The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific
and Cultural Organization

International Atomic
Energy Agency

SMR.1824 - 16

**13th International Workshop on
Computational Physics and Materials Science:
Total Energy and Force Methods**

11 - 13 January 2007

Exploring chemical reactivity in biological systems with hybrid QM-MM methods

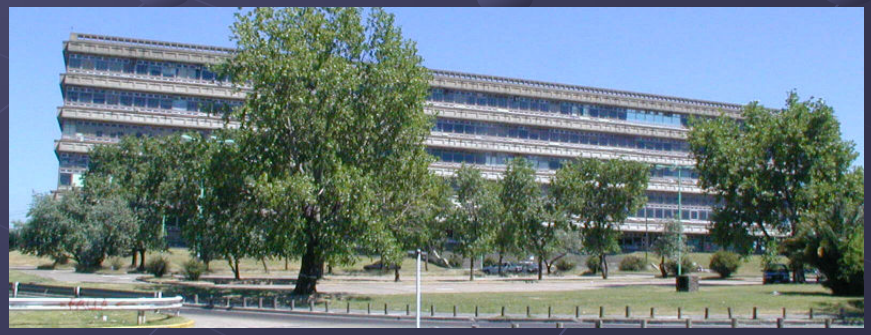
Dario A. ESTRIN
University of Buenos Aires
School of Sciences
Department of Chemistry
Pab. 2, 1424 Buenos Aires
ARGENTINA

Exploring chemical reactivity in biological systems with hybrid QM-MM methods

Darío A. Estrin

University of Buenos Aires

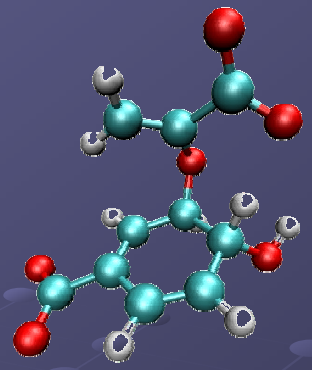
ARGENTINA



Outline:

- Introduction
- Validation: the chorismate mutase case
- Oxygen binding in globins: distal, proximal and dynamical effects
- NO detoxification by truncated hemoglobin of M. Tuberculosis
- Problems and perspectives

Computer simulation in Chemistry and Biochemistry:



- Tools based on physical-chemistry ideas

1) Models: aimed to evaluate potential energy surfaces (PES)

$$E = E(r_1, r_2, r_3, \ldots)$$

Quantum chemical and classical models

2) Schemes to extract relevant information from PES: molecular dynamics (MD), Monte Carlo (MC), optimizations.

QM models:

- Necessary to describe electronic excitations, and chemical reactions (changes in bonding patterns). Limitation in system sizes! Biomolecules: DFT method of choice

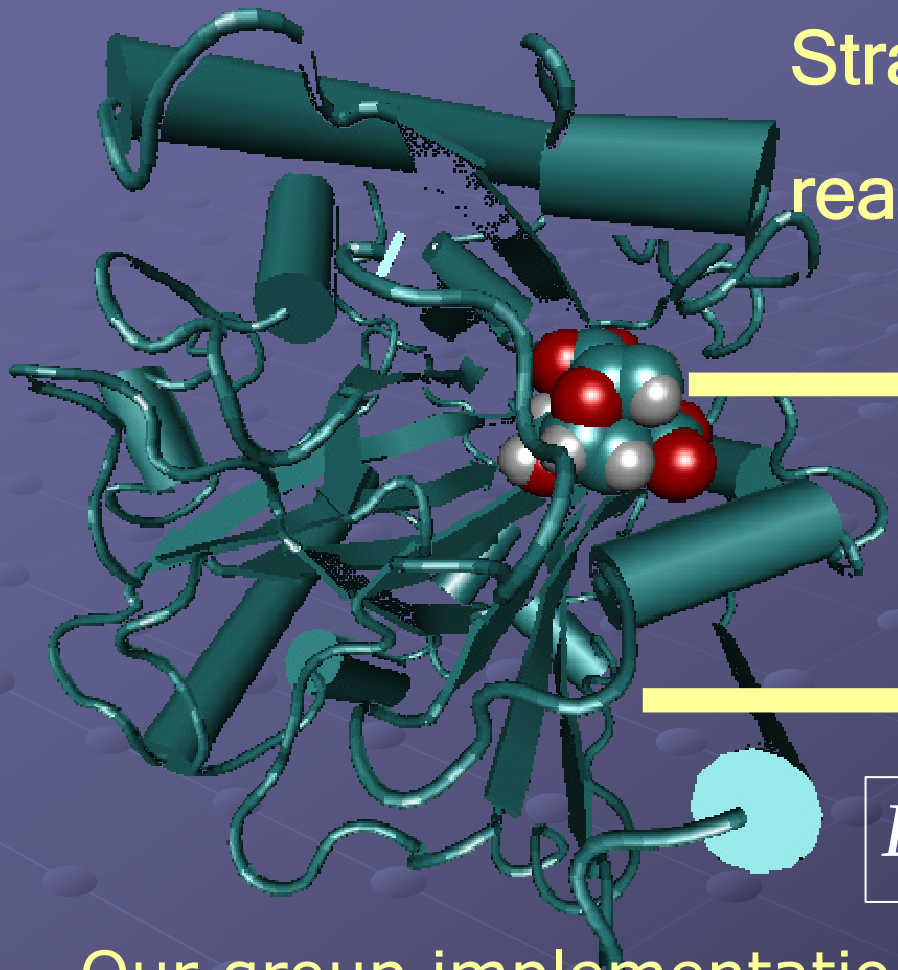
Classical models:

- based on classical ideas. Very efficient numerically. Cannot (usually) deal with reactive processes, but are extremely useful for non reactive processes (i.e. protein folding)

Amber, Charmm, Gromos, etc.

QM-MM hybrid schemes:

Strategy to be able to investigate real chemistry in biomolecules



Reactive region: QM

Environment: MM

$$E_{TOT} = E_{QM} + E_{MM} + E_{QM-MM}$$

Our group implementation: SIESTA DFT/Amber

J.Am. Chem. Soc. 127, 7721 (2005); *J. Am. Chem. Soc.* 127, 6940, (2005);
J. Phys. Chem. B 108, 18073 (2004); *J. Phys. Chem. B* 13728 (2003),
J. Am. Chem. Soc. 128, 12817 (2006); *J. Am. Chem. Soc.* 128, 12455 (2006)

Key term: QM-MM coupling

QM nuclei / environment partial charges interaction

$$E_{QM-MM} = \sum_{i=1}^{C} q_i \int \frac{\rho(\mathbf{r})}{|\mathbf{r} - \tau_i|} d\mathbf{r} + \sum_{i=1}^{C} \sum_{\alpha=1}^{A} \frac{q_i Z_\alpha}{|\mathbf{R}_\alpha - \tau_i|} + E^{LJ}$$

QM density - partial charges interaction

LJ term: short range, dispersion, etc.

Extracting valuable info from simulations:

- Geometry optimizations, constrained optimizations. energy changes, reaction profiles. NO TEMPERATURE!!

- Molecular dynamics: thermal effects. Time evolution in the nanosecond range.

- Advanced sampling techniques: umbrella sampling, multiple steering molecular dynamics (Jarzinski´s method): Free energy!

⇓

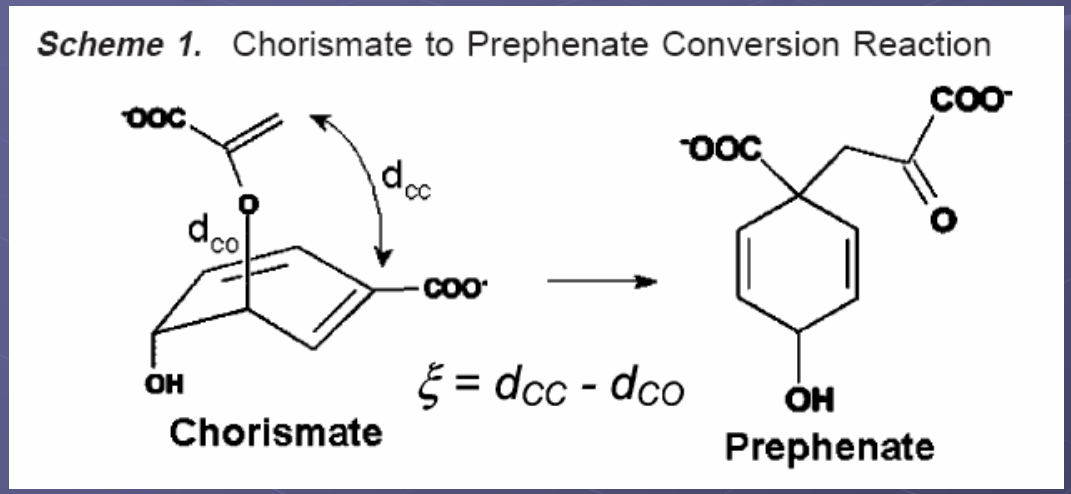
Understanding molecular determinants of a given property

Obtaining information not accesible experimentally

Experiment/Theory collaboration essential

Methodology validation: application to chorismate mutase

(Biosynthesis of aromatic aminoacids in bacteria, plants, and fungi)



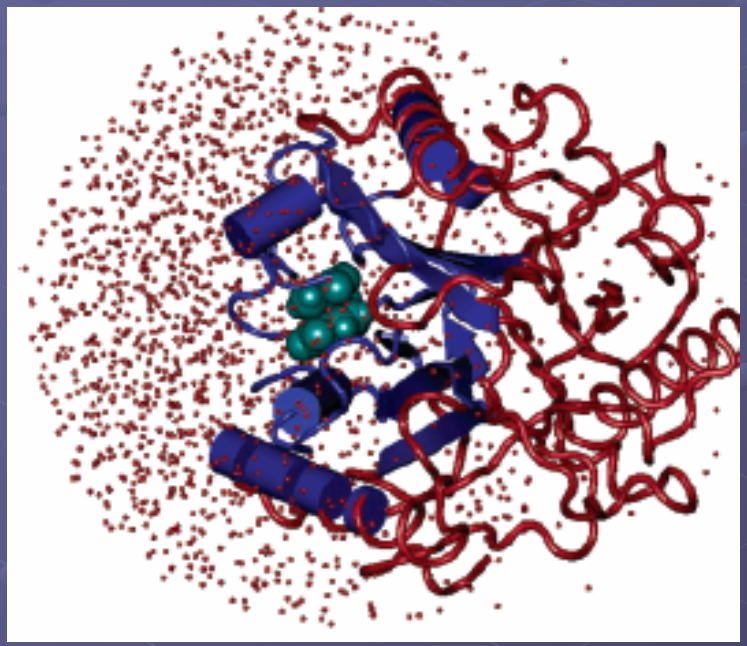
One bond forms and other breaks, simple choice of reaction coordinate

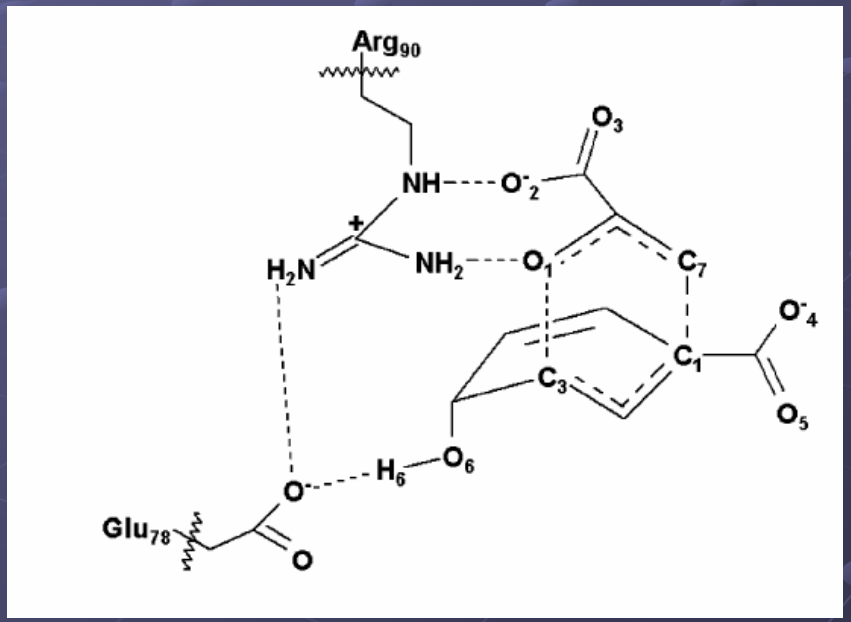
A DFT-Based QM-MM Approach Designed for the Treatment of Large Molecular Systems: Application to Chorismate Mutase **A. Crespo, D. A. Scherlis, M. A. Martí, P. Ordejón, A. E. Roitberg, D. A. Estrin, J. Phys. Chem. B. (2003) 107; 13728-13736.**

First approach: constrained optimizations

$$V_{\mathrm{R}} = \frac{1}{2}k(\xi - \xi_0)^2$$

Addition of harmonic term
Siesta PBE-DZVP/Amber level





2 possible QM subsystems: reactant, reactant plus nearby aminoacids

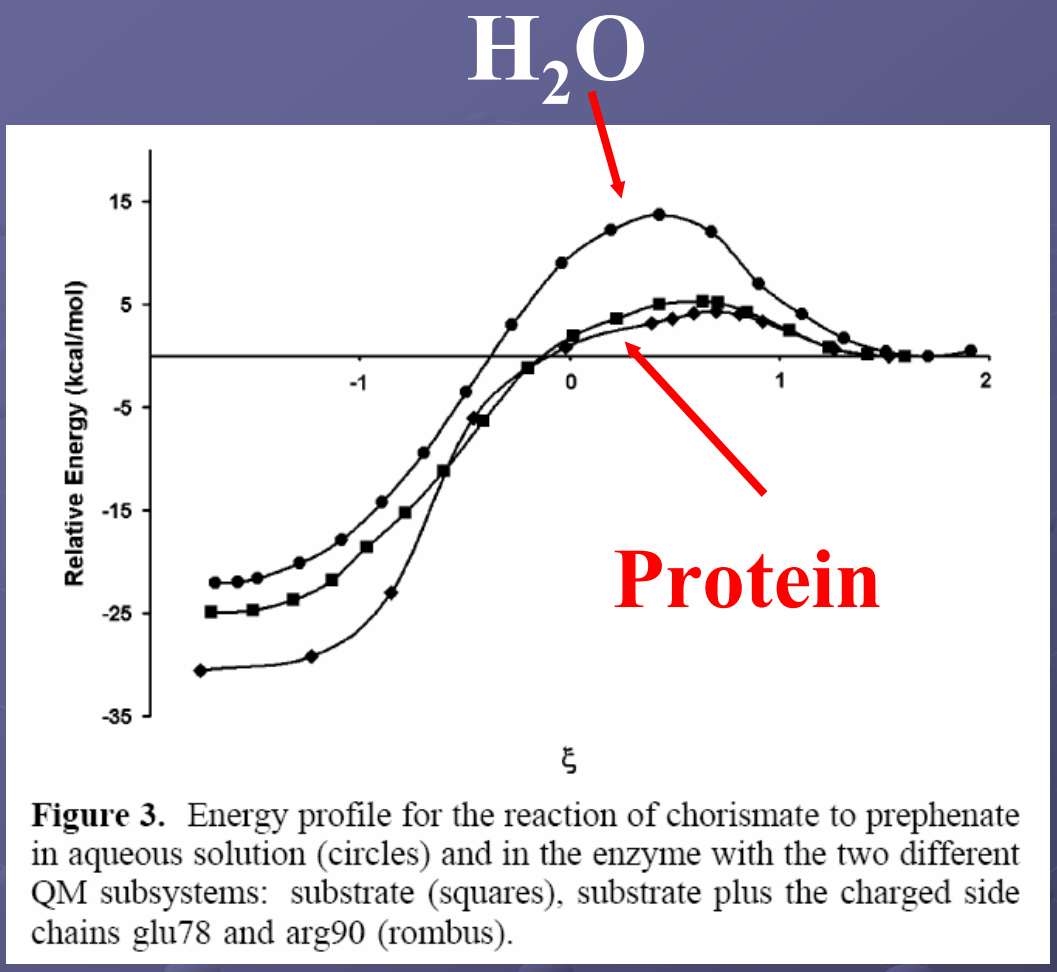


Figure 3. Energy profile for the reaction of chorismate to prephenate in aqueous solution (circles) and in the enzyme with the two different QM subsystems: substrate (squares), substrate plus the charged side chains glu78 and arg90 (rombus).

QM subsystem choice:
little influence on the profiles

The catalytic effect is reproduced

Thermal and entropic effects neglected. Is it OK?

How to incorporate thermal and entropic effects?

Free energy profiles: much more expensive!

Umbrella sampling or
Multiple Steered Molecular Dynamics MSMD:
interesting strategy

$$e^{-\beta \Delta G(\lambda)} = \langle e^{-\beta W(\lambda)} \rangle$$

Jarzynski equation

Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690.

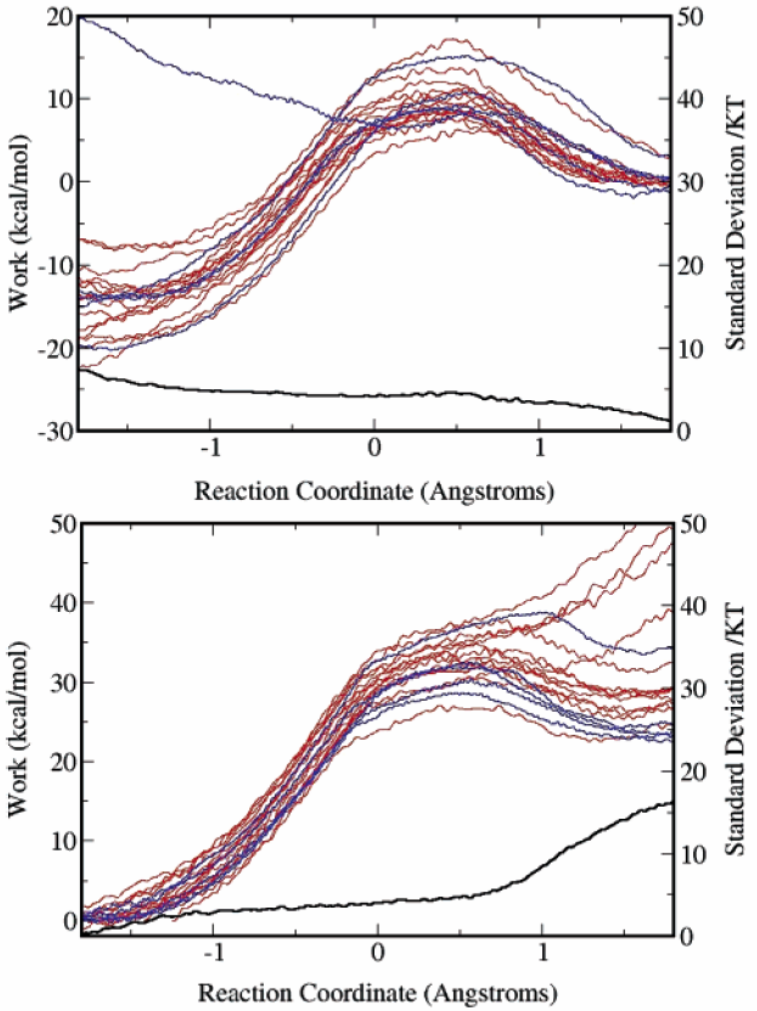


Figure 1. (top) Chorismate to prephenate work for the 20 runs (set 1: red, set 2: blue) and the standard deviation (thick black line). (bottom) Prephenate to chorismate work for the 20 runs (set 1: red, set 2: blue) and the standard deviation (thick black line).

A set of steered MD are performed. From the irreversible works the free energy change can be estimated

$$H(\mathbf{r},\lambda) = H_0(\mathbf{r}) + {}^1/_2 k\,[\lambda(r) - \lambda_0 - vt\,]^2$$

$$e^{-\beta\,\Delta G(\lambda)} = \langle\, e^{-\beta\, W(\lambda)} \,\rangle$$

Free energy profile

MSMD is more efficient than Umbrella Sampling for the user and is easily parallelized

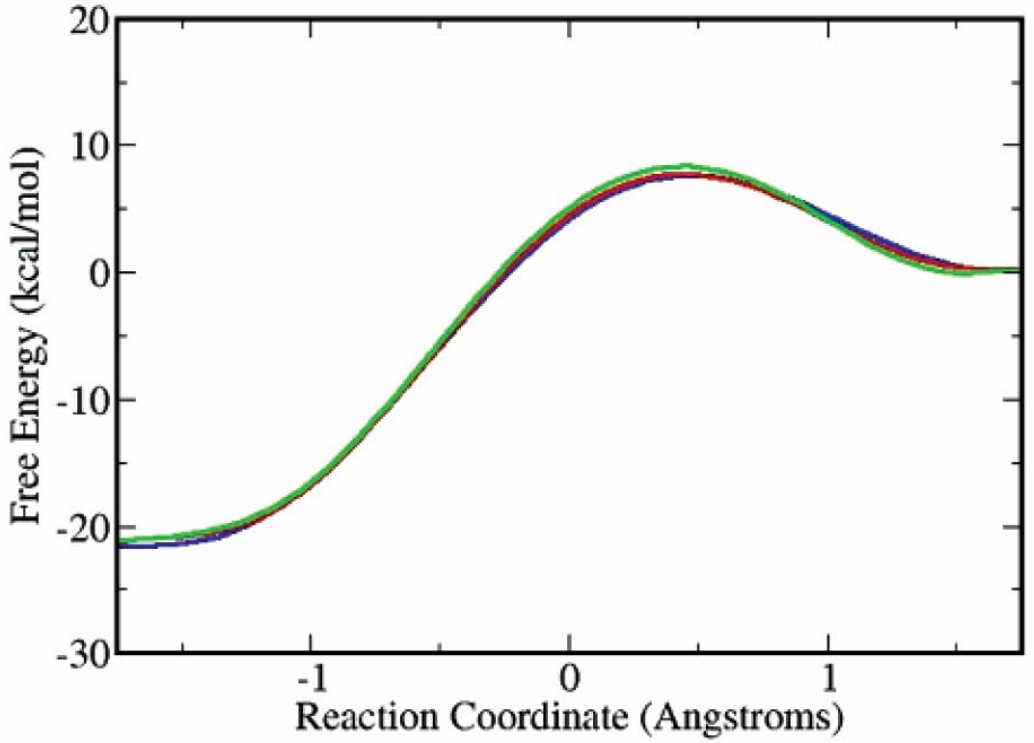


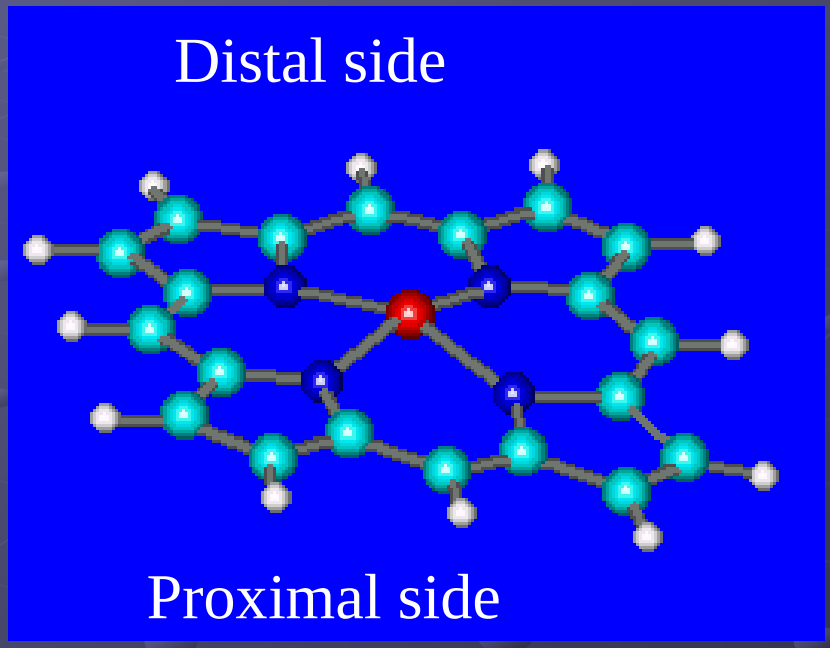
Figure 2. Free energy profile from chorismate ($\xi \approx 1.75$ Å) to prephenate ($\xi \approx -1.75$ Å), calculated using Jarzynski's equality (both forward and reverse data are used) for set 1 (red), for set 2 (green), and for umbrella sampling scheme (blue).

Multiple-Steering QM-MM Calculation of the Free Energy Profile in Chorismate Mutase A. Crespo, M.A. Marti, D.A. Estrin, A.E. Roitberg, J. Am. Chem. Soc. (2005), 127, 6940-6941.

Heme proteins

Ideal benchmarks for QM-MM

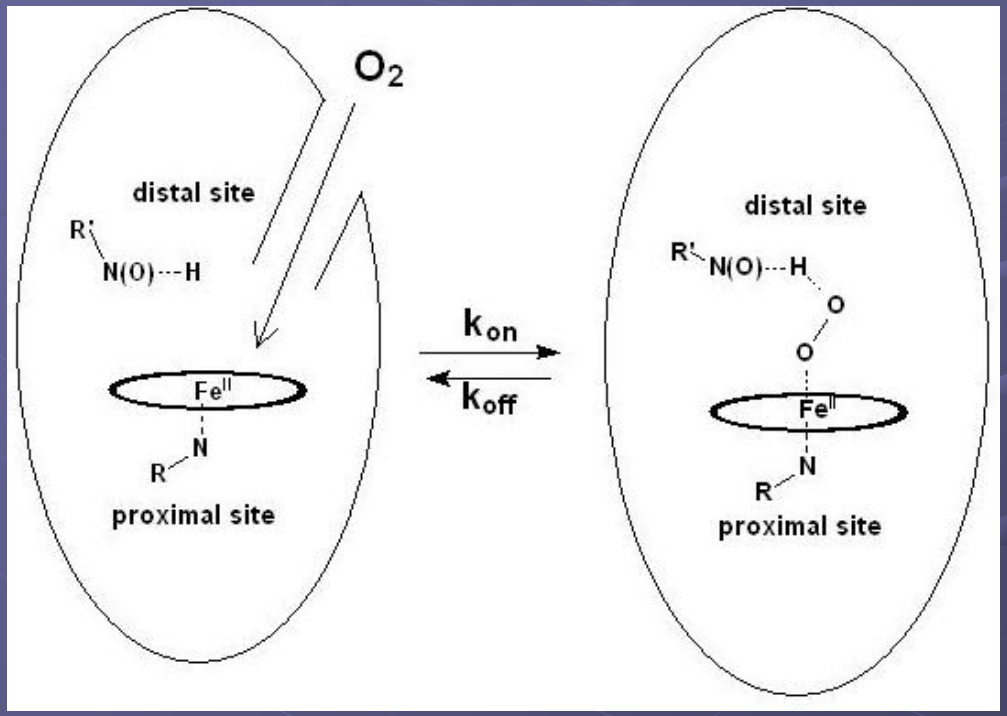
Active site: heme
Iron coordinated to a porphyrin



Very different roles:

- **O_2 transport**
- **Electron transfer**
- **Hormone biosynthesis**
- **Detoxification**

Dioxygen binding in Globins:
crucial process involved in transport and chemistry



O_2 exit(k_{off}): mainly related with bond breaking:

Distal (direct) effects
Proximal (indirect) effects

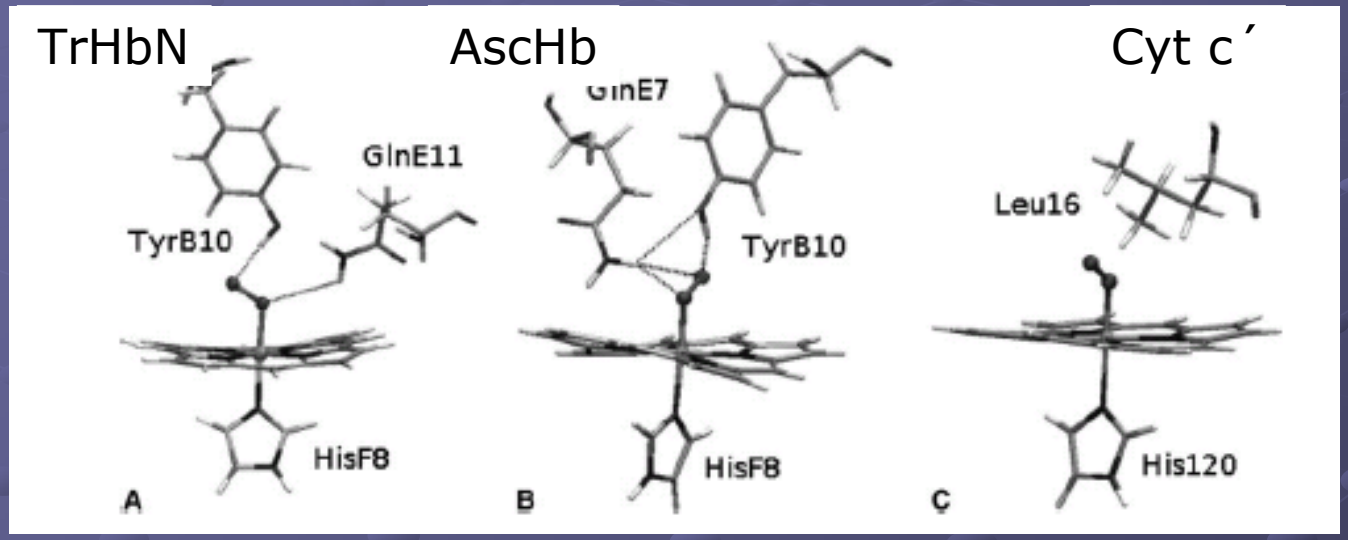
O_2 entry (k_{on}): mainly related with ligand migration

Classical MD simulations

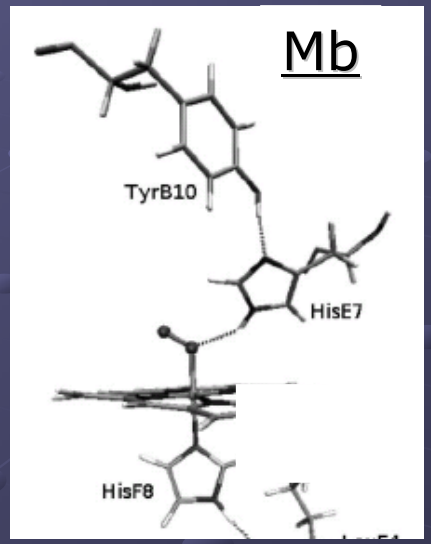
QM-MM calculations

Distal Effects:

O_2 acquires a negative charge upon binding
(electrostatic) H bonding and steric effects



37.2 kcal/mol 34.3 kcal/mol 8.4 kcal/mol



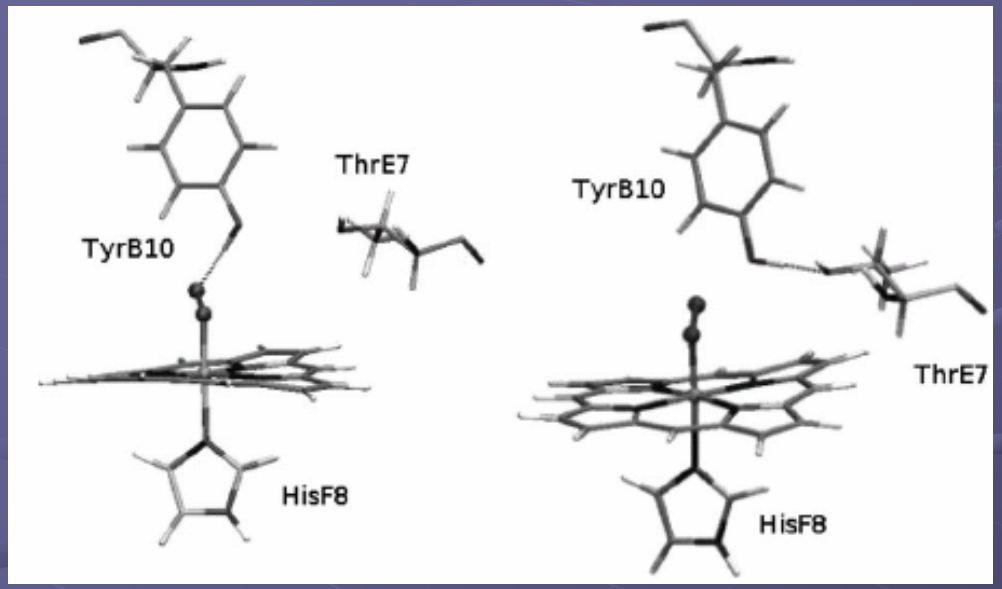
27 kcal/mol

Free heme:
22.0 kcal/mol

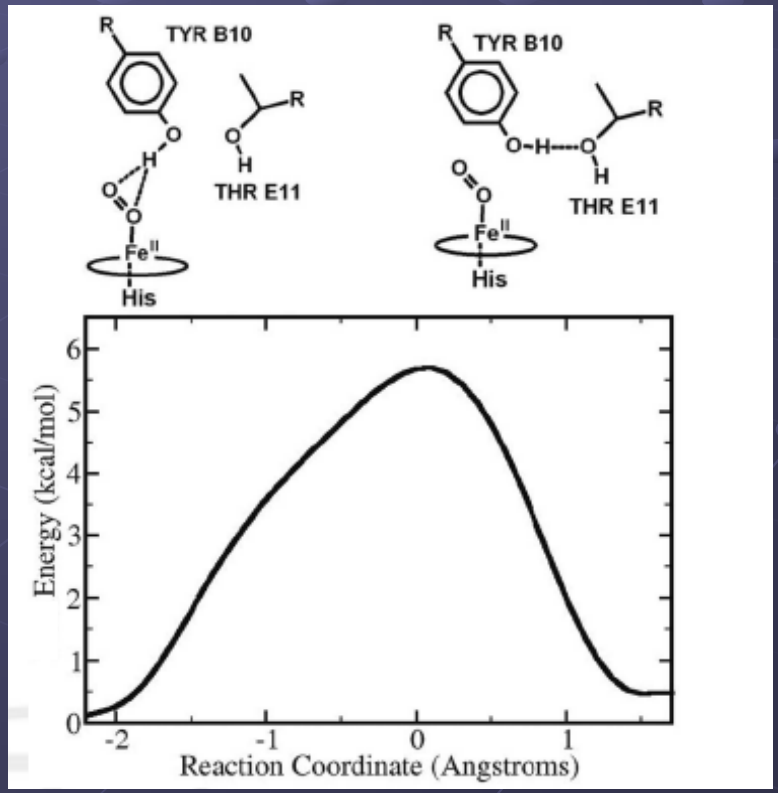
QM-MM optimizations,
DFT-PBE SIESTA level

Dioxygen affinity in heme proteins investigated by computer simulation M.A. Marti, A. Crespo, L. Capece, L. Boechi, D.E. Bikiel, D.A. Scherlis, **D.A. Estrin**, Journal of Inorganic Biochemistry, (2006), 100, 761-770.

Dynamical distal effects: multiple conformations CerHb case



QM-MM calculations, DFT SIESTA level and classical MD simulations



Two distinct distal sites modulate oxygen affinity in minihemoglobin of Cerebratulus Lacteus, M.A. Martí, D. Bikiel, A. Crespo, M. Nardini, M. Bolognesi, D.A. Estrin, Proteins (2006) 62, 641.

Other similar cases: LegHb, Hb P. caudatum Proteins, in press.

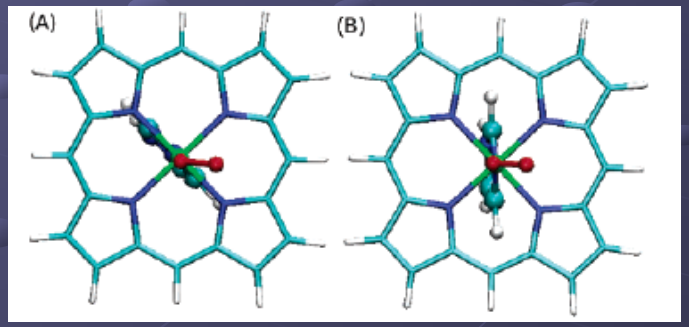
More subtle effects through the protein backbone: proximal effects

Try to analyze trends by constructing model systems in the first place:

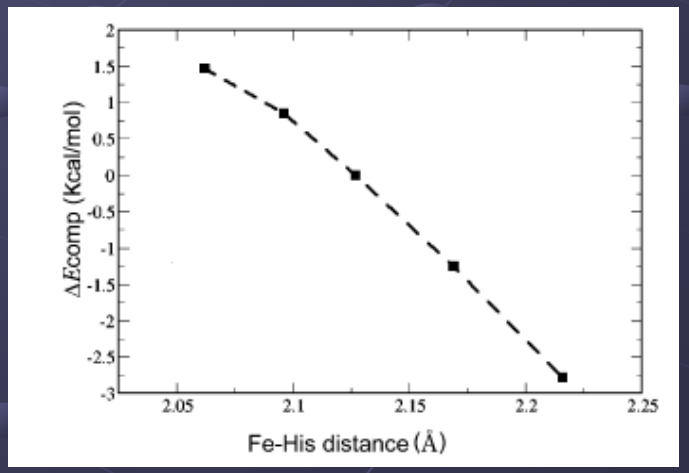
QM calculations,
DFT SIESTA level,
PBE, DZVP basis sets

- Histidine rotational position (staggered vs eclipsed)

21.4kcal/mol 24.5 kcal/mol



- Fe-Histidine bond distance

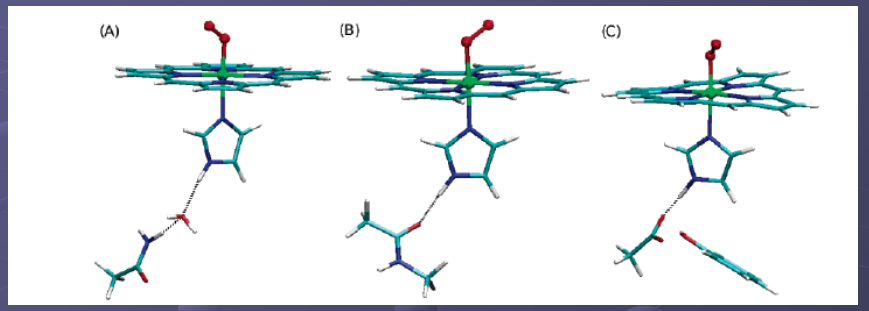


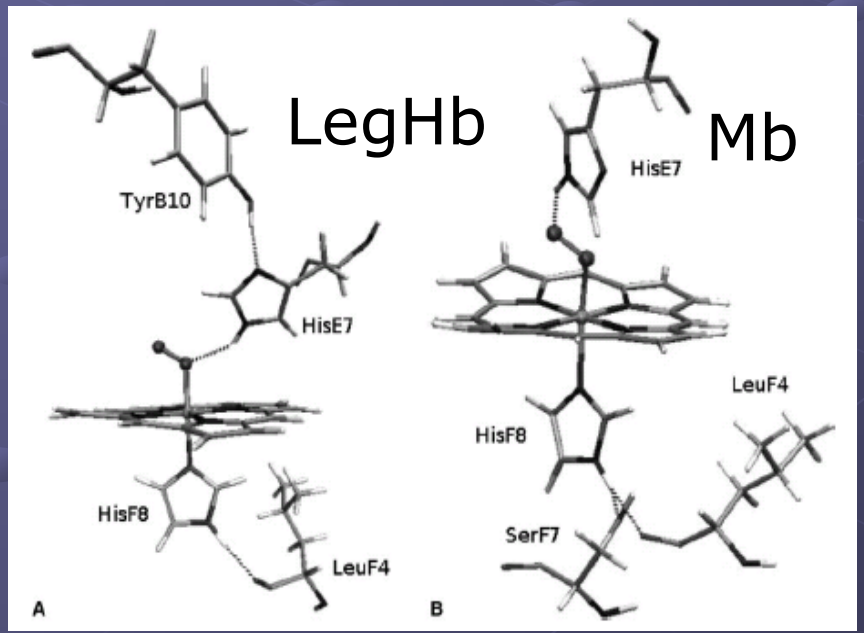
Charge transfer to histidine effects: intermediate degree enhances Fe-O_2 binding

	ΔE_{O_2} (kcal/mol)	Δq_{O_2} (e)	Δq_{prox} (e)
isolated heme−imidazole complex	22.2	−0.202	0.158
Fix-L[40]	22.0	−0.197	0.160
Hbβ[41]	24.7	−0.248	0.200
Hrp[42]	18.5	−0.259	0.325

Models of

FixL Hbβ HRP





Full protein: QM-MM calculation

LegHb: "strange" protein!

Heme protein oxygen affinity regulation exerted by proximal effects., L. Capece, M. A. Marti, A. Crespo, F. Doctorovich, D. A. Estrin, J. Am. Chem. Soc. (2006), 128, 12455.

NO detoxification by a truncated hemoglobin (trHbN) of *Mycobacterium Tuberculosis*

NO produced by the immune system, is inactivated by the oxygenated Hb:

toxic $\rightarrow$ innocuous

$$NO + Fe(II)\,O_2 \rightarrow Fe(III) + [NO_3]^-$$

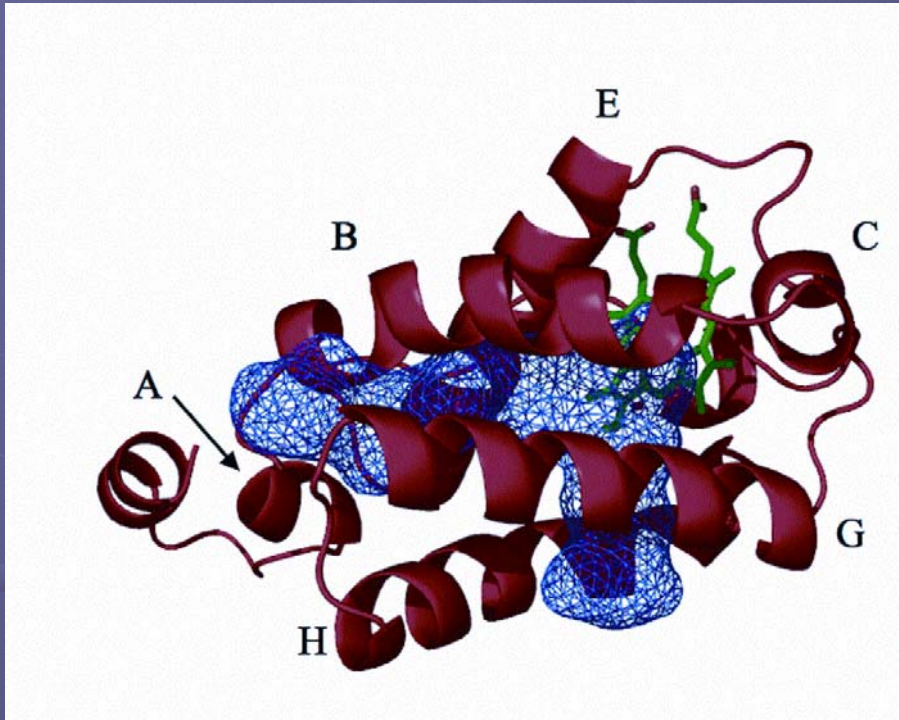
Relevant process in physiological (Hb,Mb) and pathological processes

Reaction mechanism:

1) O_2 migration (classical MD)
2) O_2 binding (QM-MM)
3) NO migration (classical MD)
4) Reaction of NO with O_2 (QM-MM)

$$NO + Fe(II)\,O_2 \rightarrow Fe(III) + [NO_3]^-$$

2 channel system proposed on the basis of x-ray results (Bolognesi' s group at Milano)
But how do O_2 and NO migrate?



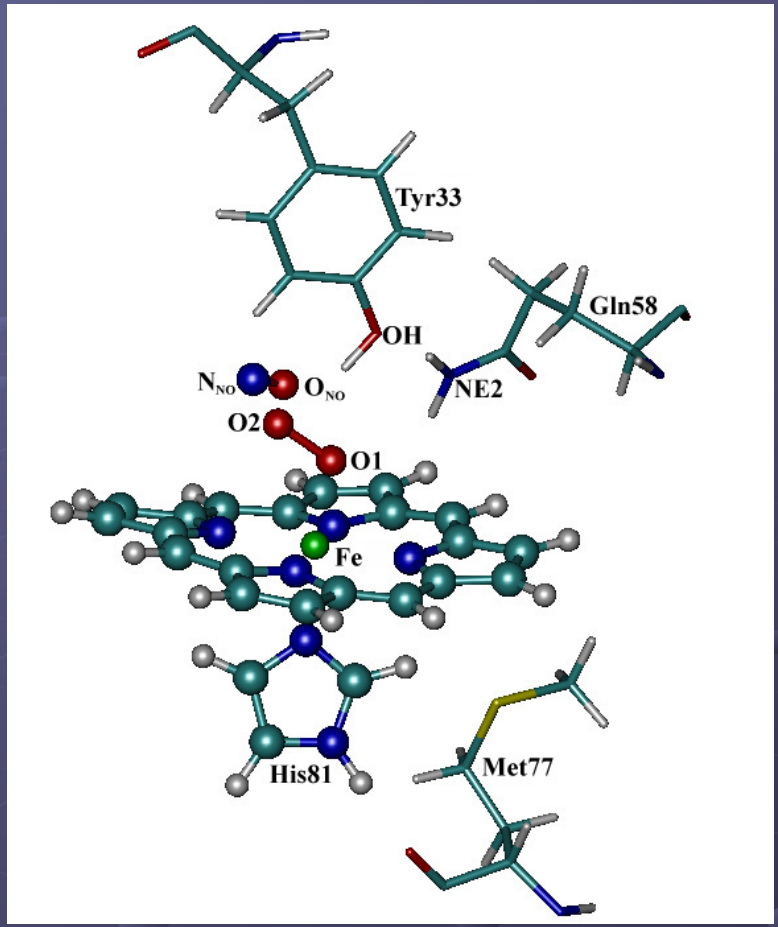
Ligand-induced dynamical regulation of NO conversion in Mycobacterium tuberculosis truncated-hemoglobin-N A. Bidon-Chanal, M. A. Martí, A. Crespo, M. Milani, M. Orozco, M. Bolognesi, F. J. Luque, D. A. Estrin, Proteins (2006), 64, 457-464.

Long (100 ns) classical MD (Amber 9) of oxy and deoxy proteins

Free energy profiles for ligand migration computed based on classical MD (no QM) (Our Jarzinski implementation available in Amber9)

What about the chemical processes?

QM-MM optimizations
Siesta PBE-DZVP/Amber level



Oxygen affinity

Hydrogen bond with OH of Tyr33 (B10)

isolated	Mb	TrHb	Tyr33→Phe33
-21.4	-25.0	-37.2	-34.3 kcal/mol

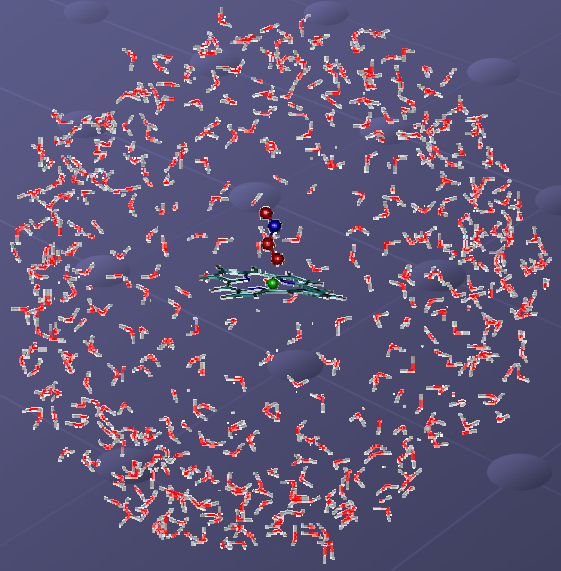
Effect of Tyr33→Phe33 mutation on k_{off} is reproduced. Affinity is large (consistent with the detoxification role).

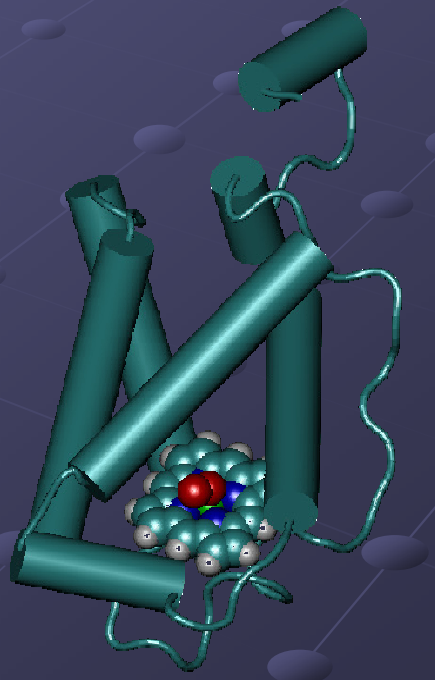
Chemical reaction:

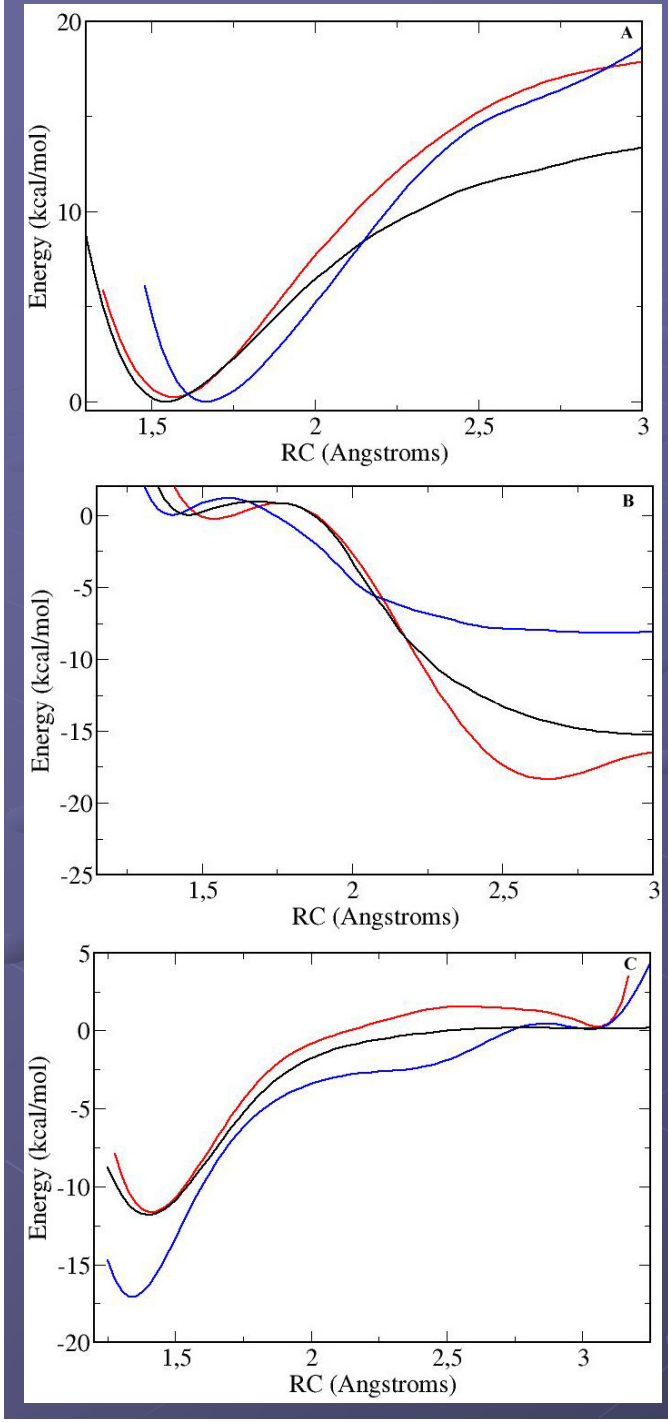
$Fe(II)\text{-}O_2 + NO \rightarrow Fe(III)\,[\text{-}OONO]$ (1) $\xi_1 = d(O_{O2}–N_{NO})$

$Fe(III)\,[\text{-}OONO] \rightarrow Fe(IV)=O + NO_2$ (2) $\xi_2 = d(O_{O1}–O_{O2})$

$Fe(IV)=O + NO_2 \rightarrow Fe(III)[\text{-}NO_3]$ (3) $\xi_3 = d(O_{O1}–N_{NO})$







Step 1

Step 2

Step 3

Vacuum
Water
Protein

Almost barrierless Reaction in protein is similar to that in water

Conclusions:

NO detoxification ability of trHbN is related to:

- Existence of an adequate ligand migration pathway
- High affinity for O_2 due to distal stabilization (H bonds)
- No significant protein effects on chemical reaction. Only spatial restriction of intermediates.

Theoretical Study of the Truncated Hemoglobin HbN: Exploring the Molecular Basis of the NO Detoxification Mechanism, A. Crespo, M.A. Martí, S.G. Kalko, A. Morreale, M. Orozco, J.L. Gelpi, F.J. Luque, D.A. Estrin J. Am. Chem. Soc. (2005), 127, 4433-4444.

- **Flaws:**

DFT at GGA level. Underestimates reaction barriers, problems with spin energetics of transition metals.

Sampling problems. In many cases optimizations are not enough!

- **Perspectives:**

GGA+ U; other functionals; making program more efficient to be able to sample more, PBC to treat solvation, etc……

In collaboration with:

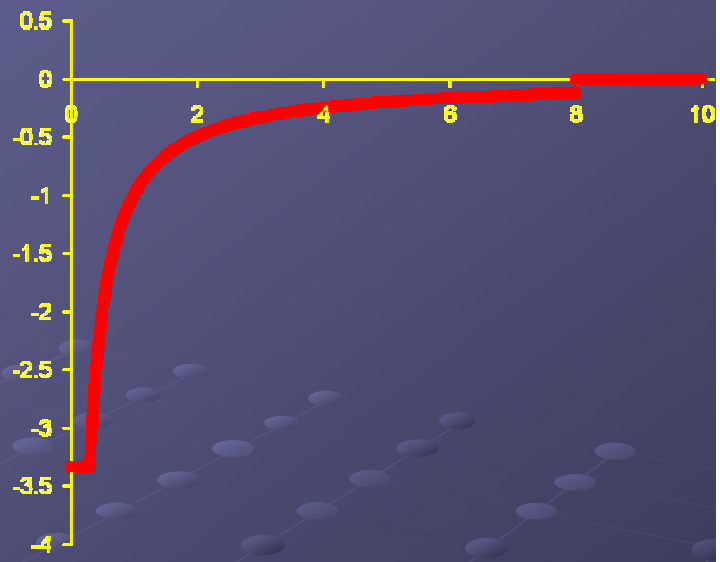
Buenos Aires' s Group

D. Elola,
D. A. Scherlis,
A. Turjanski,
M. L. Fernández,
A. Crespo,

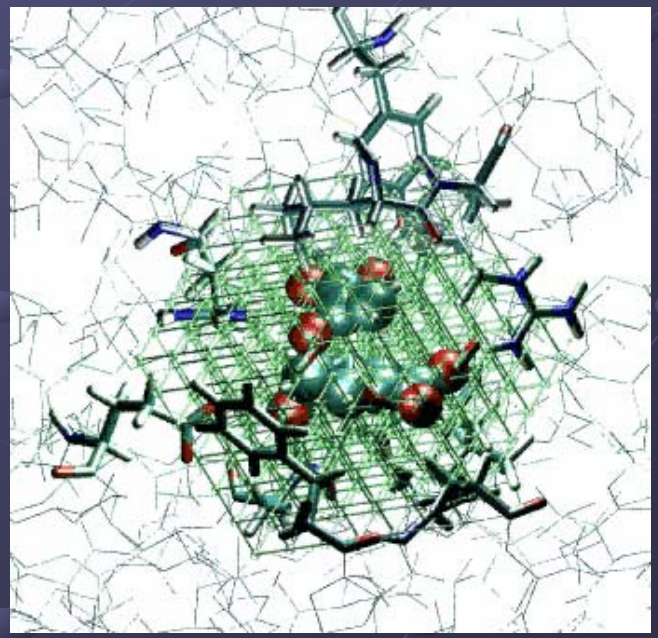
A. Nadra,
M. Martí,
L. Capece,
M. González Lebrero,
L. Perissinotti,
D. Bikiel,
L. Boechi,

J. Luque (U. Barcelona), P. Ordejón (ICMB, Barcelona), M. Bolognesi (U. Milano), A. Roitberg (U. Florida).

Charge density partial charges interaction is computed self consistently



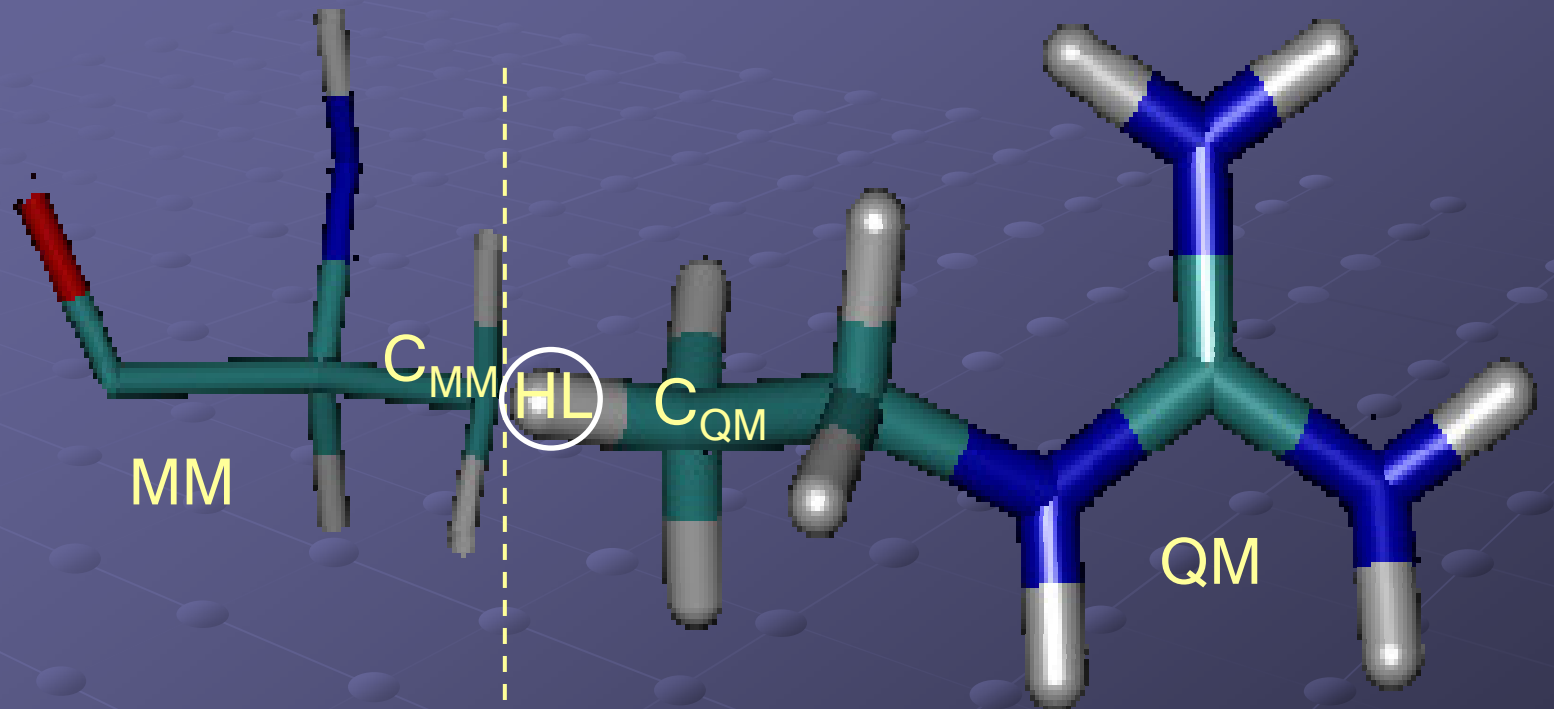
$$E = \begin{cases} \sum_{n=1}^{grilla} \sum_{i=1}^{MM} \dfrac{\rho(\mathbf{r}_n) q_i}{R_{cutQM}} & \left|\boldsymbol{\tau}_i - \mathbf{r}_n\right| \le R_{cutQM} \\ \sum_{n=1}^{grilla} \sum_{i=1}^{MM} \dfrac{\rho(\mathbf{r}_n) q_i}{\left|\boldsymbol{\tau}_i - \mathbf{r}_n\right|} & R_{cutQM} < \left|\boldsymbol{\tau}_i - \mathbf{r}_n\right| \le R_{cutQMMM} \\ 0 & \left|\boldsymbol{\tau}_i - \mathbf{r}_n\right| > R_{cutQMMM} \end{cases}$$



Generally: R_{cutQM} 0.2-0.3 Å

and $R_{cutQMMM}$ 8-10 Å.

Link atom: *frontier between QM and MM systems*



- C_{MM}-C_{QM} is broken and an H atom is added
- forces over H_L are divided, scaled, and added to C_{MM} and C_{QM}.
- classical terms involving C_{MM} are computed.