

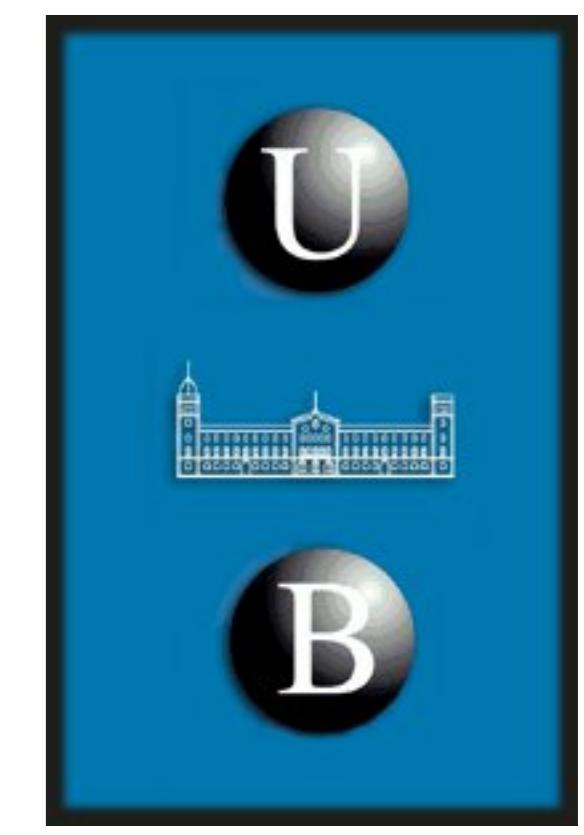
Combining Quantum Mechanics and Molecular Mechanics to Model Biochemical Reactions: Test Studies of Reactive Systems.

Xavier PRAT-RESINA¹, Gérald MONARD², Mireia GARCIA-VILOCA¹, Josep Maria BOFILL³, Àngels GONZÁLEZ-LAFONT¹, and José María LLUCH¹

¹ Departament de Química, Universitat Autònoma de Barcelona
08193 Bellaterra (Barcelona) - SPAIN

² Université Henri Poincaré - Nancy I, UMR 7565 - Groupe Chimie théorique
Faculté des Sciences - B.P. 239, F54506 Vandœuvre-lès-Nancy - FRANCE

³ Departament de Química Orgànica and Centre Especial de Recerca en Química Teòrica
Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona - SPAIN



A
Universitat
Autònoma
de Barcelona

ABSTRACT

Modeling biochemical reactions implies the use of quantum mechanics to describe the changing in electronic density along a reaction pathway, and the inclusion of the effect of the whole environment (usually represented by a protein, its substrate(s), its solvent, etc) that can affect this reaction pathway (as compared with the reaction in solution or in vacuo). Because the use of a fully quantum mechanical description would be too time consuming, the so-called QM/MM methods which combine the treatment of the reactive part by quantum mechanics whereas the rest of system is treated by molecular mechanics have emerged as solutions of choice to simulate efficiently reactions in biochemical systems. However, as some covalent bonds can be present at the separation between the quantum and the molecular mechanics parts, one has to be very careful with the description of this *frontier* bonds as well as with the interactions between the two different parts.

In this poster, we summarize our latest results concerning stationary point search on a QM/MM potential energy surface. We present some test case studies of optimized structures as well as transition states obtained with a new transition state search algorithm that was developed recently in our group. We are able to apply this procedure to enzymatic systems where thousands of atoms are moving. Comparisons are made with fully quantum computations as well as with the *link atom* approach included in the official release of ROAR, and the *Local Self Consistent Field* (LSCF) method which allows for a clear delineation between the quantum and molecular mechanics regions.

Minimization Algorithms

LM-BFGS: Limited Memory BFGS^a

It is a quasi Newton-Raphson algorithm which avoids the storage of the full approximate hessian (solve the memory problem for very large systems).

It uses a limited serie of vectors which are combined together to replace the hessian update.

RFO: Rational Function Optimization^b

It uses an augmented hessian algorithm, with a special update of the hessian that we have implemented^c

Numerical hessian is computed at the initial step.

Full diagonalization of the augmented hessian can be avoided using the Lanczos algorithm^d which scales linearly with the system size.

^aD.C. Liu, J. Nocedal *Math Programming*, **1989**, 45, 503.

^bJ. Simons, P. Jorgensen, H. Taylor, J. Ozment, *J. Phys. Chem.*, **1983**, 87, 2745.

A. Banerjee, N. Adams, J. Simons, R. Shepard, *J. Phys. Chem.*, **1985**, 89, 52.

^cX. Prat-Resina, M. Garcia-Viloca, G. Monard, A. González-Lafont, J.M. Lluch, J.M. Bofill, and J.M. Anglada, *submitted*.

^dG.H. Golub, and C.F. van Loan, *Matrix Computations*, **1983**, Johns Hopkins University Press, USA.

Transition State Search Algorithm (Small Systems)

- RFO algorithm with Powell update of the hessian
- Full diagonalization or Lanczos eigenvector search is performed on the augmented hessian.
- The initial hessian is computed numerically.
- When more than one negative eigenvalue, the hessian is corrected (full diagonalization only).
- Comparisons between full diagonalization and Lanczos algorithm have been performed on small systems.

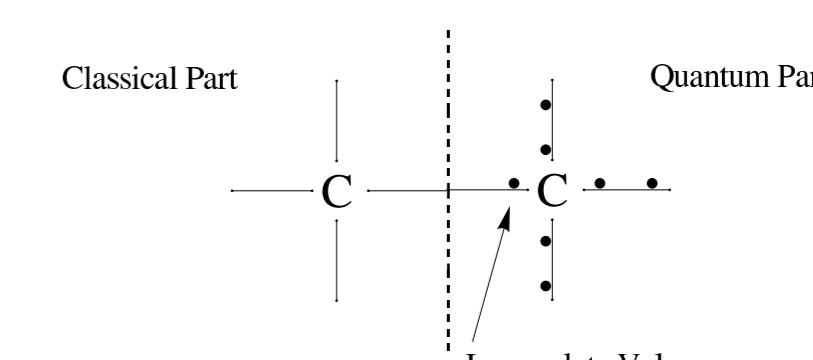
QM/MM Methodology

On a reactive chemical system:

Active part → Modelled with Quantum Mechanics (QM)

Non Reactive part → Modelled with Molecular Mechanics (MM)

$$E = E_{QM} + E_{MM} + E_{QM/MM}$$



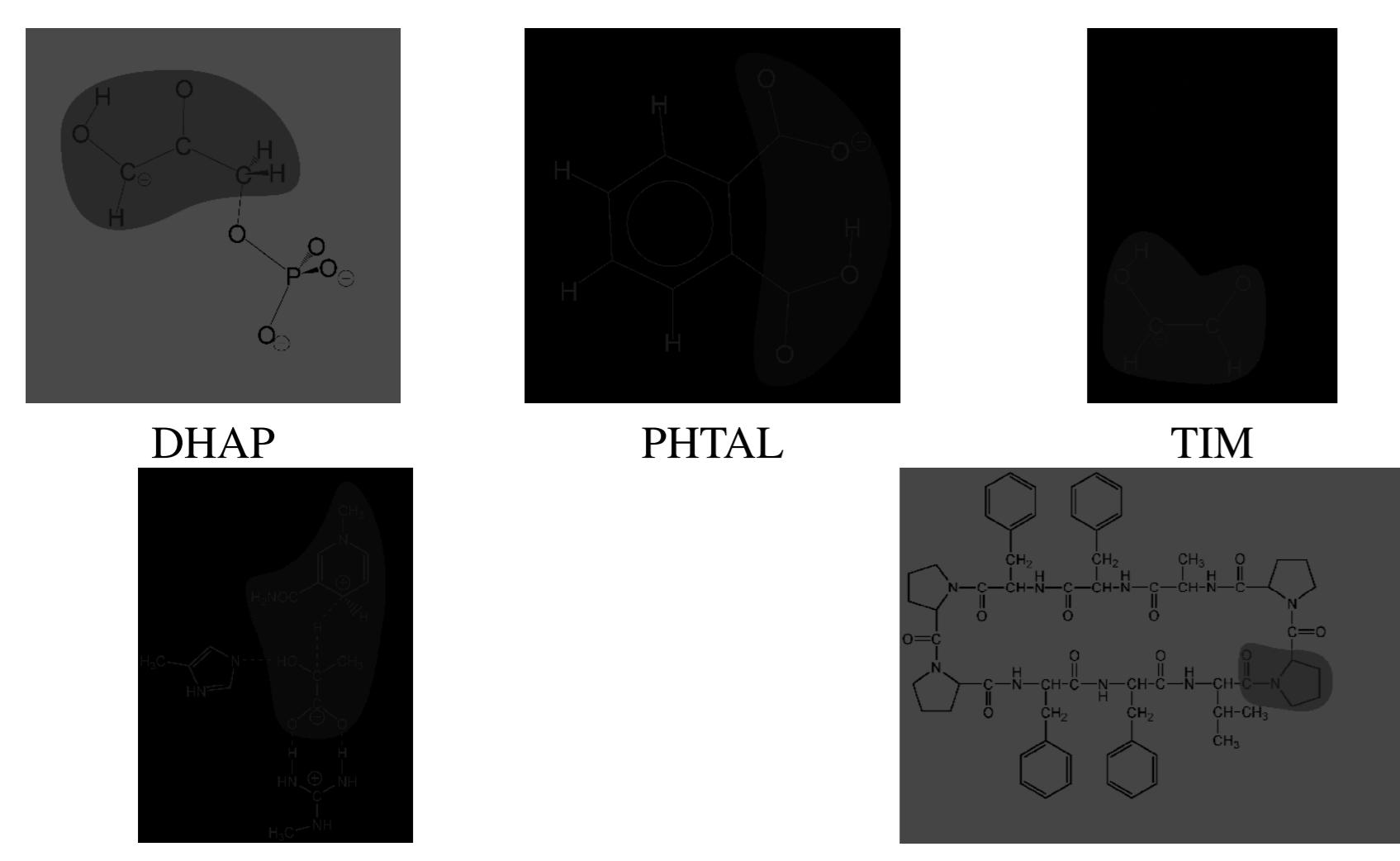
Two main solutions are tested here:

- “Link Atom” approach
- Local Self Consistent Field

Transition State Search Algorithm (Large Systems)

- RFO algorithm with Powell update of the hessian is performed on a defined core region (*i.e.*, the chemical active region).
- LM-BFGS minimization is performed on the environment region (*i.e.*, the non chemical active region).
- Both procedures are done self-consistently until mutual convergence.
- Lanczos eigenvector search is used because it allows the treatment of large core region. It is a linear scaling procedure whereas the traditional full diagonalization procedure scales cubically.
- Avoiding the storage of the full hessian would be an improvement. We are currently working on a solution to this problem.

Test Systems



Minimization LM-BFGS vs. RFO

System	Starting from MM minimized Structures				Starting with RMS gradient < 0.01				
	Initial gradient ^a	RFO	LM-BFGS	RFO	LM-BFGS	RFO	LM-BFGS	RFO	LM-BFGS
DHAP	34.556	224 ^b	346	255	353	4	61	1	failed ^c
PHTAL	25.422	9	55	183	55	(229)	(351)	(1)	(375)
TIM	18.061	29	150	136	158	(11)	(59)	(1)	(184)
LDH	13.739	684	1142	681	1297	(31)	(161)	(6)	(157)
ANTA	5.850	128	2065	518	2318	(699)	(1178)	(51)	(1329)

^agradient norm is given in Kcal/mol/Å

^bthe above value stands for the number of steps required to reach convergence, whereas the number into brackets is the total energy and gradient evaluations

^cthe failed searches are due to LM-BFGS when it is not able to go in a lower gradient zone

Transition State Search Full diagonalization vs. Lanczos approach

Systems	QM/MM		Full QM	
	Full Diag.	Lanczos	Full Diag.	Lanczos
DHAP	66 ^a	30	116	failed ^b
PHTAL	3	3	failed	57
TIM	259	282	223	614
ANTA	22	28	19	17

^athe above value stands for the number of steps required to reach convergence, whereas the number into brackets is the total energy and gradient evaluations

^bThe failed searches take place when the starting point structure is too far from the TS.

Frontier Approach Comparisons

DHAP	Minimum			Transition State					
	ΔE^\ddagger	f	d1	d2	d3	f	d1	d2	d3
Full QM	23.265	1.373	0.951	2.260	2.792	1.372	1.225	1.231	2.303
Link Atoms	22.361	1.353	0.952	2.212	2.761	1.351	1.222	1.228	2.297
LSCF	23.312	1.456	0.952	2.239	2.775	1.454	1.226	1.232	2.231

f: Frontier bond distance

d1: O₁-H distance (QM active center)

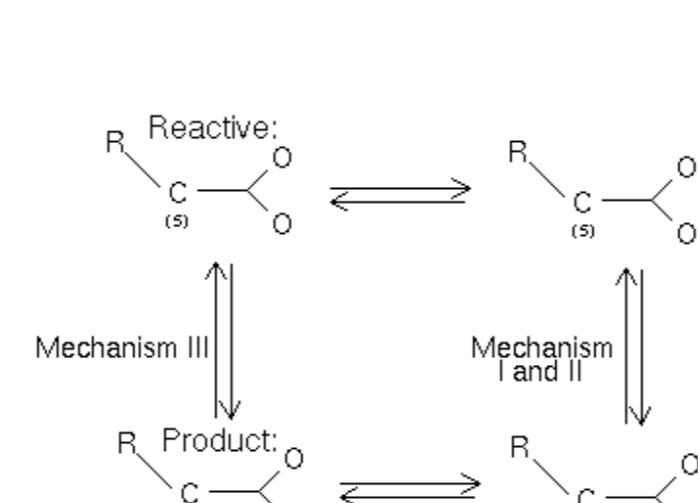
d2: O₂-H distance (QM active center)

d3: O₁-O₂ distance (QM active center)

Application to Enzyme Reactivity 1/2

We have used our algorithms to elucidate the reactivity of Mandelate racemase. Our model system is constituted of 3963 atoms. An active sphere of 15 Å around the active center is considered. It includes 1299 moving atoms, with 81 atoms treated with a PM3 semiempirical hamiltonian while the rest of the system is described thanks to the AMBER force field.

Two substrates are studied, propargyl glycolate(R:ethyne) and mandelate(R:phenyl).



The transition states here located have been found with our LM-BFGS/RFO algorithm. Thanks to the Lanczos approximation a big size for the core can be taken.

Application to Enzyme Reactivity 2/2

Structures	Mandelate			Propargyl		
	Mech. I ^a	Mech. II	Mech. III	Mech. I	Mech. II	Mech. III
S	0.0	0.0	0.0	0.0	0.0	0.0
TS1	18.2	17.4	1	19.7	11.2	1
1	7.8	4.7	1	6.0	0.5	1
TS2	19.6	14.1	1	19.9	15.3	1
2	18.3	12.3	1	19.7	15.3	1
TS3	20.1	14.5	1	22.2	17.0	1
3	19.1	7.9	1	20.2	15.9	1
TS4	22.6	1	18.6	24.2	19.1	21.9
4	22.3	1	1	21.2	15.4	1
TS5	25.7	23.4	1	22.7	20.1	1
5	18.6	18.3	1	15.7	15.5	1
TS6	27.3	27.8	1	24.4	24.6	1
R	6.7	6.7	6.7	3.3	3.3	3.3

^aenergy differences are given in Kcal/mol