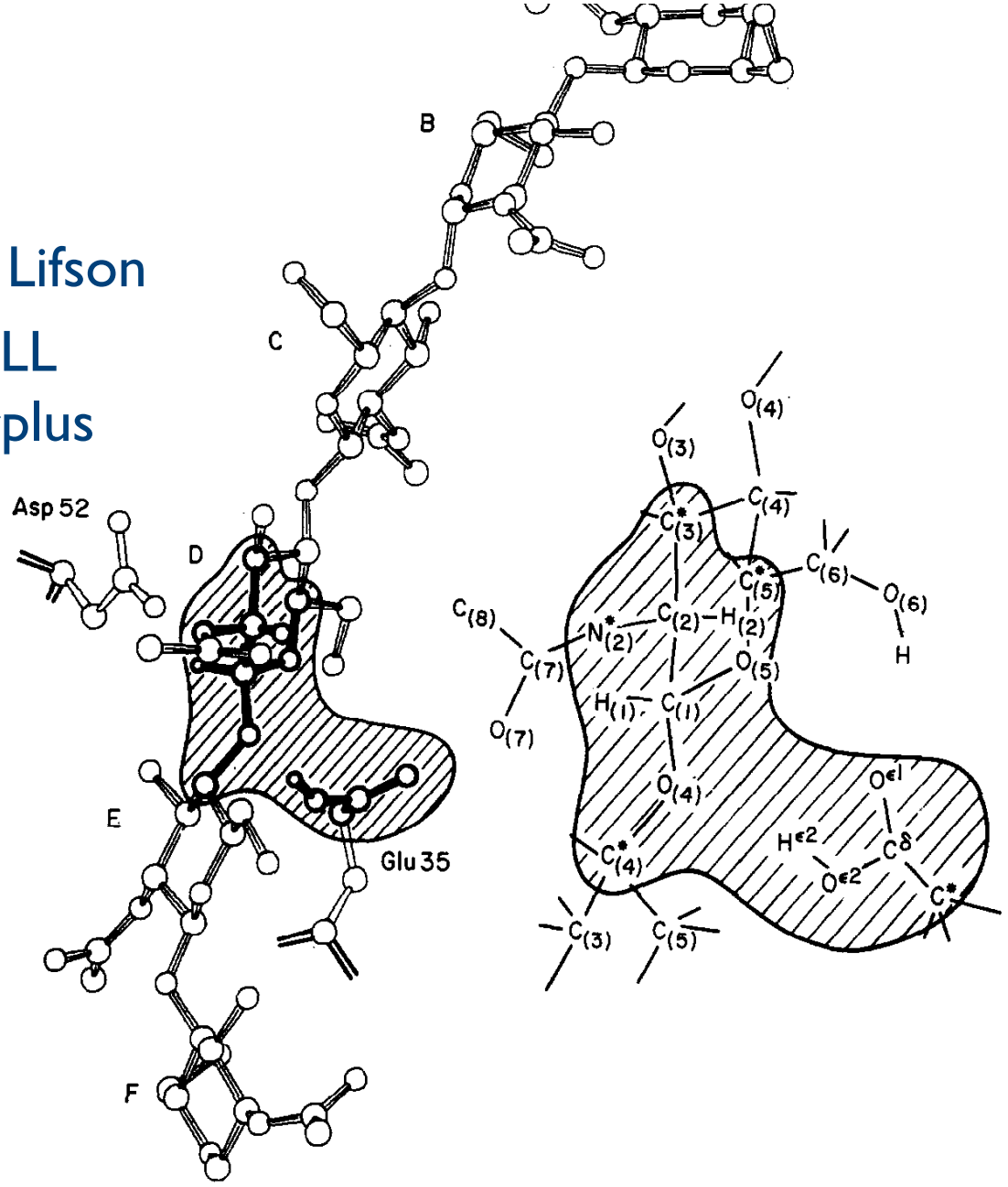
MM: forcefield of Levitt and Lifson

QM: semiempirical QCFF/ALL method of Warshel and Karplus

Uses frozen hybrid orbitals for covalent bonds across QM-MM boundary

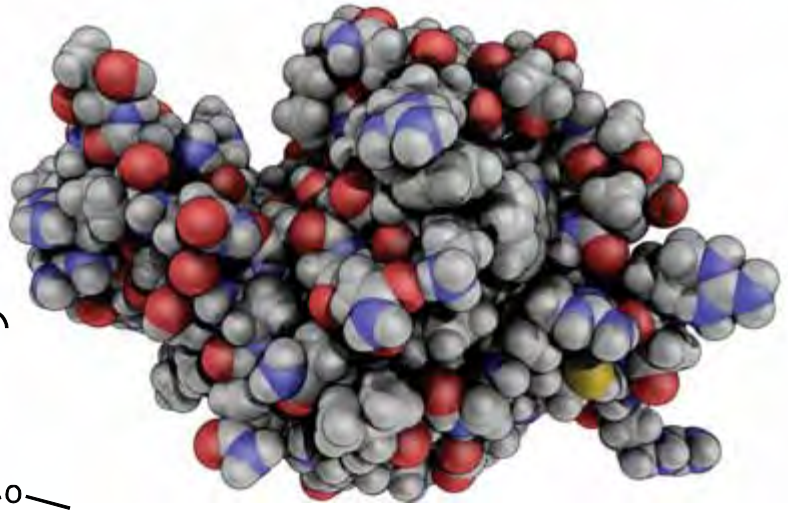
MM atoms are polarizable

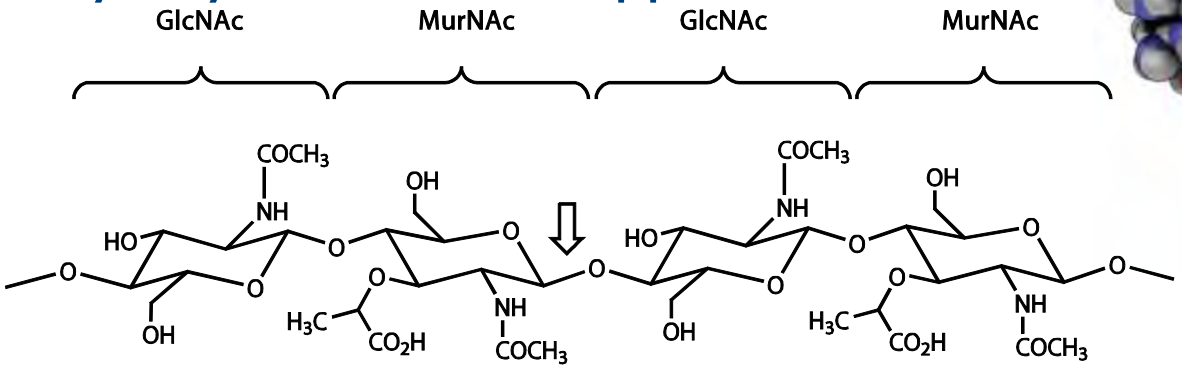
Surrounding water treated as a quasi-continuum: this was a QM/MM/continuum study!

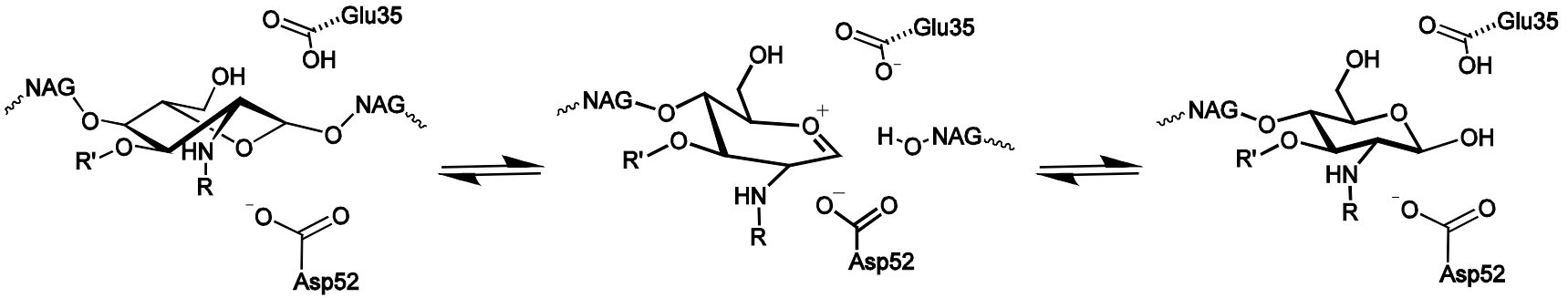


Warshel – Levitt QM/MM Aims

The aim was to characterize the mechanism of hydrolysis of peptidoglycan by lysozyme, in the Philipps mechanism.







KU LEUVEN

QM/MM Results

Distortion of ring and C-O heterolysis treated by 2-D adiabatic mapping:

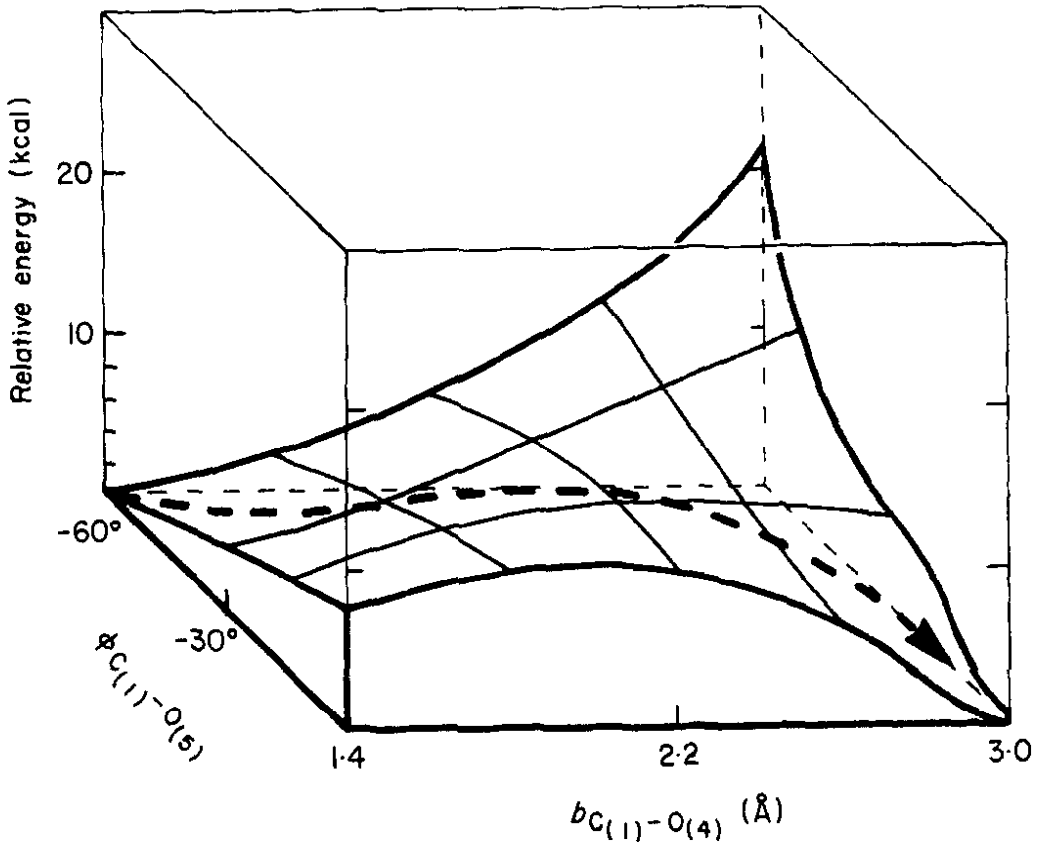


FIG. 6. The adiabatic potential energy surface of sugar D for torsion around the $C_{(1)}$—$O_{(5)}$ torsional angle and stretching of the $C_{(1)}$—$O_{(4)}$ glycosidic bond. The proton $H^{\varepsilon 2}$ has been transferred from Glu35 to $O_{(4)}$, and the $H^{\varepsilon 2}$—$O_{(4)}$ distance constrained to remain at 1·1 Å, so that the sugar remains protonated.

KU LEUVEN

Warshel/Levitt Summary

We have applied the method to study the factors that influence the stability of the carbonium ion intermediate in the reaction of lysozyme.

"The method" is remarkably 'modern' with some features that have not become standard even 40 years later!

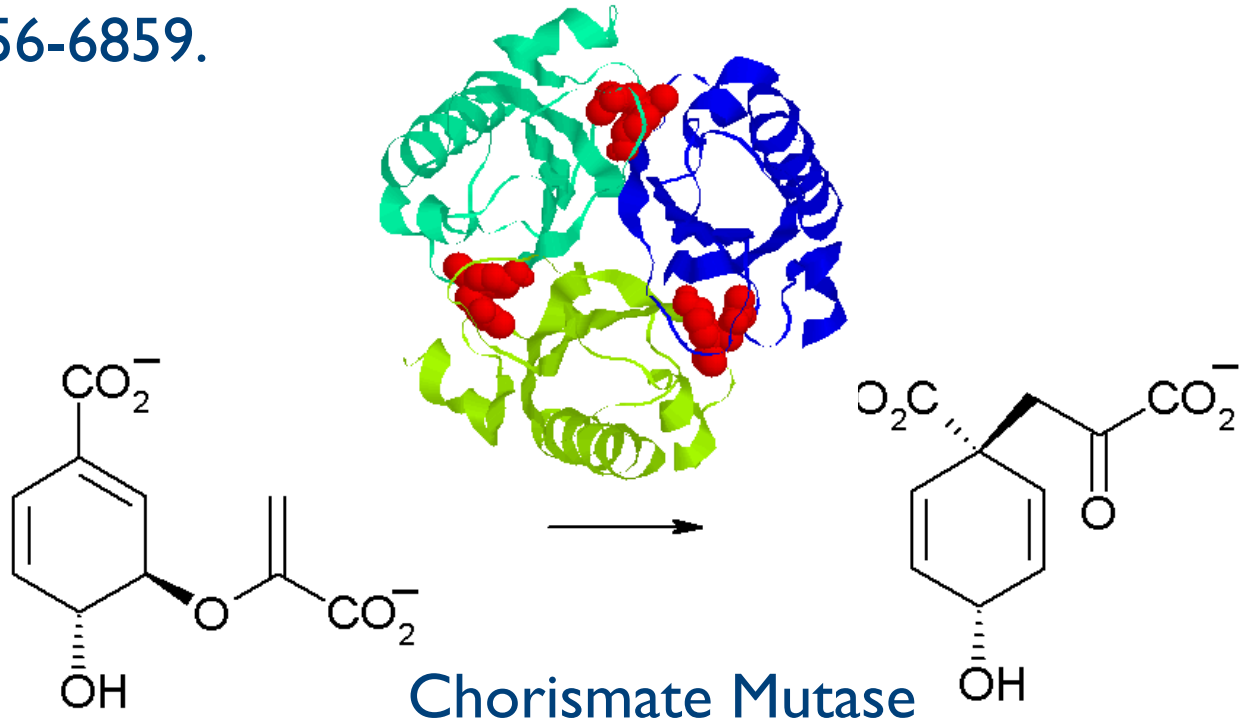
The mechanism studied is now not believed not to be the one followed by the enzyme, but some of the general conclusions remain topical even now:

3×10^6. In general, it seems much easier to control the reaction rate by electrostatic interactions which vary slowly with distance, than by steric interactions which vary rapidly with distance and can be relaxed by small atomic movements. As the charge distribution on the substrate changes significantly during the lysozyme reaction (Table 5), the enzyme can best control the reaction rate by having charges and effective local dielectric designed to interact well with the charge distribution of the transition state. Although the role of strain in the lysozyme mechanism has been

II. QM/MM with Accurate Correlated Methods

High-Accuracy Computation of Energy Barriers in Enzymes, F. Claeyssens, J. N. Harvey, F. R. Manby, R. A. Mata, A. J. Mulholland, K. E. Ranaghan, M. Schütz, S. Thiel, W. Thiel and H.-J. Werner, *Angew. Chem. Int. Ed. 2006*, **45**, 6856-6859.

$$\frac{\partial V_{\mathrm{QM}}}{\partial X_{\mathrm{MM}}}$$

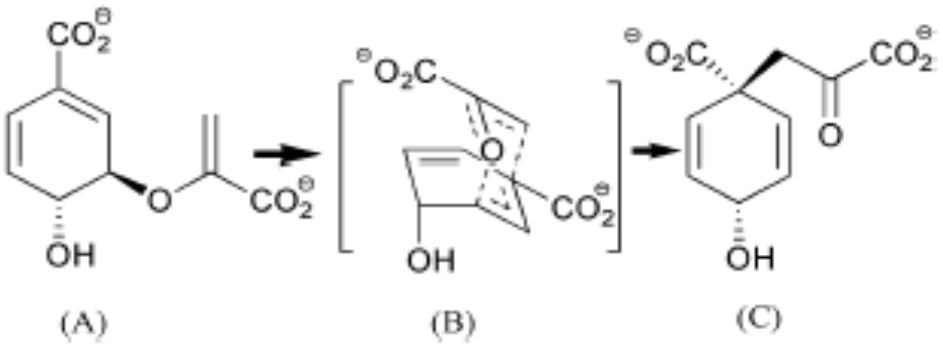


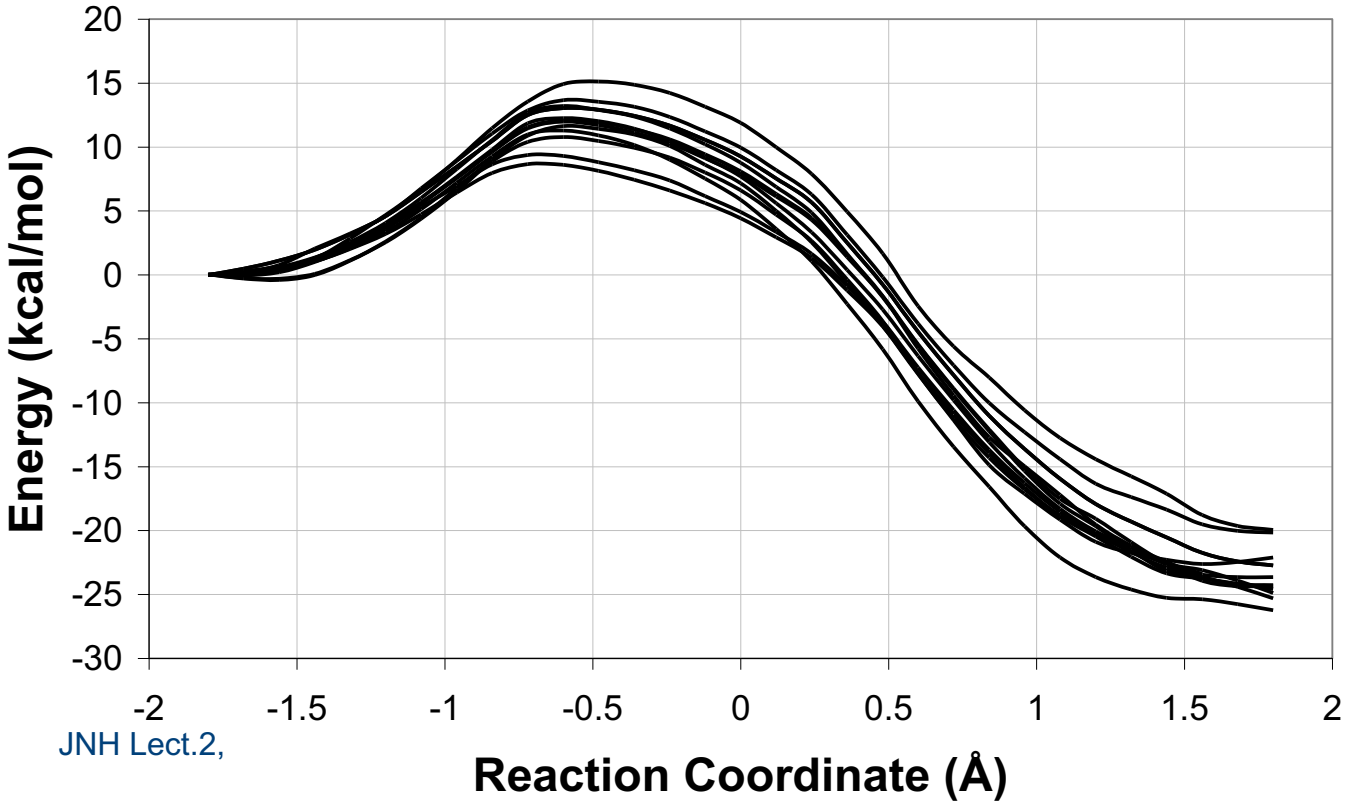
Chorismate Mutase

KU LEUVEN

Multiple Pathways

Initial AM1 or PM3 QM/MM MD yields starting points which are different in detail though all similar in QM region.
Average barrier is 12 kcal/mol.



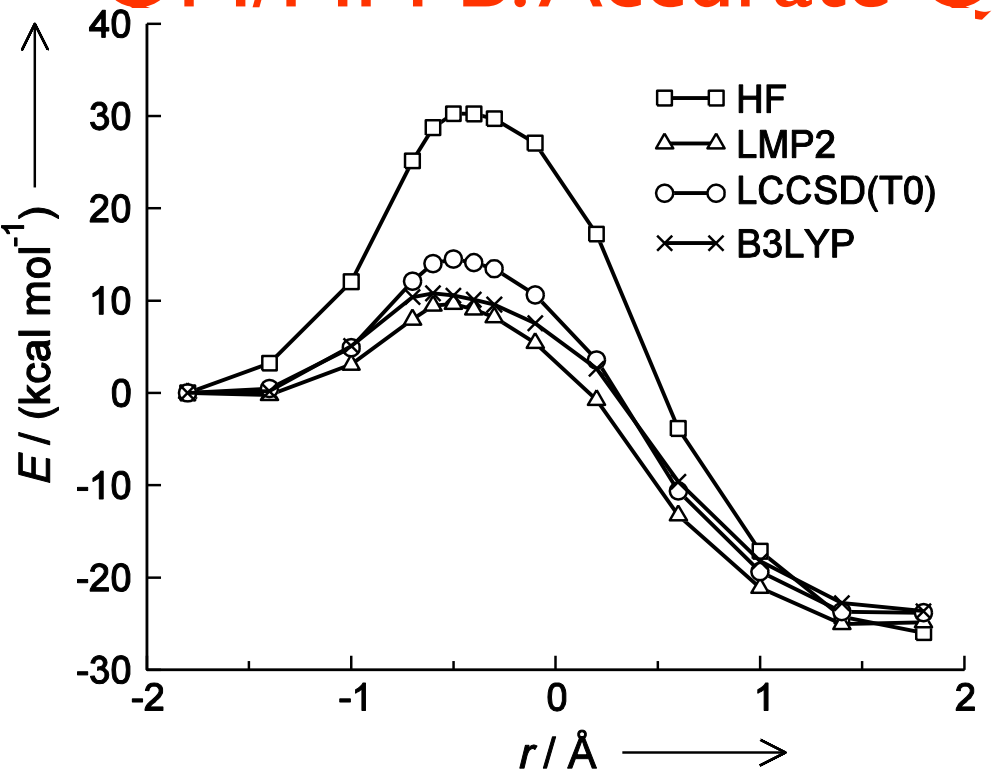


B3LYP/6-31G*:: CHARMM22

Claeyssens, Ranaghan, Manby, Harvey, Mulholland, *Chem. Commun.* 2005, 5068

KU LEUVEN

QM/MM B: Accurate QM



Method	CM
DF-MP2/[aug]-cc-pVTZ[b]	12.3
DF-MP2/aug-cc-pVTZ	11.8
DF-MP2/aug-cc-pVQZ	12.2
DF-MP2/aug-cc-pV5Z	12.3
DF-MP2-F12/aug-cc-pVTZ[c]	12.3

QM: Molpro; MM: Tinker

$\Delta H_{calc}^{\ddagger}$ {LCCSD(T0)/(aug-)cc-pVTZ(CHARMM)} **13.1** kcal/mol (exp. **12.7**)

$\Delta G_{calc}^{\ddagger}$ (AM1 QM/MM umbrella sampling correction): **15.6** kcal/mol (exp **15.4**)

Claeyssens, Harvey, Manby, Mata, Mulholland, Ranaghan, Schütz, S. Thiel, W. Thiel, Werner, *Angew. Chem. Int. Ed.* 2006, **45**, 6856.

KU LEUVEN

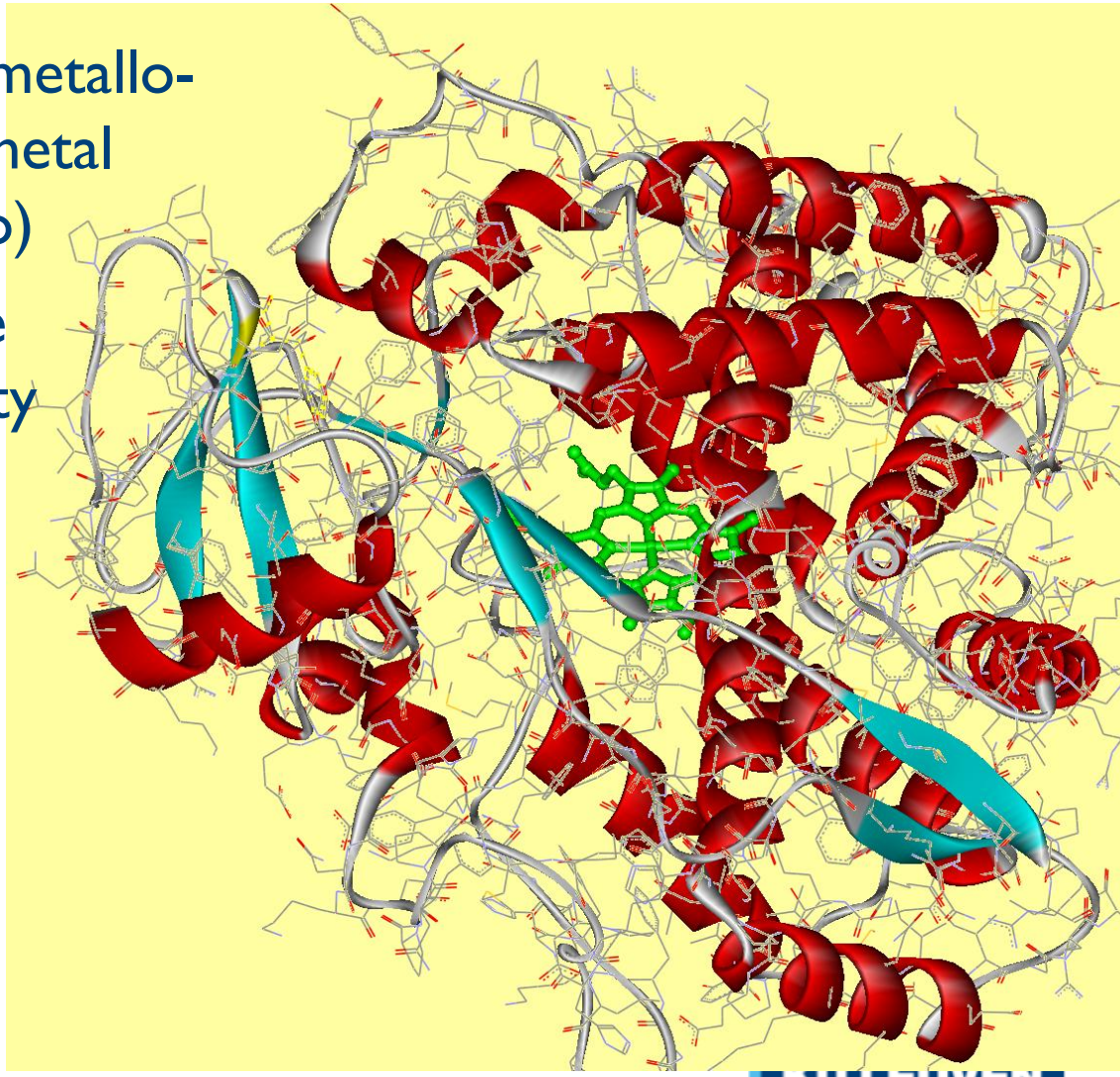
III. A recent application: Drug Metabolism by Cyp450

The cytochrome P450s are metallo-enzymes i.e. they contain a metal cofactor: Fe (also Cu, Zn, Mo)

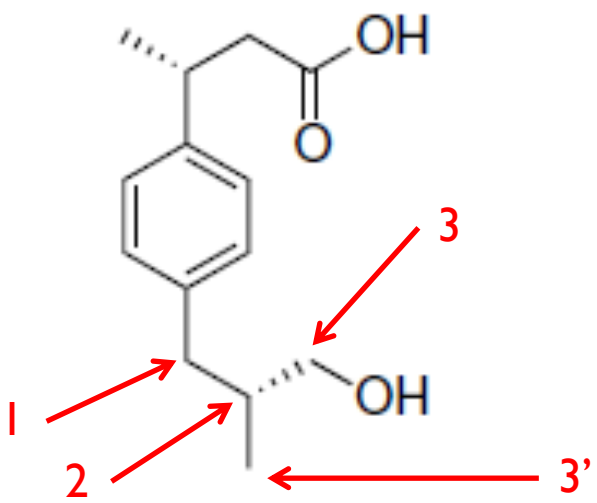
Metal centres display flexible bonding and diverse reactivity

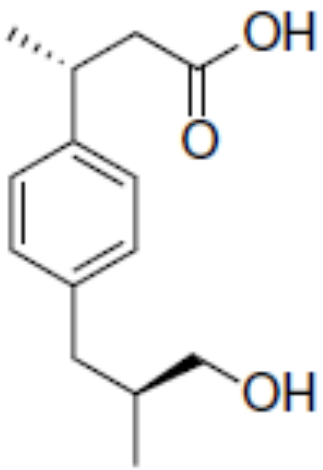
Many metalloenzyme mech-anisms are still not well understood

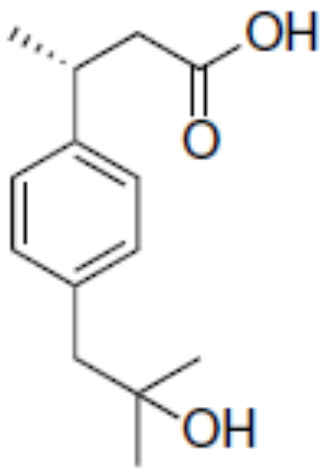
Electronic structure aspects are non-trivial



Ibuprofen Oxidation







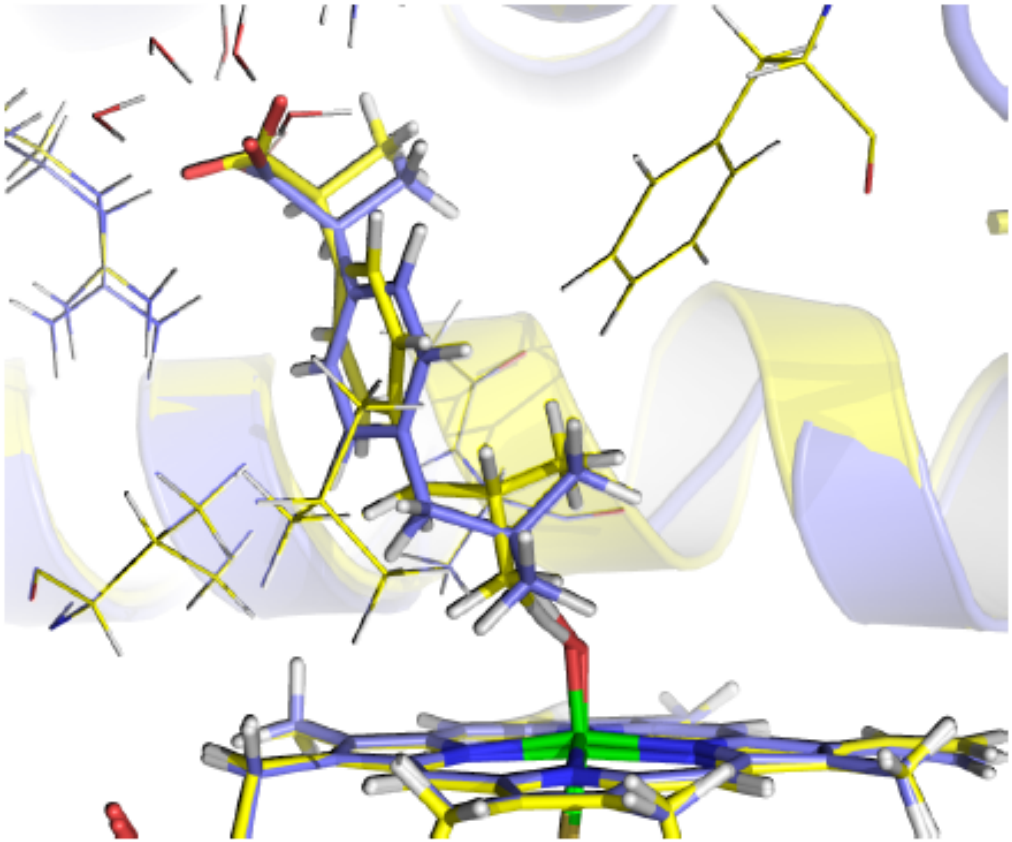
Oxidation in 1, 2 and 3 position should be competitive – the 1 and 2 positions are *intrinsically* much more reactive than the 3 (and 3') position.

Only oxidation in the 3 (and 3') position is observed experimentally

Ibuprofen TSs

MD simulations show that the 2- and 3- C–H bonds approach the Fe–O moiety in the enzyme-substrate complex.

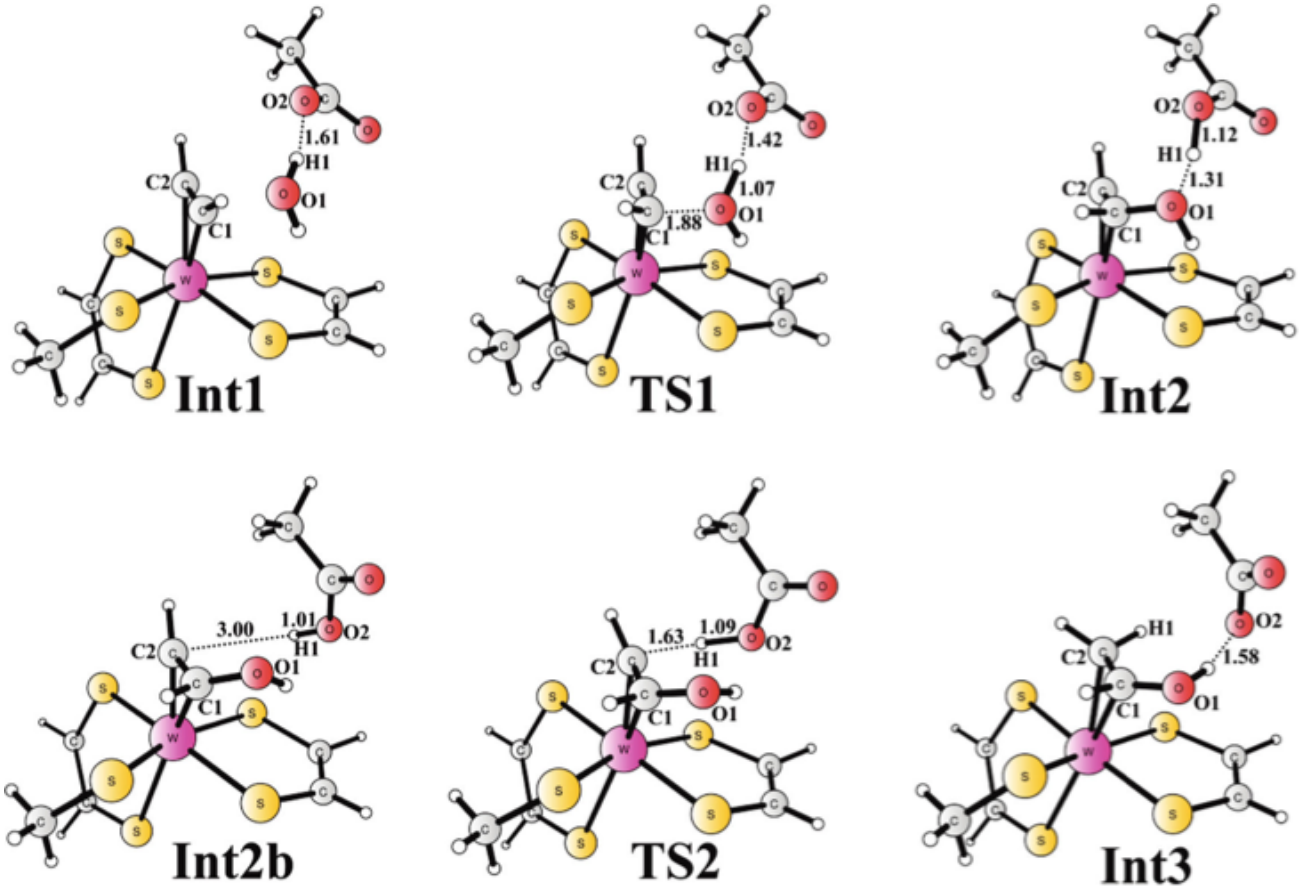
The mean barrier height for oxidation of the 3- position is 19 kcal/mol – for the 2- position, it is 27 kcal/mol



Lonsdale, Houghton, Żurek, Bathelt, Foloppe, de Groot, Harvey, Mulholland, *JACS* 2013, **135**, 8001

IV. Convergence of QM/MM: Energy Profiles in Acetylene Hydratase

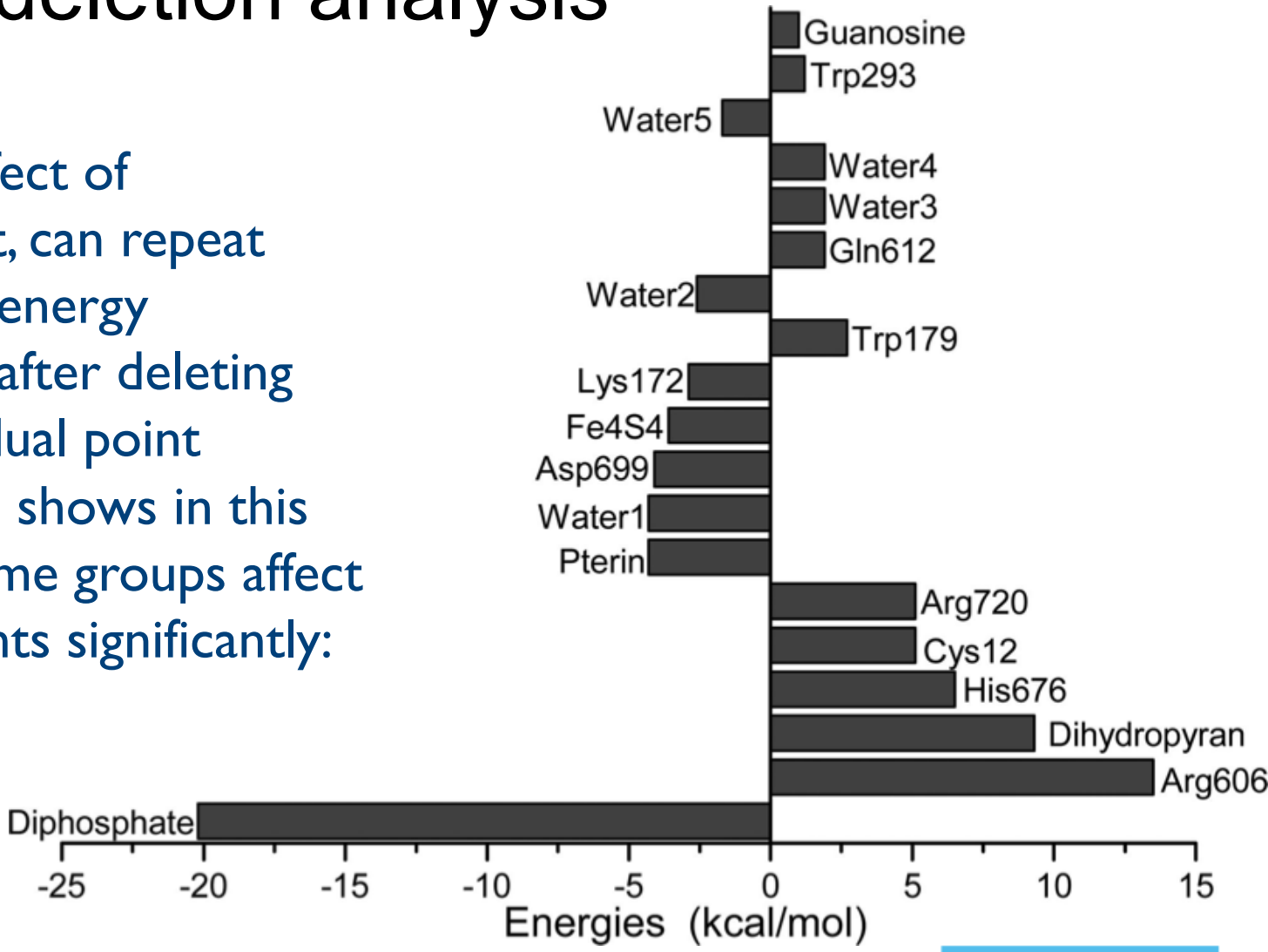
Comparison of QM-Only and QM/MM Models for the Mechanism of Tungsten-Dependent Acetylene Hydratase, R.-Z. Liao and W. Thiel, *J. Chem. Theory Comput.* 2012, **8**, 3793.



KU LEUVEN

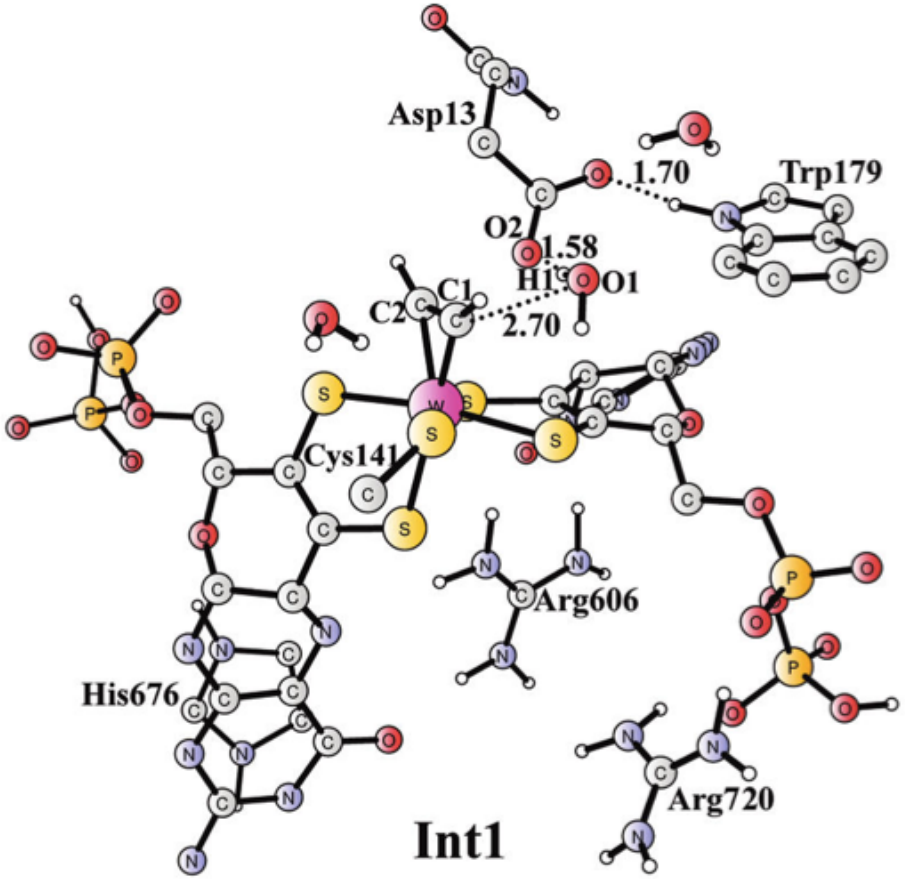
Charge deletion analysis

To assess effect of environment, can repeat single-point energy calculations after deleting some individual point charges. This shows in this case that some groups affect barrier heights significantly:



KU LEUVEN

Large QM Region

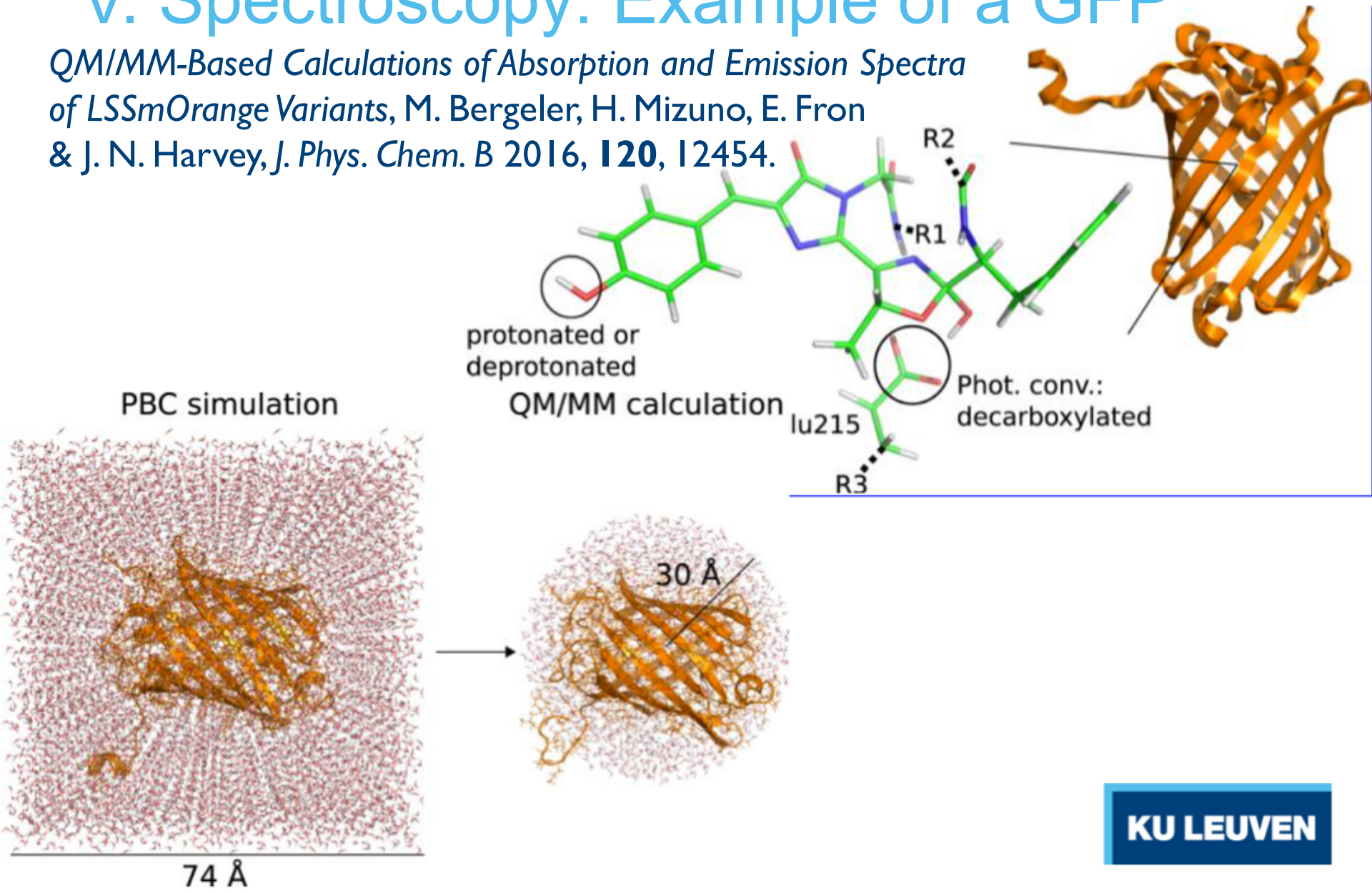


A large QM region yields much better consistency with experiment and with previous QM-only calculations, due to better treatment of electrostatics

KU LEUVEN

V. Spectroscopy: Example of a GFP

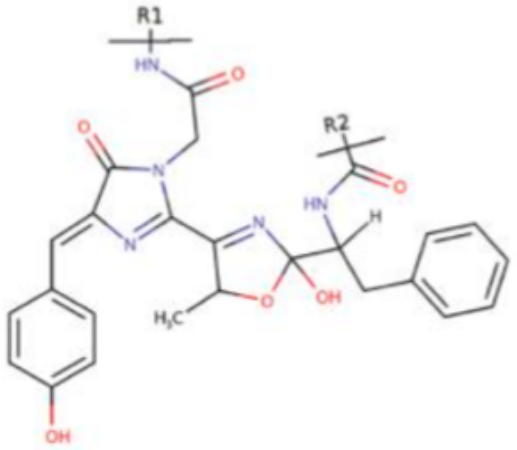
QM/MM-Based Calculations of Absorption and Emission Spectra of LSSmOrange Variants, M. Bergeler, H. Mizuno, E. Fron & J. N. Harvey, *J. Phys. Chem. B* 2016, **120**, 12454.

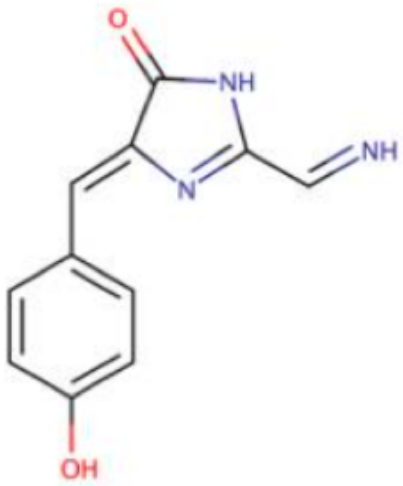


KU LEUVEN

Methodology: Testing Excitation Energy

Method	λ_{calc} nm	ΔE kcal/mol	f
CAM-B3LYP/6-31G(QM/MM)	340.7	83.9	0.79
CAM-B3LYP/6-31G*(QM/MM)	344.5	83.0	0.81
CAM-B3LYP/6-311G*(QM/MM)	347.7	82.2	0.81
CAM-B3LYP/Def2TZVP(QM/MM)	352.2	81.2	0.81
ωB97XD/Def2TZVP(QM/MM)	350.0	81.7	0.79
ZIndo(QM+PCM)	399.7	71.5	0.99
RICC2/Def2TZVP	361.5	79.1	0.87
CASPT2(6,6)/Def2TZVP	323.7	88.3	N/A
CASPT2(8,8)/Def2TZVP	334.7	85.4	N/A
CAM-B3LYP/6-31G	349.0	81.9	0.85
CAM-B3LYP/6-31G*	354.3	80.7	0.83
CAM-B3LYP/6-311G*	357.8	79.9	0.82
CAM-B3LYP/Def2TZVP	364.6	78.4	0.81
ωB97XD/Def2TZVP	362.5	78.9	0.80
LC-ωPBE/Def2TZVP	340.6	83.9	0.86
ZIndo(no PCM)	387.6	73.8	1.01





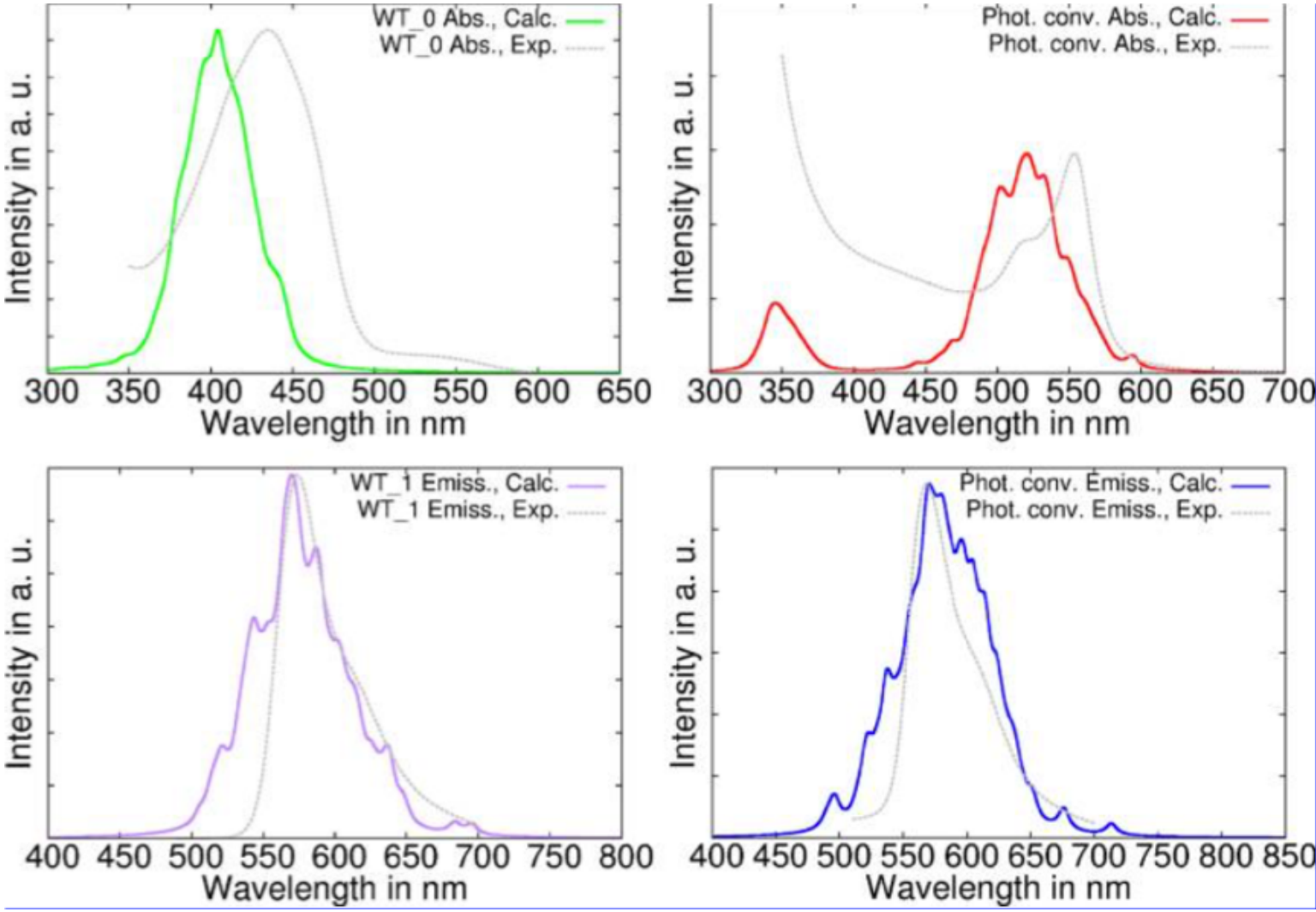
KU LEUVEN

Methodology: QM Region Size Effect

	CAM-B3LYP + PCM		CAM-B3LYP/MM	
QM model(charge)	λ_{calc}	f	λ_{calc}	f
chromophore(0)	365.6(78.2)	0.93	344.5(83.0)	0.81
1.5 Å(0)	364.5(78.4)	0.82	346.2(82.6)	0.73
1.8 Å(−2)	355.3(80.5)	0.78	350.2(81.7)	0.83
2.0 Å(−1)	366.0(78.1)	0.94	353.1(81.0)	0.84
2.5 Å(−1)	363.9(78.6)	0.80	356.2(80.3)	0.78
3.0 Å(0)	363.0(78.8)	0.91	358.8(79.7)	0.76
4.0 Å(0)	363.5(78.7)	0.91	360.3(79.4)	0.69
5.0 Å(−2)	364.9(78.4)	0.49*	360.6(79.3)	0.68
C_144_bb(0)	374.5(76.4)	1.01	349.2(81.9)	0.93
C_144(−1)	378.4(75.6)	1.04	349.6(81.8)	0.95
C_215_sc(−1)	350.7(81.5)	0.84	343.8(83.2)	0.78

QM region size ranges from 63/64 atoms (chromophore only) to 650 (selecting residues within 5 Å)

Spectra



MD simulations are used to sample hundreds of structures for the ground and excited state, and vertical excitation energies computed with QM/MM; convoluted to form spectra

KU LEUVEN

Lecture 11 Conclusions

- The QM/MM method is extensively used to study reactivity and other properties in biochemical systems
- The tone was set already by the impressive Warshell & Levitt study
- Obtaining accurate results is challenging, due to issues with QM level accuracy, force-field used for initial structure generation, size of QM region, and sampling method