2^{nd} Penn State Bioinorganic Workshop
May/June 2012

Quantum Chemistry at Work

(1) Introduction, General discussion, Methods, Benchmarking

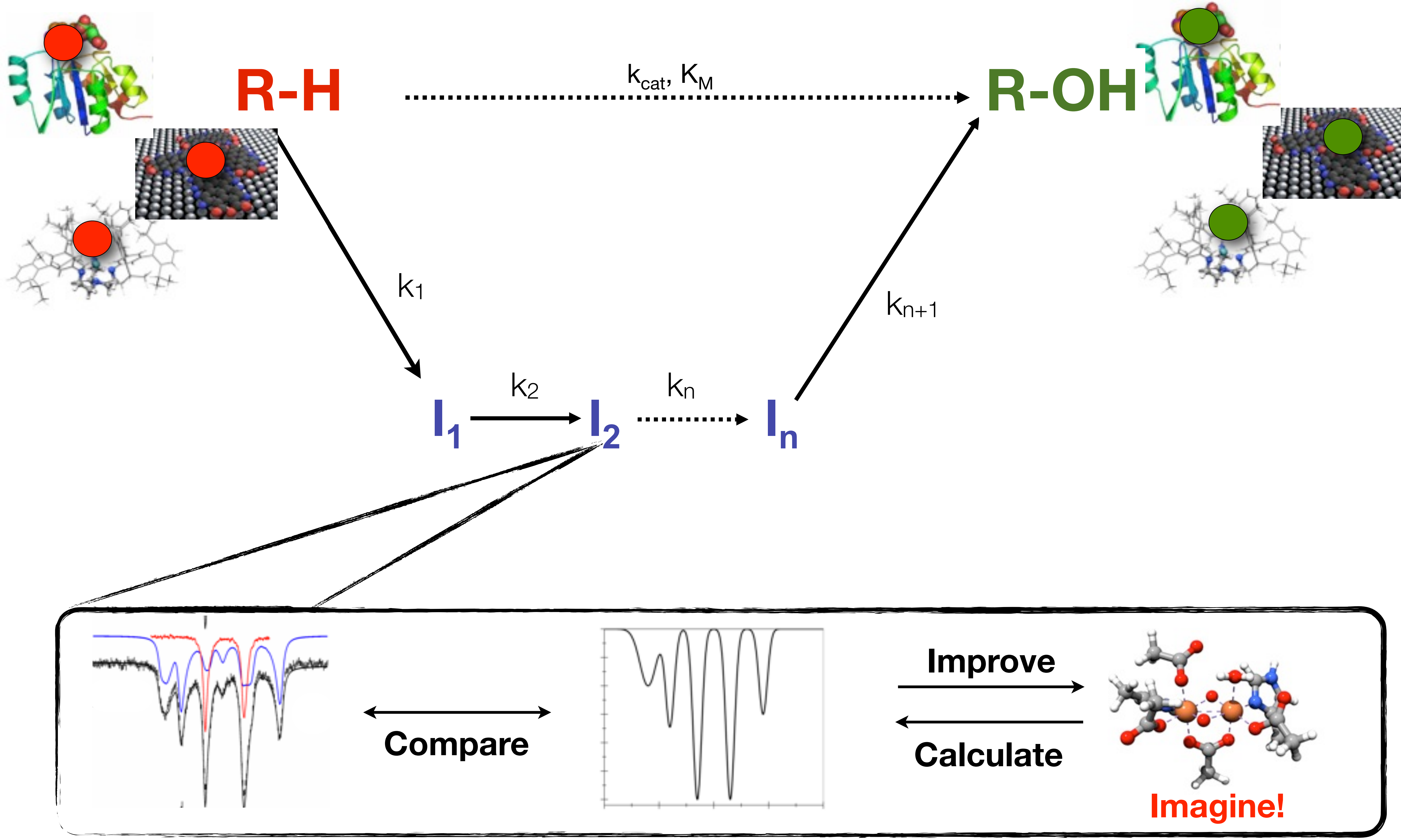
Frank Neese

Max-Planck Institut für Bioanorganische Chemie
Stiftstr. 34-36
D-45470 Mülheim an der Ruhr
Germany

Frank.Neese@mpi-mail.mpg.de

MAX-PLANCK-INSTITUT
FÜR BIOANORGANISCHE CHEMIE

Mechanisms-Intermediates-Spectra-Calculations



R-H $\xrightarrow{k_{cat},\ K_M}$ **R-OH**

k_1

$I_1 \xrightarrow{k_2} I_2 \xrightarrow{k_n} I_n$

k_{n+1}

Compare

Improve

Calculate

Imagine!

The ORCA Project

Hartree-Fock

RHF,UHF,ROHF,CASSCF
Direct, Semidirect, Conventional,
RI-Approx., Newton-Raphson

Semiempirical

INDO/S,MNDO,AM1,PM3,NDDO/1

Density Functional

LDA, GGA, Hybrid Functionals
Double hybrid functionals,
RI-Approx., Newton-Raphson
RKS,UKS,ROKS

Electron Correlation

MP2/RI-MP2
CCSD(T),QCISD(T),CEPA,CPF
(all with and without RI, Local)
MR-MP2, MR-MP3, MR-MP4(SD)
MR-CI, MR-ACPF, MR-AQCC

Join ~10,000 users
FREE Download
http://www.thch.
uni-bonn.de /tc/orca

Excited States

TD-DFT/CIS+gradients
MR-CI/DDCI/SORCI

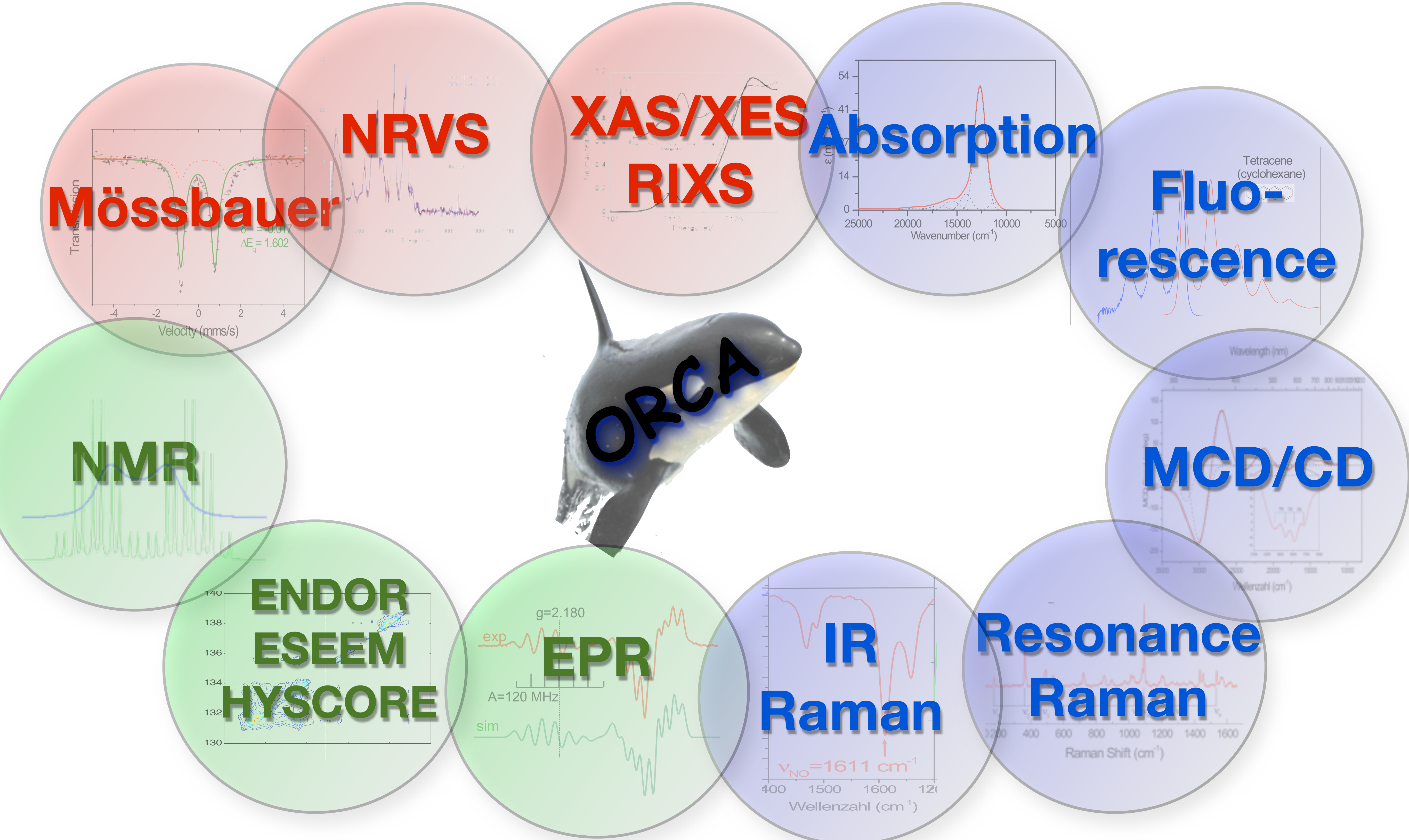
Relativistic Methods

1st-5th Order Douglas-Kroll-Hess
Zero'th Order Regular Approximation (ZORA)
Infinite Order Regular Approximation (IORA)
Picture Change Effects, All electron basis sets,
(Effective core potentials)

Molecular Properties

Analytical Gradients(HF,DFT,MP2) + Geometries + Trans. States
Polarizabilities, Magnetizabilities (Coupled-Perturbed HF/KS)
COSMO Solvation Model Throughout
IR, Raman and Resonance Spectra (Numerical Frequencies)
EPR-Parameters (g,A,D,J,Q)
Mössbauer-Parameters $(\delta,\Delta E_Q)$
ABS,CD,MCD Spectra
Population Analysis, NBOs, Localization, Multipole Moments,...

Advanced Theoretical Spectroscopy with ***ORCA***



Mössbauer

NRVS

XAS/XES
RIXS

Absorption

Fluo-rescence

NMR

MCD/CD

ENDOR
ESEEM
HYSCORE

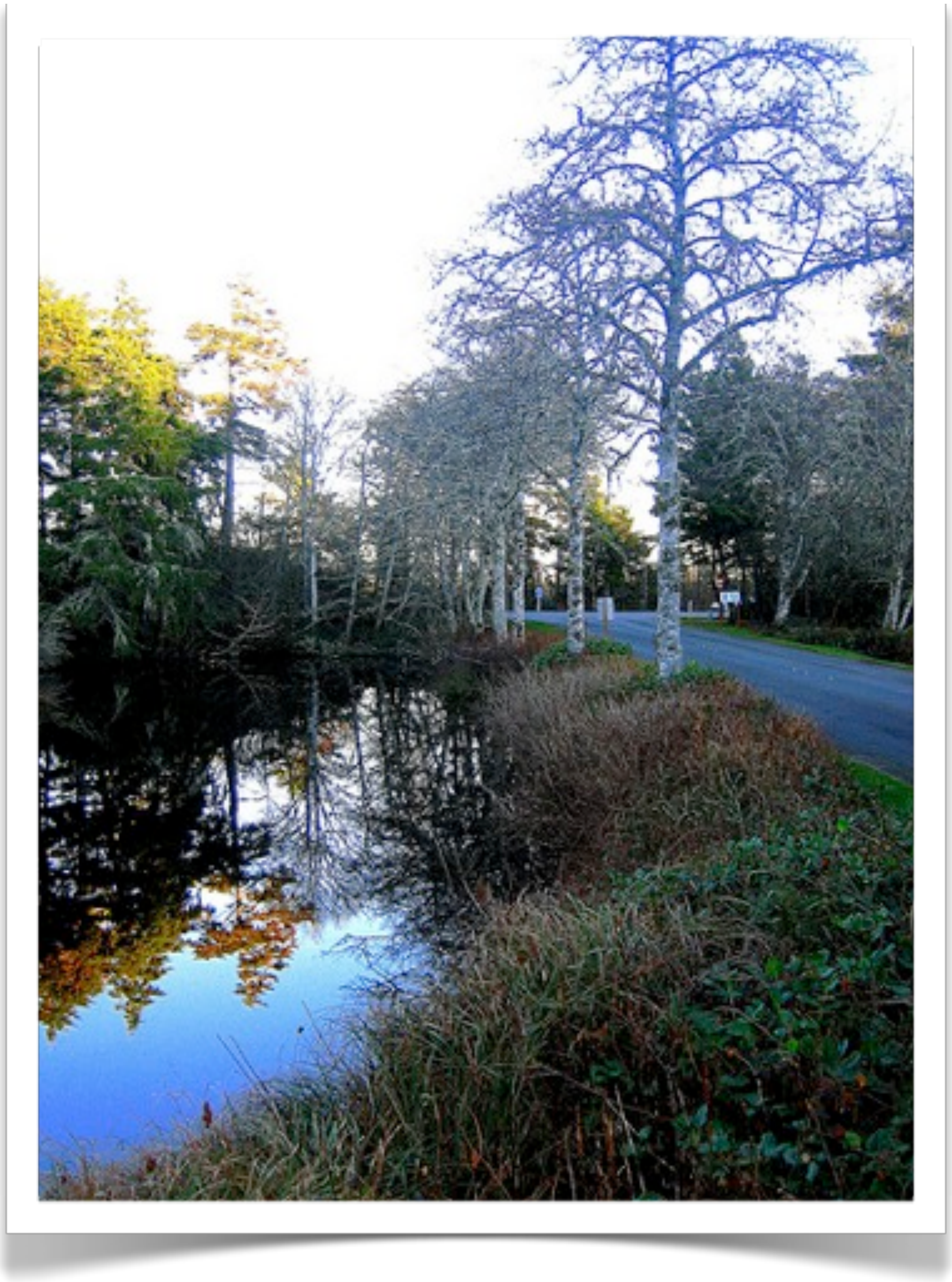
EPR

IR
Raman

Resonance
Raman

ORCA

Designing a Computational Chemistry Project



Many pathways to happiness ...

... Very little (if any) generally agreed-upon wisdom ...

... Some things to think about ...

Designing a Computational Chemistry Project

(1) Have a well defined question!

- ✓ Are you doing a collaboration with an experimental group?
  - *What do they know?*
  - *What do they want to know?*
  - *Can you provide this information?*
- ✓ Are you doing a collaboration with another theoretical group?
  - *Are you using compatible methods?*
  - *Are you exchanging data?*
  - *Do you agree on file and data formats?*
- ✓ Are you benchmarking theoretical methods?
  - *What is your reference data?*
  - *Experiment?*
  - *Higher level calculations?*
- ✓ Are you aiming at reproducing experiments or predicting the outcome of possible experiments?
- ✓ Are you puzzled by seemingly conflicting observations?
- ✓ Are you trying to bring order and meaning to a series of observations?
- ✓ ... there is an infinite number of good reasons to do a computational chemistry study

Designing a Computational Chemistry Project

(2) Be aware of the experimental and theoretical literature

✓ What is known for sure experimentally?
✓ Are there error bars on experimental numbers?
✓ What has been speculated upon and is not known for sure?

... it is neither forbidden nor a bad idea to talk to experimentalists of all flavors

✓ Are there other experiments outside your field of specialty that shine more light on the subject?

... even when your are trying to answer an NMR question, there might be useful insights from, e.g. PES or IR or ....

✓ What theoretical work has been published prior to your study?

... Theoretical work is not automatically good if a big computer has worked for a long time and not automatically bad if it is performed at a (seemingly) lower level of theory.

Designing a Computational Chemistry Project

(3) What is the timescale for your project?

- ✓ Very important question when working with experimentalists. They want to have an answer soon and not in one or two years!
- ✓ Very important question when you want to get out of gradschool!

(4) How accurate does your result have to be to be useful?

- ✓ Ideally we would always solve the relativistic many particle Schrödinger equation combined with quantum dynamics for a the entire system including its environment at finite temperature and inclusion of radiative corrections ...

  .... but we cannot do that

  .... neither do we *need* to do that to answer many useful chemical questions

Designing a Computational Chemistry Project

(5) Choose the right tools to approach the problem

- ✓ Choose a theoretical method:
  - Density functional theory?
  - MP2?
  - More accurate *ab initio*?
  - Are benchmark data available for the kind of problem you are studying?
- ✓ Are relativistic effects important?
- ✓ Are dynamics important?
- ✓ Are solvent effects important?
- ✓ Make a concious choice of basis set

(6) Choose the right software

- ✓ Which programs do the things best that you want to do?
- ✓ Are they available?
- ✓ Are they user friendly enough?
- ✓ What do they cost?

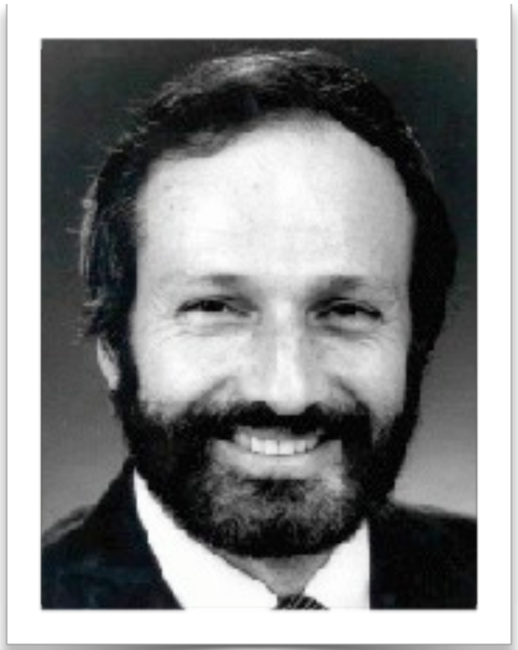
.... investing a little time to learn something new may save you a lot of time later if you don‘t just stick to what you have done all the time

Designing a Computational Chemistry Project

(7) Talk chemistry to chemists

- ✓ Try to go beyond tabulating numbers!
- ✓ Be interested in the electronic structure of the systems that you are studying
  - Analyze orbitals, densities, states, populations
  - If you feel comfortable explore NBOs, AIM, ELF, ...
  - **BUT**
- ✓ Try to interpret the results in terms of a language familiar to the chemists working in your branch of chemistry (e.g. Hückel theory, ligand field theory, ...)
- ✓ Try to understand what is characteristic for the specific molecule you are studying vs the class of compounds that are under investigation

„... the best calculations are those that, after the fact, I realize I wouldn‘t have needed“



(Mike Zerner)

Ernest R Davidson

„... Computers don‘t solve problems - people do!“

Designing a Computational Chemistry Project

(8) Separate fact from fiction

- ✓ Be clear on what is an experimental observable and what is not.
    - Observables have unambigous values and can be measured (e.g. spectroscopic transition energies and intensities, thermondynamic or kinetic quantities)
    - The total energy is, in principle, an observable. In practice it hardly is.
- ✓ Non-observable properties are „interpretation aids“ that help us understand and be creative. They have no unambigous definition
    - Partial charges, spin populations, ...
- ✓ Quite typically experimentally working colleagues are more interested in the non-observable properties. Be careful in explaining the difference.

Disgression: If you choose to do so, you can engage in vicious fights about non-observable properties (e.g. the interpretation of the rotational barrier in the ethane molecule). However, be aware that neither of you is „more right“ - it is useless to argue whether red or green is more beautiful. It subjective! At the end of day it matters what helps us designing new experiments, new molecules, new methods. Different people are inspired by different pictures. These pictures are good, they are necessary for chemistry - but there is no objective truth in them.

Designing a Computational Chemistry Project

(9) Seek feedback from experiment

- ✓ Calculations are becoming more accurate and reliable. Yet, there is every reason to *not* just hit the enter button and believe everything the computer tells you.
  - Your theoretical method may be dead wrong (happens!)
  - Your system setup might be inadequate
  - You might have converged to a wrong electronic state
  - You might have converged to a wrong minimum on the PES
  - You might have missed alternative reaction pathways
  - ...
- ✓ In all these situations comparing theory and experiment is a highly useful way to strengthen everybodies (e.g. yourself, advisor, experimental collaborator, referee, competitor) faith in your results by comparing as many observables to experiment as possible. This may include:
  - Geometries
  - Thermodynamic data (reaction energies, isomerization energies, ...)
  - Kinetic data (rates, isotope effects)
  - Spectroscopic properties of all kind
  - ...

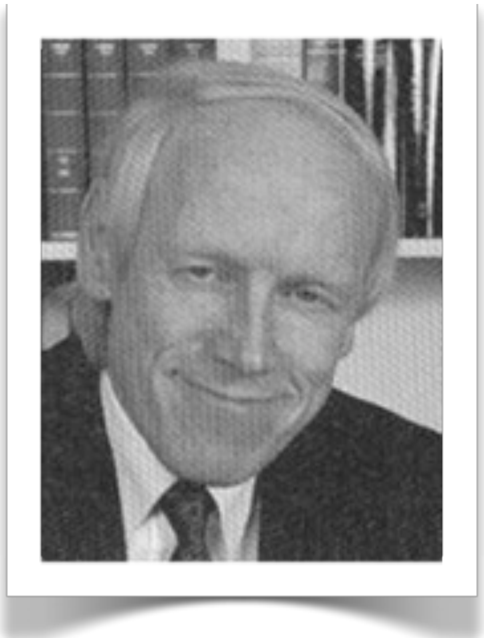
... we will come back to this in lecture 3 ...

Designing a Computational Chemistry Project

(10) Avoid „buffet theory“

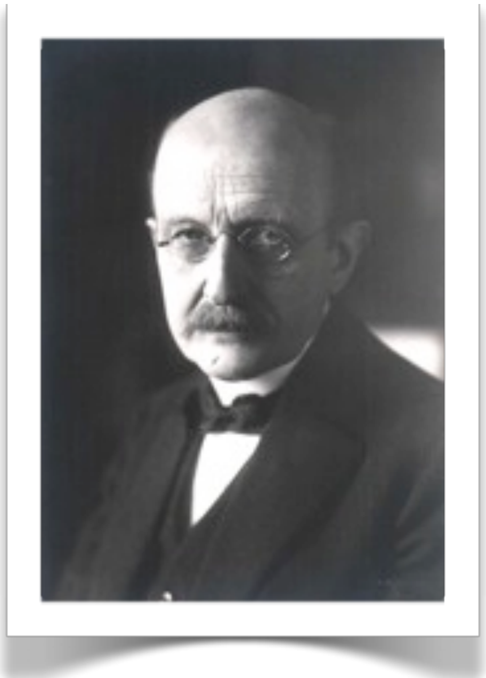
- ✓ *„this looks good so I will put it on the plate, this looks bad so I will leave it on the table*“
  - ... a well worked out negative result might be as useful as a fantastic positive result.

Remember ....



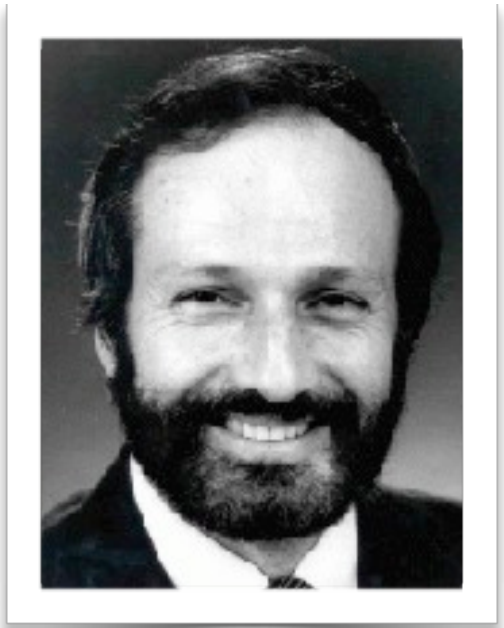
„Experimentalists are working for eternity - their measurements must remain true forever. Theoreticians are working for tomorrow - the interpretation of the facts may (and likely will) change over time“

(Günther Wächsterhäuser)



„The only source of knowledge is experiment. The rest is poetry, imagination“

(Max Planck)



„Let‘s face it - somebody will do a better calculation tomorrow.“

(Mike Zerner)

Practical Aspects of Working with **ORCA**

The Computational Environment

In order to run calculations with **ORCA**, three things are necessary:

1. Installation of the *ORCA* program
2. Using a text-editor to specify the calculation details, i. e. the ***input file***
3. Running the *ORCA* program (in a cluster environment possibly controlled by a batch system)

NOTE: **ORCA** is available for all popular platforms:

- ★ Windows,
- ★ MacOS,
- ★ Linux

The Computational Environment

- ★ ***ORCA*** is available for the popular Windows, Mac OS X, and Linux platforms.

- ★ ***ORCA*** is distributed as an ***archive*** for all platforms. In this archive resides a **directory** with all **executables**. There are plenty of programs for extracting files from archives on all platforms (gzip, tar, zip/unzip,...).

- ★ On the follwing slides it will be shown how to run a calculation on Windows Vista, Mac OS X, and Linux.

- ★ It is assumed, that the orca executables reside in a directory ‘**orca**’

Tasks to be Performed

- Calculation of single point energies
  - Hartree-Fock calculations: RHF, UHF and ROHF
  - Density functional calculations
- Optimization of molecular structures
  - Equilibrium geometries
  - Transition states and reaction rates
- Calculation of vibrational frequencies
  - Characterization of stationary points
  - Thermodynamic properties
  - Vibrational spectra
- Calculation of ground state properties
  - Looking at charge distributions and orbitals
  - IR+Raman spectra
  - NMR spectra
  - EPR spectra and exchange couplings
  - Mössbauer spectra
- Calculation of excited states and their properties
  - DFT calculation of absorption and CD spectra
  - *Ab initio* calculation of absorption and CD spectra
  - Advanced: Resonance-Raman, X-Ray absorption, forbidden transitions, MCD, Excited state geometry optimizations, ...

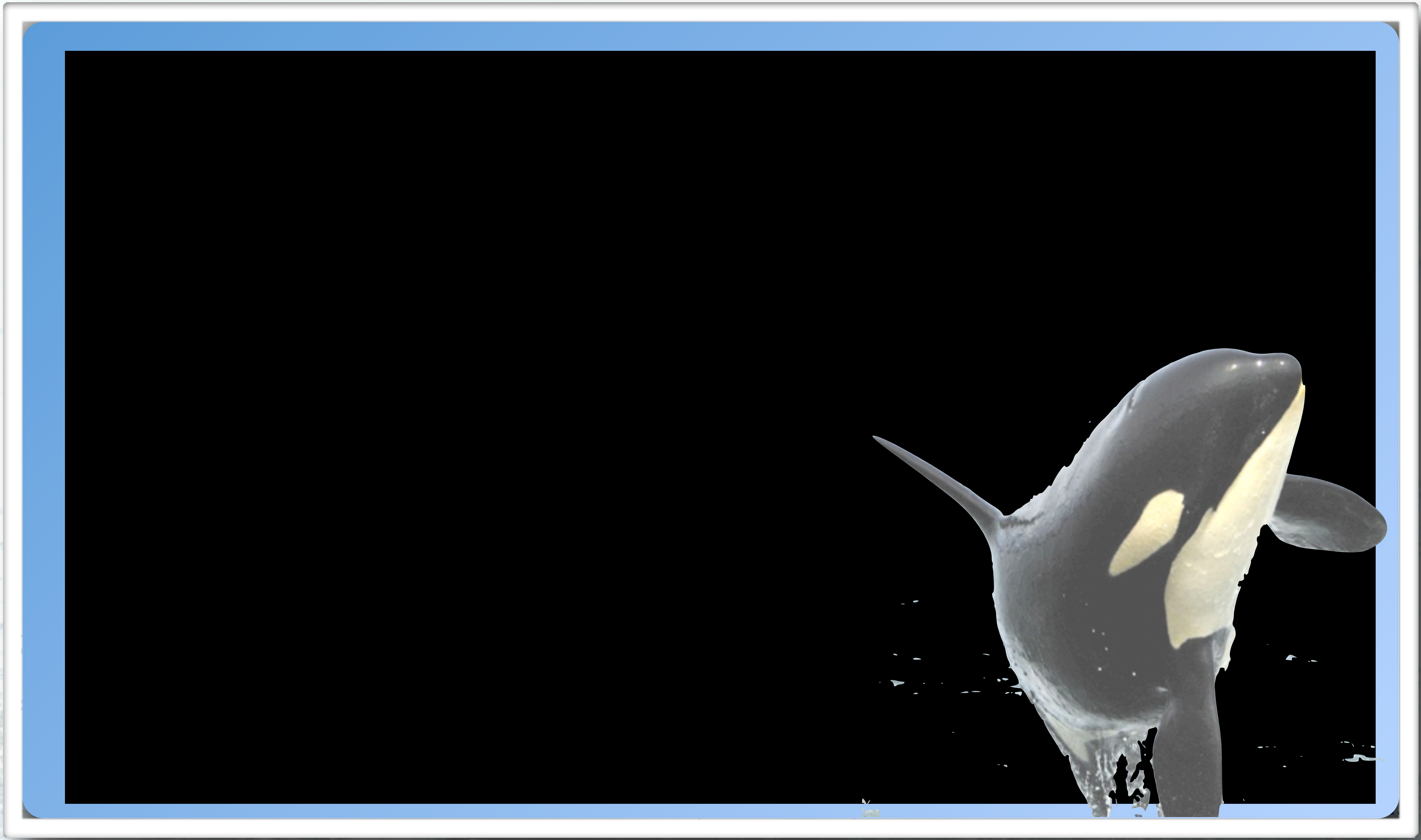
Philosophy of the **ORCA** Project

General goal: Create a powerful tool to allow connection between theory and experiment. → Observables! (Spectra, Structures, Energies)

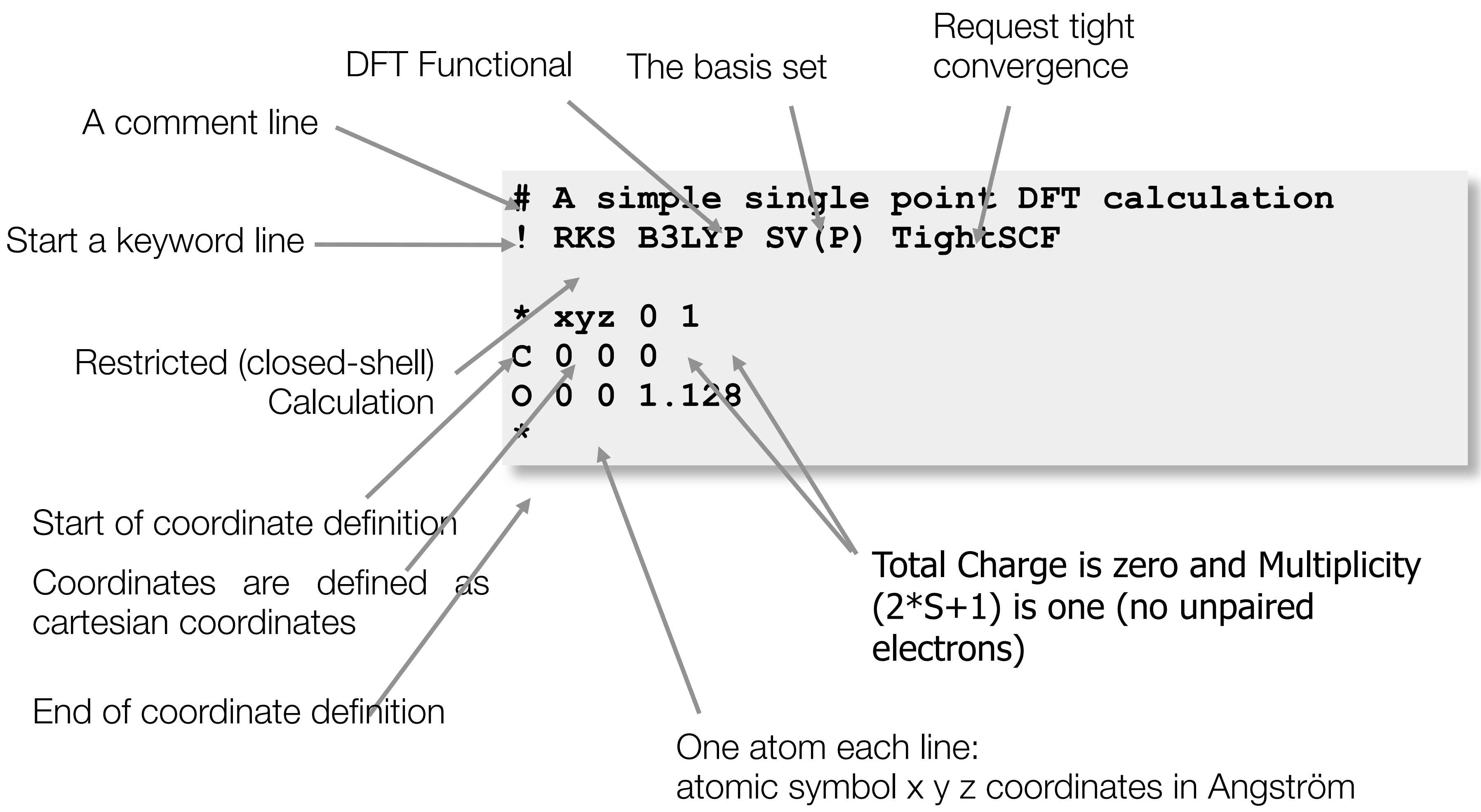
Design principles:

1. The program should be as flexible as possible
2. The program should be as efficient as possible (parallel, efficient algorithms, BLAS libraries)
3. The program should be as comprehensive as possible
4. The program should be as user friendly as possible
5. The program should be easily extendable (highly modular)
6. The source code should be as clean and well structured as possible (C++ rather than Fortran).
7. The program should be platform independent

The ORCA Project



A First **ORCA** Job



```
# A simple single point DFT calculation
! RKS B3LYP SV(P) TightSCF

* xyz 0 1
C 0 0 0
O 0 0 1.128
*
```

A comment line

Start a keyword line

DFT Functional

The basis set

Request tight convergence

Restricted (closed-shell) Calculation

Start of coordinate definition

Coordinates are defined as cartesian coordinates

End of coordinate definition

One atom each line:
atomic symbol x y z coordinates in Angström

Total Charge is zero and Multiplicity (2*S+1) is one (no unpaired electrons)

Output of Single Points

We first echo the input file and some references to the basis sets used. Then you get information on the job-type, the input coordinates in various formats and the basis set

```
                       ****************************
                       * Single Point Calculation *
                       ****************************

---------------------------------
CARTESIAN COORDINATES (ANGSTROEM)
---------------------------------
  C      0.000000    0.000000    0.000000
  O      0.000000    0.000000    1.130000

---------------------
BASIS SET INFORMATION
---------------------
There are 2 groups of distinct atoms

 Group   1 Type C   : 7s4p1d contracted to 3s2p1d pattern {511/31/1}
 Group   2 Type O   : 7s4p1d contracted to 3s2p1d pattern {511/31/1}
```

Next the one-electron integrals are calculated (and perhaps also the two-electron integrals if **conv** if requested for “conventional SCF”)

```
-------------------------------------------------------------------------
                        ORCA GTO INTEGRAL CALCULATION
-------------------------------------------------------------------------

                      BASIS SET STATISTICS AND STARTUP INFO

 # of primitive gaussian shells          ...   24
 # of primitive gaussian functions       ...   48
 # of contracted shell                   ...   12
 # of contracted basis functions         ...   28
 Highest angular momentum                ...    2
 Maximum contraction depth               ...    5
 Integral threshhold             Thresh  ... 1.000e-010
 Primitive cut-off               TCut    ... 3.000e-012
```

Next the SCF program is taking over and commences with giving all details about the SCF settings

```
------------
SCF SETTINGS
------------
Hamiltonian:
 Density Functional     Method          .... DFT(GTOs)
 Exchange Functional    Exchange        .... B88
   X-Alpha parameter    XAlpha          ....  0.666667
   Becke's b parameter  XBeta           ....  0.004200
 Correlation Functional Correlation     .... LYP
 LDA part of GGA corr.  LDAOpt          .... VWN-5
 Gradients option       PostSCFGGA      .... off
 Hybrid DFT is turned on
   Fraction HF Exchange ScalHFX         ....  0.200000
   Scaling of DF-GGA-X  ScalDFX         ....  0.720000
   Scaling of DF-GGA-C  ScalDFC         ....  0.810000


General Settings:
 Integral files         IntName         .... JOB-01
 Hartree-Fock type      HFTyp           .... RHF
 Total Charge           Charge          ....    0
 Multiplicity           Mult            ....    1
 Number of Electrons    NEL             ....   14
 Basis Dimension        Dim             ....   28
 Nuclear Repulsion      ENuc            ....     22.4778902655 Eh


Convergence Tolerance:
 Energy Change          TolE            .... 1.000e-008 Eh
 Max Density Change     TolMaxP         .... 1.000e-007
 RMS Density Change     TolRMSP         .... 1.000e-008
 DIIS Error             TolErr          .... 1.000e-007


Diagonalization of the overlap matrix:
Smallest eigenvalue                        ... 1.911e-002
Time for diagonalization                   ...    0.010 sec
Time for construction of square roots      ...    0.030 sec
Total time needed                          ...    0.040 sec
```

The integration grid is produced and the initial guess performed

```
-------------------
DFT GRID GENERATION
-------------------

General Integration Accuracy     IntAcc      ...   4.010

Total number of grid points                  ...     5057

------------------------------
INITIAL GUESS: MODEL POTENTIAL
------------------------------
Loading Hartree-Fock densities                         ... done
....etc
Now organizing SCF variables                           ... done
                      ------------------
                      INITIAL GUESS DONE
                      ------------------
```

Now we are ready to start the SCF iterations

```
--------------
SCF ITERATIONS
--------------
               ***  Starting incremental Fock matrix formation  ***
                               ***Turning on DIIS***

                         ----------------------------
                         !        ITERATION     0   !
                         ----------------------------
   Total Energy        :    -112.951951547431 Eh
   Energy Change       :    -112.951951547431 Eh
   MAX-DP              :       0.674569966353
   RMS-DP              :       0.053899776162
   Actual Damping      :       0.0000
   Int. Num. El.       :    14.00002071 (UP=    7.00001035
   Exchange            :   -10.98287769
   Correlation         :    -0.58429246
   DIIS-Error          :       0.450135783168
```

Which hopefully eventually converge and:

```
          *****************************************************
          *                     SUCCESS                       *
          *           SCF CONVERGED AFTER  11 CYCLES          *
          *****************************************************
```

The total energy and its components are printed

```
----------------
TOTAL SCF ENERGY
----------------

Total Energy       :         -113.17273851 Eh          -3079.50944 eV

Components:
Nuclear Repulsion  :           22.47789027 Eh            611.63913 eV
Electronic Energy  :         -135.65062877 Eh          -3691.14856 eV

One Electron Energy:         -198.09397475 Eh          -5390.27572 eV
Two Electron Energy:           62.44334598 Eh           1699.12715 eV

Virial components:
Potential Energy   :         -225.62136955 Eh          -6139.31540 eV
Kinetic Energy     :          112.44863104 Eh           3059.80596 eV
Virial Ratio       :            2.00643945
```

Next you get the orbital energies printed:

```
----------------
ORBITAL ENERGIES
----------------

  NO   OCC          E(Eh)            E(eV)
   0   2.0000     -19.243280      -523.6231
   1   2.0000     -10.301611      -280.3141
   2   2.0000      -1.152619       -31.3636
   3   2.0000      -0.568518       -15.4698
   4   2.0000      -0.476766       -12.9731
   5   2.0000      -0.476766       -12.9731
   6   2.0000      -0.373742       -10.1698
   7   0.0000      -0.026298        -0.7156
   8   0.0000      -0.026298        -0.7156
   9   0.0000       0.221613         6.0303
  10   0.0000       0.400346        10.8937
  11   0.0000       0.457384        12.4458
  12   0.0000       0.457384        12.4458
...
```

Now comes the “soft science” (e.g. population analysis of the SCF density matrix

```
              ********************************
              * MULLIKEN POPULATION ANALYSIS *
              ********************************

-----------------------
MULLIKEN ATOMIC CHARGES
-----------------------
   0 C :    0.016669
   1 O :   -0.016669
Sum of atomic charges:   -0.0000000

--------------------------------
MULLIKEN REDUCED ORBITAL CHARGES
--------------------------------
  0 C s       :     3.834569  s :     3.834569
      pz      :     1.009723  p :     2.073008
      px      :     0.531642
      py      :     0.531642
      dz2     :     0.027088  d :     0.075754
      dxz     :     0.024333
      dyz     :     0.024333
      dx2y2   :     0.000000
      dxy     :     0.000000
  1 O s       :     3.737862  s :     3.737862
      pz      :     1.381818  p :     4.256855
      px      :     1.437518
      py      :     1.437518
      dz2     :     0.008940  d :     0.021953
      dxz     :     0.006506
      dyz     :     0.006506
      dx2y2   :     0.000000
      dxy     :     0.000000

------------------------
MULLIKEN OVERLAP CHARGES
------------------------
B(  0-C ,   1-O ) :   1.3340
```

The Löwdin analysis contains a detailed breakdown of the MOs in terms of angular momentum components of each atom. This helps a lot when you select the orbitals for plotting.

```
                     *******************************
                     * LOEWDIN POPULATION ANALYSIS *
                     *******************************

----------------------
LOEWDIN ATOMIC CHARGES
----------------------
   0 C :   -0.050462
   1 O :    0.050462

 etc.
------------------------------------------
LOEWDIN REDUCED ORBITAL POPULATIONS PER MO
------------------------------------------
THRESHOLD FOR PRINTING IS 0.1%
                      0         1         2         3         4         5
ORB-EN           -19.24328 -10.30161  -1.15262  -0.56852  -0.47677  -0.47677
OCC                2.00000   2.00000   2.00000   2.00000   2.00000   2.00000
                  --------  --------  --------  --------  --------  --------
 0 C  s               0.0      99.5      18.6      11.5       0.0       0.0
 0 C  pz              0.1       0.0      16.7       4.8       0.0       0.0
 0 C  px              0.0       0.0       0.0       0.0       6.9      20.9
 0 C  py              0.0       0.0       0.0       0.0      20.9       6.9
 0 C  dz2             0.1       0.0       2.4       0.0       0.0       0.0
 0 C  dxz             0.0       0.0       0.0       0.0       0.5       1.5
 0 C  dyz             0.0       0.0       0.0       0.0       1.5       0.5
 1 O  s              99.8       0.1      54.7      20.8       0.0       0.0
 1 O  pz              0.0       0.4       7.3      62.3       0.0       0.0
 1 O  px              0.0       0.0       0.0       0.0      17.4      52.4
 1 O  py              0.0       0.0       0.0       0.0      52.4      17.4
 1 O  dz2             0.0       0.1       0.3       0.6       0.0       0.0
 1 O  dxz             0.0       0.0       0.0       0.0       0.1       0.3
 1 O  dyz             0.0       0.0       0.0       0.0       0.3       0.1

                      6         7         8         9        10        11
                  -0.37374  -0.02630  -0.02630   0.22161   0.40035   0.45738
                   2.00000   0.00000   0.00000   0.00000   0.00000   0.00000
                  --------  --------  --------  --------  --------  --------
 0 C  s              49.2       0.0       0.0      33.5      60.9       0.0
 0 C  pz             39.3       0.0       0.0      46.5      35.3       0.0
 0 C  px              0.0      15.2      54.7       0.0       0.0      15.5
 0 C  py              0.0      54.7      15.2       0.0       0.0      81.4
 0 C  dz2             0.8       0.0       0.0       7.8       1.2       0.0
 0 C  dxz             0.0       0.4       1.5       0.0       0.0       0.1
 0 C  dyz             0.0       1.5       0.4       0.0       0.0       0.3
 1 O  s               2.0       0.0       0.0      12.0       0.6       0.0
 1 O  pz              8.8       0.0       0.0       0.2       1.0       0.0
 1 O  px              0.0       5.9      21.2       0.0       0.0       0.3
 1 O  py              0.0      21.2       5.9       0.0       0.0       1.5
 1 O  dz2             0.0       0.0       0.0       0.0       0.9       0.0
 1 O  dxz             0.0       0.2       0.9       0.0       0.0       0.1
 1 O  dyz             0.0       0.9       0.2       0.0       0.0       0.7
```

Looking at Orbitals

There is a utility program `orca_plot` which lets you generate graphics information. The information about the MOs, the geometry and the basis set is stored in the so called gbw-File (‚geometry-basis-wavefunction‘).

To generate the plot information interactively use:

```
orca_plot myjob.gw -i
```

You will get a „stone-age“ menu which you can use to generate the necessary files.

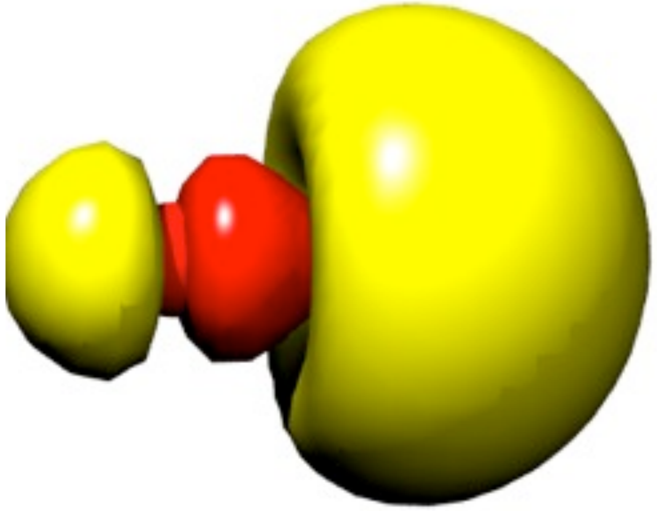
```
1.  Press 5 ENTER to choose the output formation. (press 7 ENTER for
    gaussian cube, the preferred format)
2.  Press 4 ENTER to choose the number of grid intervals. Something like 40
    will be o.k. For high resolution on larger molecules choose 65-75.
3.  To plot an orbital from a closed shell calculation press 3 ENTER and
    choose 0 ENTER. For spin-up from UHF/UKS the same. For spin down from
    UHF/UKS choose 1 ENTER.
4.  Press 2 ENTER and enter the number of the MO that you want to plot.
    NOTE THAT COUNTING STARTS WITH 0!!!
5.  Press 10 ENTER to generate the output file.
6.  Repeat steps 4 and 5 until all desired files are produced
```

The `orca_plot` program automatically produces an **‚xyz‘ file** which contains the molecular coordinates.

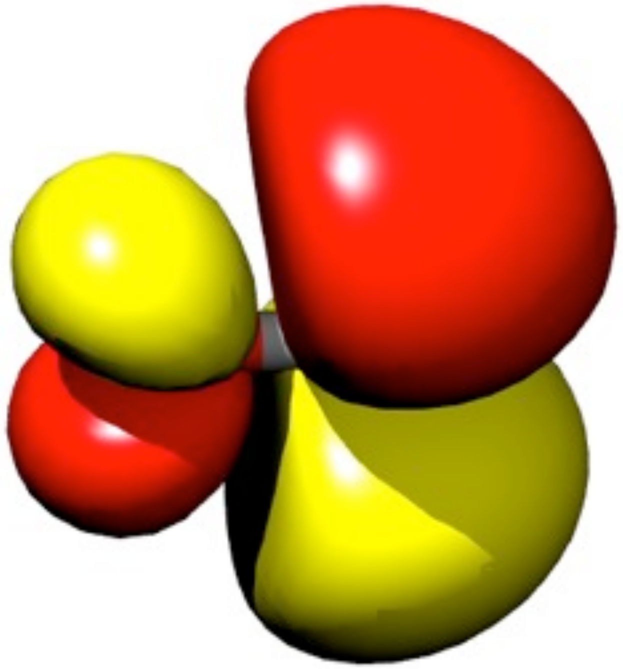
We presently like the ***Chimera*** program for vizualization. There are many free alternatives such as *Molekel, gOpenMol, Molden*,... any program that reads .xyz and .cube files

With **Chimera**, do the following:

1. Start Chimera
2. Choose „**File Open**“ and navigate to open the .xyz file of interest
3. Choose „**Presets → Publication 1**“
4. Choose „**Actions → Atoms&Bonds → ball & stick**“
5. Choose „**Actions → Color → By element**“
6. Choose „**Tools → Volume Data → Volume Viewer**“
7. Choose „**File → Open Map**“ and **Gaussian Cube format** and open the cube file of interest
8. Enter in **„Level**“ 0.03 and press ENTER. Then click on **Color** and choose red (or whatever you like)
9. Hold the Ctrl-key and click on the negative part of the contours. Enter in „**Level**“ -0.03 and press ENTER, then go to **color** and choose yellow (or whatever you like)
10. Orient the contour in the way you like and go to „**File → Save Image**“. Choose PovRay „true“ and go ahead.



HOMO
(MO#6)
of CO



LUMO
(MO#7)
of CO

Efficient DFT: The RI Approximation

As long as there is NO Hartree-Fock exchange present (no hybrid functionals), a very efficient approximation can be used to speed up (factor 10-100) DFT calculations: the **RI approximation** (also called **density fitting** method)

In this method the electron density is fit to an auxiliary basis set which must be provided by the user. The effect of the approximation on structures and frequencies is barely visible. Absolute energies are affected to a few kcal/mol, relative energies much less.

```
! RKS BP86 RI SV(P) SV/J TightSCF Opt

* int 0 1
C 0 0 0  0      0   0
O 1 0 0  1.15   0   0
H 1 2 0  1.00 115   0
H 1 2 3  1.00 115 180
*
```

Auxiliary basis appropriate for SV(P) (equivalently **Def-2** uses SV(P) and SV/J together; **Def-3** is equivalent to TZVP and TZV/J)

Use the RI approximation

The job leads to an energy of **-114.37494125 Eh**

Without RI and SV/J **-114.37466595 Eh**

Error **0.00028 Eh = 0.17 kcal/mol**

TIP: For geometry and frequency calculations the BP86 and PBE functionals together with the RI approximation is recommended. Basis sets of TZVP quality are appropriate for good accuracy! SV(P) is already good enough for a first orientation. Use keywords **QuickOpt, NormalOpt or GoodOpt**!
For energy calculations I recommend the B3LPY or PBE0 functionals and larger basis sets (TZVPP if possible or even aug-TZVPP). Use Keyword **DFTEnergy**

Efficient HF and Hybrid DFT: RIJCOSX

Unfortunately, RI does not smoothly carry over to Hartree-Fock and hybrid DFT calculations. One attempt to do so is the **RI-JK** approximation that needs to be invoked together with ‚JK‘ fitting bases. RI-JK cannot be used for optimizations

```
! B3LYP RI-JK def2-SVP def2-SVP/JK TightSCF
```

A more efficient approximation that leads to large speedups is the **RIJCOSX** approximation. it uses **RI-J** for the Coulomb part and a special approximation ‚**COSX**‘ for the exchange part. It is available throughout the program and leads to large speedups at very little loss in accuracy

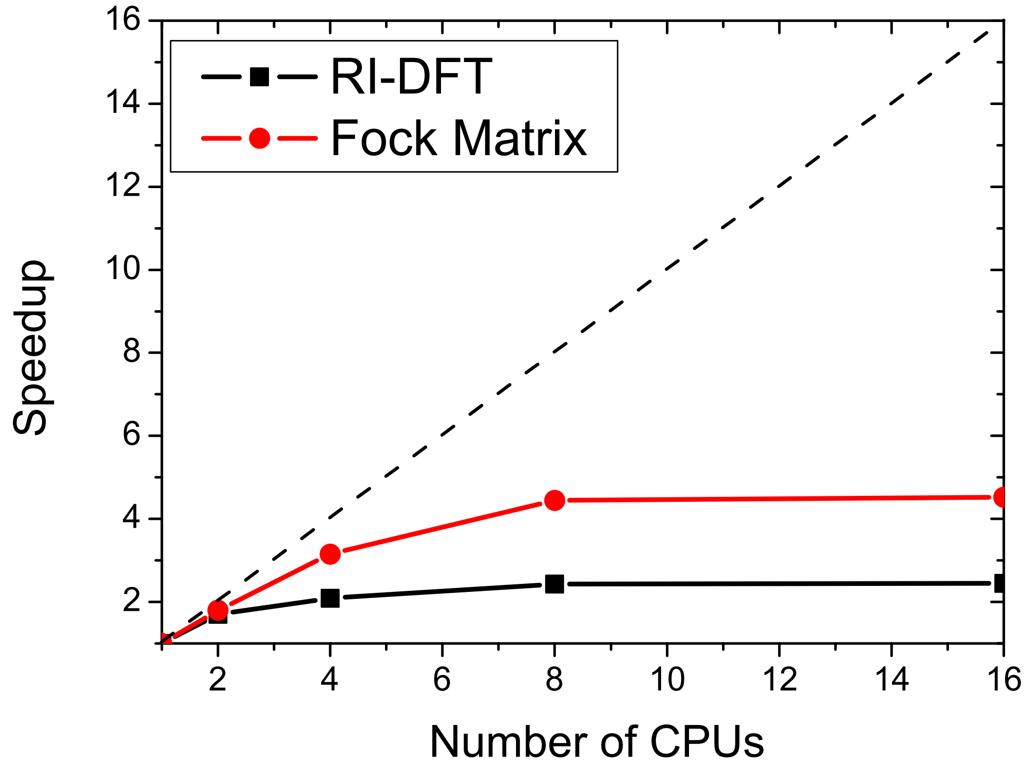
```
! B3LYP RIJCOSX def2-SVP def2-SVP/J TightSCF
```

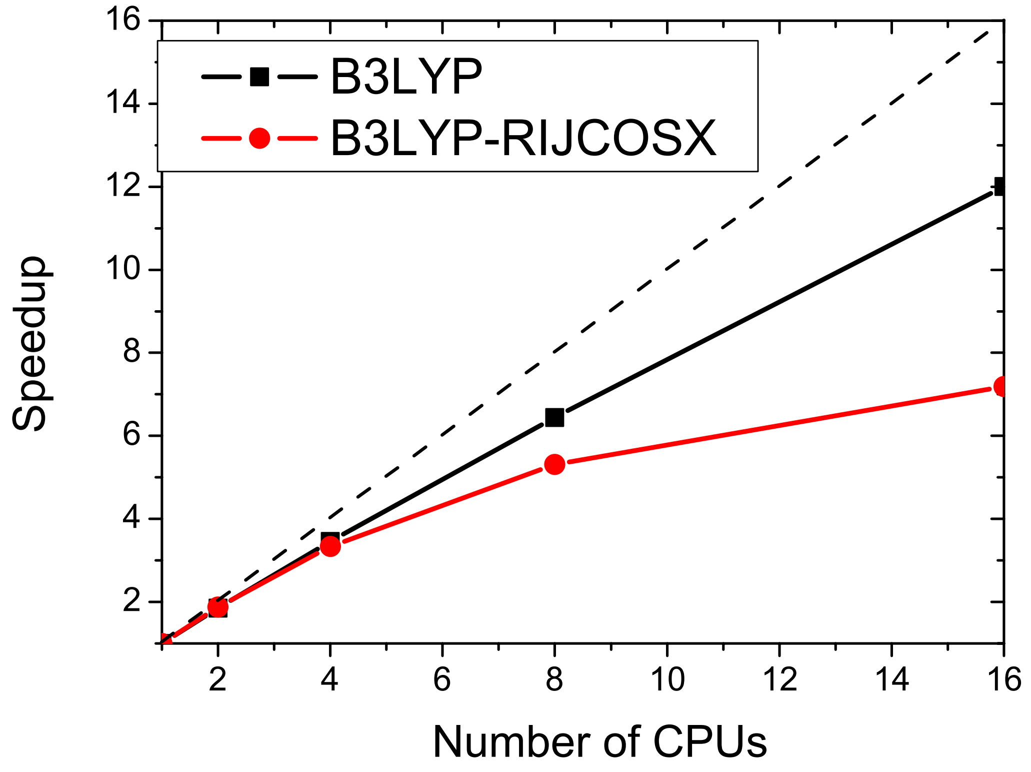
If combined with **RI-MP2** you need to give two auxiliary basis sets for optimal speed and accuracy:

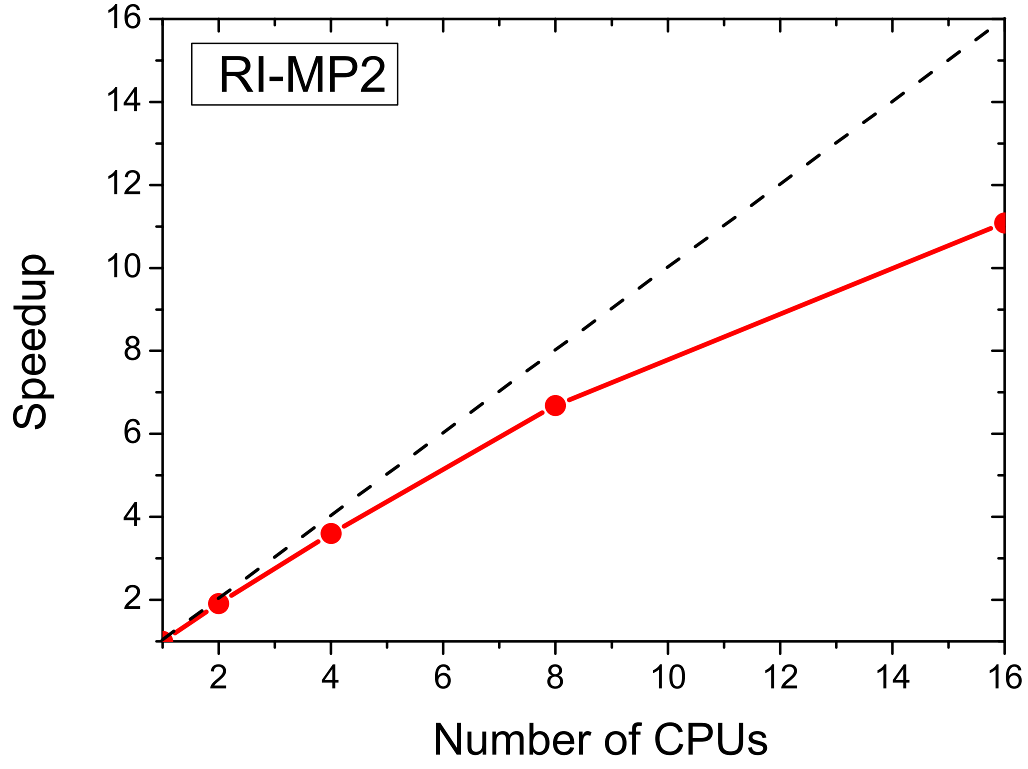
```
! RI-MP2 RIJCOSX def2-SVP def2-SVP/J def2-SVP/C TightSCF
```

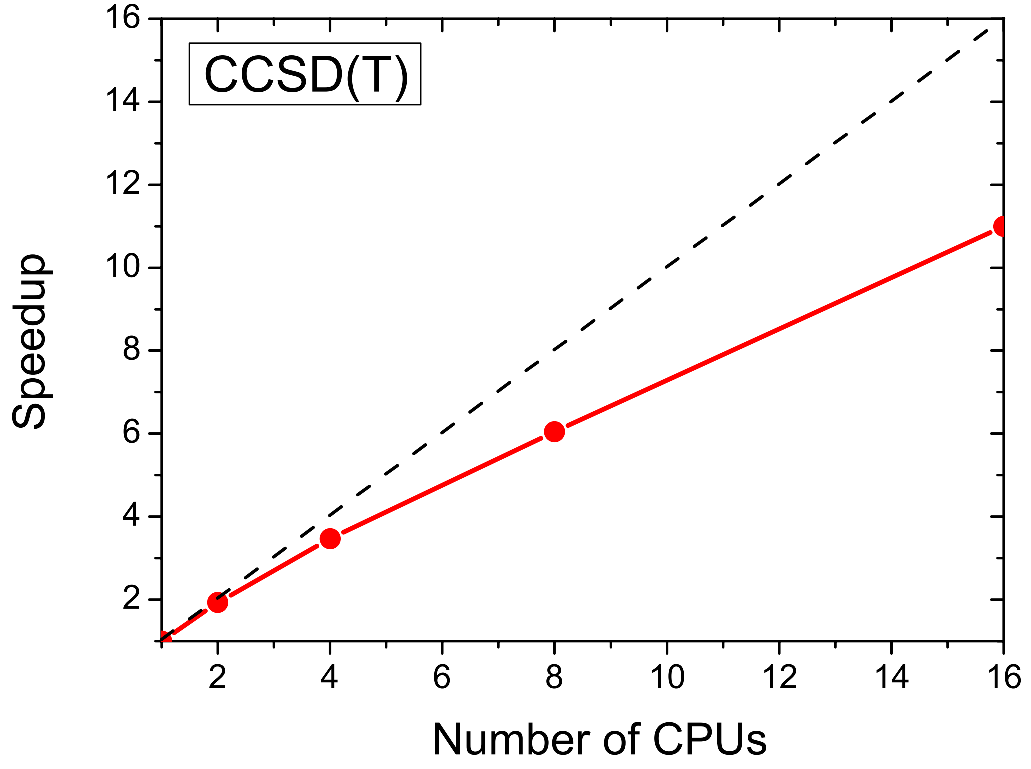
Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. (**2009**) *Chem. Phys.*, 356, 98–109

Speedup through Parallelization









Transition Metal Calculations

For open-shell transition metals convergence of the SCF is often a problem. There are a few ways in ORCA to help this process:

```
# Help a transition metal calculation to converge
! UKS B3LYP SV(P) TightSCF SlowConv

%scf  shift shift 0.1 erroff 0 end
      damp fac 0.80 erroff 0.001 end
      end

* xyz -2 2
Cu   0 0 0
F    1.9 0 0
F   -1.9 0 0
F    0   1.9 0
F    0  -1.9 0
*
```

This tells the program that you expect slow-convergence

Use level-shifting. Never turn it off

Use damping. Turn it off after convergence of the DIIS procedure to 0.001 has been achieved

... Actually, this particular job converges without any damping or shifting. However, in many situations, the damping and shifting is necessary as you will undoubtedly find out yourself once you start calculations in the „real world“.

In the present example the convergence „aids“ acctually strongly slow down convergence. It is quite uniform since the energy is always decreasing. Thus, these „helpers“ more or less „babysit“ the job to a converged solution.

Restarting Calculations

In many cases it is a very good idea to start from the convergence MOs of a previous job.

Use previous MOs

Name of the GBW file
NOTE: must not have the same name as your present input file!

```
# Start calculation from MOs of a previous one
! UKS B3LYP TZVP TightSCF SlowConv
! moread
%moinp “JOB-02.gbw”

%scf  guessmode cmatrix
      end

* xyz -2 2
Cu  0 0 0
F   1.9 0 0
F  -1.9 0 0
F   0   1.9 0
F   0  -1.9 0
*
```

Large basis set calculation started from MOs of a small basis set calculation

„Play safe“ if restarting calculations on anions

Note that the calculation can be started from a gbw file which is from a nearby geometry, uses a different basis set of theoretical method or HFType (UHF,RHF or ROHF). The program will take care to translate the orbitals to the present situation.

Multiple Job Steps

You can also run the two jobs from one input file.

```
# Run a two-step job. The first one is “cheap” and provides
# input orbitals for the second one.Note that we also re-
# read the geometry in the second job-step
! UKS SV(P) B3LYP TightSCF SlowConv XYZFile
%base "JOB_a"
%scf shift shift 0.1 erroff 0 end
     damp damp 0.8 erroff 0.001 end
     end

* xyz -2 2
Cu   0 0 0
F    1.9 0 0
F   -1.9 0 0
F    0   1.9 0
F    0  -1.9 0
*

$new_job
# Now this job should converge rather well. Turn off all
# damping and shifting
! UKS TZVP B3LYP TightSCF NoDamp NoLShift
! moread
%moinp "JOB_a.gbw"
%base "JOB_b"

%scf guessmode cmatrix
     end

* xyzfile -2 2 JOB_a.xyz
```

Calculating EPR Parameters

The calculation of EPR parameters is controlled via the EPRNMR block:

```
# A simple EPR job
! UKS EPR-II B3LYP TightSCF SlowConv

* xyz 0 3
N 0 0 0
H 0 0 1.0
*
%eprnmr dtensor SSandSO  # request calculation of the ZFS tensor
                         # using both Spin-Spin (SS) and Spin-Orbit
                         # (SO) contribvutions
        dsoc cp          # linear response treatment of SO contrib.
                         # alternative is PK (Pederson-Khanna)
        dss uno          # us the spin-restricted density for SS part
                         # alternative is „direct"
        gtensor true     # calculate the g-tensor using linear response
        # nuclear properties. Note that aorb is expensive and should
        # only be applied to heavier nuclei like metals. For ligand
        # nuclei aorb is small. The other properties are simple
        # expectation values
        nuclei = all H { aiso, adip, aorb, fgrad, rho}
        nuclei = all N { iaos, adip, aorb, fgrad, rho}
        # printlevel 3 provides a detailed analysis of all properties
        # the default is to print only a minimum amount of information.
        printlevel 3
        end
```

Geometry Optimization

To optimize the geometry of the molecule, simply include the keyword `Opt`

```
! RKS SV(P) B3LYP TightSCF Opt

* int 0 1
C 0 0 0  0      0   0
O 1 0 0  1.15   0   0
H 1 2 0  1.00 115   0
H 1 2 3  1.00 115 180
*
```

The program will first produce a set of „redundant internal coordinates“ which are used in the calculation.

```
--------------------------------------------------------------
                        Redundant Internal Coordinates

--------------------------------------------------------------
      Defintion               Initial Value    Approx d2E/dq
--------------------------------------------------------------
    1. B(O   1,C   0)                1.1500        1.351281
    2. B(H   2,C   0)                1.0000        0.501167
    3. B(H   3,C   0)                1.0000        0.501167
    4. A(H   2,C   0,O   1)      115.0000        0.425466
    5. A(H   3,C   0,O   1)      115.0000        0.425466
    6. A(H   3,C   0,H   2)      130.0000        0.323418
    7. I(O   1,H   3,H   2,C   0)   0.0000        0.151694
--------------------------------------------------------------
```

TIP: Always use **TightSCF** or **VeryTightSCF** in geometry optimizations. Otherwise the gradients are somewhat noisy.

After calculating the SCF energy and the gradient of the energy, a relaxation is step is carried out:

```
-----------------------------------------------------------------------------
                        ORCA GEOMETRY RELAXATION STEP
-----------------------------------------------------------------------------

Number of atoms                             ....    4
Number of internal coordinates              ....    7
Current Energy                              ....  -114.317745134 Eh
Current gradient norm                       ....     0.207887808 Eh/bohr
Maximum allowed component of the step       ....  0.300
Current trust radius                        ....  0.300
Evaluating the initial hessian              ....  (Almloef) done
Projecting the Hessian                      .... done
Forming the augmented Hessian               .... done
Diagonalizing the augmented Hessian         .... done
Last element of RFO vector                  ....  0.957975075
Lowest eigenvalues of augmented Hessian:
 -0.057174708  0.151693870  0.360949845  0.425465740  0.501166791
Length of the computed step                 ....  0.299435193
The final length of the internal step       ....  0.299435193
Converting the step to cartesian space:
Transforming coordinates:
 Iter   0:  RMS(Cart)=    0.0824560429 RMS(Int)=    0.1127292561
 Iter   1:  RMS(Cart)=    0.0024951498 RMS(Int)=    0.0033197589
 Iter   2:  RMS(Cart)=    0.0002432861 RMS(Int)=    0.0003067902
 Iter   3:  RMS(Cart)=    0.0000185328 RMS(Int)=    0.0000231161
 Iter   4:  RMS(Cart)=    0.0000012690 RMS(Int)=    0.0000015779
 Iter   5:  RMS(Cart)=    0.0000000832 RMS(Int)=    0.0000001034
 Iter   6:  RMS(Cart)=    0.0000000054 RMS(Int)=    0.0000000067
done
Storing new coordinates                     .... Done
```

The status of the geometry convergence is printed:

```
                           .--------------------.
----------------------|Geometry convergence|-------------------------
Item                value                   Tolerance     Converged
---------------------------------------------------------------------
RMS gradient        0.07883145              0.00010000      NO
MAX gradient        0.14233649              0.00030000      NO
RMS step            0.11317586              0.00200000      NO
MAX step            0.18511086              0.00400000      NO
...................................................
Max(Bonds)      0.0980      Max(Angles)    4.89
Max(Dihed)        0.00      Max(Improp)    0.00
---------------------------------------------------------------------
```

```
The optimization has not yet converged - more geometry cycles are needed
```

And a new geometry is proposed:

```
    ---------------------------------------------------------------------------
                        Redundant Internal Coordinates
                           (Angstroem and degrees)

        Definition                    Value    dE/dq     Step     New-Value
    ---------------------------------------------------------------------------
     1. B(O   1,C   0)                1.1500 -0.142336  0.0535    1.2035
     2. B(H   2,C   0)                1.0000 -0.103355  0.0980    1.0980
     3. B(H   3,C   0)                1.0000 -0.103355  0.0980    1.0980
     4. A(H   2,C   0,O   1)          115.00 -0.017685    2.44    117.44
     5. A(H   3,C   0,O   1)          115.00 -0.017685    2.44    117.44
     6. A(H   3,C   0,H   2)          130.00  0.035370   -4.89    125.11
     7. I(O   1,H   3,H   2,C   0)      0.00 -0.000000    0.00      0.00
    ---------------------------------------------------------------------------

         *************************************************************
         *                GEOMETRY OPTIMIZATION CYCLE   2            *
         *************************************************************
```

Then the next SCF is done and the next gradient calculated, a new geometry is proposed until (hopefully) finally:

```
***********************HURRAY********************
***          THE OPTIMIZATION HAS CONVERGED     ***
*************************************************
```

Following this statement one more energy calculation is performed in order to make sure that the energy and properties are really done at the stationary point of the PES.

Constraints and Relaxed Scans

You can „freeze“ certain geometrical parameters in an optimization:

```
! RKS B3LYP/G SV(P) TightSCF Opt
%geom Constraints
        { B 0 1 1.25 C }
        { A 2 0 3 120.0 C }
        end
      end

* int 0 1
    C  0 0 0 0.0000   0.000    0.00
    O  1 0 0 1.2500   0.000    0.00
    H  1 2 0 1.1075 122.016    0.00
    H  1 2 3 1.1075 122.016  180.00
*
```

... Or freeze some and vary others (one frequently used possibility is to only optimize hydrogen positions **`OptimizeHydrogens true`**). Constrained surfaces are calculated as:

```
%geom Scan
        B 0 1 = 1.35, 1.10, 12 # C-O distance that will be scanned
        end
      end
```

Frequency Calculations

There are several good reasons for calculating the harmonic frequencies:

1. Characterize stationary points as minima (no negative frequencies), transition states (one negative frequency) or higher-order saddle point (more negative frequencies
2. Predict vibrational spectra (IR, Raman)
3. Calculate thermodynamic properties (zero-point energy, finite temperature correction)

ORCA presently calculates harmonic frequencies through (one- or two-sided) numeric differentiation of analytic frequencies

```
! RKS BP86 RI SV(P) SV/J Grid4 TightSCF SmallPrint
! TightOpt NumFreq

%freq CentralDiff true
      Increment 0.005
      end

* xyz 0 1
 C   0.000000    0.000000   -0.533905
 O   0.000000    0.000000    0.682807
 H   0.000000    0.926563   -1.129511
 H   0.000000   -0.926563   -1.129511
*
```

Run a numerical frequency calculation

Two sided differences (twice as expensive but more accurate! Be careful – numerical frequencies can be quite noisy)

Increment for displacements (in Bohrs)

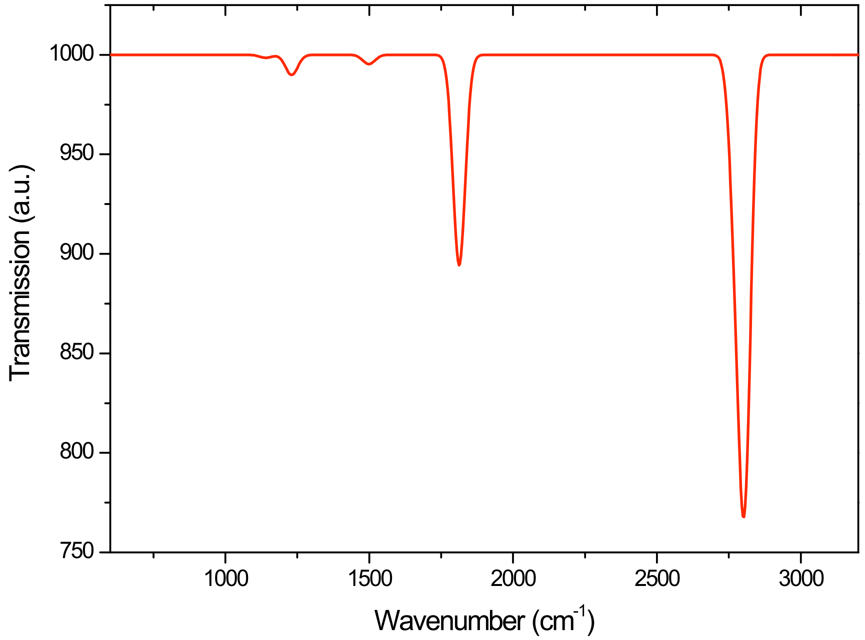
The first thing that is printed are the vibrational frequencies. The first six modes are translations and rotations and these are zero because they are projected out.

```
-----------------------
VIBRATIONAL FREQUENCIES
-----------------------

   0:         0.00 cm**-1
   1:         0.00 cm**-1
   2:         0.00 cm**-1
   3:         0.00 cm**-1
   4:         0.00 cm**-1
   5:         0.00 cm**-1
   6:      1140.72 cm**-1
   7:      1230.49 cm**-1
   8:      1498.92 cm**-1
   9:      1812.75 cm**-1
  10:      2773.72 cm**-1
  11:      2805.59 cm**-1
```

Then the program prints the normal modes. This is usually not very revealing. The program produces a **`BaseName.hess`** file which you can run through **`orca_vib`** to get additional information.
XYZ files to be used for animation of vibrational modes are produces by the program **`orca_pltvib`**.
Then you get the IR spectrum:

```
-----------
IR SPECTRUM
-----------

 Mode    freq (cm**-1)   T**2         TX         TY         TZ
-------------------------------------------------------------------
   6:      1140.72    1.465985  ( -1.210739  -0.008963  -0.004000)
   7:      1230.49   10.082152  (  0.004428  -3.175237  -0.001184)
   8:      1498.92    4.648016  ( -0.000197   0.001760  -2.155925)
   9:      1812.75  105.821353  ( -0.000043   0.002847 -10.286950)
  10:      2773.72   73.358541  (  0.000042  -0.020336  -8.564936)
  11:      2805.59  205.230431  ( -0.000359 -14.325864   0.007790)
```

The IR spectrum can be plotted through the orca_mapspc utility (**`orca_mapspc BaseName.out IR`**)



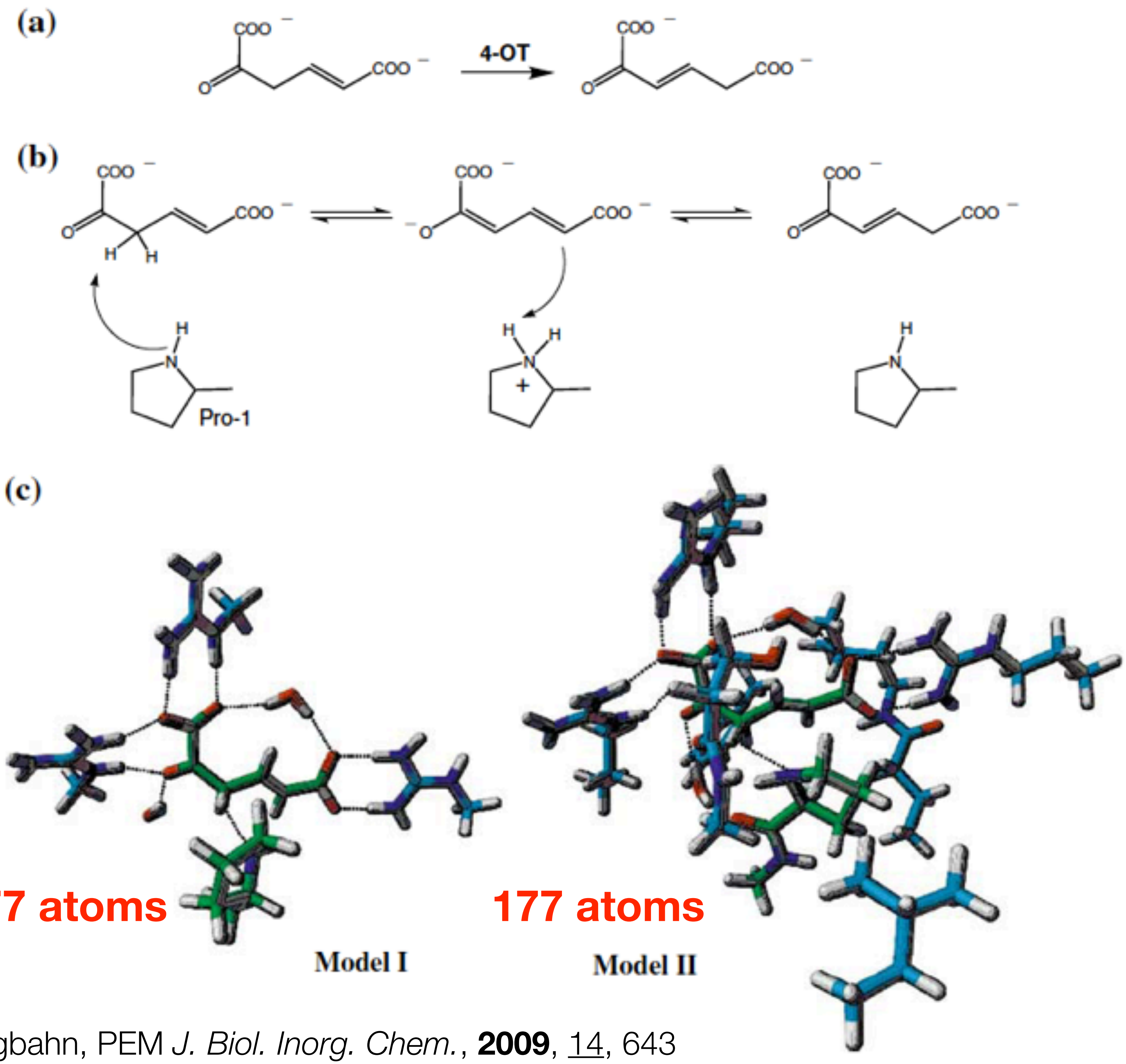
Finally, the thermodynamic properties at 298.15 K are printed (assuming ideal gas behaviour)

```
Electronic energy                ...  -114.41435858 Eh
Zero point energy                ...     16.10 kcal/mol
Thermal vibrational correction   ...      0.03 kcal/mol
Thermal rotational correction    ...      0.89 kcal/mol
Thermal translational correction ...      0.89 kcal/mol
-----------------------------------------------------------
Total thermal energy                  -114.38582646 Eh

Then enthalpy+entropy and finally the free energy:

Total enthalpy               -114.38488225 Eh
Total entropy correction      -15.36 kcal/mol
-------------------------------------------
Final Gibbs free enthalpy -114.40935858 Eh
G-E(el) = 0.00500000 Eh = 3.14 kcal/mol
```

Metalloproteins: Cluster vs QM/MM Models

Stages of Model Building

Cluster Models are often sufficient



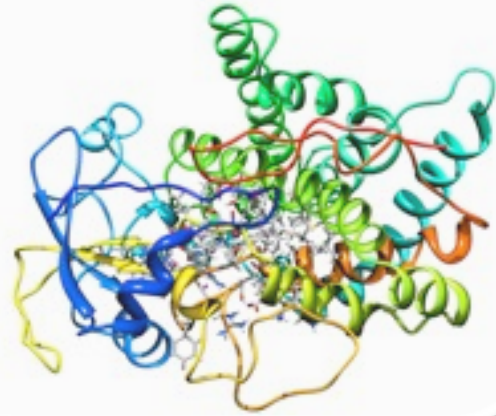
77 atoms

177 atoms

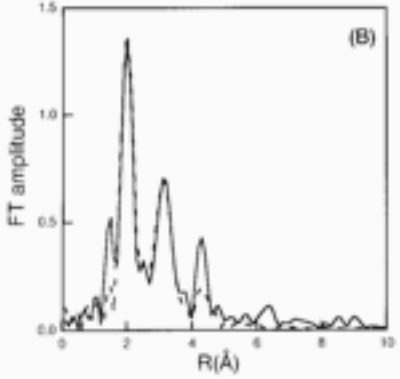
Himo, F.; Siegbahn, PEM *J. Biol. Inorg. Chem.*, **2009**, 14, 643

... But Protein effects can be very subtle

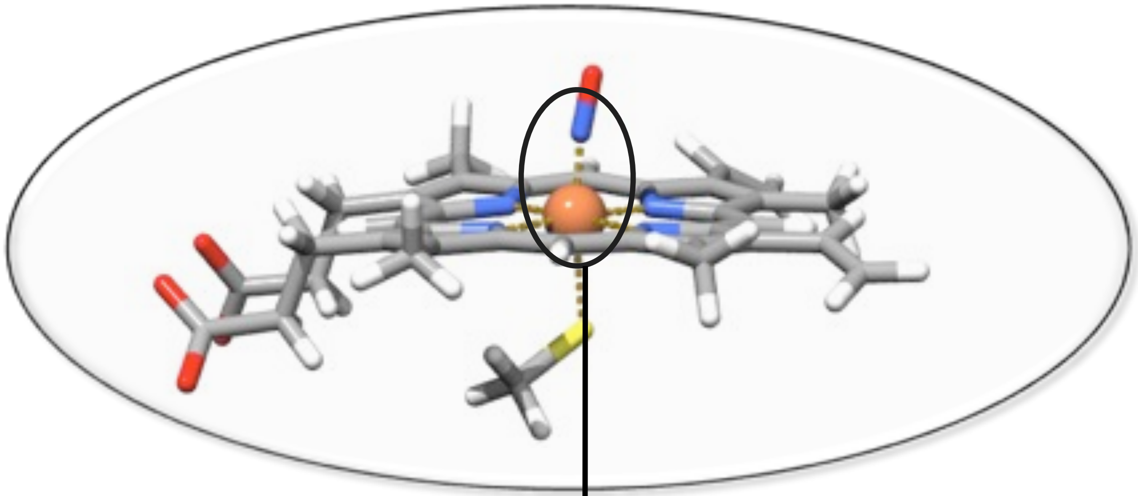
P450nor



EXAFS



Fe-N: 1.66 Å



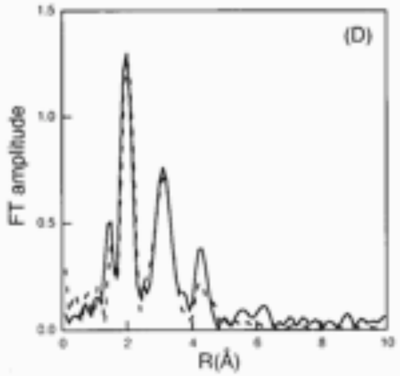
Fe-N bond length?

Same active site, different Fe-N distance, different reactivity.

→ Protein Effect?

P450cam

EXAFS



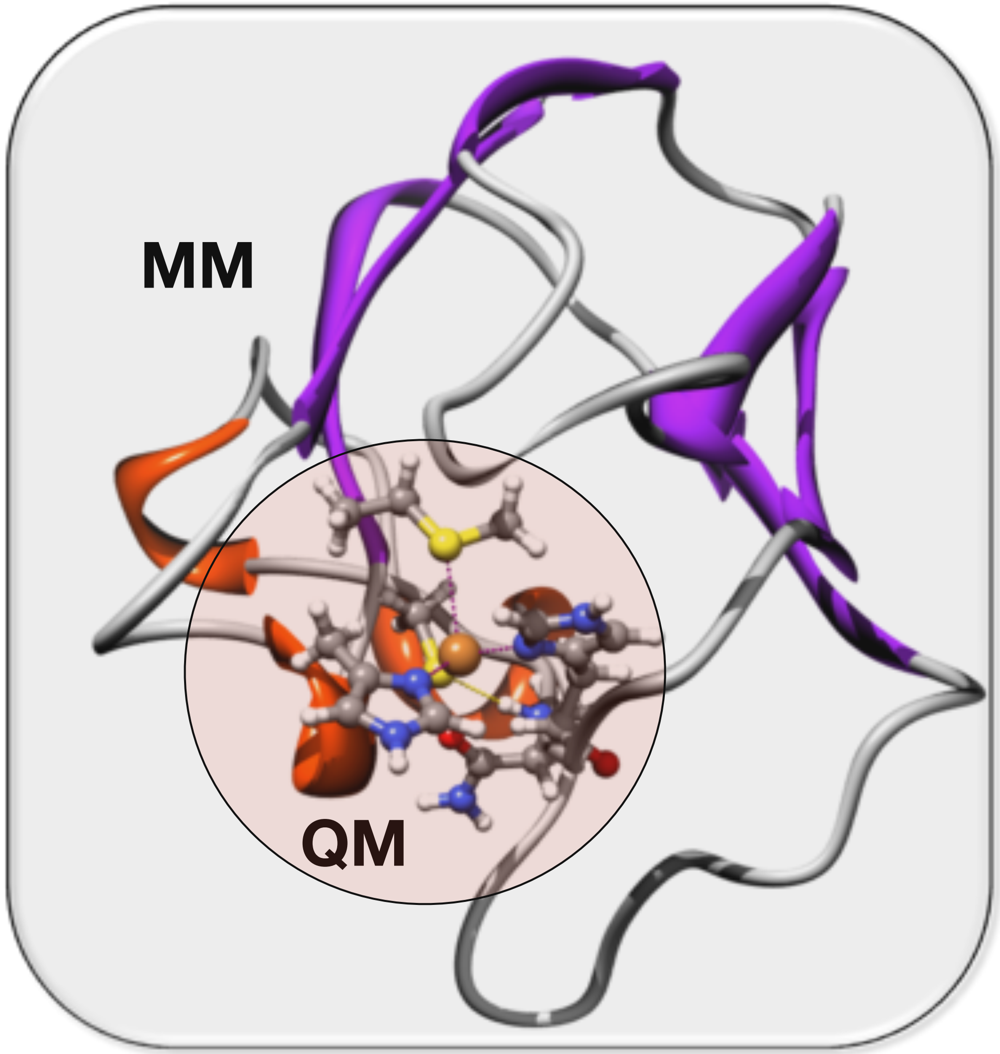
Fe-N: 1.76 Å

Motivation for QM/MM: QM/MM

- Combine the advantages of both methods

- **QM/MM level:**
  - ✓ Small and chemically important part of the system (e.g. active site): **QM**
  - ✓ Large and less important part of the system: **MM**

→ Realistic chemistry of real systems

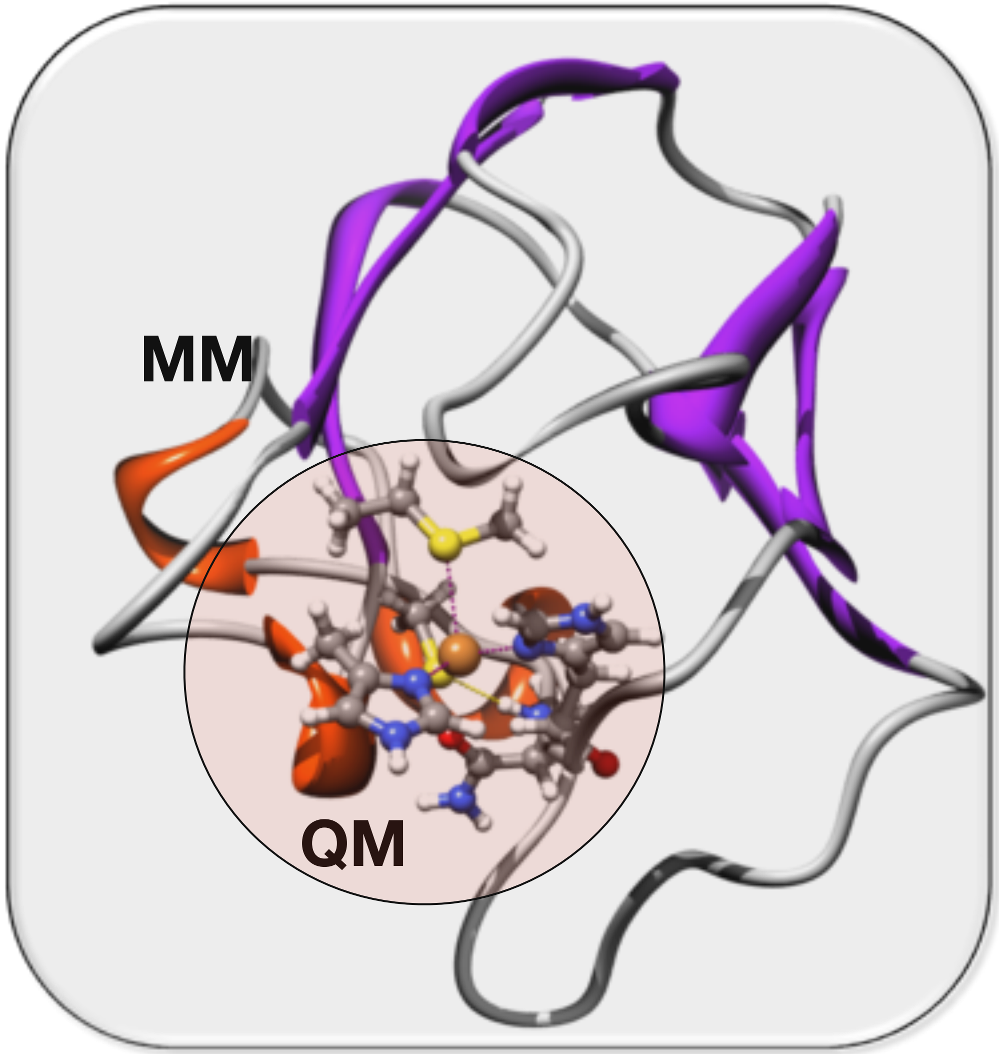


Separation of a QM/MM system

- ✓ **Primary-** (QM-) **subsystem**: a localized region, where a process of interest takes place (e.g. enzymatic reaction, charge transfer process, electronic excitation, ...).

- ✓ **Outer-** (MM-) **subsystem**: the environment of the QM-subsystem. This part plays an important role for the inner subsystem (e.g. a protein which imposes steric constraints on the active site, specific hydrogen bonds and salt bridges).

- ✓ **Boundary region**: is of importance if both subsystems are connected via bonds.



QM/MM Energy

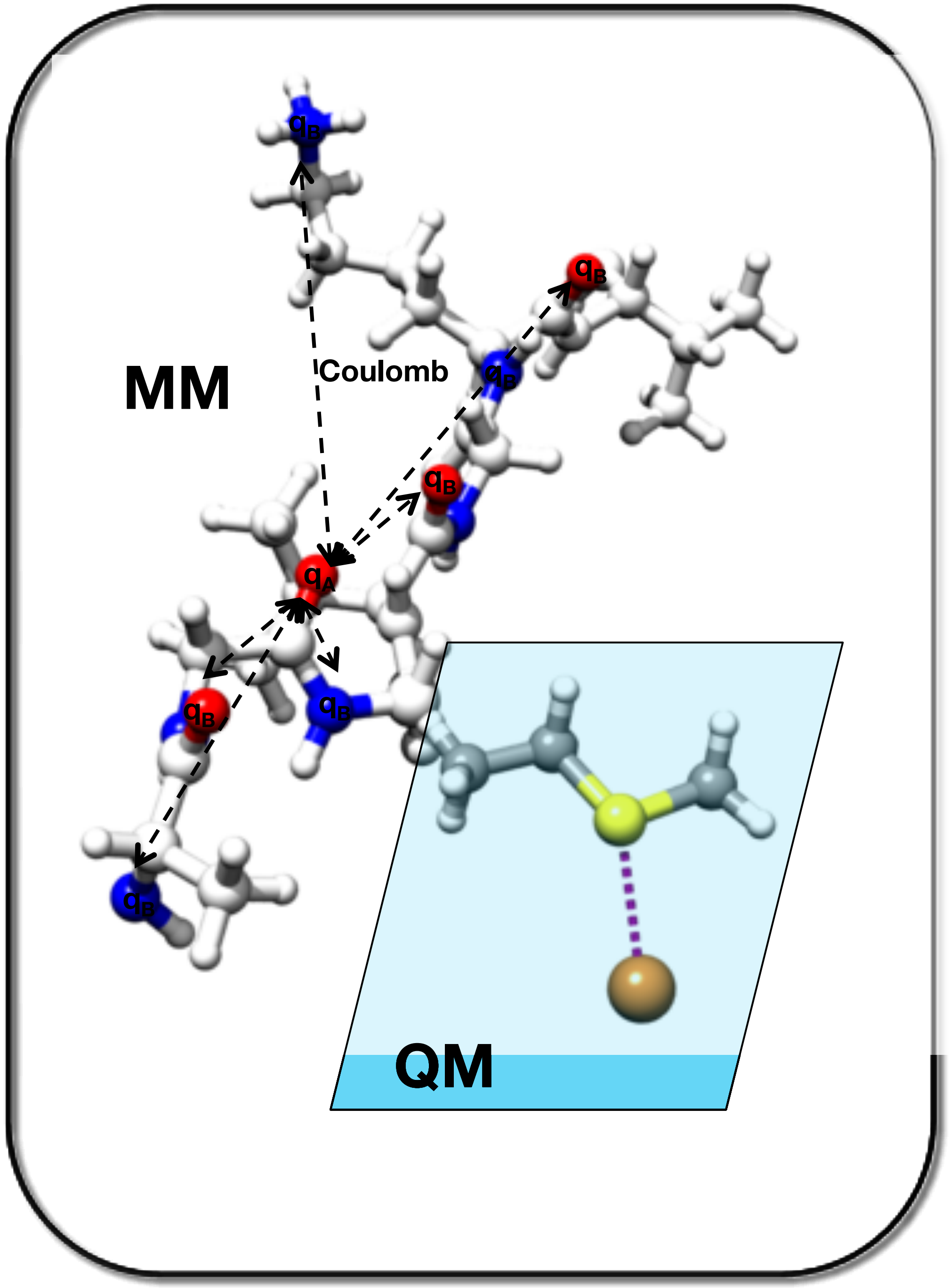
$$E_{QM/MM}(S) = E_{MM}(O) + E_{QM}(P+L) + E_{QM-MM}(P,O)$$

- $\boldsymbol{E_{MM}}$: MM energy of the outer subsystem.
- $\boldsymbol{E_{QM}}$: QM energy of the (capped) primary subsystem.
- $\boldsymbol{E_{QM\text{-}MM}}$: QM-MM interaction calculation between primary and outer subsystem.

MM energy E_{MM}

- Bonded and nonbonded interactions.

$$
\begin{aligned}
E_{MM} = {} & \sum_{\text{bonds}} k_d \left(d - d_0\right)^2 \\
& + \sum_{\text{angles}} k_\theta \left(\theta - \theta_0\right)^2 \\
& + \sum_{\text{dihedrals}} k_\phi \left[1 + \cos\left(n\phi + \delta\right)\right] \\
& + \sum_{\substack{\text{non-bonded} \\ \text{pairs AB}}} \varepsilon_{AB} \left[\left(\frac{\sigma_{AB}}{r_{AB}}\right)^{12} - \left(\frac{\sigma_{AB}}{r_{AB}}\right)^{6}\right] \\
& + \sum_{\substack{\text{non-bonded} \\ \text{pairs AB}}} \frac{1}{4\pi\varepsilon_0} \frac{q_A q_B}{r_{AB}}
\end{aligned}
$$



MM force fields

✓ **Force field**: collection of MM parameters.

✓ **MM parameters** (σ_{AB}, k_d, …) are atomtype specific (e.g. a hydrogen in a methyl group has a different atom type than a hydrogen atom in a benzene).

✓ **Standard force fields** in biochemical research have parameters for biomolecules (standard amino acids, DNA, lipids, sugars), but e.g. not for metal atoms.

✓ **... If no parameters are available for a molecule ...**

➡ Parametrization necessary *or*

➡ Assign parameters from similar chemical patterns.

QM/MM Boundary

- If bonds are cut between the primary and the outer subsystem, the **cleaved bond has to be saturated** (→← homolytic/heterolytic cleavage).
- Several procedures were developed to handle this problem. The most popular one is the usage of **link atoms**:
  - A hydrogen atom caps the cleaved bond.

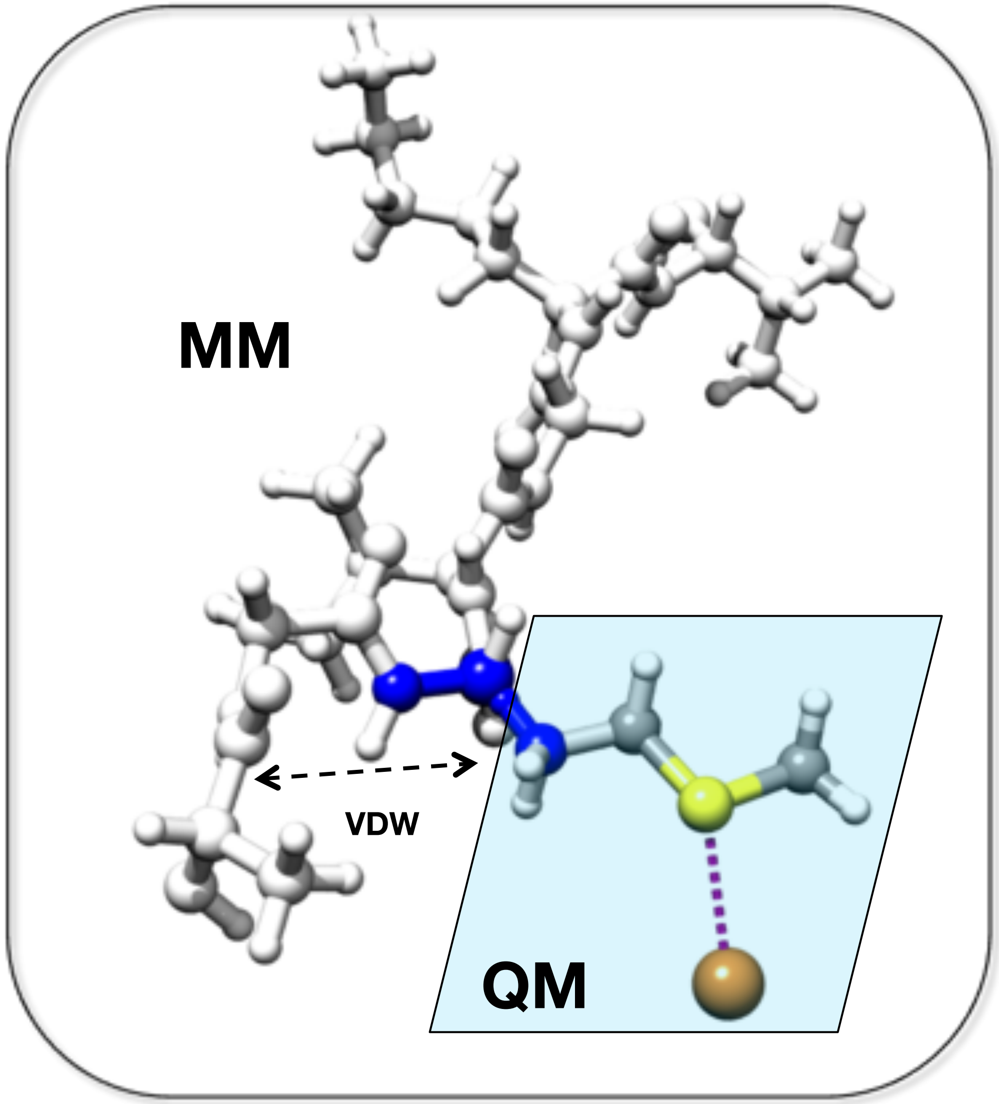
QMMM interaction energy $E_{QM\text{-}MM}$

- Interactions between primary and outer subsystem:

$$E_{QM-MM} = E_{bonded}(P,O) + E_{VDW}(P,O) + E_{el}(P,O)$$

- $\boldsymbol{E_{bonded}}$ and E_{VDW} always calculated on MM level.

- E_{el} dependent on embedding scheme.
    - Mechanical embedding
    - Electrostatic embedding



QMMM interaction energy $E_{QM\text{-}MM}$

- Interactions between primary and outer subsystem:

$$E_{QM-MM} = E_{bonded}(P,O) + E_{VDW}(P,O) + \boxed{E_{el}(P,O)}$$

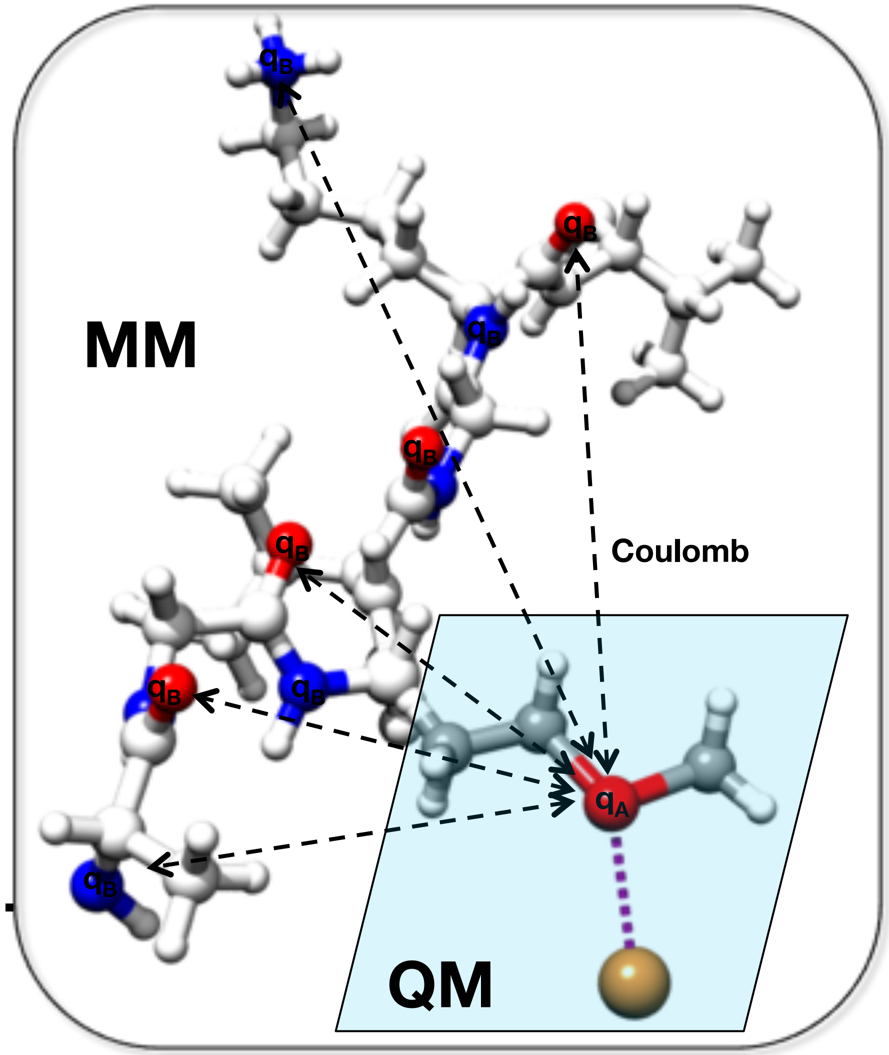
- E_{bonded} and E_{VDW} always calculated on MM level.
- $\boldsymbol{E_{el}}$ dependent on embedding scheme:
  - Mechanical embedding
  - Electrostatic embedding

Mechanical embedding

$$E_{QM-MM} = E_{bonded}(P,O) + E_{VDW}(P,O) + \boxed{E_{el}(P,O)}$$

- ✓ Electrostatic Interaction between primary and outer subsystem is **calculated on the MM level.**

$$E_{el} = \sum_{\substack{\text{non-bonded} \\ \text{pairs AB}}} \frac{1}{4\pi\varepsilon_0} \frac{q_{\mathrm{A}} q_{\mathrm{B}}}{r_{\mathrm{AB}}}$$



- ✓ Drawbacks:
  - QM charge density is mimicked by point charges.
  - **QM electron density** is ***not* polarized** by the MM-point charges.

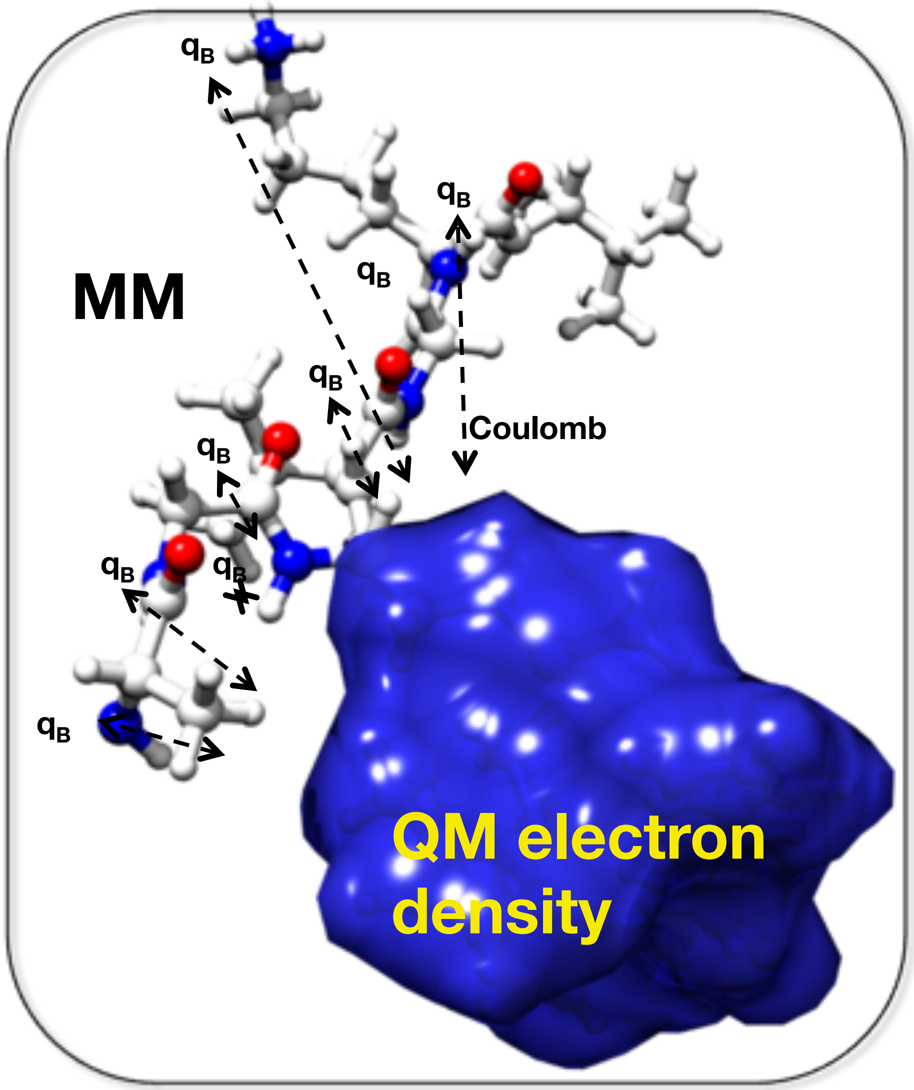
Electrostatic embedding

$$E_{QM-MM} = E_{bonded}(P,O) + E_{VDW}(P,O) + \boxed{E_{el}(P,O)}$$

- ✓ Electrostatic Interaction between primary and outer subsystem is **calculated on the QM level**.

$$\hat{H}^{el}_{QM-MM} = -\sum_{i}^{electrons}\sum_{B\in O}\frac{q_B}{|r_i - R_B|} + \sum_{C\in(I+L)}\sum_{B\in O}\frac{q_M Z_C}{|R_C - R_B|}$$

- ✓ Advantage: **QM electron density** is **directly polarized** by the MM-point charges.
- ✓ Drawbacks:
  - Might become expensive (long range interaction).
  - MM charges may not be well constructed to interact with the QM density.



Scheme of a QM/MM Study:

Scheme of a QMMM calculation

- Constructing missing MM parameters (ESP charges, prodrg)
- Structure validation
- Adding hydrogens
- Solvation
- Energy Minimization
- Equilibration / MD → snapshots

MM

- Optimization of snapshots
- Property calculations on optimized geometries

QM/ MM

Scheme of a QMMM calculation

- Constructing missing MM parameters (ESP charges, prodrg)
- Structure validation
- Adding hydrogens
- Solvation
- Energy Minimization
- Equilibration / MD → snapshots

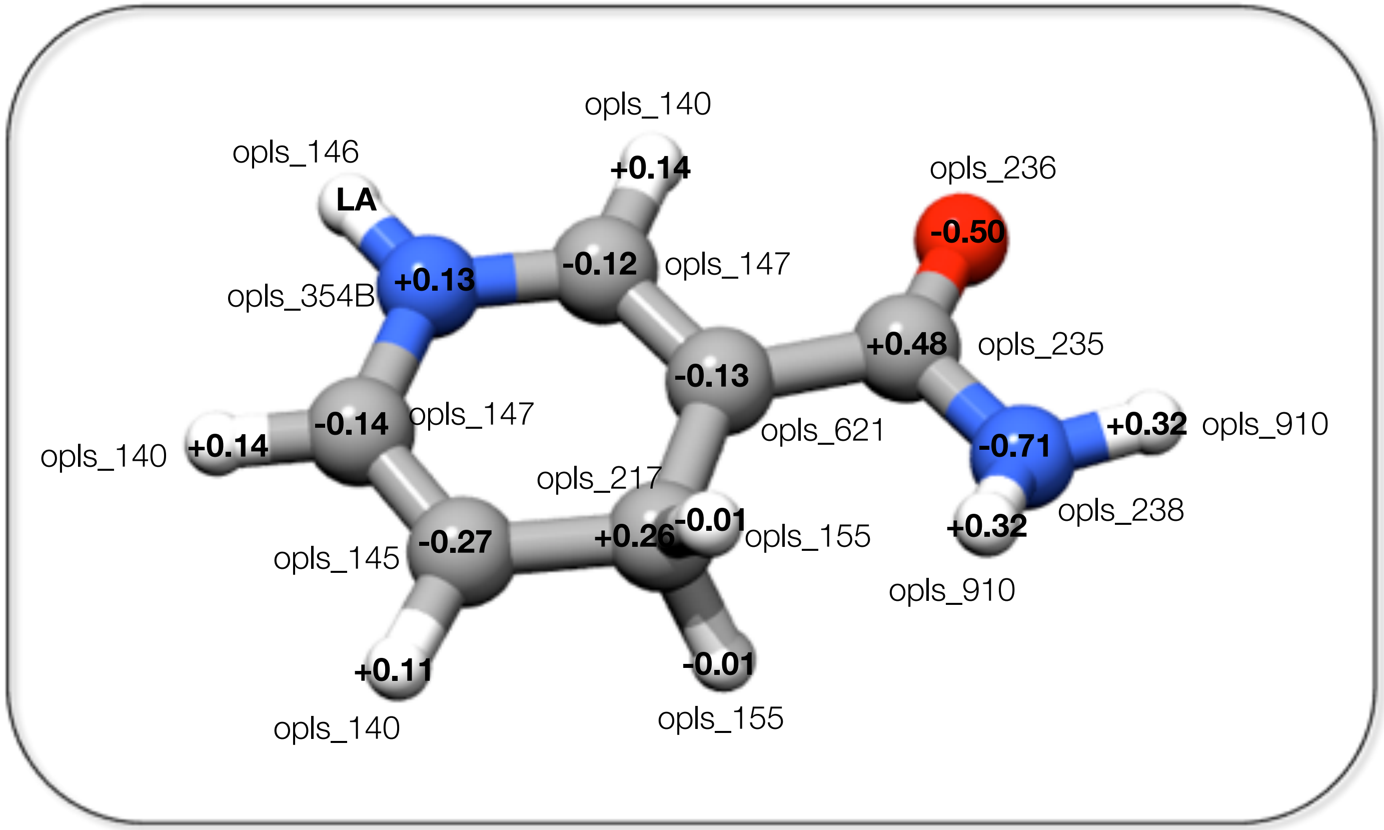
MM

- Optimization of snapshots
- Property calculations on optimized geometries

QM/MM

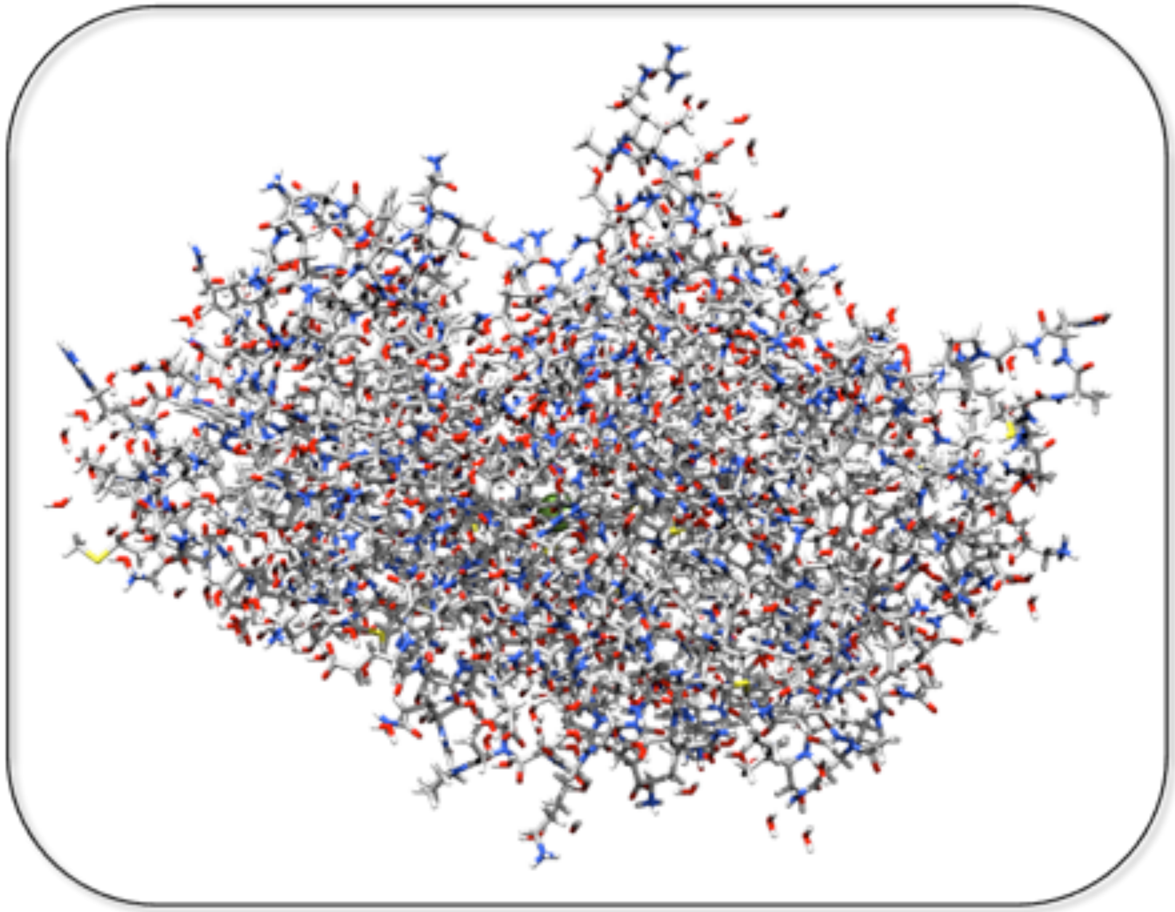
Scheme of a QMMM calculation

- Constructing missing MM parameters (ESP charges, prodrg)
  - E.g. Nicotinamide:



Scheme of a QMMM calculation

- Constructing missing MM parameters (ESP charges, prodrg)
- Structure validation



- Adding hydrogens
- Solvation
- Energy Minimization
- Equilibration / MD → snapshots

MM

- Optimization of snapshots
- Property calculations on optimized geometries

QM/MM

Scheme of a QMMM calculation

- Constructing missing MM parameters (ESP charges, prodrg)
- Structure validation
- Adding hydrogens
- Solvation
- Energy Minimization
- Equilibration / MD → snapshots

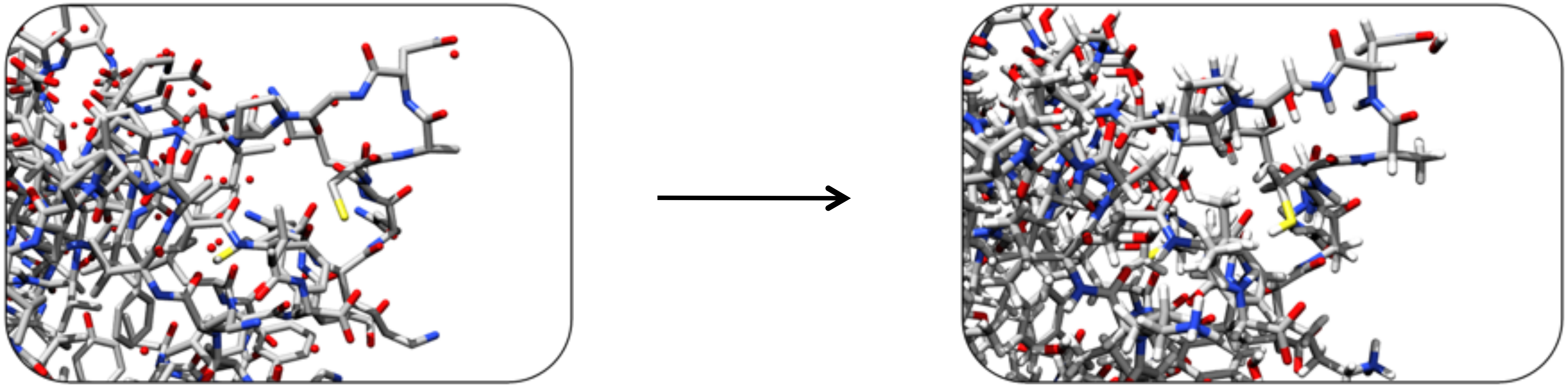
MM

- Optimization of snapshots
- Property calculations on optimized geometries

QM/MM

Scheme of a QMMM calculation

- Adding hydrogens



PDB structure without hydrogens

Structure with hydrogens (dep. on pk_A)

Scheme of a QMMM calculation

- Constructing missing MM parameters (ESP charges, prodrg)
- Structure validation
- Adding hydrogens
- Solvation
- Energy Minimization
- Equilibration / MD → snapshots

MM

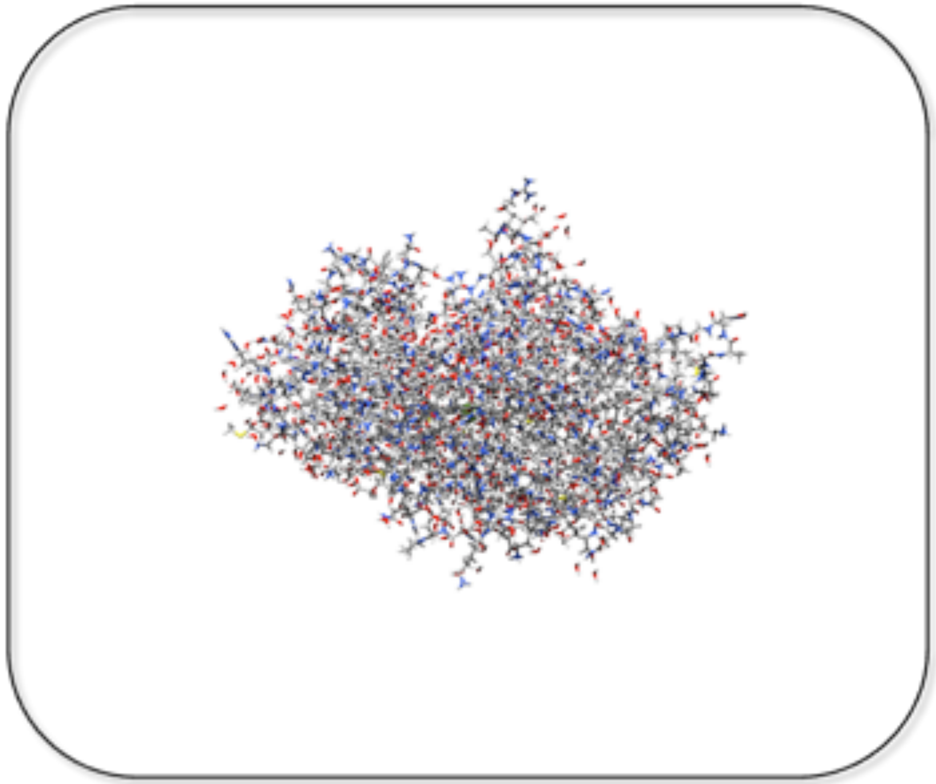
- Optimization of snapshots
- Property calculations on optimized geometries

QM/MM

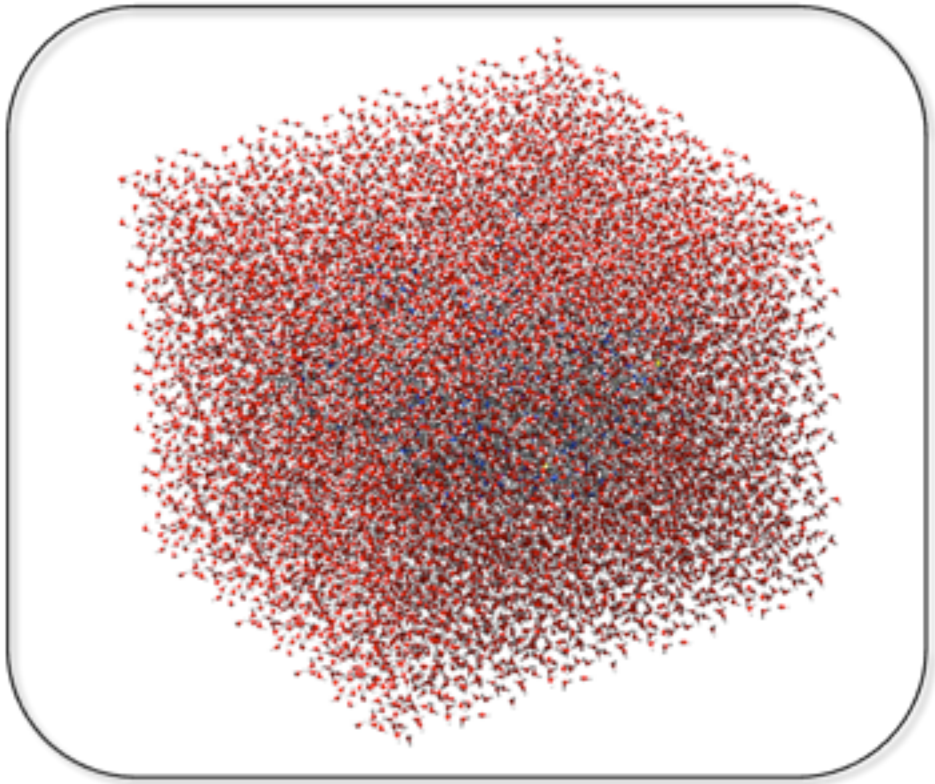
Scheme of a QMMM calculation

- Solvation

Unsolvated protein

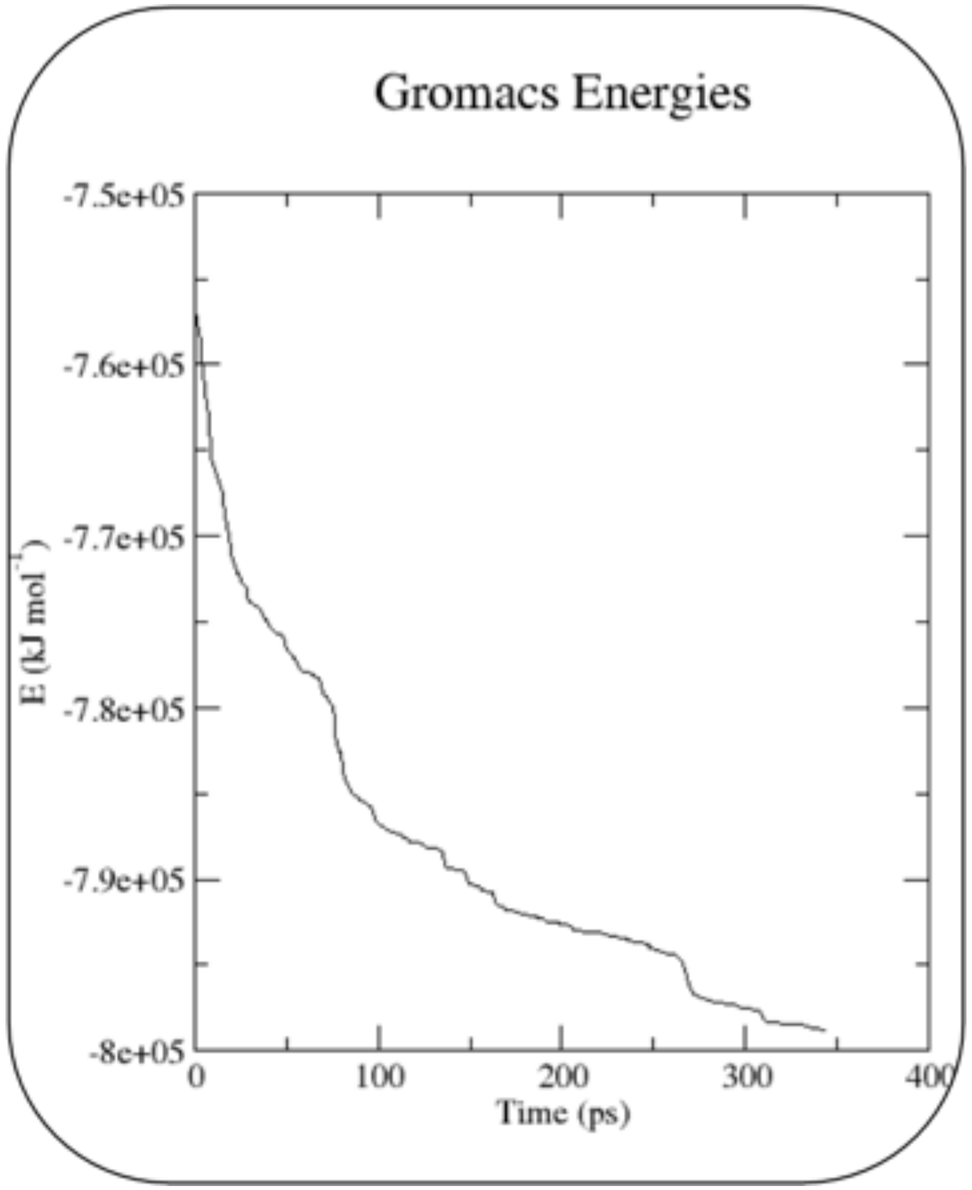


Solvated protein in a box



Scheme of a QMMM calculation

- Constructing missing MM parameters (ESP charges, prodrg)
- Structure validation
- Adding hydrogens
- Solvation
- Energy Minimization
- Equilibration / MD → snapshots



MM

- Optimization of snapshots
- Property calculations on optimized geometries

QM/ MM

Scheme of a QMMM calculation

- Constructing missing MM parameters (ESP charges, prodrg)
- Structure validation
- Adding hydrogens
- Solvation
- Energy Minimization
- Equilibration / MD → snapshots

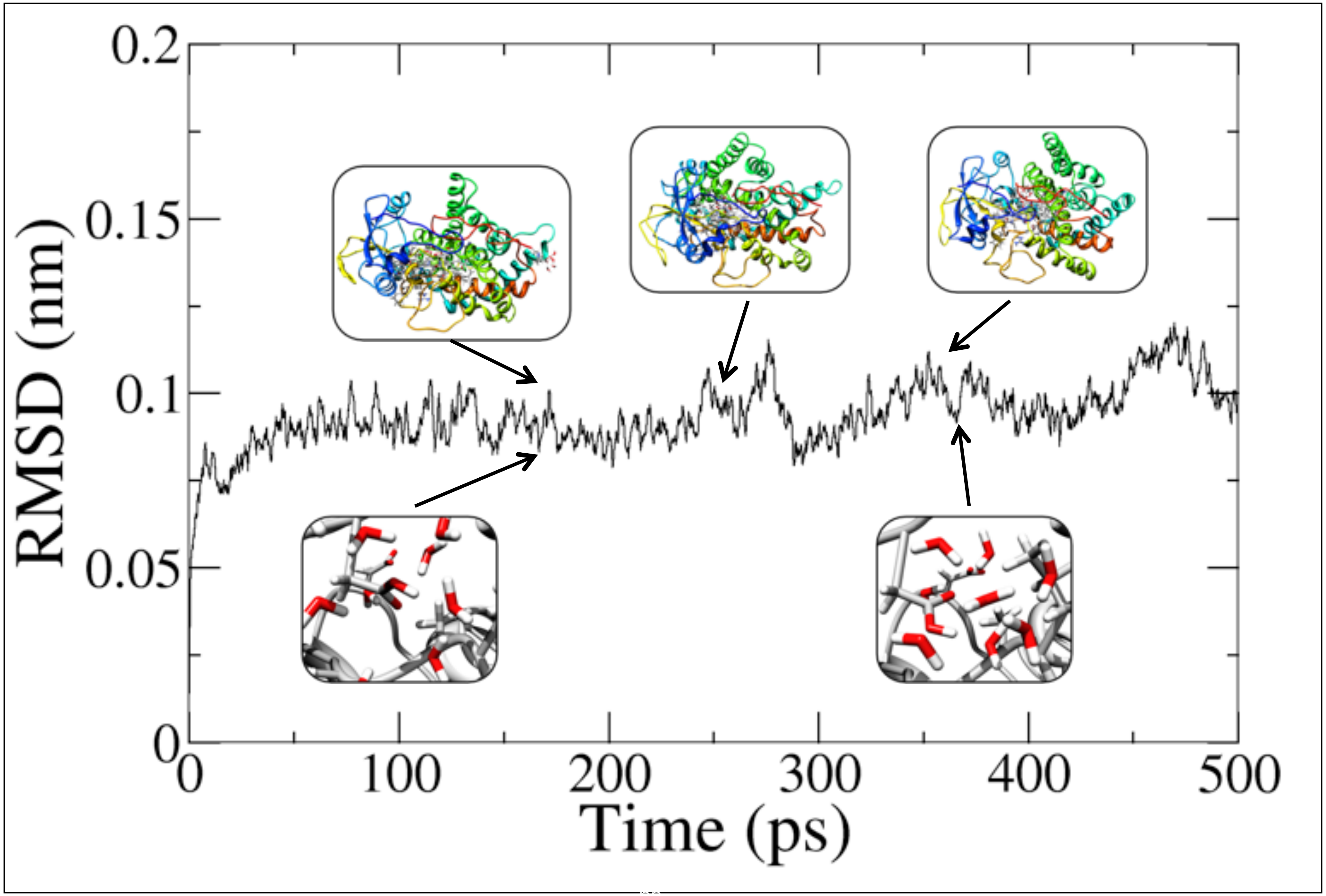
MM

- Optimization of snapshots
- Property calculations on optimized geometries

QM/MM

Scheme of a QMMM calculation

- Equilibration / MD → snapshots



Scheme of a QMMM calculation

- Constructing missing MM parameters (ESP charges, prodrg)
- Structure validation
- Adding hydrogens
- Solvation
- Energy Minimization
- Equilibration / MD → snapshots

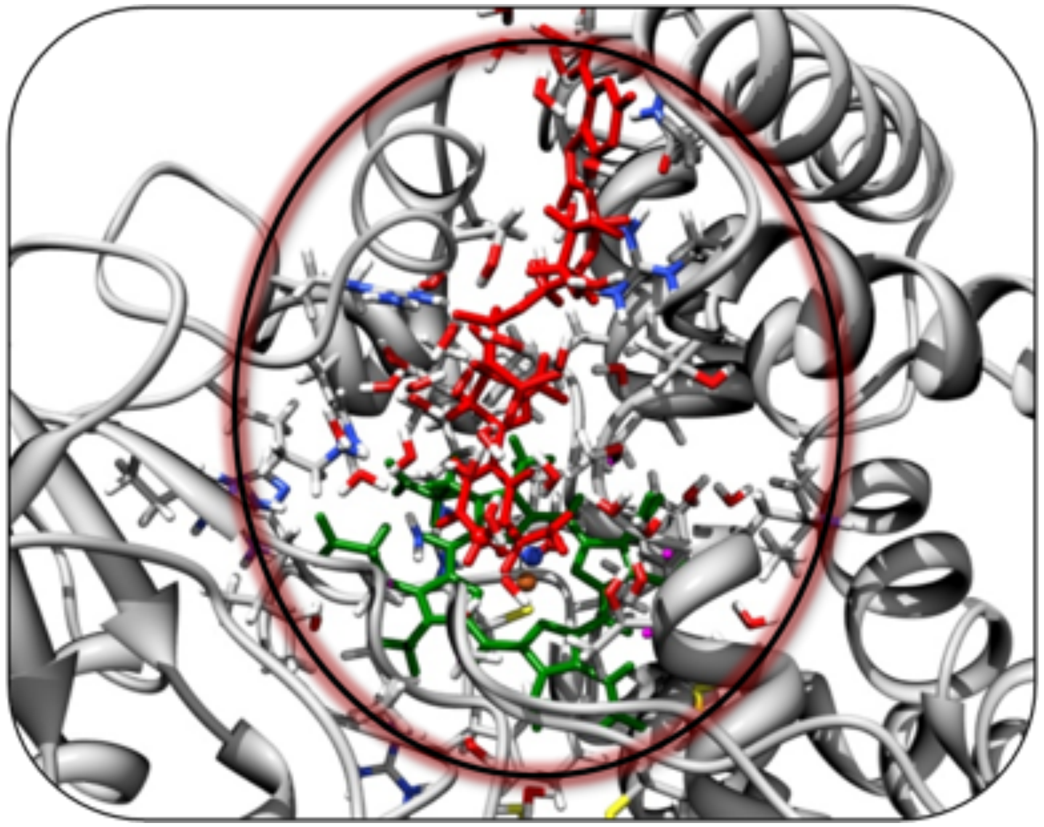
MM

- Optimization of snapshots
- Property calculations on optimized geometries

QM/ MM

Scheme of a QMMM calculation

- Optimization of snapshots:
  - A sphere around the QM-region is optimized: usually about 1000 atoms



Scheme of a QMMM calculation

MM

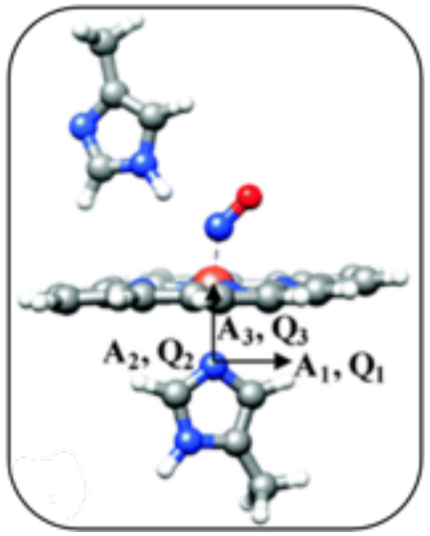
- Constructing missing MM parameters (ESP charges, prodrg)
- Structure validation
- Adding hydrogens
- Solvation
- Energy Minimization
- Equilibration / MD → snapshots

QM/MM

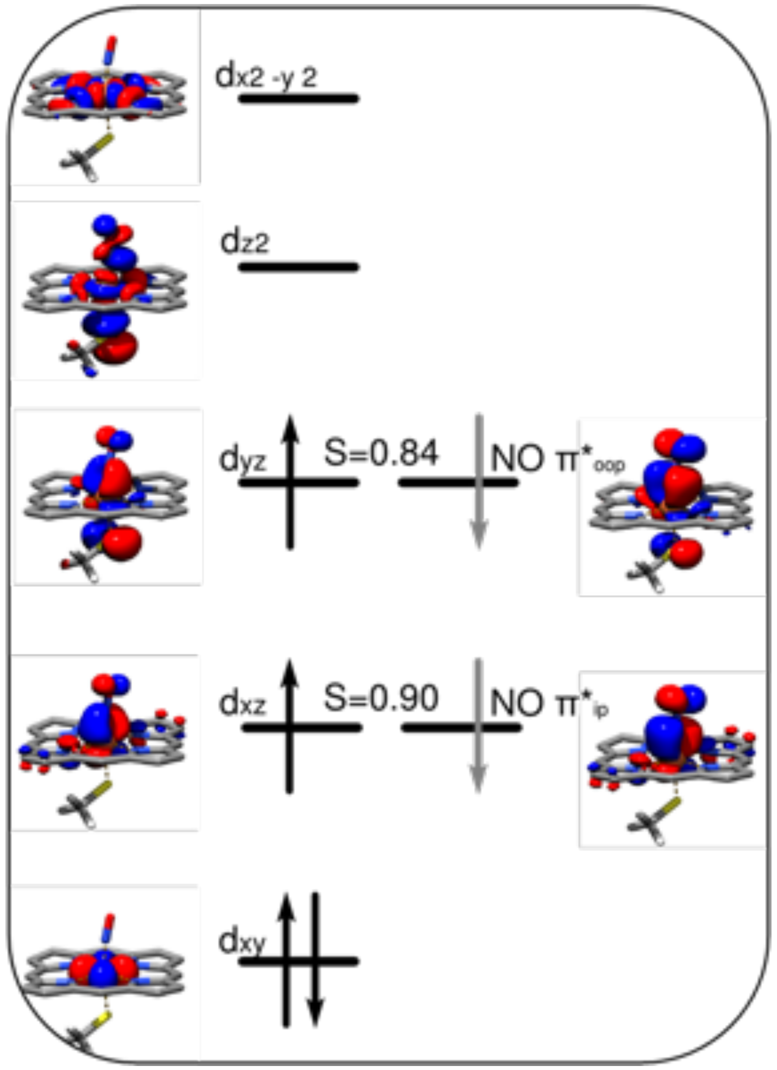
- Optimization of snapshots
- Property calculations on optimized geometries

Scheme of a QMMM calculation

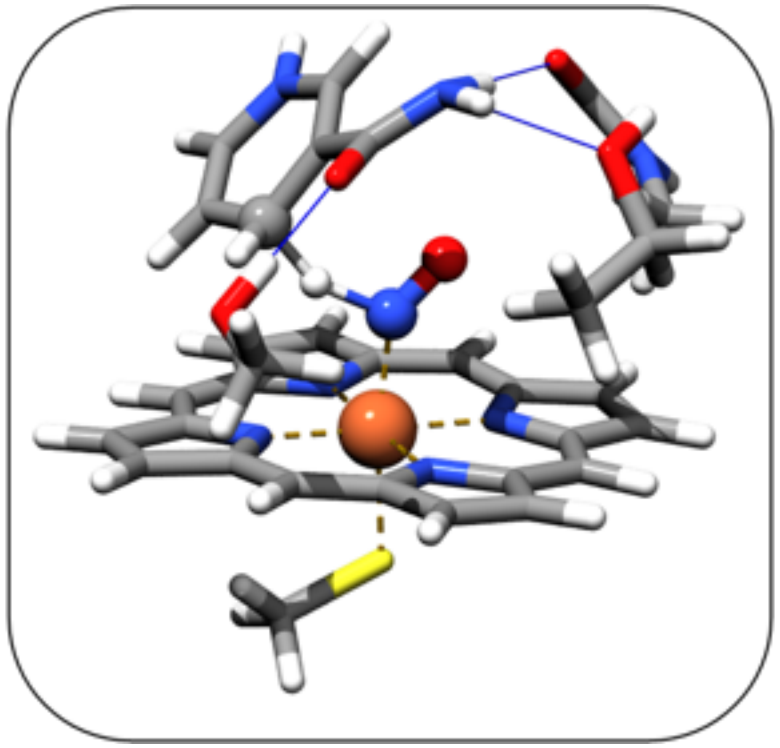
- Property calculations on optimized geometries:



EPR

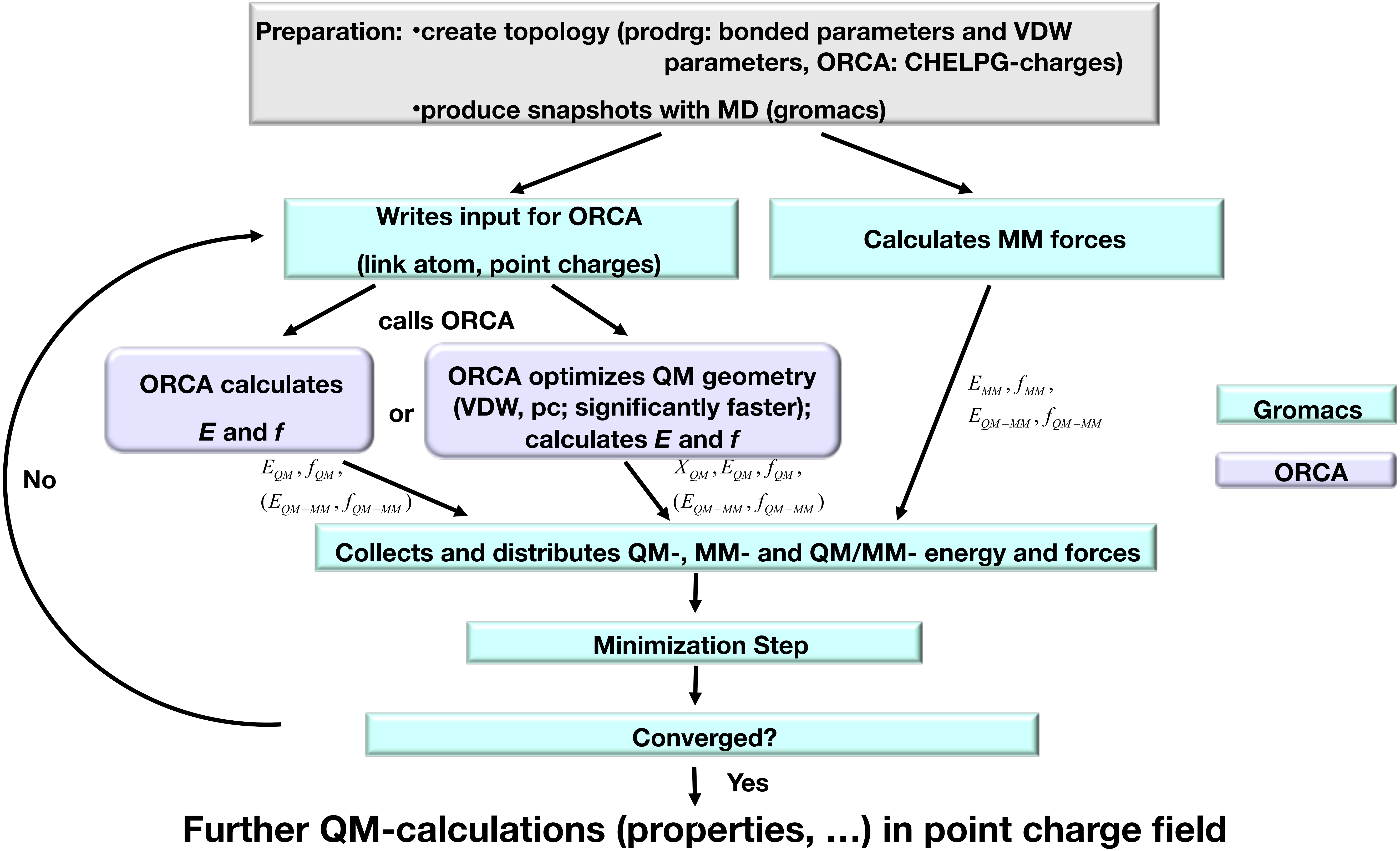


MO-analysis



Ab initio calculations → more accurate energies

Example Setup: QM/MM with Gromacs/ORCA



Preparation: •create topology (prodrg: bonded parameters and VDW parameters, ORCA: CHELPG-charges)

•produce snapshots with MD (gromacs)

Writes input for ORCA

(link atom, point charges)

Calculates MM forces

calls ORCA

ORCA calculates

E* and *f

or

ORCA optimizes QM geometry (VDW, pc; significantly faster); calculates *E* and *f*

$E_{MM}, f_{MM},$

E_{QM-MM}, f_{QM-MM}

$E_{QM}, f_{QM},$

(E_{QM-MM}, f_{QM-MM})

$X_{QM}, E_{QM}, f_{QM},$

(E_{QM-MM}, f_{QM-MM})

Collects and distributes QM-, MM- and QM/MM- energy and forces

Minimization Step

Converged?

No

Yes

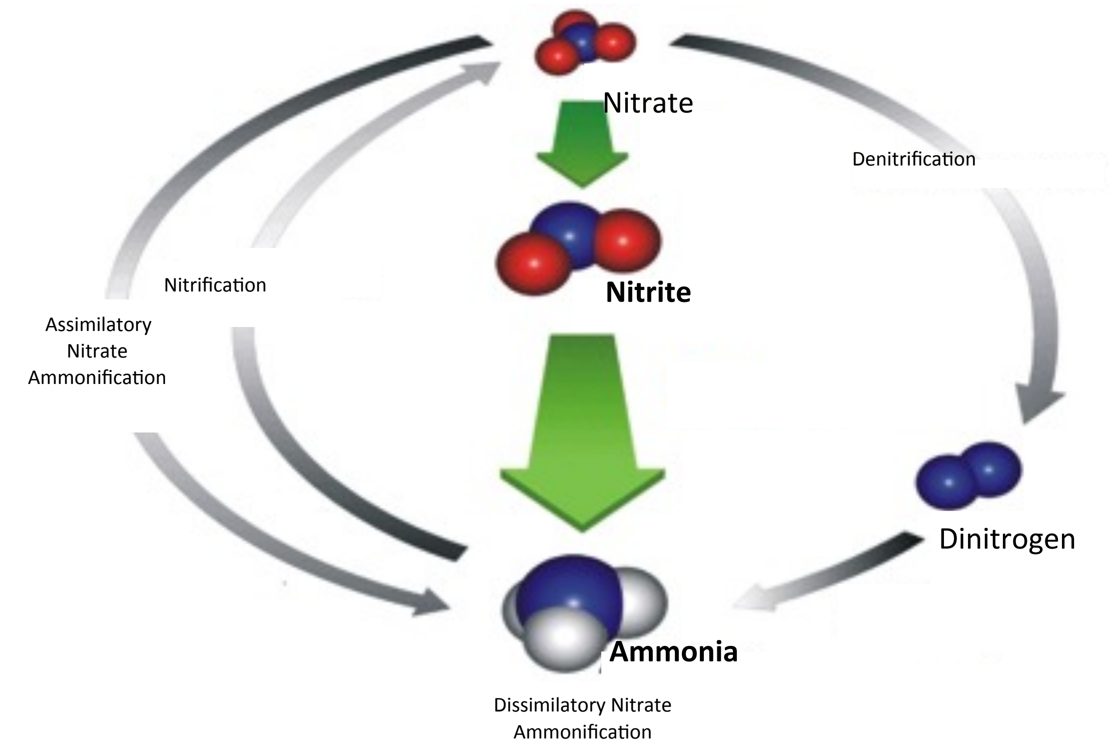
Further QM-calculations (properties, …) in point charge field

Gromacs

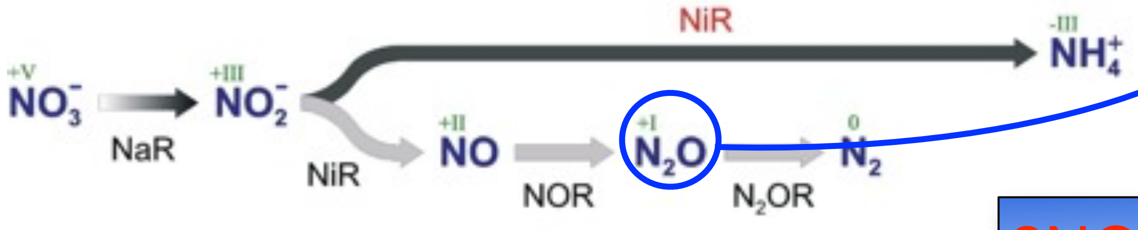
ORCA

QM/MM Example: P450 NO Reductase

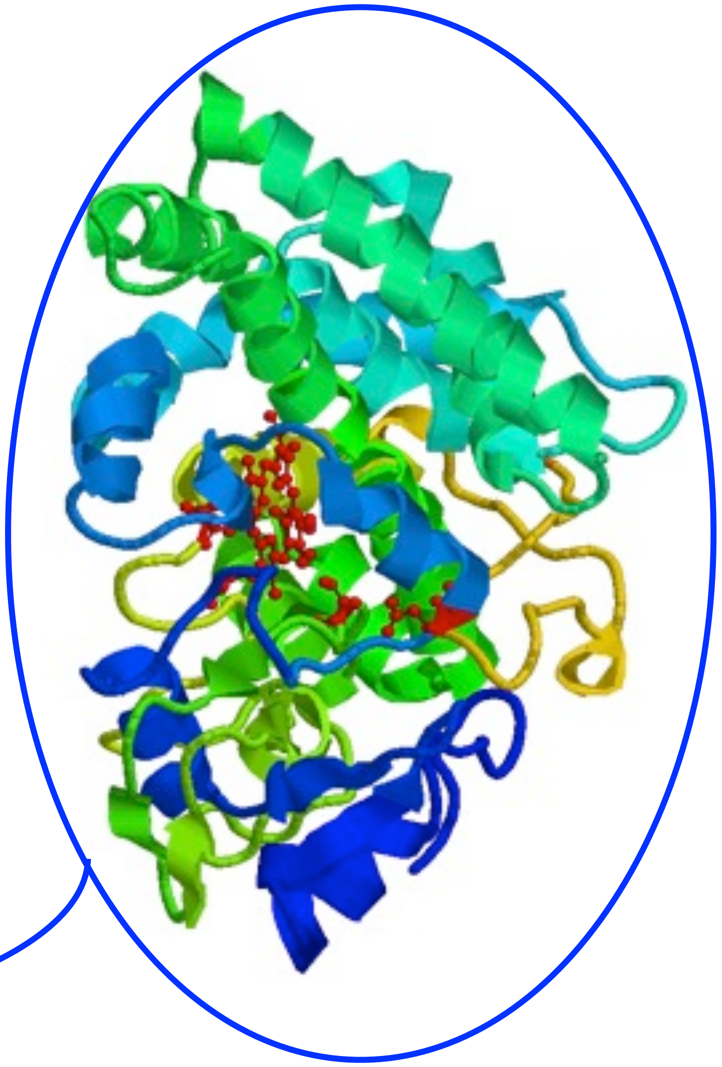
Characterizing Intermediates - P450$_{\text{nor}}$



Nitrate Ammonification



Denitrification



$2NO + 2e^- + 2H^+ \rightarrow N_2O + H_2O$

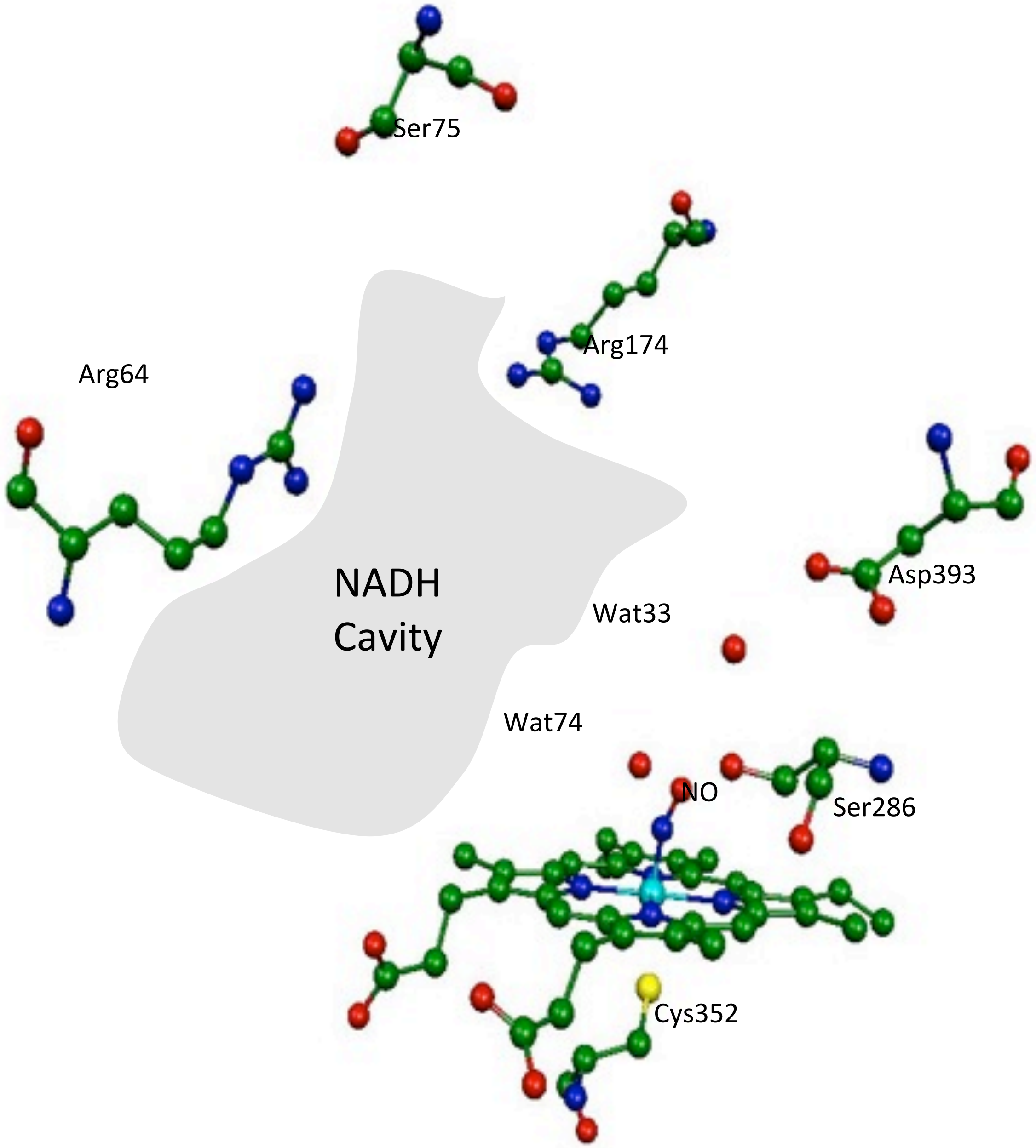
Hirofumi Shoun

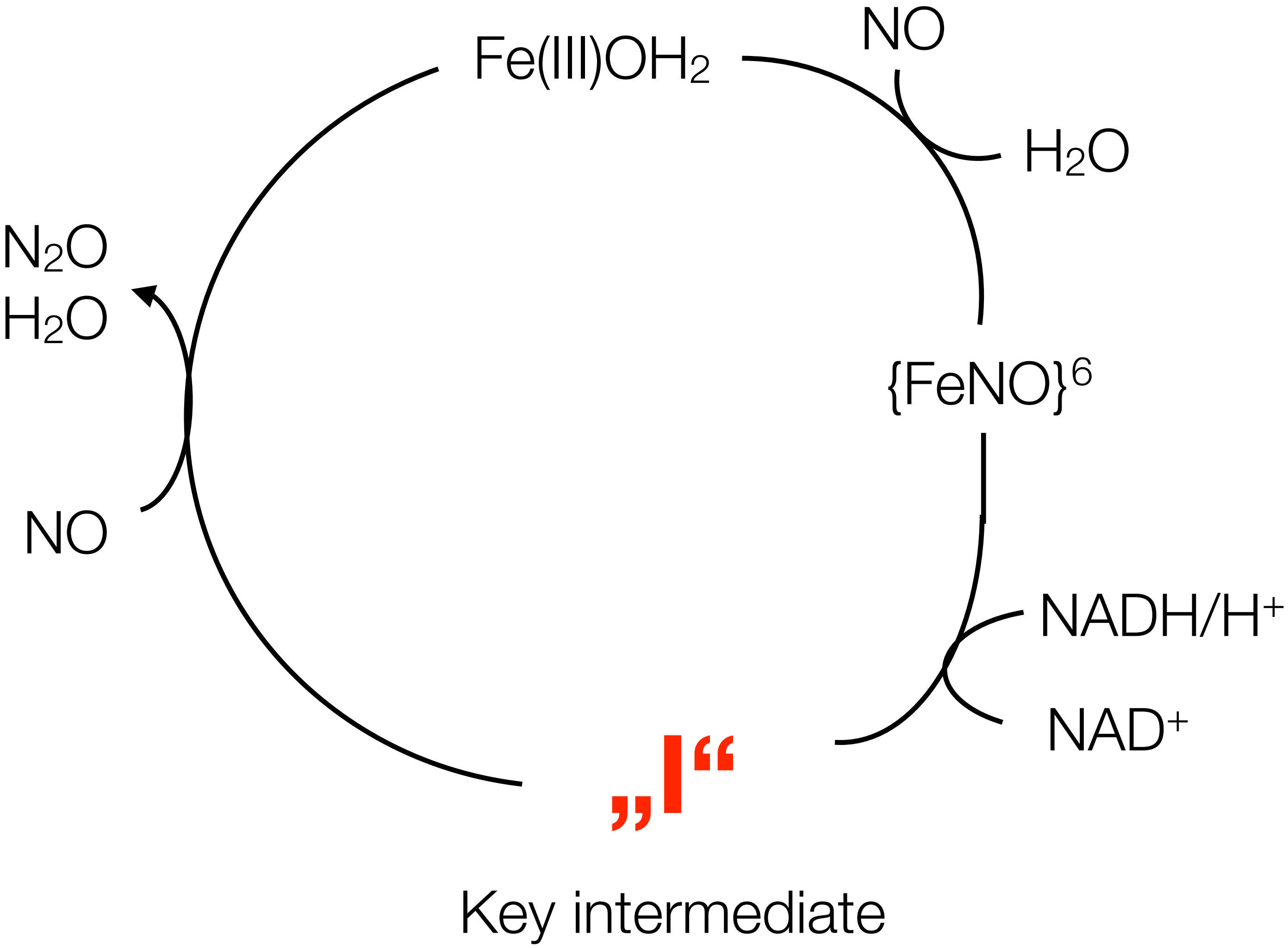
Christoph Riplinger

Eckhard Bill
Bernd Mienert
Marion Stapper

Park, S.Y. *et al.* (**1997**) *Nat. Struc. Biol.* , 4, 827

Structure and Mechanism of P450$_{nor}$





Fe(III)OH_2

NO

H_2O

$\{FeNO\}^6$

NADH/H^+

NAD^+

„I“

Key intermediate

NO

N_2O
H_2O

Shimizu, H. *et al.* (**2000**) *J. Biol. Chem.* , 275, 4816

(a) Shiro, Y. et al. (**1994**) *Biochemistry*, 33, 8673 (b) Daiber, A. *et al.* (**2002**) *J. Inorg. Biochem.*, 88, 343

Nature of Intermediate „I“ ?

$$\mathrm{Fe(III)NO} \xrightarrow[\text{? } H^+]{2e^-} \text{„I“}$$

- $[Fe\text{-}NO]^+$ $(=\{FeNO\}^8)$
- $[Fe\text{-}NOH]^{2+}$ or $[Fe\text{-}NHO]^{2+}$
- $[Fe\text{-}NHOH]^{3+}$
- $[Fe\text{-}NOH_2]^{3+}$
- $[Fe\text{-}N]^{3+}$ *(Compare $[Fe\text{-}N]^{2+}$)*

Spectroscopic Characterization of „I“

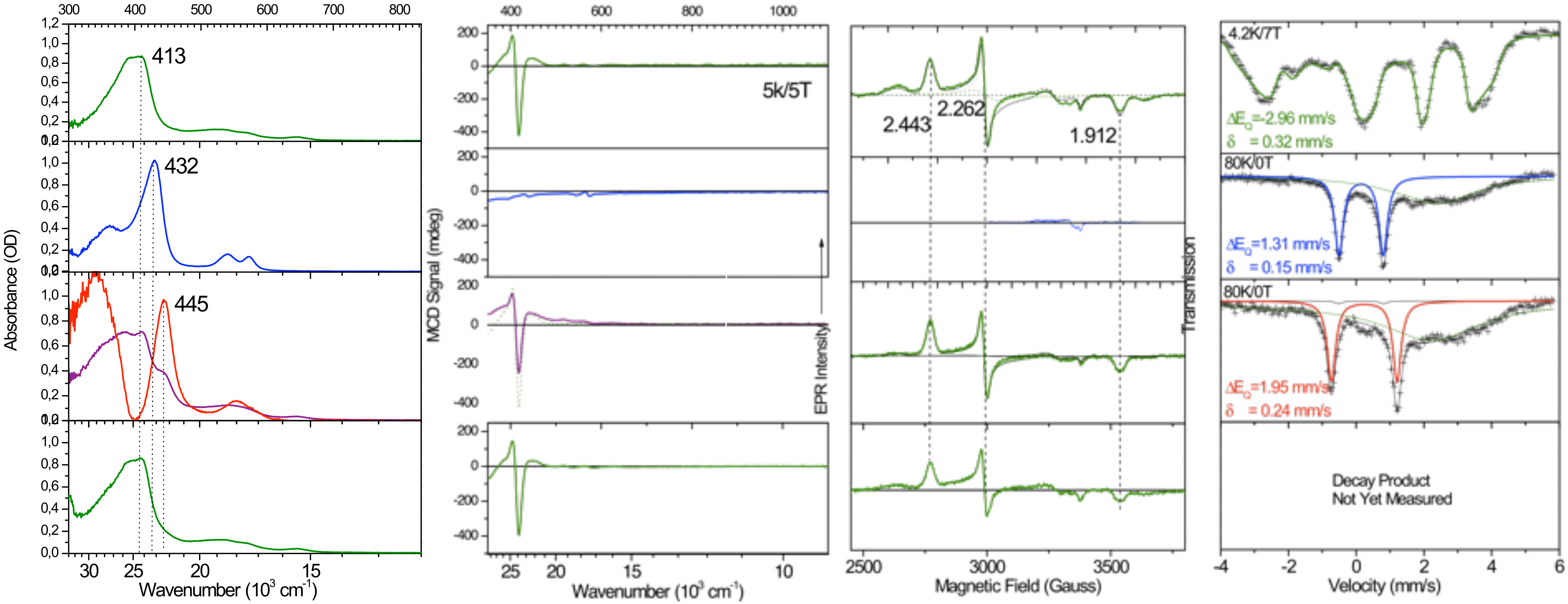
ABS **MCD** **EPR** **MB**

Native

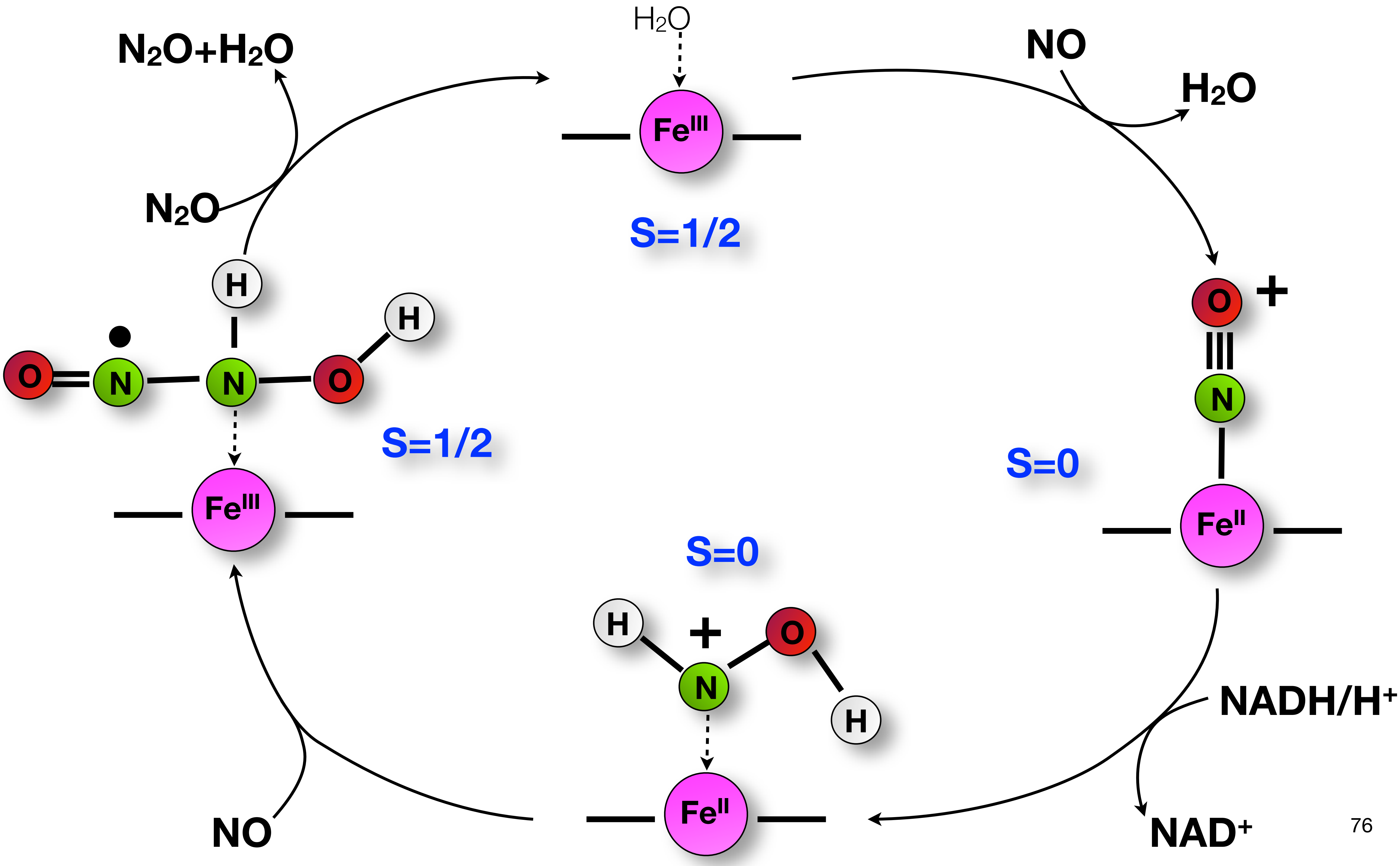
+NO

+NADH

+24h

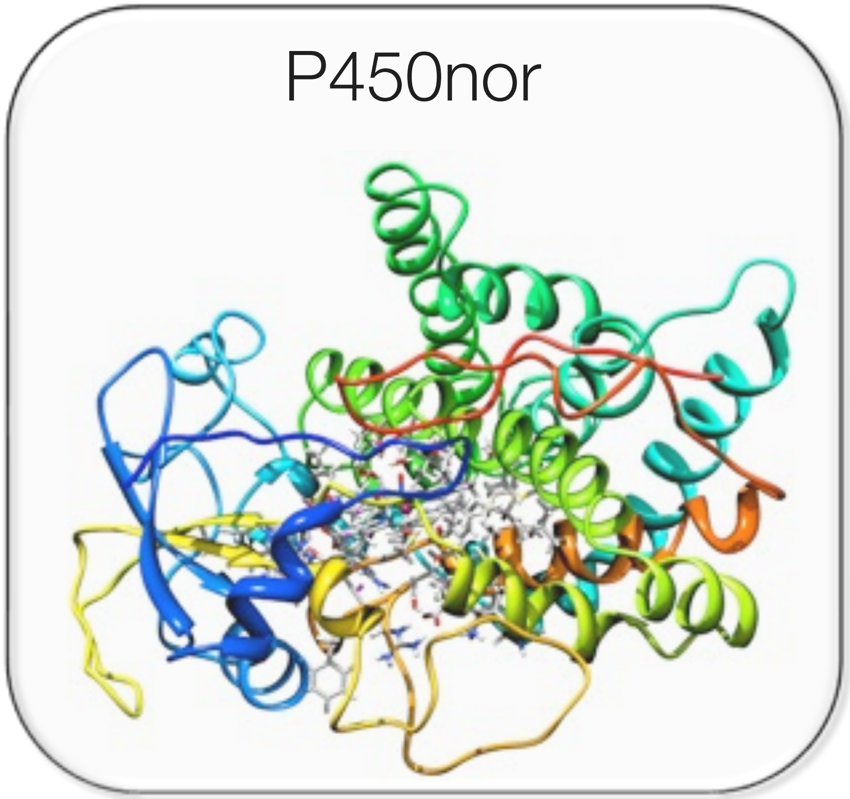


Proposed Mechanism of P450$_{nor}$

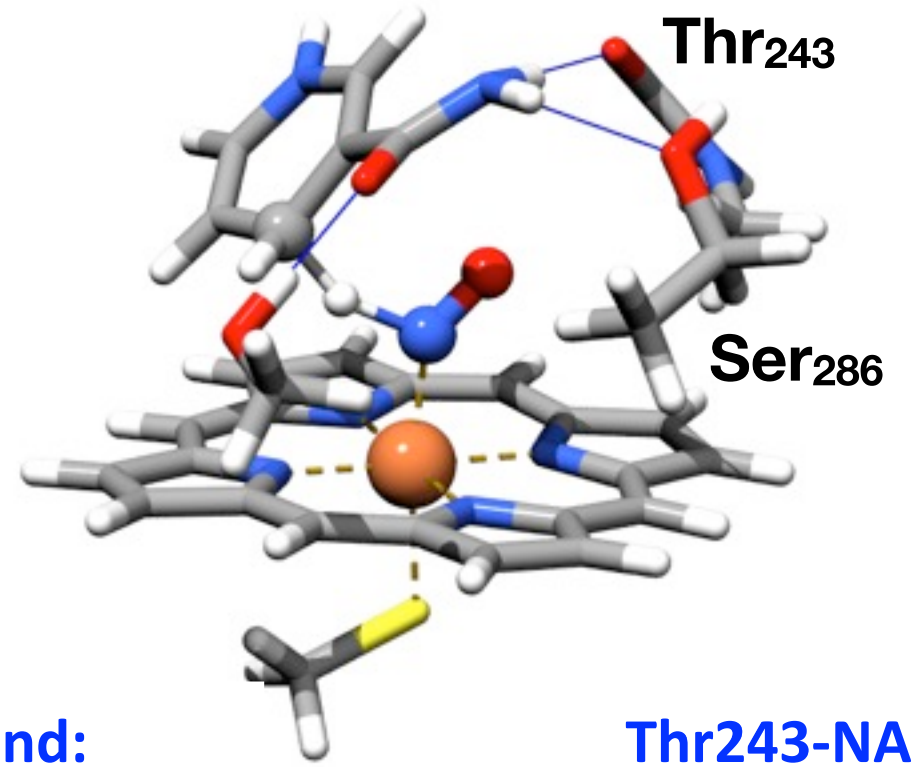


P450nor- System setup - Summary

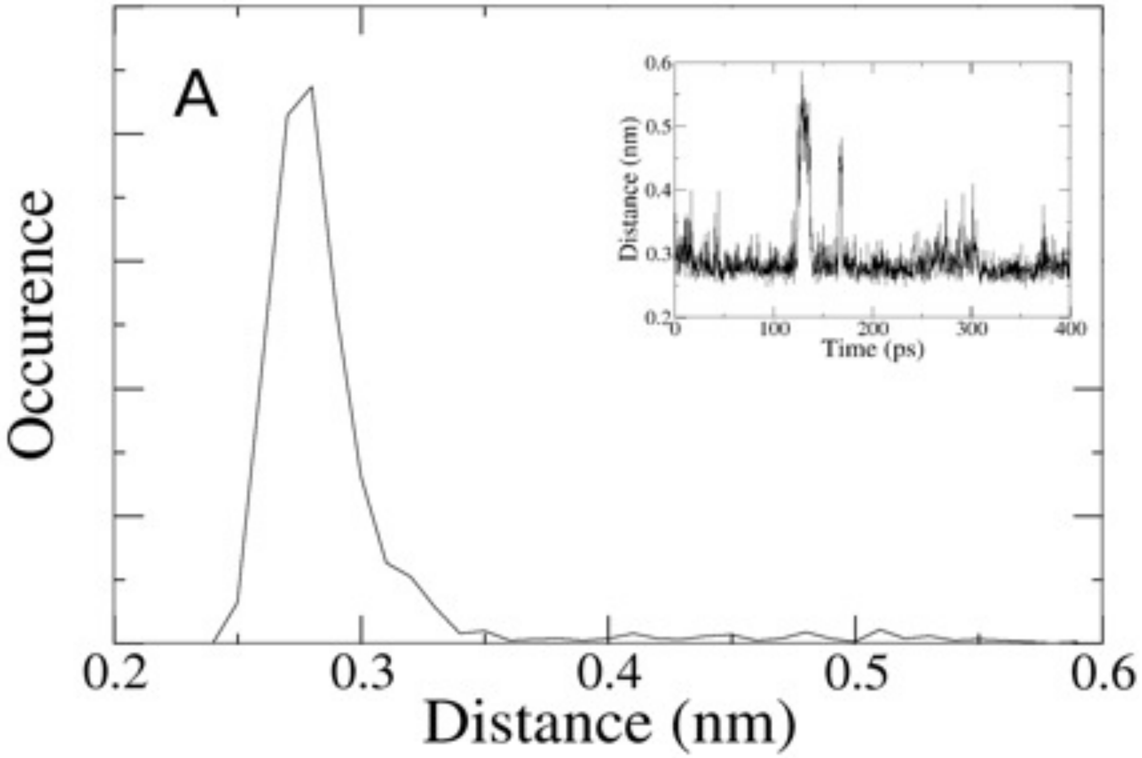
- P450nor from *Fusarium oxysporum*.
- QM/MM geometry optimizations performed at RIJCOSX-B3LYP/SV(P) / OPLSaa-level.
- Total system: 45500 atoms.
- 60-140 QM atoms
- About 1000 atoms optimized.
- Moessbauer calculations performed with B3LYP*/Fe CP(PPP)/others TZVP, with ZORA.



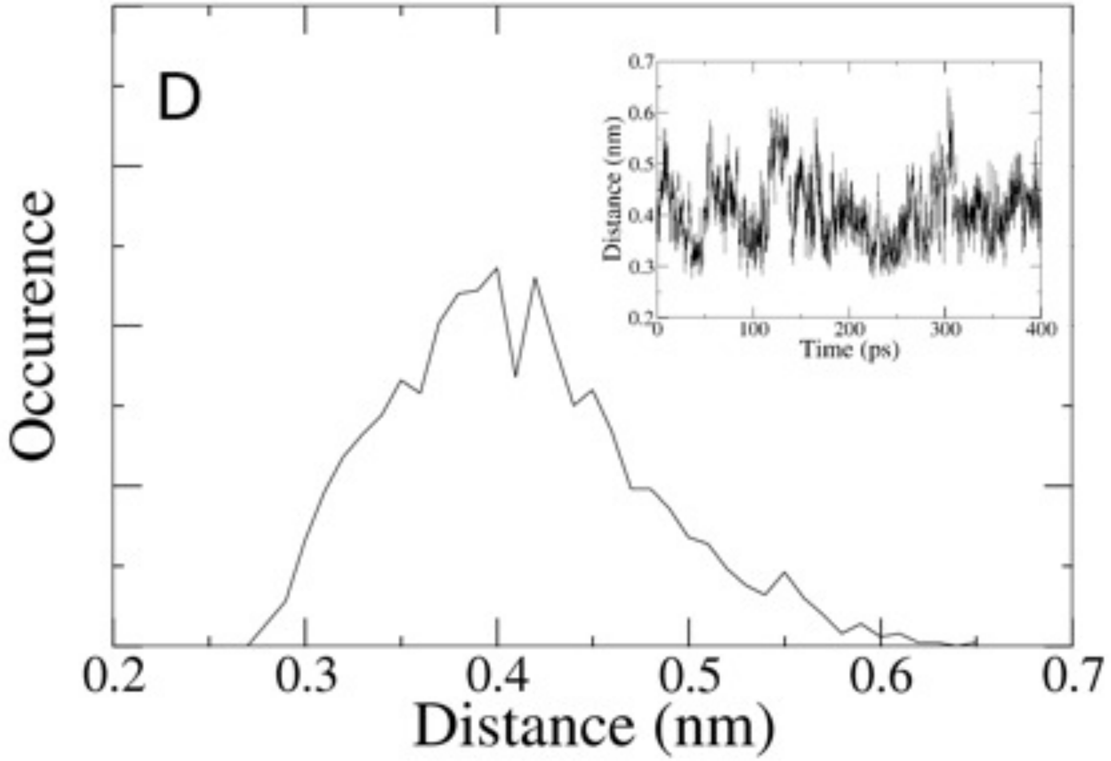
NADH Binding Interactions



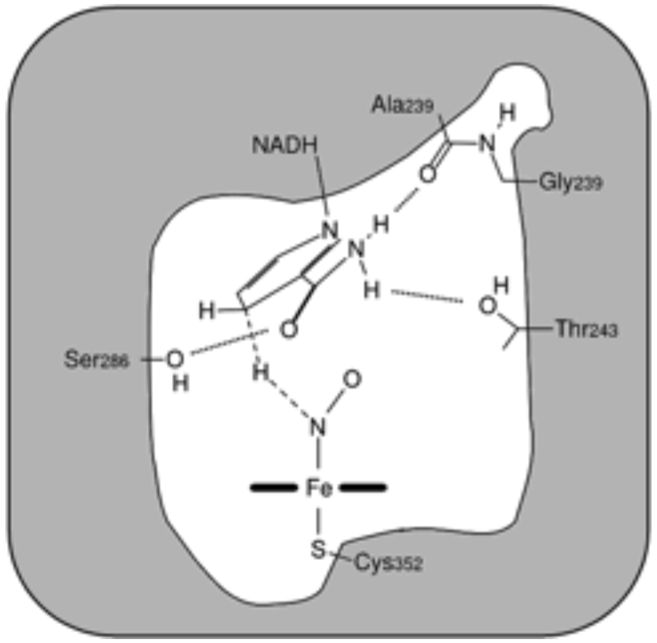
Ser286-NADH hydrogen bond:



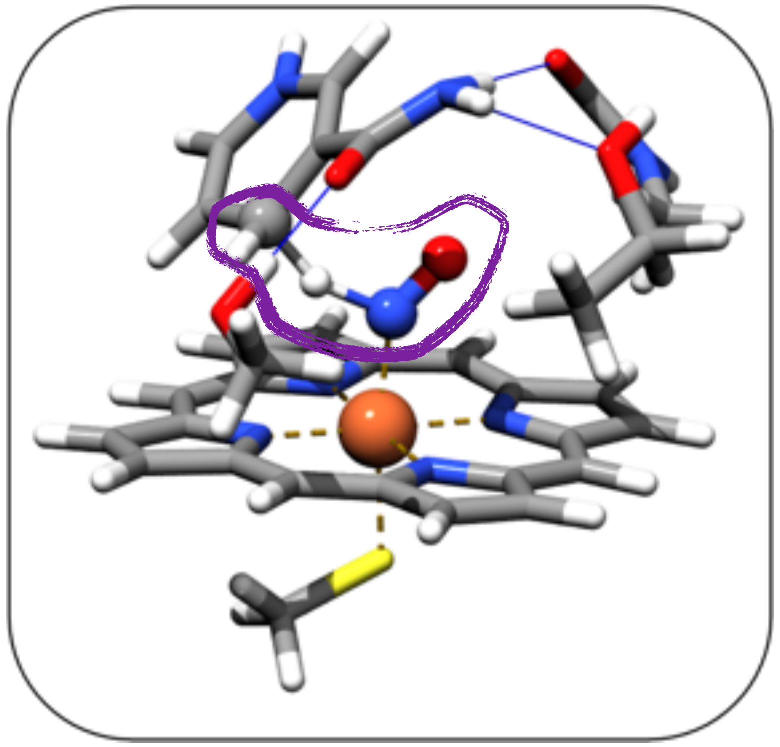
Thr243-NADH hydrogen bond:

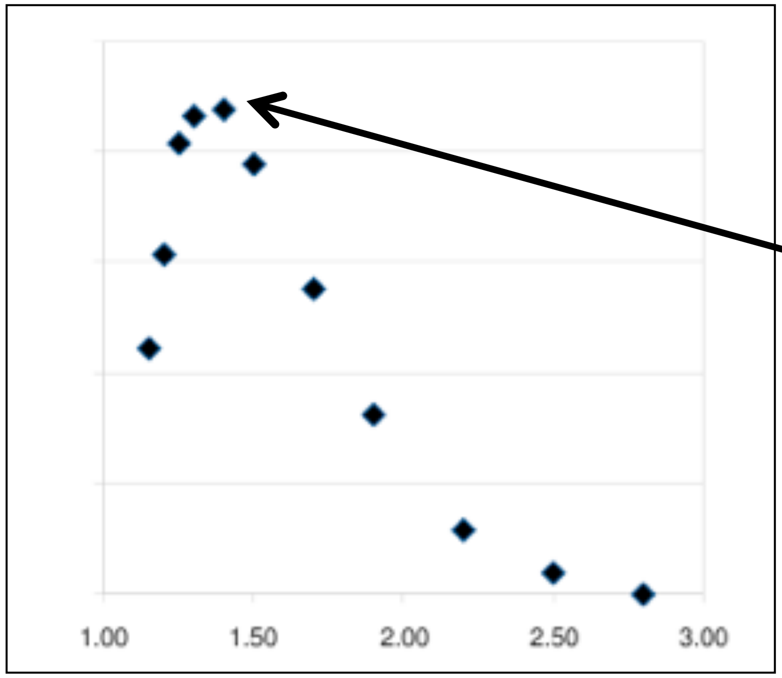


Hydride transfer – QM/MM Surface Scan

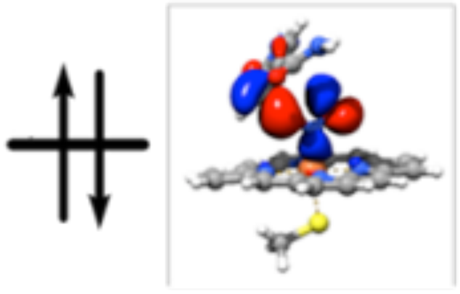


Pull hydrogen from NADH to NO



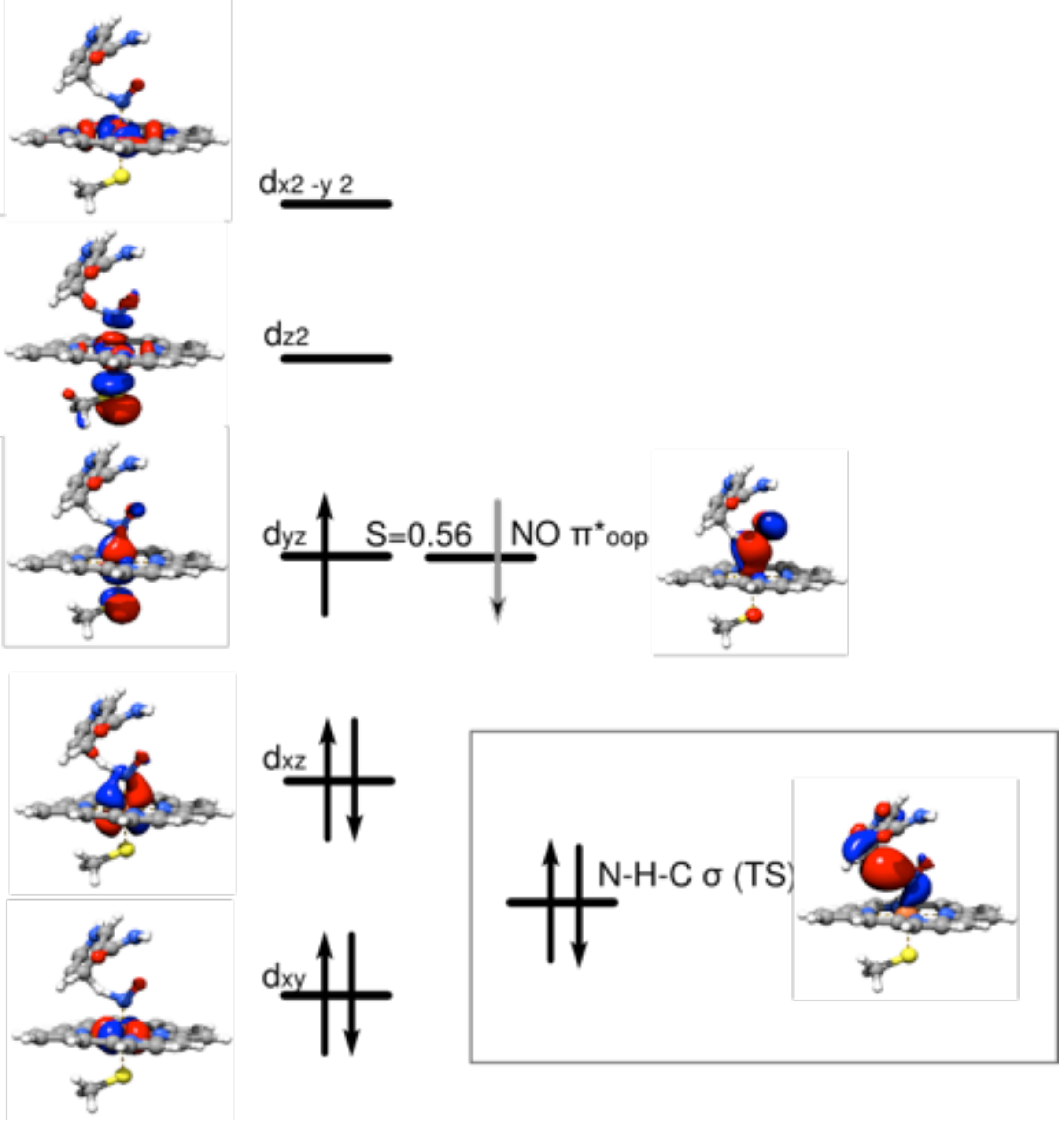


N-H-C TS MO

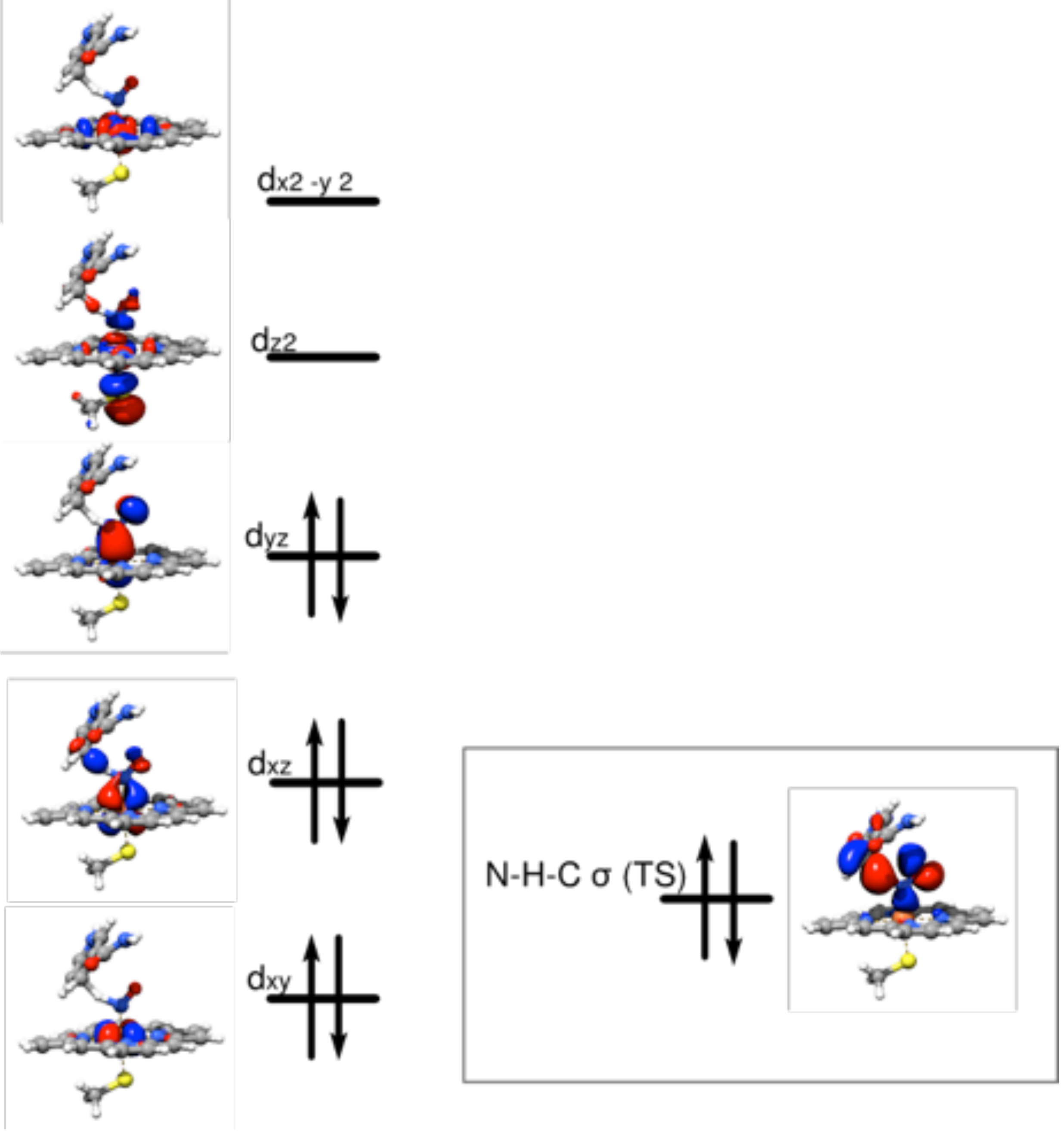


Hydride Transfer Transition State

Two different low-lying electronic states were found that are both kinetically competent:

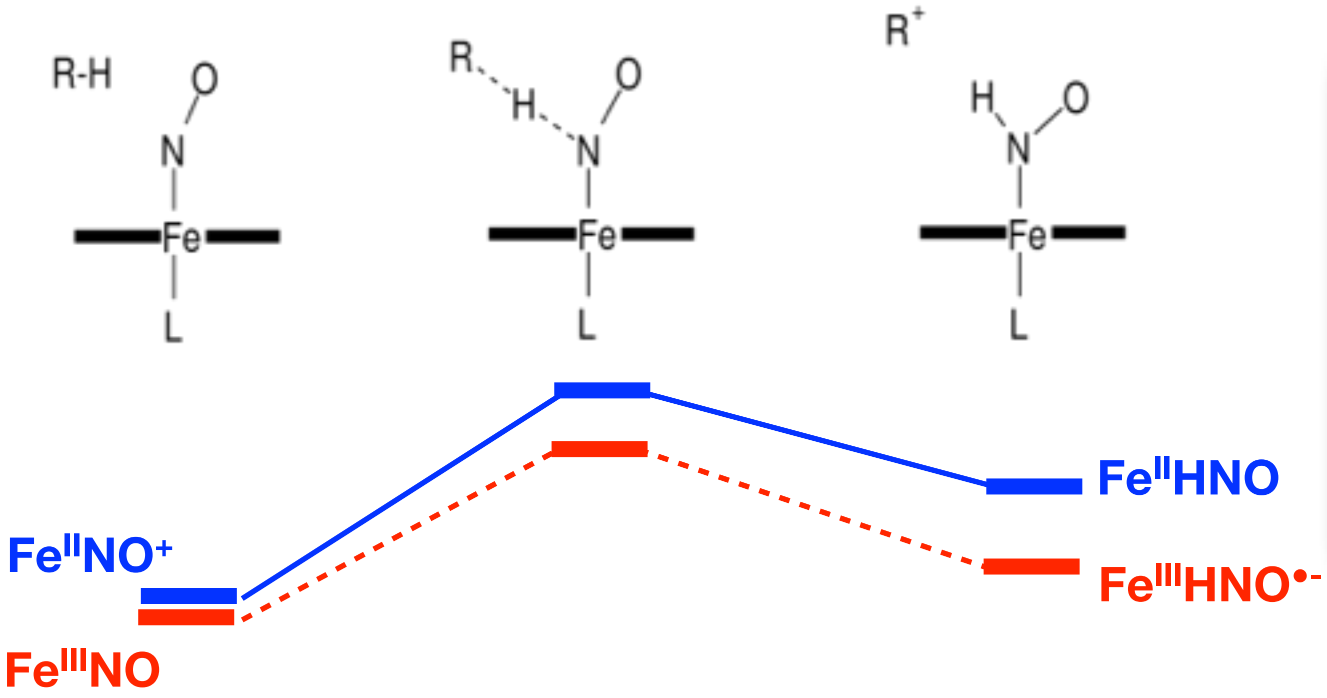


Fe(III)-species



Fe(II)-species

Hydride Transfer Transition State



Reduction of Fe(III)-NO is the rate limiting step:

- **Direct** hydride transfer from NADH to NO
- Loss of hydrogen bonding to Ser286 and Thr243 highly increases the activation barrier

		NADH + $\{FeNO\}^6$	TS	NAD^+ + $\{FeNHO\}^8$	KIE
Snapshot 1	Fe(II)	0.0	13.2	9.2	3.9
	Fe(III)	0.1	10.1	5.0	3.5
Snapshot 2	Fe(II)	0.0	9.3	8.5	3.8
	Fe(III)	0.1	6.5	4.3	3.2
Snapshot 3	Fe(II)	0.0	10.8	5.9	3.8
	Fe(III)	0.1	8.6	2.0	3.1
Exp.			**≈ 8-9**[a]		**2.7±0.4**[b]

[a] Shiro et al. (**1995**) 270, 1617

[b] Daiber et al. (**2002**) 88, 343

Hydride transfer – Link to Experiment

Comp. activation barrier: 6.5 – 9.3 kcal/mol

agrees well with

Exp. barrier: 8-9 kcal/mol

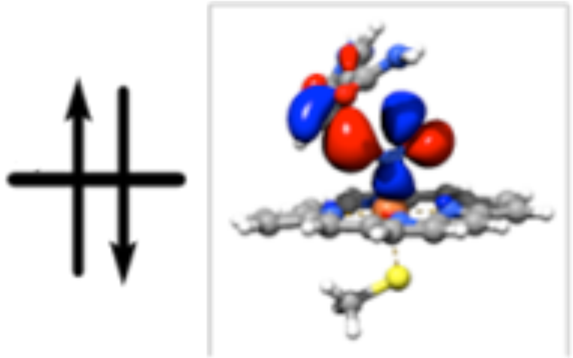
Comp. KIE: 3.1 – 3.5

agrees well with

Exp. KIE: 2.7 ±0.4

TS-Energy

N-H-C TS MO



Frequency Calculation

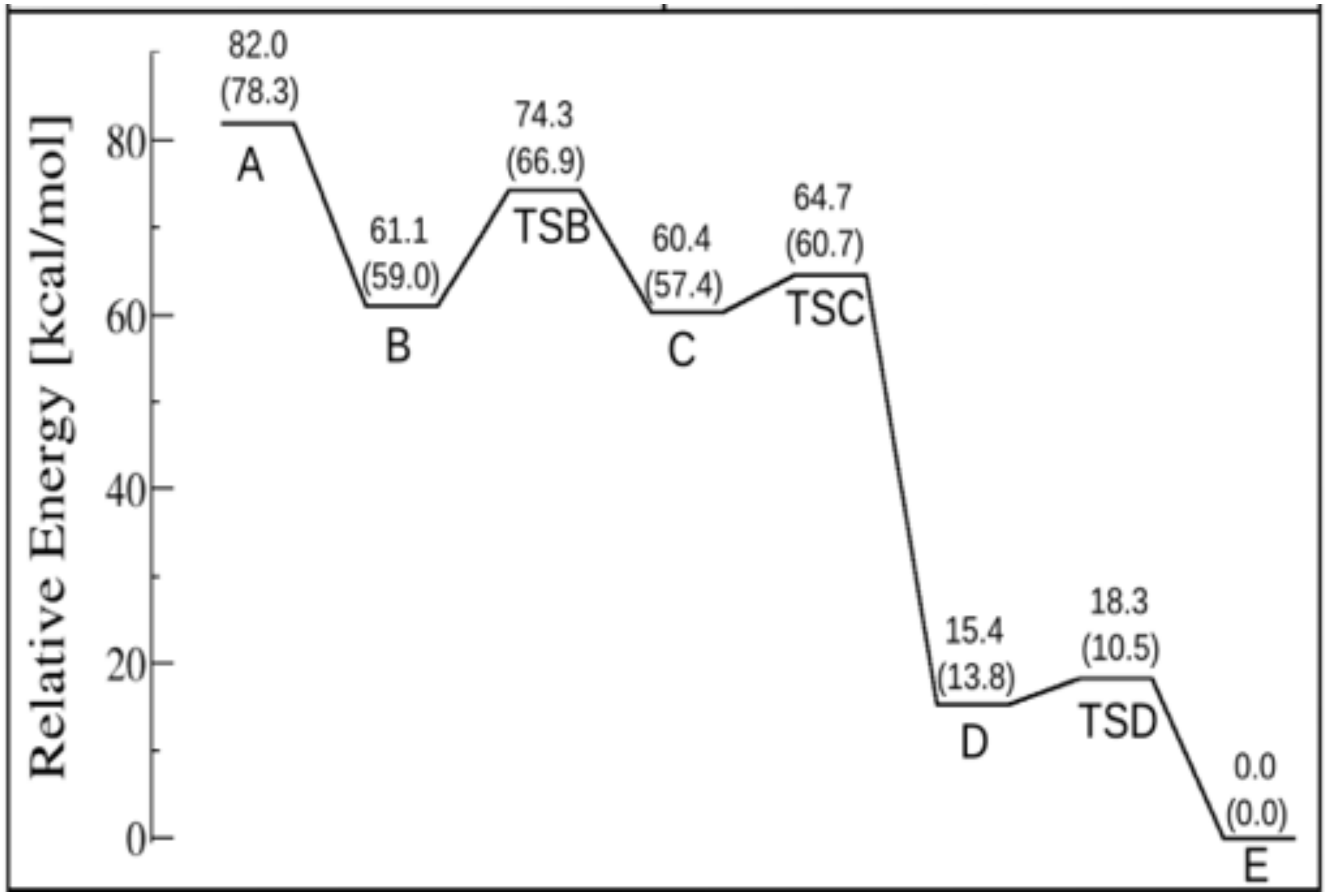
Concerted Proton and two-electron transfer

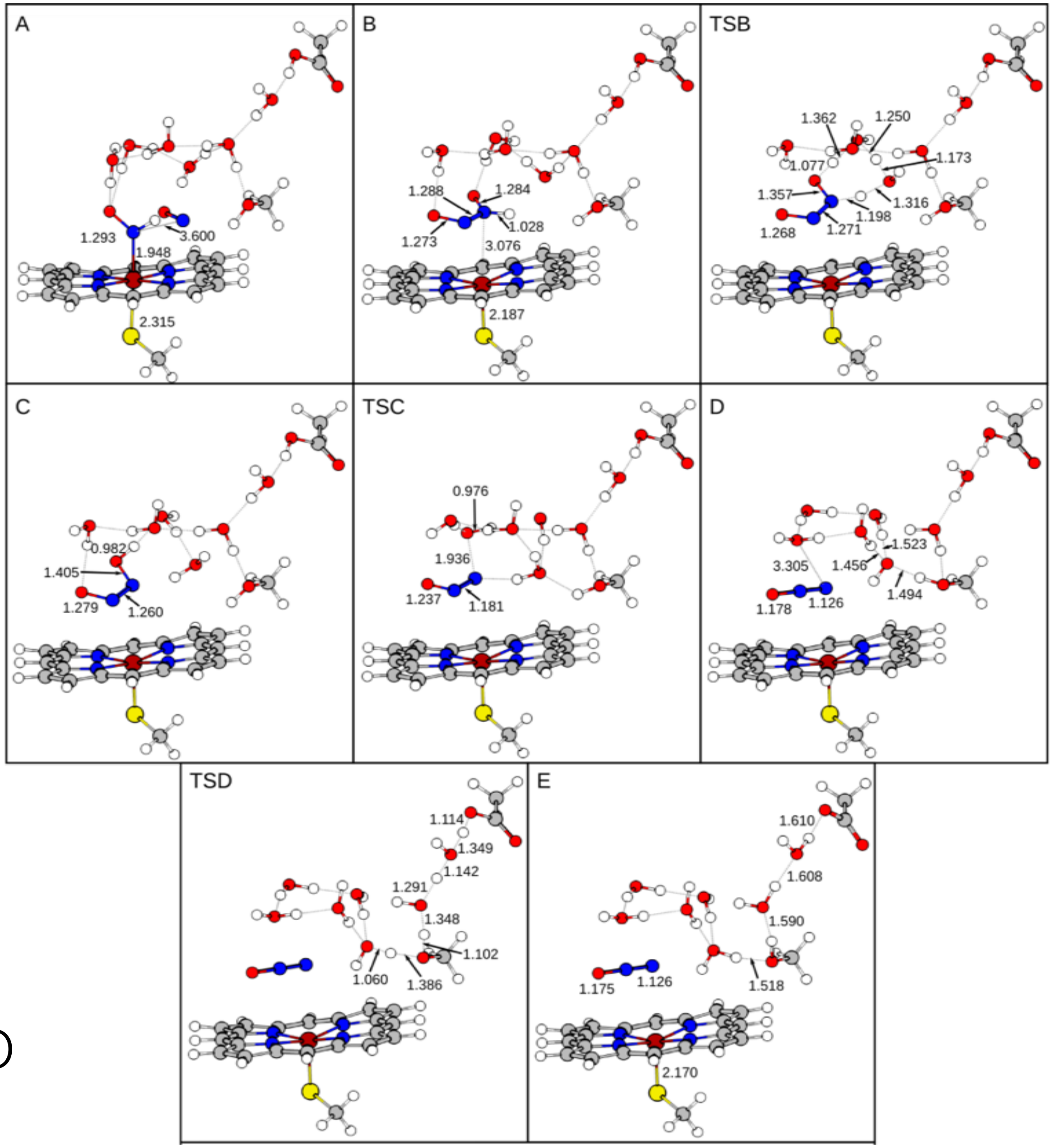
Protonation State of Intermediate I

	Fe^{III}-Water.	$\{FeNO\}^6$	Intermediate I	
			$\{FeHNO\}^8$	$\{FeHNOH\}^8$
Fe^{II}	0.31/-2.90	0.04/0.76	0.23/-1.84	0.08/-2.88
Fe^{III}		0.10/1.20	0.27/-1.68	0.22/-2.47
Exp.	0.32/-2.96	0.15/1.31	0.24/1.95	

- Calculations are in excellent agreement with experiment for the two well characterized species.
- The Mössbauer data is consistent with mono- or diprotonated species

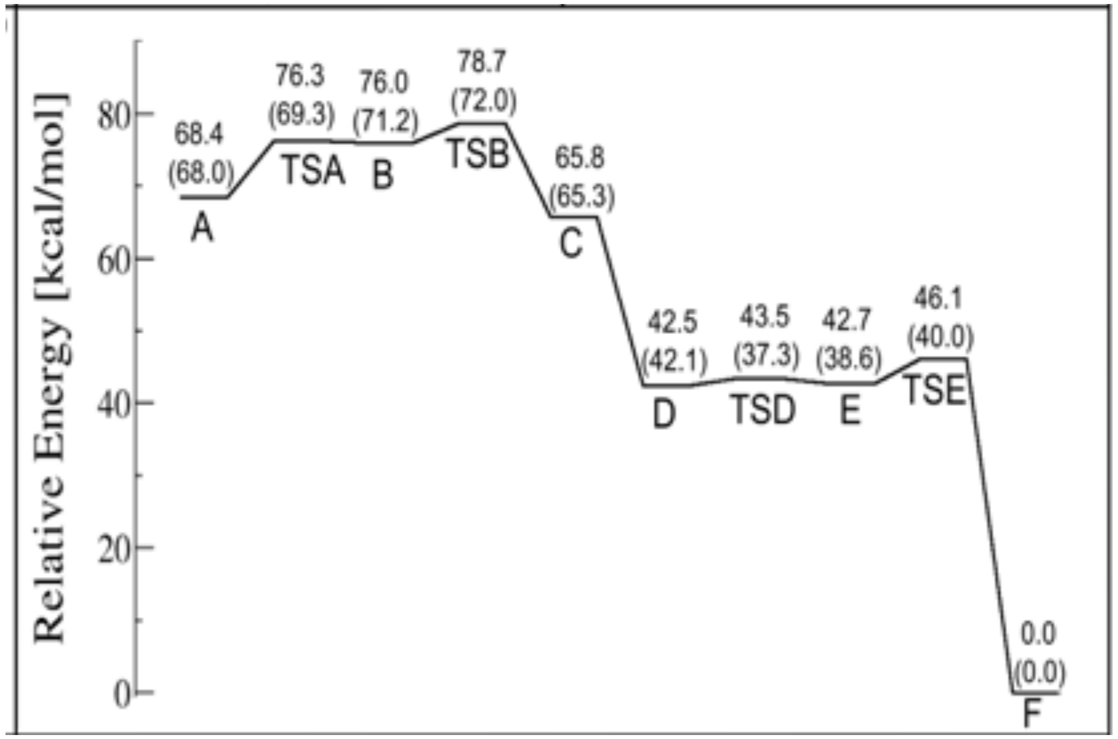
Second Half Reaction via HNO

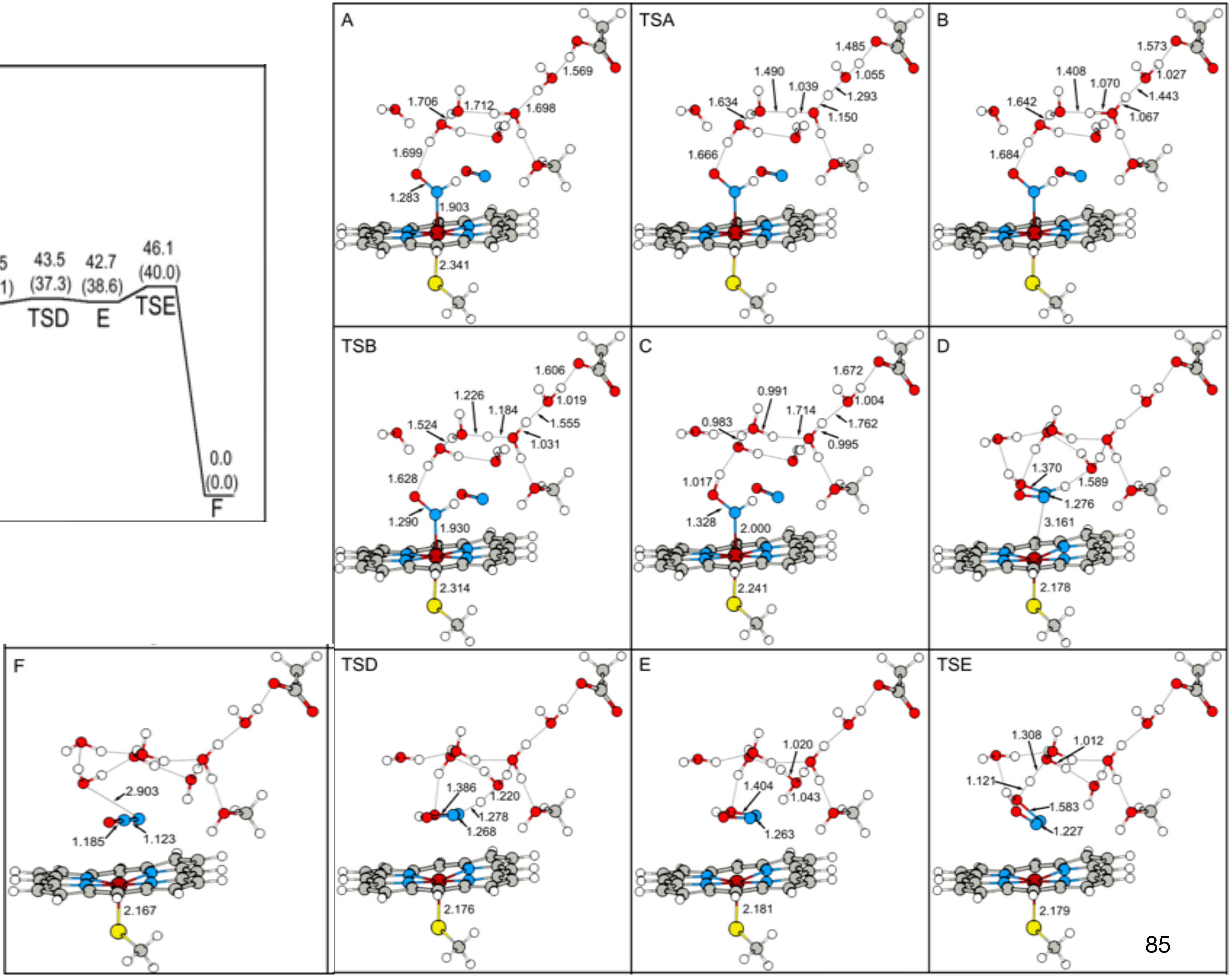




- ✓ Barrierless N-N coupling reaction with second molecule of NO in the diradical state $Fe^{III}HNO^{\bullet -}$
- ✓ Spontaneous decomposition into N_2O and H_2O after protonation from Asp

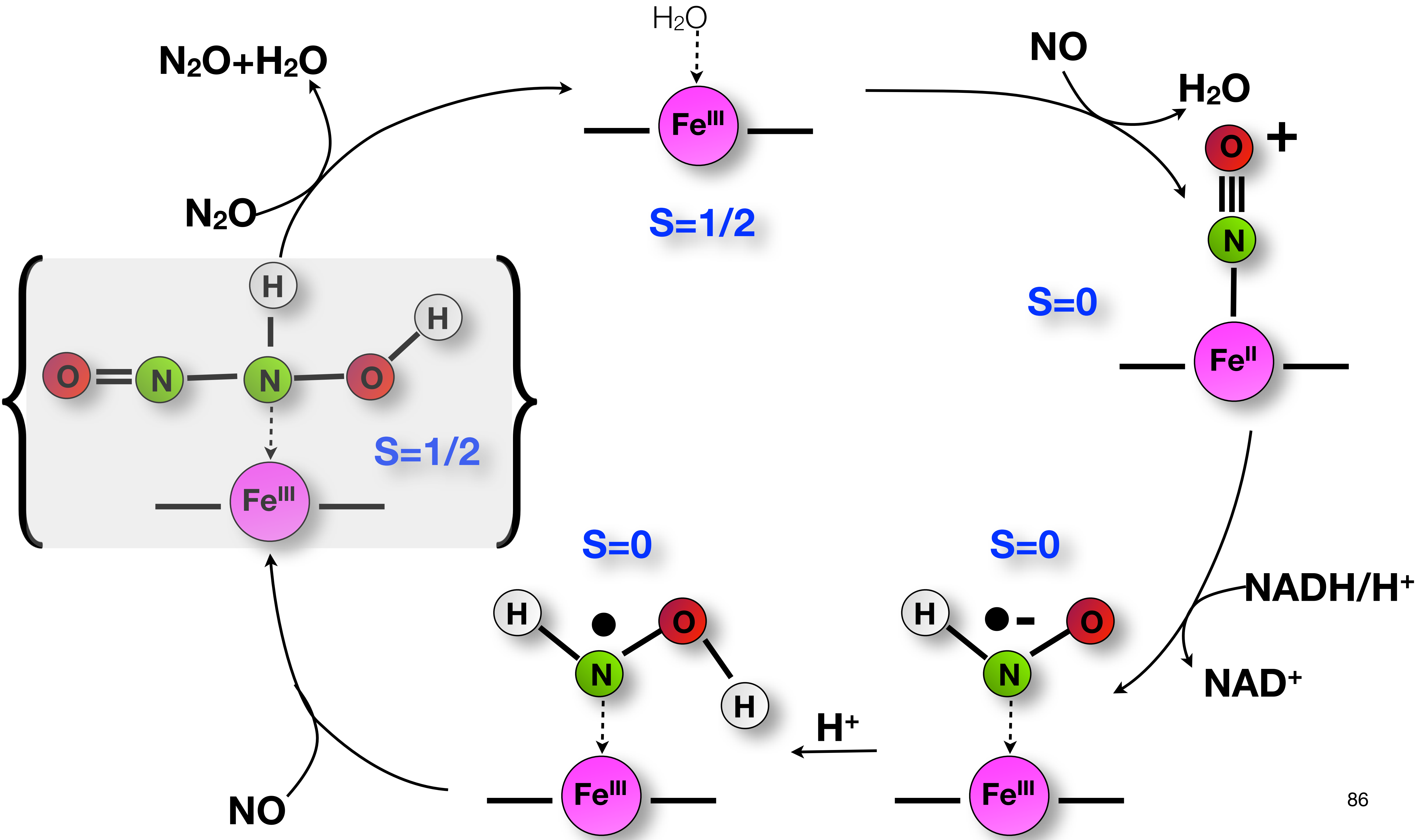
Second Half Reaction via HNOH





- ✓ HNOH Intermediate formed from HNO
- ✓ Energy barriers ~3 kcal/mol lower
- ✓ Barriers slightly lower than in the HNO case

Revised Mechanism of $P450_{nor}$



Summary

- Successfully performing computational chemistry projects requires careful planing, awareness of methodological issues and knowledge of available experimental data.
- Many computational projects can be properly addressed with DFT once properly calibrated.
- When applicable (single-reference, affordable cost), coupled-cluster methods are preferable.
- Multireference problems are abundant in chemistry and need to be carefully addressed. Stay Single-reference as along as you can.
- **Computers don‘t solve problems - people do!**

Have fun with

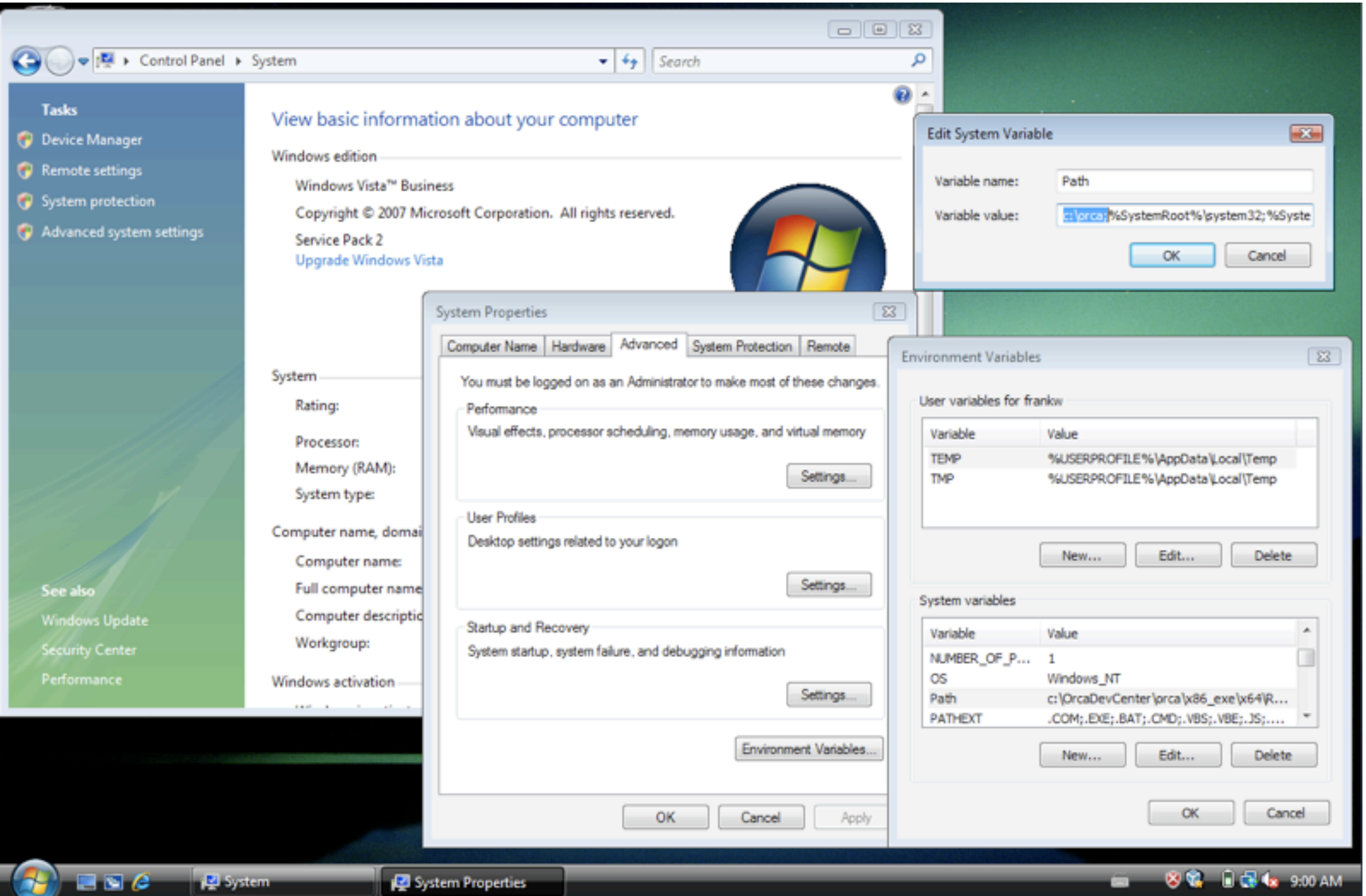
.... ORCA

http://www.mpibac.mpg.de (download area)

Appendix: More information on Computational Methods

Configuring ***ORCA*** under Windows Vista

Setting the **PATH variable** to the ***ORCA*** directory



View basic information about your computer

Windows edition

Windows Vista™ Business

Copyright © 2007 Microsoft Corporation. All rights reserved.

Service Pack 2

Upgrade Windows Vista

System

Rating:

Processor:

Memory (RAM):

System type:

Computer name, domai

Computer name:

Full computer name

Computer descriptic

Workgroup:

Windows activation

System Properties

You must be logged on as an Administrator to make most of these changes.

Performance

Visual effects, processor scheduling, memory usage, and virtual memory

User Profiles

Desktop settings related to your logon

Startup and Recovery

System startup, system failure, and debugging information

Edit System Variable

Variable name: Path

Variable value: c:\orca;%SystemRoot%\system32;%Syste

Environment Variables

User variables for frankw

Variable	Value
TEMP	%USERPROFILE%\AppData\Local\Temp
TMP	%USERPROFILE%\AppData\Local\Temp

System variables

Variable	Value
NUMBER_OF_P...	1
OS	Windows_NT
Path	c:\OrcaDevCenter\orca\x86_exe\x64\R...
PATHEXT	.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;....

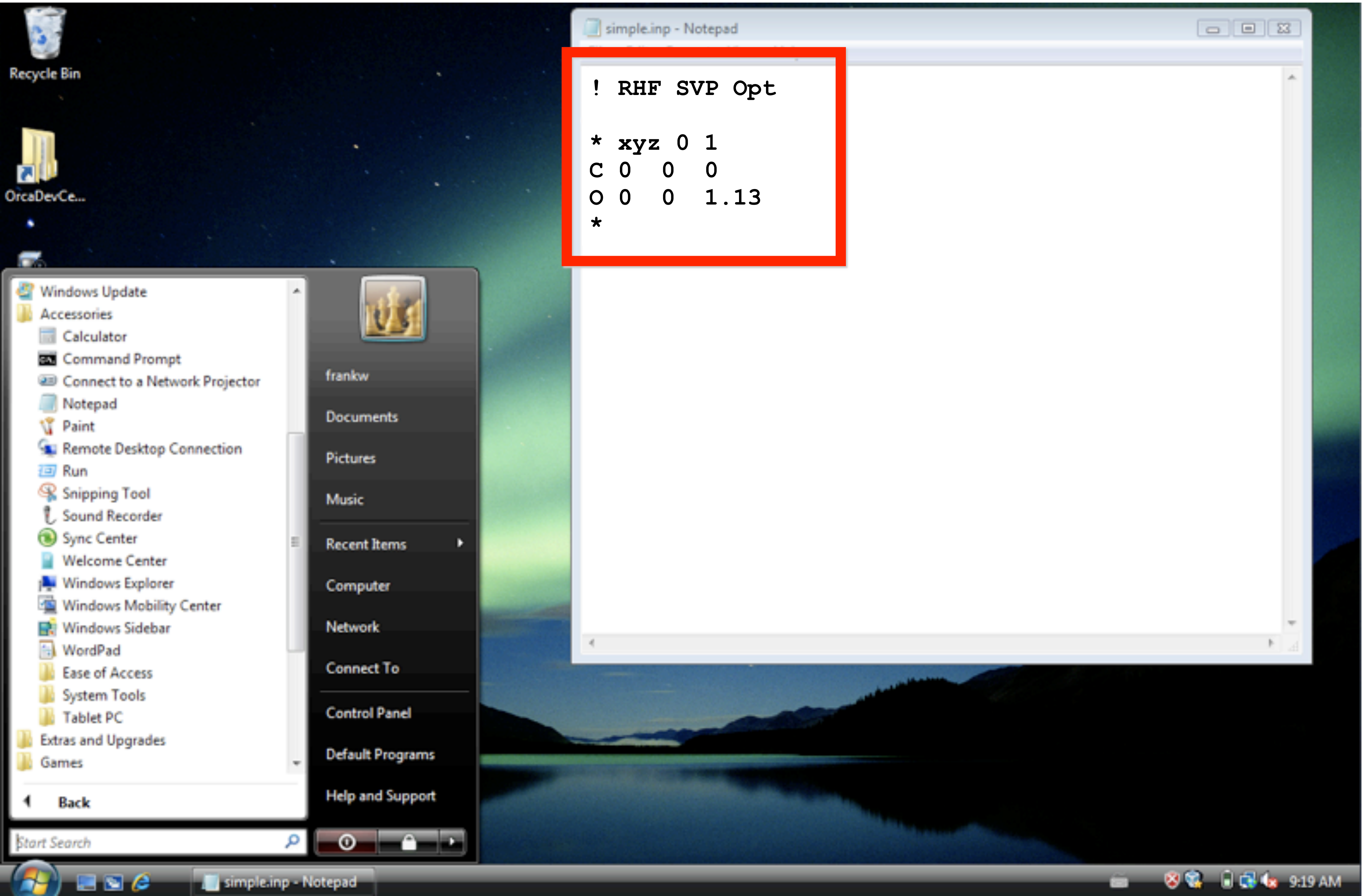
The Path Variable can be accessed via

Control Panel > System > Advanced system Settings

Here it is assumed that ORCA resides in **c:\orca**

Editing Textfiles under Windows Vista

Editing Text using **Notepad**



```
! RHF SVP Opt

* xyz 0 1
C 0   0   0
O 0   0   1.13
*
```

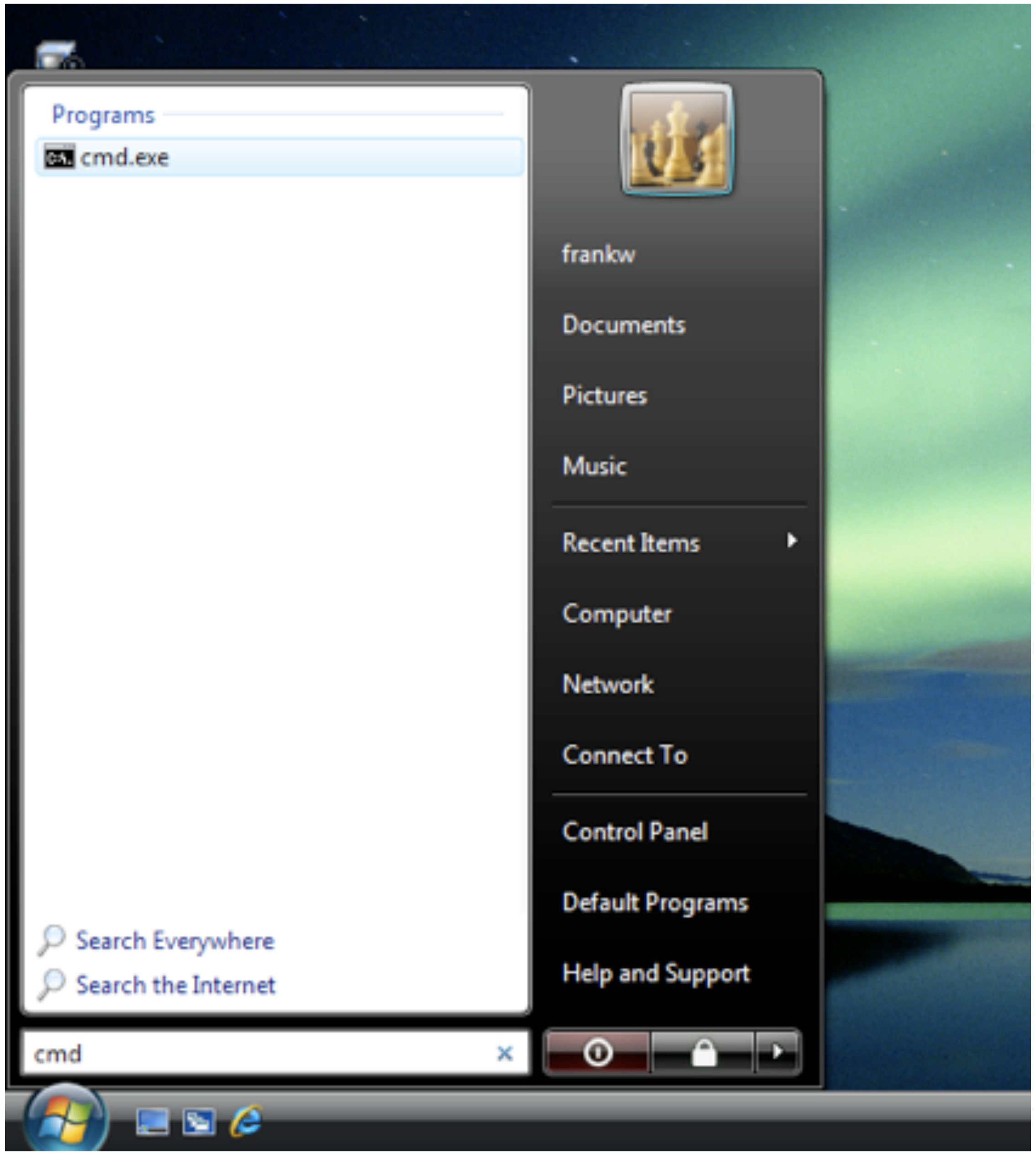
The **Notepad** application can be found under **Accessories**

Of course all other Text-editors can be used on Windows.

(But not Word!) Textfiles have to be plain ASCII !

Running ***ORCA*** under Windows Vista

Opening a command window:



Enter

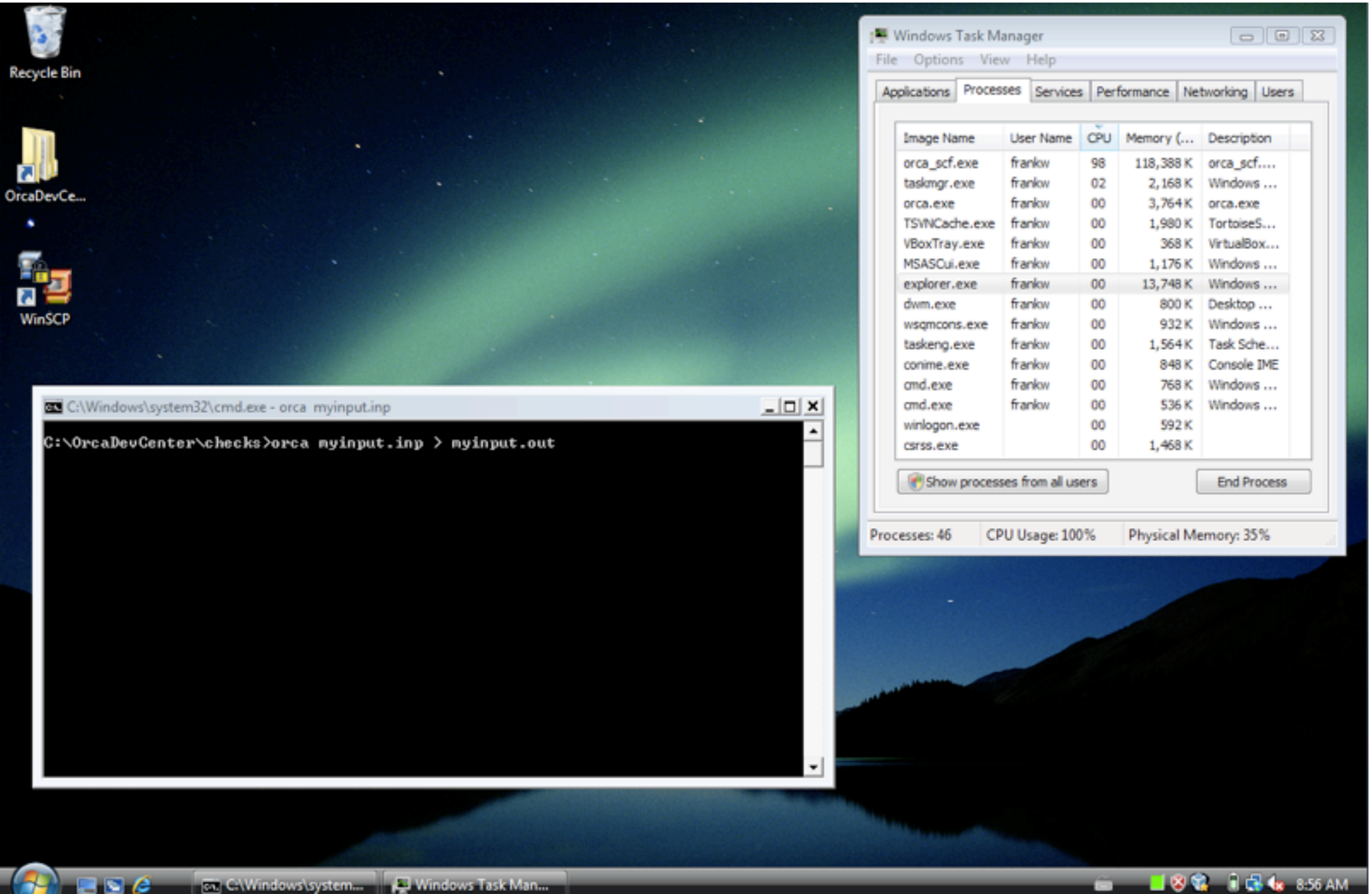
‘**cmd**’

into the **Search line,**

then cmd.exe will be executed
upon pressing **RETURN**.

Running *ORCA* under Windows Vista

Executing the *ORCA* program



```
C:\OrcaDevCenter\checks>orca myinput.inp > myinput.out
```

Image Name	User Name	CPU	Memory (...	Description
orca_scf.exe	frankw	98	118,388 K	orca_scf....
taskmgr.exe	frankw	02	2,168 K	Windows ...
orca.exe	frankw	00	3,764 K	orca.exe
TSVNCache.exe	frankw	00	1,980 K	TortoiseS...
VBoxTray.exe	frankw	00	368 K	VirtualBox...
MSASCui.exe	frankw	00	1,176 K	Windows ...
explorer.exe	frankw	00	13,748 K	Windows ...
dwm.exe	frankw	00	800 K	Desktop ...
wsqmcons.exe	frankw	00	932 K	Windows ...
taskeng.exe	frankw	00	1,564 K	Task Sche...
conime.exe	frankw	00	848 K	Console IME
cmd.exe	frankw	00	768 K	Windows ...
cmd.exe	frankw	00	536 K	Windows ...
winlogon.exe		00	592 K	
csrss.exe		00	1,468 K	

Processes: 46 CPU Usage: 100% Physical Memory: 35%

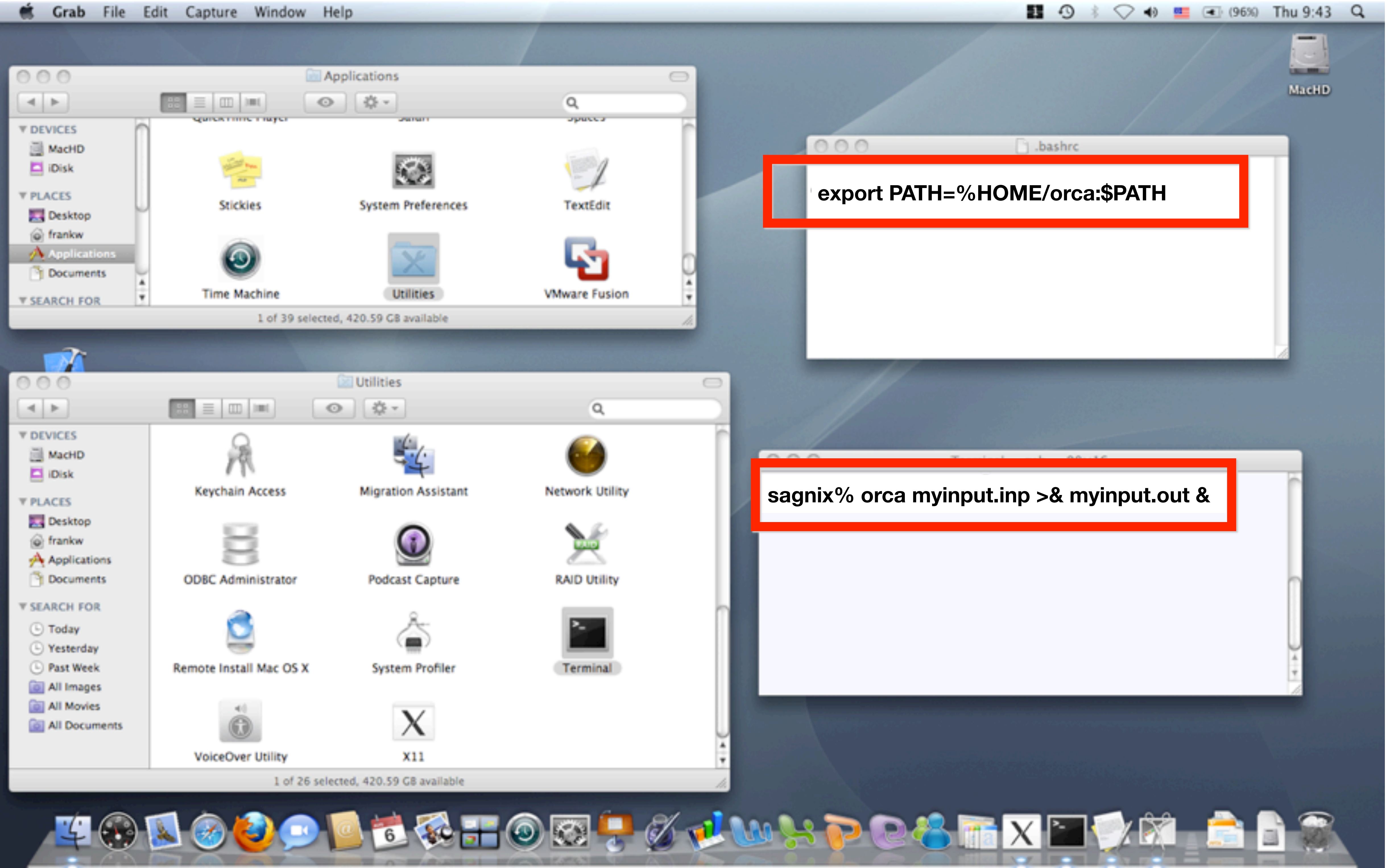
Once *ORCA* is in the **PATH**, you can run jobs from any directory you want

Using the sequence **orca myinput.inp > myinput.out’** the output will be put in a file named like that.

The **TaskManager** is a handy tool to check if the calculation is still running.

Getting Started with ***ORCA*** under Mac OS X

TextEdit & Terminal

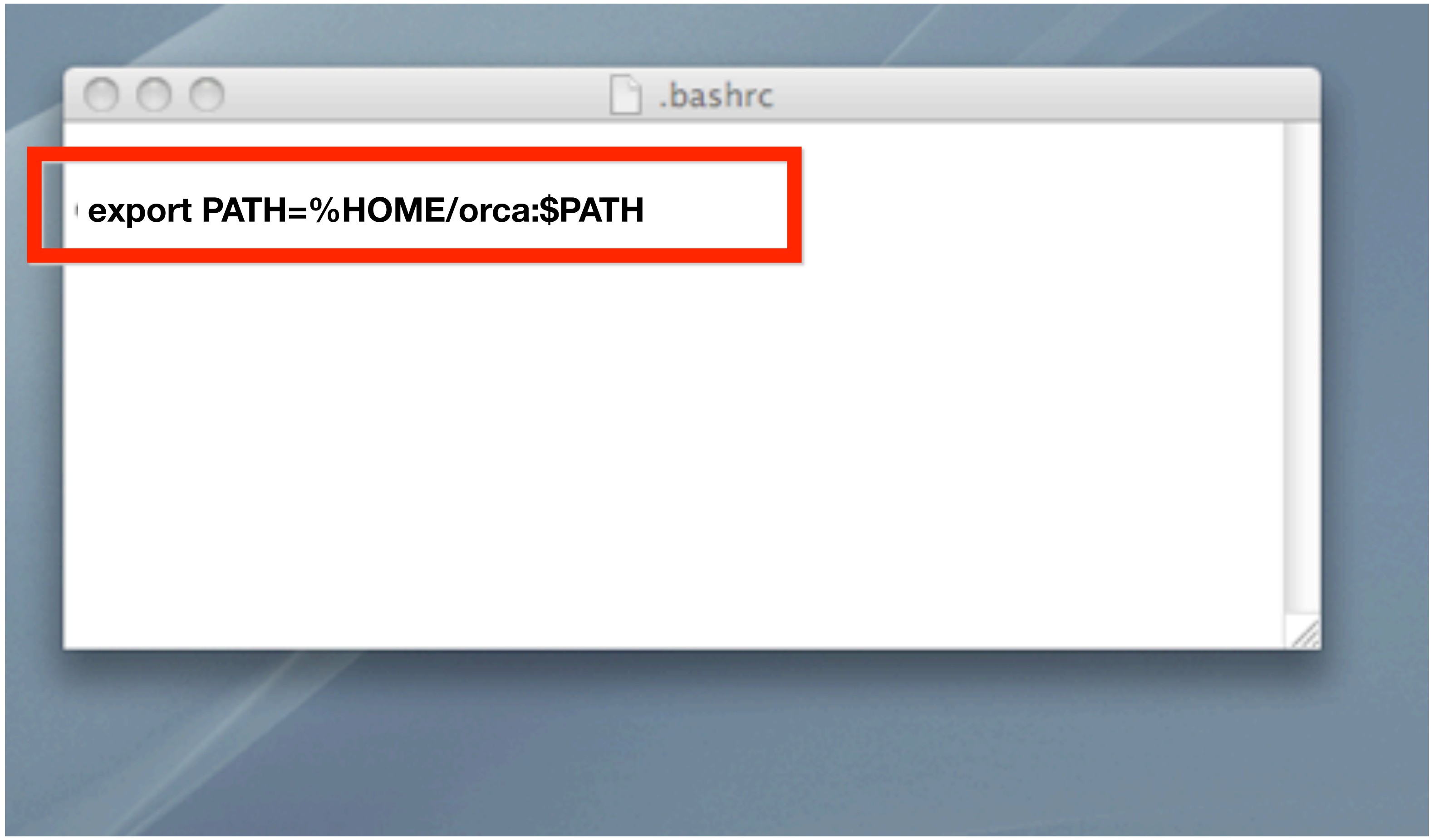


```
export PATH=%HOME/orca:$PATH
```

```
sagnix% orca myinput.inp >& myinput.out &
```

Configuring **ORCA** under Mac OS X

TextEdit & Terminal



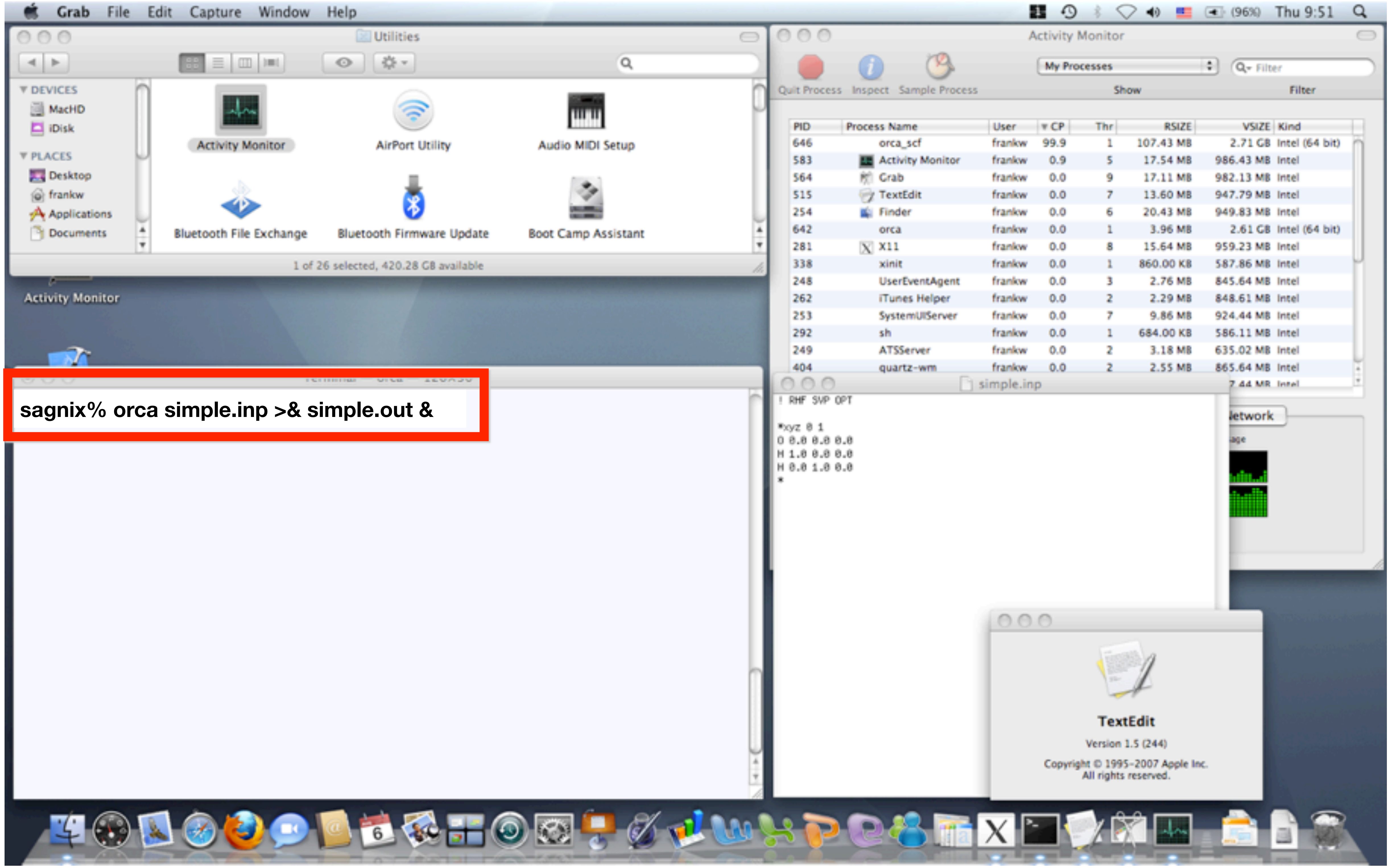
```
export PATH=%HOME/orca:$PATH
```

ORCA path entered in the file

‘**.bashrc**’

using TextEdit

Running *ORCA* under Mac OS X



```
sagnix% orca simple.inp >& simple.out &
```

PID	Process Name	User	CP	Thr	RSIZE	VSIZE	Kind
646	orca_scf	frankw	99.9	1	107.43 MB	2.71 GB	Intel (64 bit)
583	Activity Monitor	frankw	0.9	5	17.54 MB	986.43 MB	Intel
564	Grab	frankw	0.0	9	17.11 MB	982.13 MB	Intel
515	TextEdit	frankw	0.0	7	13.60 MB	947.79 MB	Intel
254	Finder	frankw	0.0	6	20.43 MB	949.83 MB	Intel
642	orca	frankw	0.0	1	3.96 MB	2.61 GB	Intel (64 bit)
281	X11	frankw	0.0	8	15.64 MB	959.23 MB	Intel
338	xinit	frankw	0.0	1	860.00 KB	587.86 MB	Intel
248	UserEventAgent	frankw	0.0	3	2.76 MB	845.64 MB	Intel
262	iTunes Helper	frankw	0.0	2	2.29 MB	848.61 MB	Intel
253	SystemUIServer	frankw	0.0	7	9.86 MB	924.44 MB	Intel
292	sh	frankw	0.0	1	684.00 KB	586.11 MB	Intel
249	ATSServer	frankw	0.0	2	3.18 MB	635.02 MB	Intel
404	quartz-wm	frankw	0.0	2	2.55 MB	865.64 MB	Intel

simple.inp

```
! RHF SVP OPT

*xyz 0 1
O 0.0 0.0 0.0
H 1.0 0.0 0.0
H 0.0 1.0 0.0
*
```

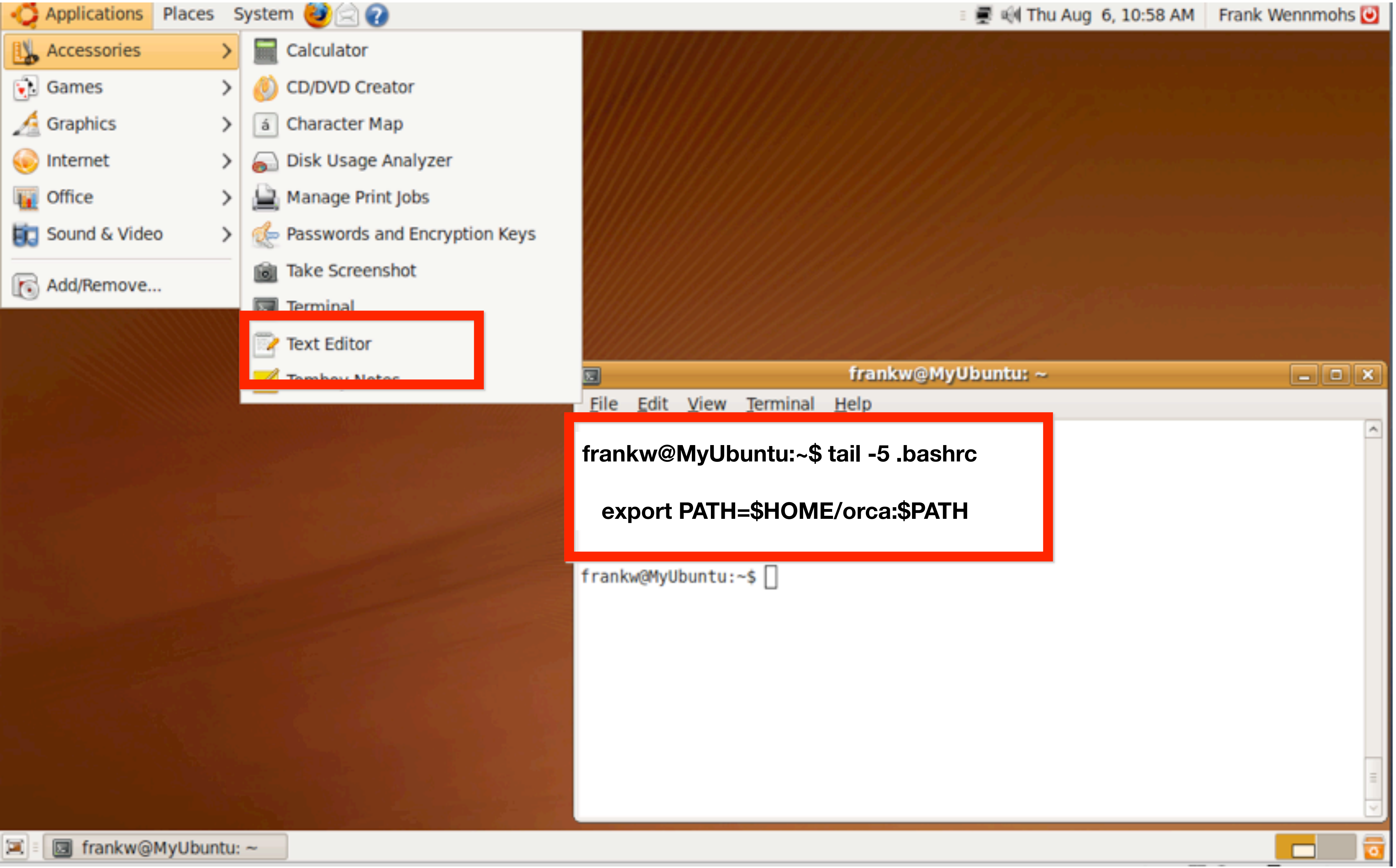
TextEdit

Version 1.5 (244)

Copyright © 1995-2007 Apple Inc.
All rights reserved.

Configuring *ORCA* under Linux

Linux (Ubuntu 9.10): Setting the PATH variable



```
frankw@MyUbuntu:~$ tail -5 .bashrc

  export PATH=$HOME/orca:$PATH

frankw@MyUbuntu:~$ 
```

Setting the **PATH** variable is done in the file

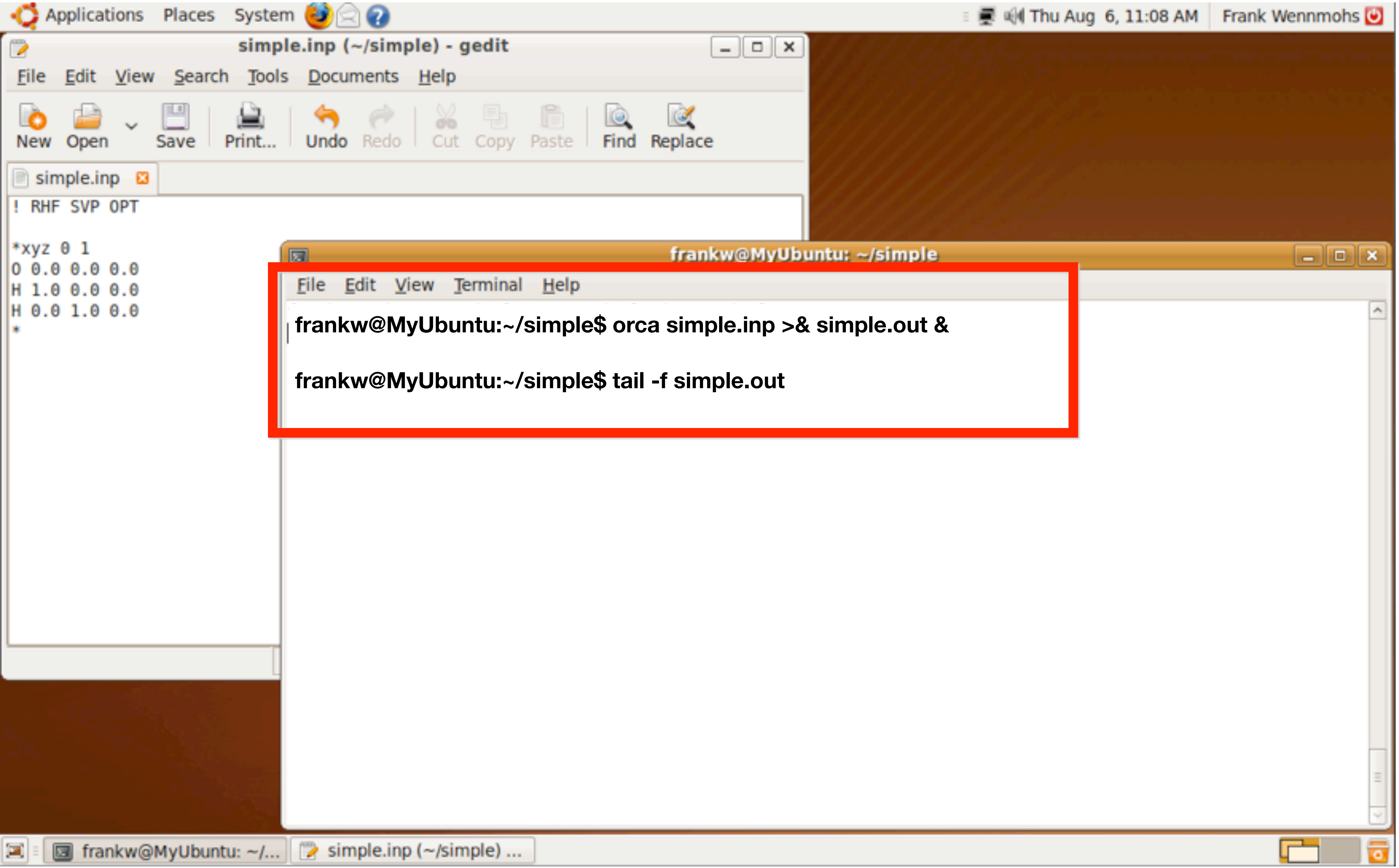
'**.bashrc**'

You can use the '**Text Editor**'

to add the line shown.

Running *ORCA* under Linux

Linux (Ubuntu 9.10): Text Editor (gedit) & ORCA execution



```
! RHF SVP OPT

*xyz 0 1
O 0.0 0.0 0.0
H 1.0 0.0 0.0
H 0.0 1.0 0.0
*
```

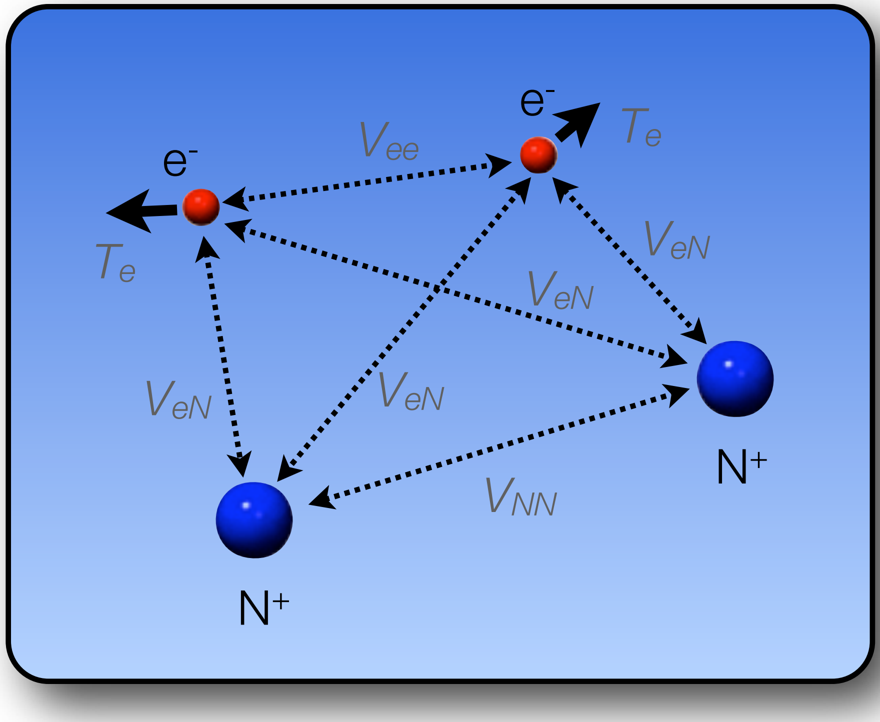
```
frankw@MyUbuntu:~/simple$ orca simple.inp >& simple.out &

frankw@MyUbuntu:~/simple$ tail -f simple.out
```

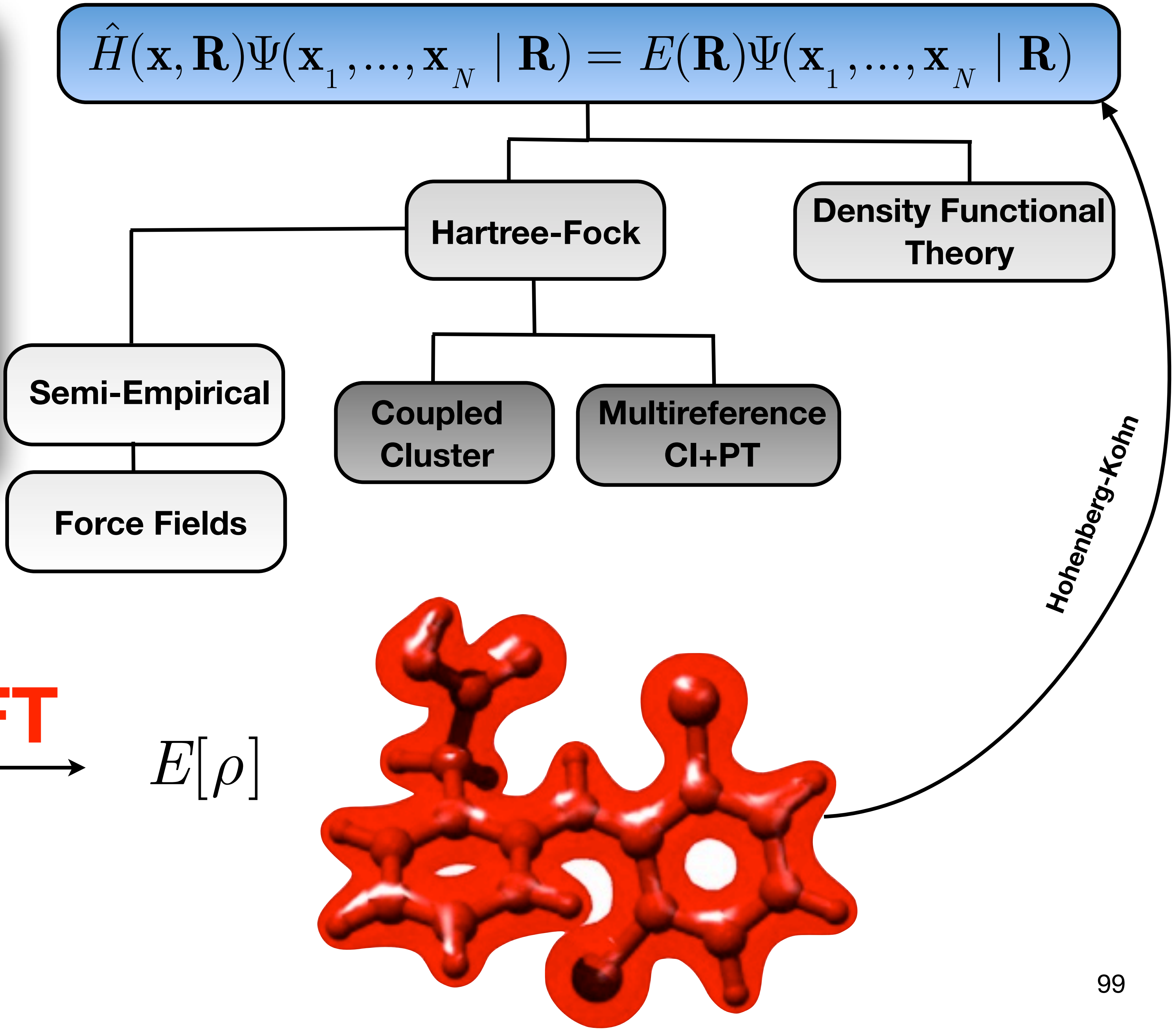
Especially under Linux there is an uncountable number of text-editors.

Very popular ones are **Vim**, **Emacs**, **NEdit**, **Kate**, etc.

Appendix: More information on Computational Methods

Method Overview





$$\hat{H}(\mathbf{x},\mathbf{R})\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N \mid \mathbf{R}) = E(\mathbf{R})\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N \mid \mathbf{R})$$

Hartree-Fock

Density Functional Theory

Semi-Empirical

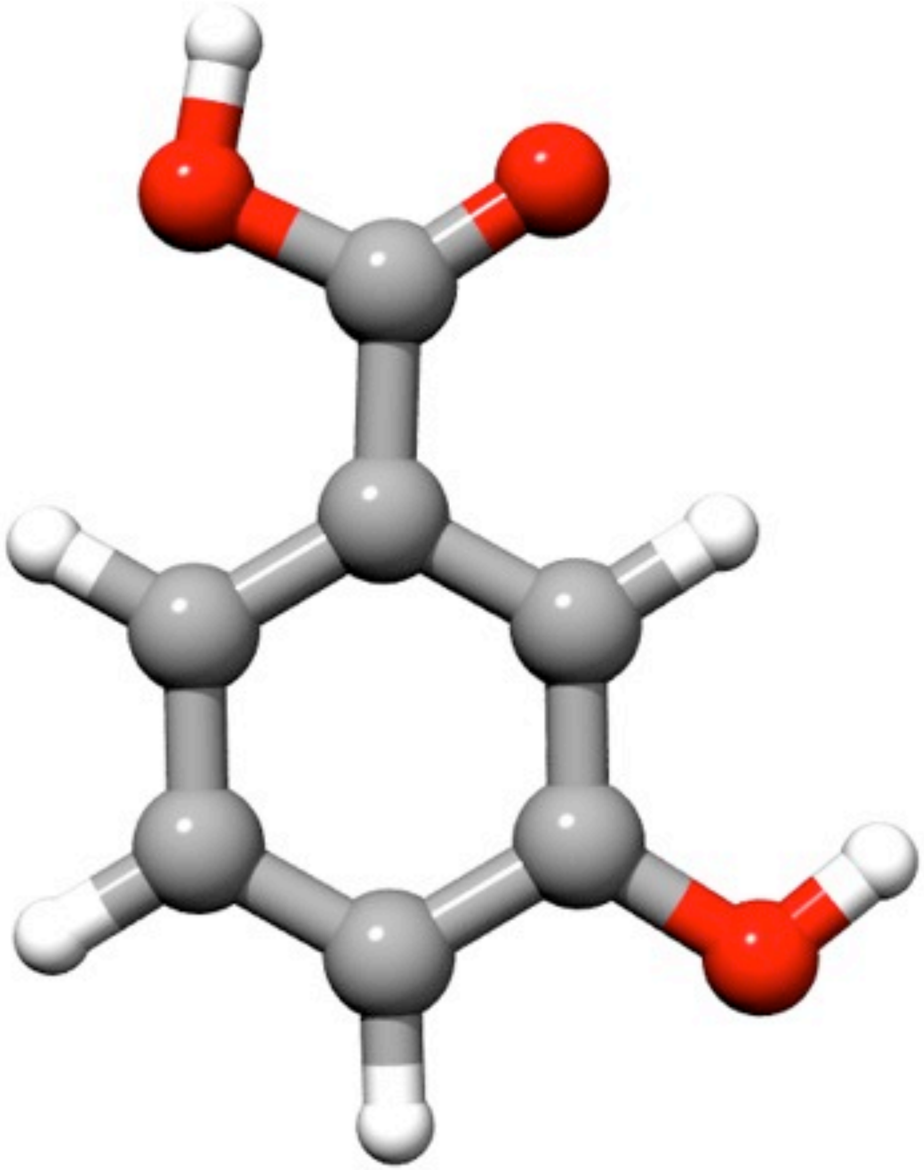
Force Fields

Coupled Cluster

Multireference CI+PT

$$E[\Psi] = \frac{\langle \Psi \mid H \mid \Psi \rangle}{\langle \Psi \mid \Psi \rangle} \xrightarrow{\textbf{DFT}} E[\rho]$$

Computational Cost



Meta Hydroxy Benzoic Acid

16 Atoms, 72 Electrons
346 Basis functions (def2-TZVP; ~cc-pVTZ)

		Energy		Gradient
PBE	:	~50		~80
B3LYP	:	~1300		~720
B2PYLP(~MP2)	:	~1325	(RI!)	~2300
CCSD(T)	:	~90000		B.P.
LPNO-CCSD	:	~3100		B.P.
CASSCF(8,8)	:	~1800		~440
NEVPT2	:	~60		B.P.

B.P.=Big project

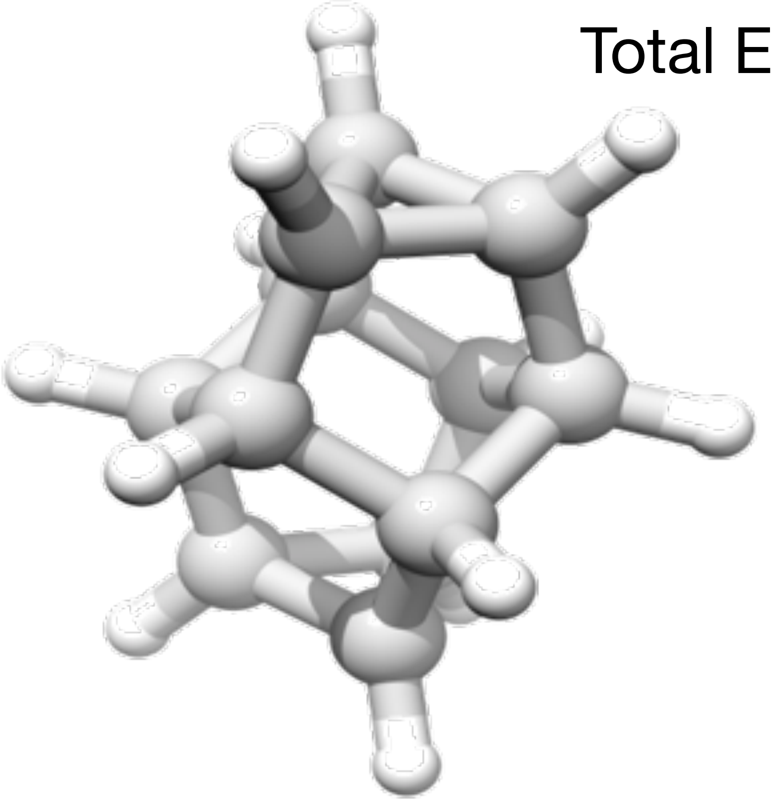
How accurately do we have to calculate?

An error of only **1.3 kcal/mol** is equivalent to:

- pK_a-value : **One log-unit**
- Redox Potential : **56 mV**
- Reaction Rate : **Factor 10**

Accuracy of ~1 kcal/mol required

Total Energy ~ 290668 kcal/mol



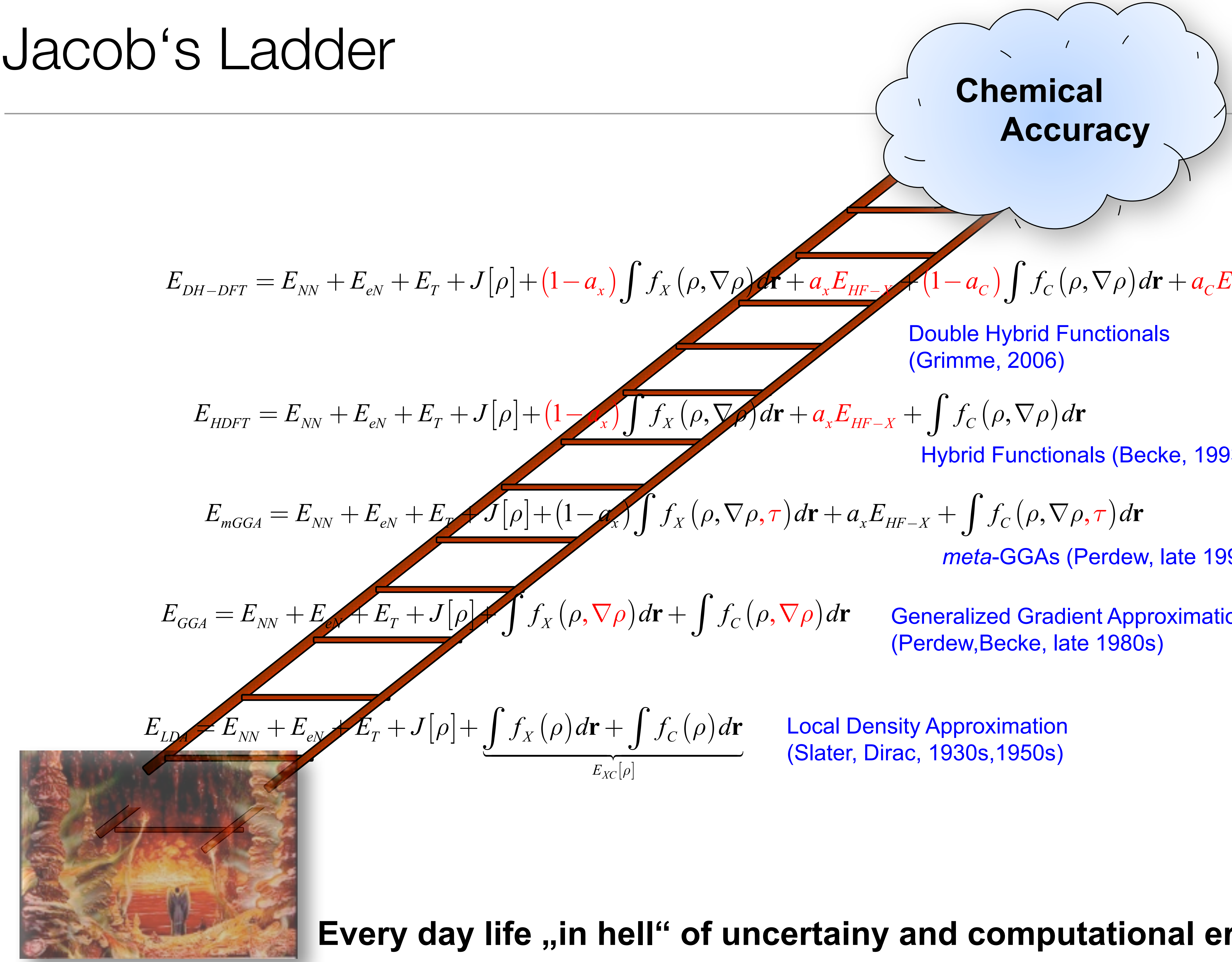
Possible Consequences:

a) Need extremely accurate theory

b) Need very good error compensation

c) Care other properties than E_{tot}

Jacob‘s Ladder



Chemical Accuracy

$$E_{DH-DFT} = E_{NN} + E_{eN} + E_T + J[\rho] + (1-a_x)\int f_X(\rho,\nabla\rho)\,d\mathbf{r} + a_x E_{HF-X} + (1-a_C)\int f_C(\rho,\nabla\rho)\,d\mathbf{r} + a_C E_{MP2}$$

Double Hybrid Functionals
(Grimme, 2006)

$$E_{HDFT} = E_{NN} + E_{eN} + E_T + J[\rho] + (1-a_x)\int f_X(\rho,\nabla\rho)\,d\mathbf{r} + a_x E_{HF-X} + \int f_C(\rho,\nabla\rho)\,d\mathbf{r}$$

Hybrid Functionals (Becke, 1993)

$$E_{mGGA} = E_{NN} + E_{eN} + E_T + J[\rho] + (1-a_x)\int f_X(\rho,\nabla\rho,\tau)\,d\mathbf{r} + a_x E_{HF-X} + \int f_C(\rho,\nabla\rho,\tau)\,d\mathbf{r}$$

meta-GGAs (Perdew, late 1990s)

$$E_{GGA} = E_{NN} + E_{eN} + E_T + J[\rho] + \int f_X(\rho,\nabla\rho)\,d\mathbf{r} + \int f_C(\rho,\nabla\rho)\,d\mathbf{r}$$

Generalized Gradient Approximation
(Perdew,Becke, late 1980s)

$$E_{LDA} = E_{NN} + E_{eN} + E_T + J[\rho] + \underbrace{\int f_X(\rho)\,d\mathbf{r} + \int f_C(\rho)\,d\mathbf{r}}_{E_{XC}[\rho]}$$

Local Density Approximation
(Slater, Dirac, 1930s,1950s)

Every day life „in hell“ of uncertainy and computational errors

The Functional Zoo

Name	**Type**	**Comments**
BLYP	GGA	One of the earliest GGA functionals. Usually inferior to BP86 and PBE. Predicts too long bonds.
BP86	GGA	Excellent geometries and vibrational frequencies. Energetics is usually not highly accurate but performs often well in spectroscopic investigations.
PW91	GGA	One of the older GGA functionals with excellent accuracy for exchange couplings.
PBE	GGA	A GGA version designed to replace PW91. Very popular in physics. Often similar to BP86.
OLYP	GGA	Violates the uniform electron gas limit but gives improved results for molecules
B3LYP	Hybrid	De facto standard in chemistry for structures, energies and properties. See discussion in the text.
PBE0	Hybrid	Excellent accuracy; competitive with B3LYP
TPSS	Meta-GGA	Improvement over PBE. Includes the kinetic energy density and obeys more constraints known from rigorous theory.
TPSSh	Hybrid meta-GGA	Probably improvement over PBE0; perhaps increase fraction of HF to 25% (TPSS0)
B2PLYP	Double hybrid	First (and prototypical) member of the double hybrid class of functionals. So far been proven excellent for energies and geometries. More exploration needed.

Total Energies

Total, correlation and exchange energies of the Neon atom using the *ab initio* CCSD(T) method and various standard functionals (deviations from the wavefunction results in mEh).

	E_{tot}	E_{corr}	E_x
CCSD(T)	-128.9260	-0.379	-12.098
	-129.0640 (DKH2)		
BP86	-128.9776 (-52)	-0.388 (- 9)	-12.104 (-6)
PBE	-128.8664 (+60)	-0.347 (+32)	-12.028 (+70)
BLYP	-128.9730 (-47)	-0.383 (- 4)	-12.099 (-1)
TPSS	-128.9811 (-55)	-0.351 (+28)	-12.152 (-54)
B3LYP	-128.9426 (-17)	-0.452 (-73)	-12.134 (-36)
B2PLYP	-128.9555 (-30)	-0.392 (-13)	-12.103 (- 5)
Exp	**-129.056**		

➡ Wavefunction theory is very accurate (but also very expensive). DFT results vary widely among different functionals and either over- or undershoot.

total energies are not important in chemistry – relative energies matter.

DFT Energies - Benchmarking (1)

Today the accuracy of a given density functional is no longer assessed by doing a few illustrative calculations or studying the „G2 set“.

(Even worse: atomization energies. There is no correlation between the performance of a method for atomization energies and its performance in chemistry)

	BP86	**PBE**	**TPSS**	**TPSSh**	**B3LYP**	**PBE0**
(kcal/mol)	*Small molecule test set*					
Mean error	0,4	0,2	0,7	0,4	0,3	-0,3
Mean Abs. error	2,8	2,9	3,1	2,7	**2,1**	**2,2**
Max error	24,2	25,9	21,7	19,8	14,5	14,7
	Large molecule test set					
Mean error	-4,6	-2,7	-4,5	-3,6	-6,9	-0,9
Mean Abs. error	8,8	7,7	8,5	7,0	**8,5**	**4,6**
Max error	87,4	79,0	70,3	52,9	77,9	36,9

optimistic

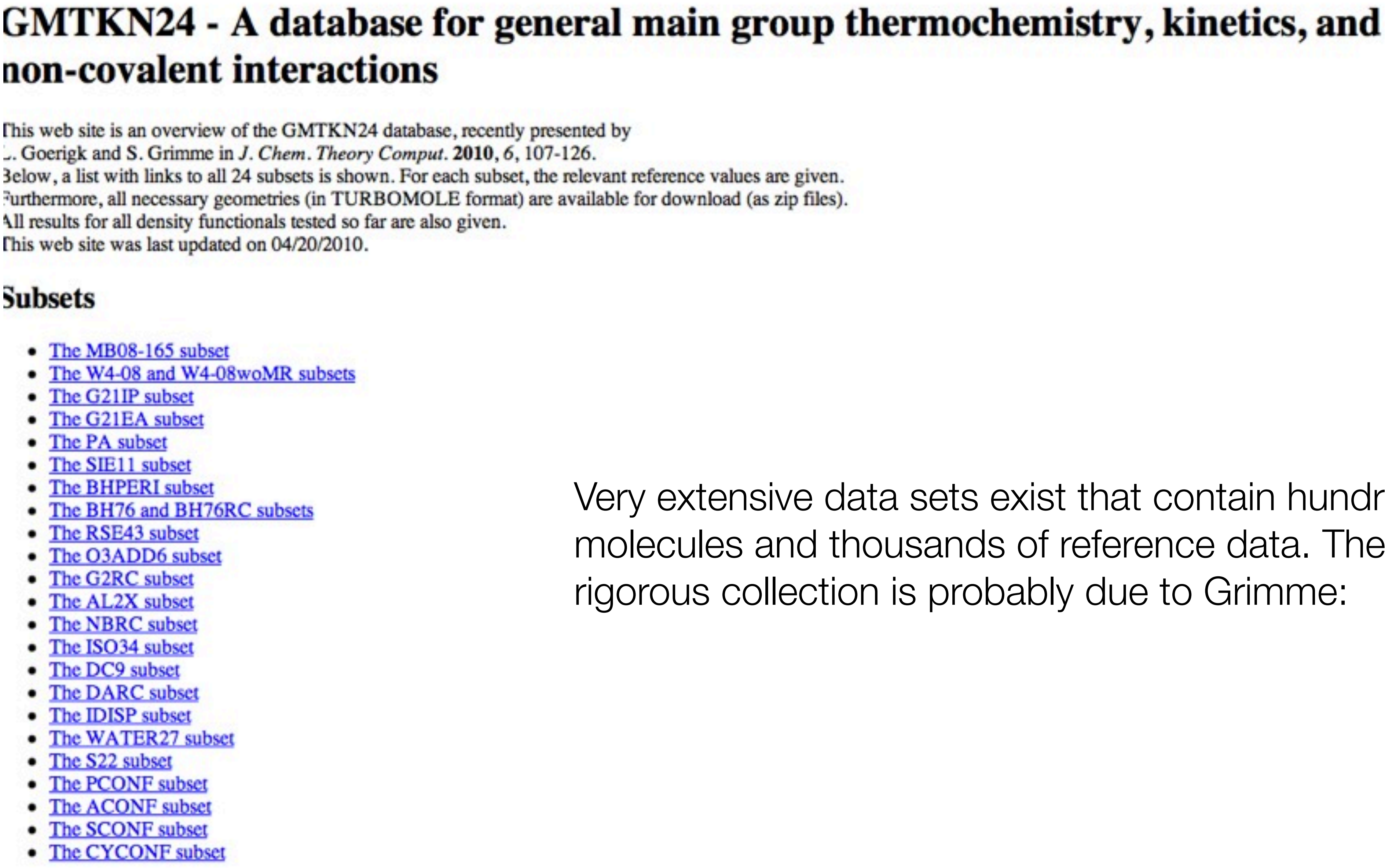
realistic

Grimme, S. *J. Phys. Chem. A* **2005**, 109, 3067-3077

DFT Energies - Benchmarking (2)

Very extensive data sets exist that contain hundreds of molecules and thousands of reference data. The most rigorous collection is probably due to Grimme:

http://toc.uni-muenster.de/GMTKN/GMTKN24/GMTKN24main.html



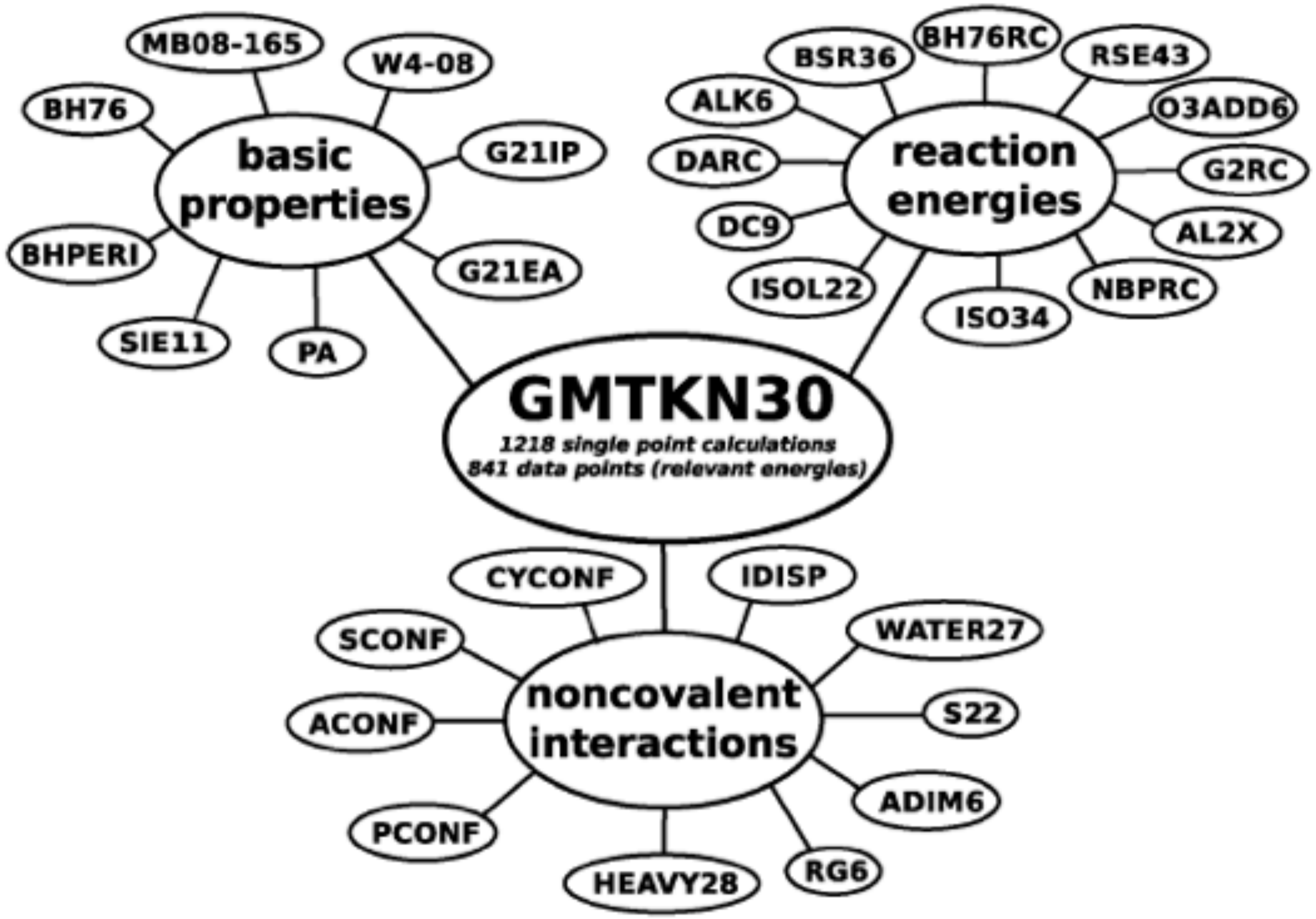
GMTKN24 - A database for general main group thermochemistry, kinetics, and non-covalent interactions

This web site is an overview of the GMTKN24 database, recently presented by
L. Goerigk and S. Grimme in *J. Chem. Theory Comput.* **2010**, 6, 107-126.
Below, a list with links to all 24 subsets is shown. For each subset, the relevant reference values are given.
Furthermore, all necessary geometries (in TURBOMOLE format) are available for download (as zip files).
All results for all density functionals tested so far are also given.
This web site was last updated on 04/20/2010.

Subsets

- The MB08-165 subset
- The W4-08 and W4-08woMR subsets
- The G21IP subset
- The G21EA subset
- The PA subset
- The SIE11 subset
- The BHPERI subset
- The BH76 and BH76RC subsets
- The RSE43 subset
- The O3ADD6 subset
- The G2RC subset
- The AL2X subset
- The NBRC subset
- The ISO34 subset
- The DC9 subset
- The DARC subset
- The IDISP subset
- The WATER27 subset
- The S22 subset
- The PCONF subset
- The ACONF subset
- The SCONF subset
- The CYCONF subset

DFT Energies - Benchmarking (3)

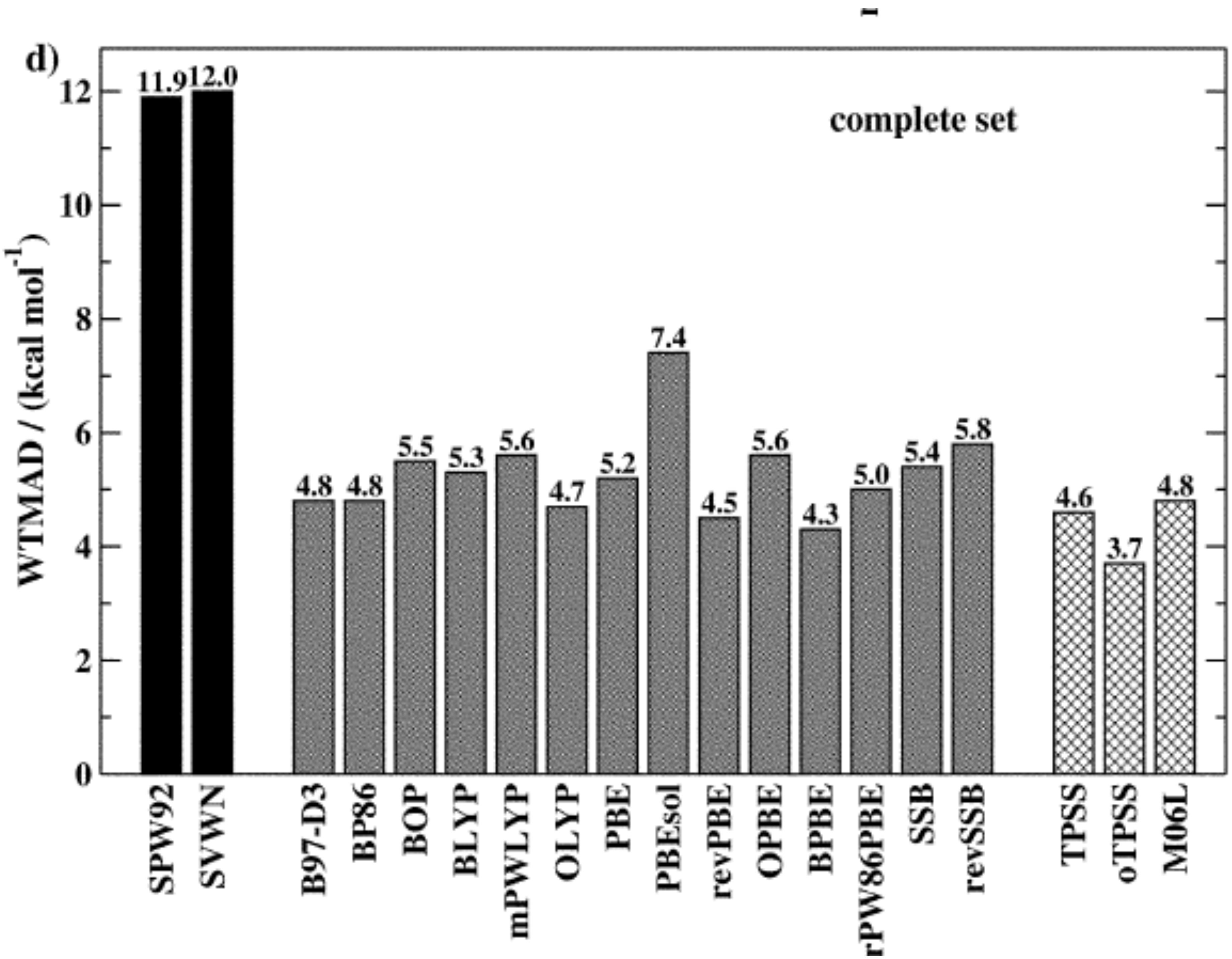


NOTE:
- Requires 3218 single point calculations
- Compare 841 data points
- Covers:
  * *Reaction energies*
  * *Isomerization energies*
  * *Weak interactions*

BUT:
- Still only closed-shell organic/main group chemistry!
- This is *NOT* transferable to transition metal chemistry or open shells!

Göricke, L.; Grimme, S. PCCP, **2011**, *13*,6670

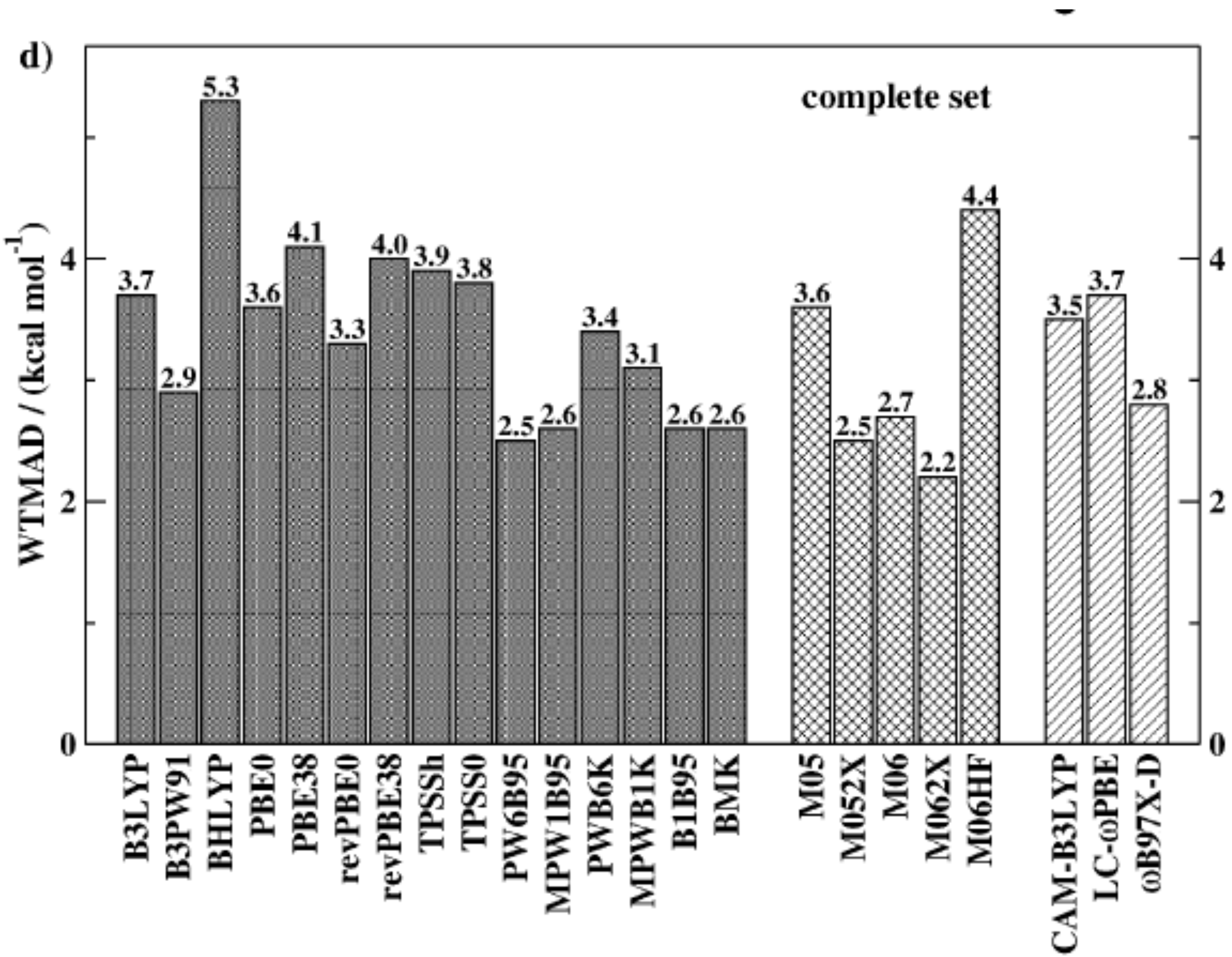
DFT Energies - Benchmarking (4)



- LDA is useless
- Rather similar behavior for GGAs
- No or only minor improvements for meta-GGAs

Göricke, L.; Grimme, S. PCCP, **2011**, *13*,6670

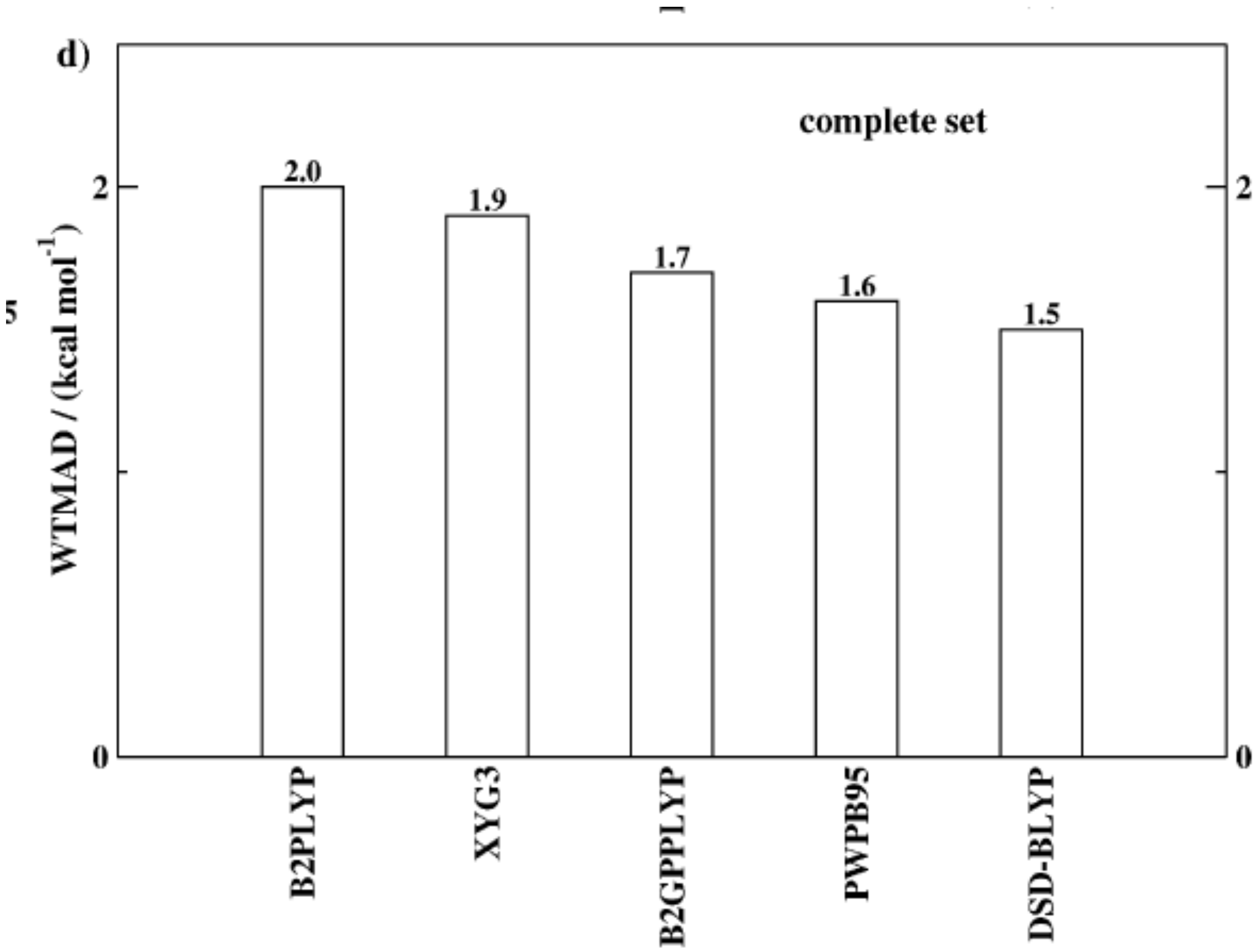
DFT Energies - Benchmarking (5)



- Rather similar behavior for differenty hybrids with B3LYP not the best.
- Minnesota functionals accurate (M06-2X) but not stable
- No improvement for range-corrected functionals

Göricke, L.; Grimme, S. PCCP, **2011**, *13*,6670

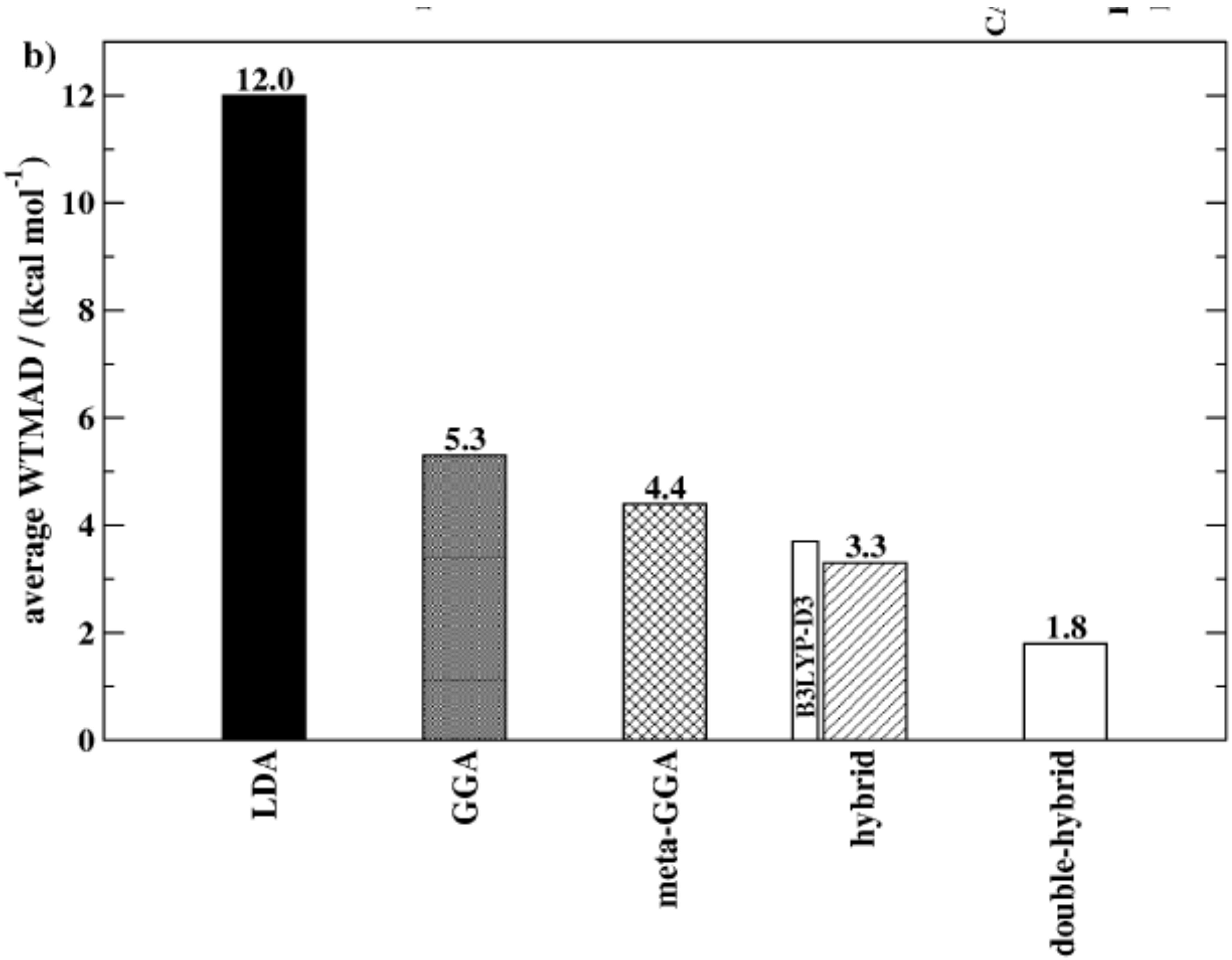
DFT Energies - Benchmarking (6)



- DHDFs are all very accurate.

Göricke, L.; Grimme, S. PCCP, **2011**, *13*,6670

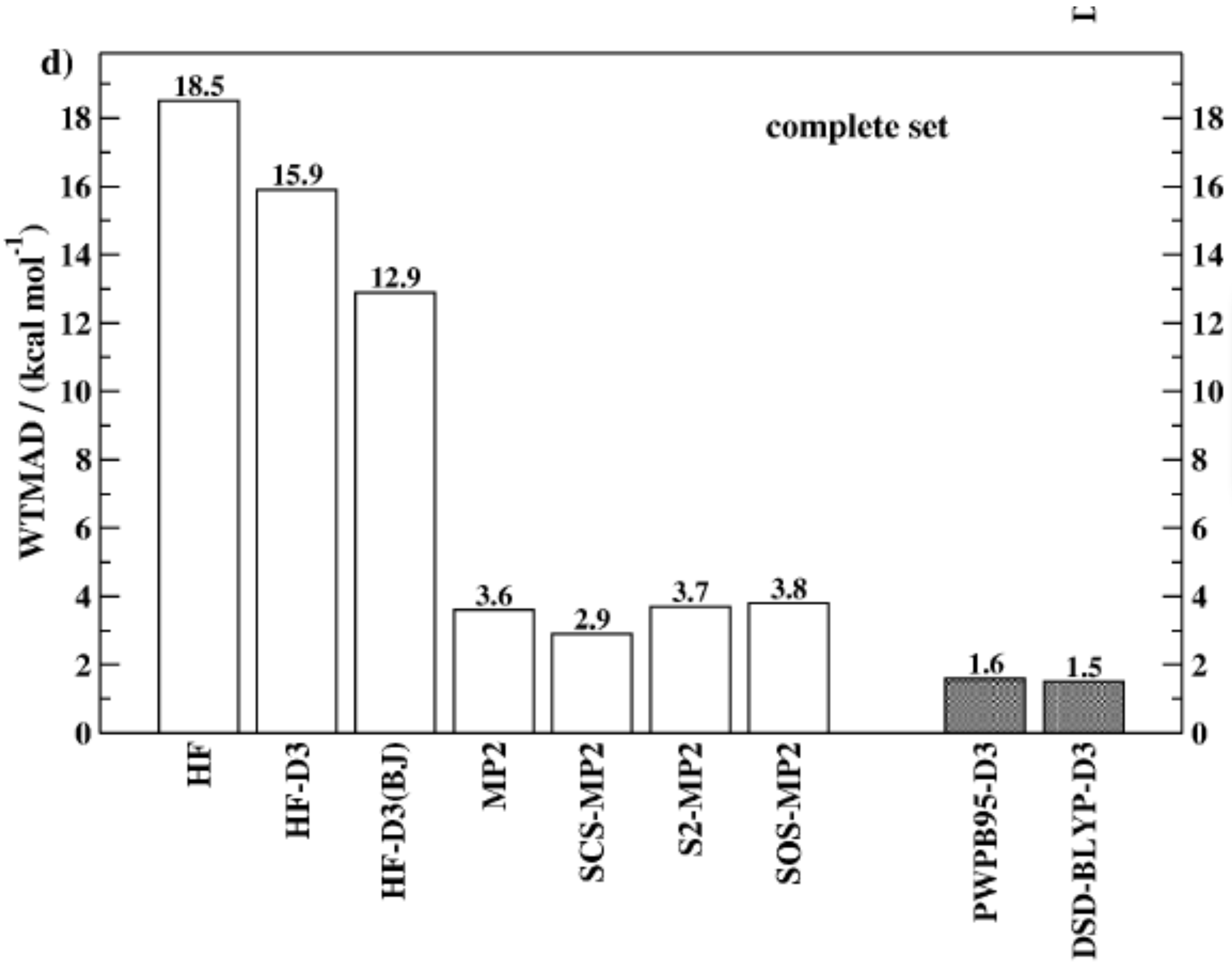
DFT Energies - Benchmarking (6)



- Extensive modern benchmarking confirms the existence of Jacob‘s ladder (at least for the thermochemistry of main group compounds)

Göricke, L.; Grimme, S. PCCP, **2011**, *13*,6670

DFT Energies - Benchmarking (7)



- DHDFs generally perform better than MP2 variants of any kind.

Göricke, L.; Grimme, S. PCCP, **2011**, *13*,6670

DFT Energies - Summary

Overall recommendations by Göricke and Grimme (2011):

LDA: Not recommended for chemistry

GGA: **B97-D3** is the most accurate functional
BLYP-D3 is the second best

Hybrid: **PW6B95** is the most accurate functional
Minnesota functional (**M062X**) are as good but are not robust
Range separated functionals do not represent an improvement
„**B3LYP-D3** *is not the overall applicable functional as many user may still believe Surprisingly it is even worse than the average hybrid. Particularly for reaction energies it was the worst of all 23 tested hybrids.*“

DHDF: **PWPB95** is the most accurate functional
All DHDFs outperform all other functionals. Their basis set dependence is higher
DHDFs are more accurate than MP2 or SCS-MP2 except in cases with large SIE

Göricke, L.; Grimme, S. PCCP, **2011**, *13*,6670

Recommendations: Density Functional Theory

- Use a functional that is **well tested** for your domain of application (... often this will come out to default to B3LYP ...)
- Always use **empirical dispersion corrections** (D3 of Grimme). They come for free, almost never make your results worse and are frequently essential for correct chemistry.
- Whenever you can: **use GGAs** - they can be evaluated extremely efficiently. Invest the time saved into a more realistic model of your system or its environment, better basis sets, ...
- Use the **largest basis sets** that you can afford. Studying basis set artifacts is boring. Be careful with special properties or anions. The basis set of the Karlsruhe group are particularly consistent, accurate and efficient.
- In particular GGA, **geometries** are very good. There is little (if any) point in spending all the extra time for getting MP2 or CCSD geometries (CCSD(T) would be slightly better but comes at humungous cost)
- When dealing with **heavy elements** (e.g. beyond Ca):Scalar relativistic corrections (ZORA, DKH, ...) are more rigorous than ECPs and spin-free calculations are not much more expensive.
- When studying systems in the **condensed phase** (particularly anions), some model of the environment must be included. At least COSMO/PCM.

Example: 3d+4d+5d Transition metal bondlengths

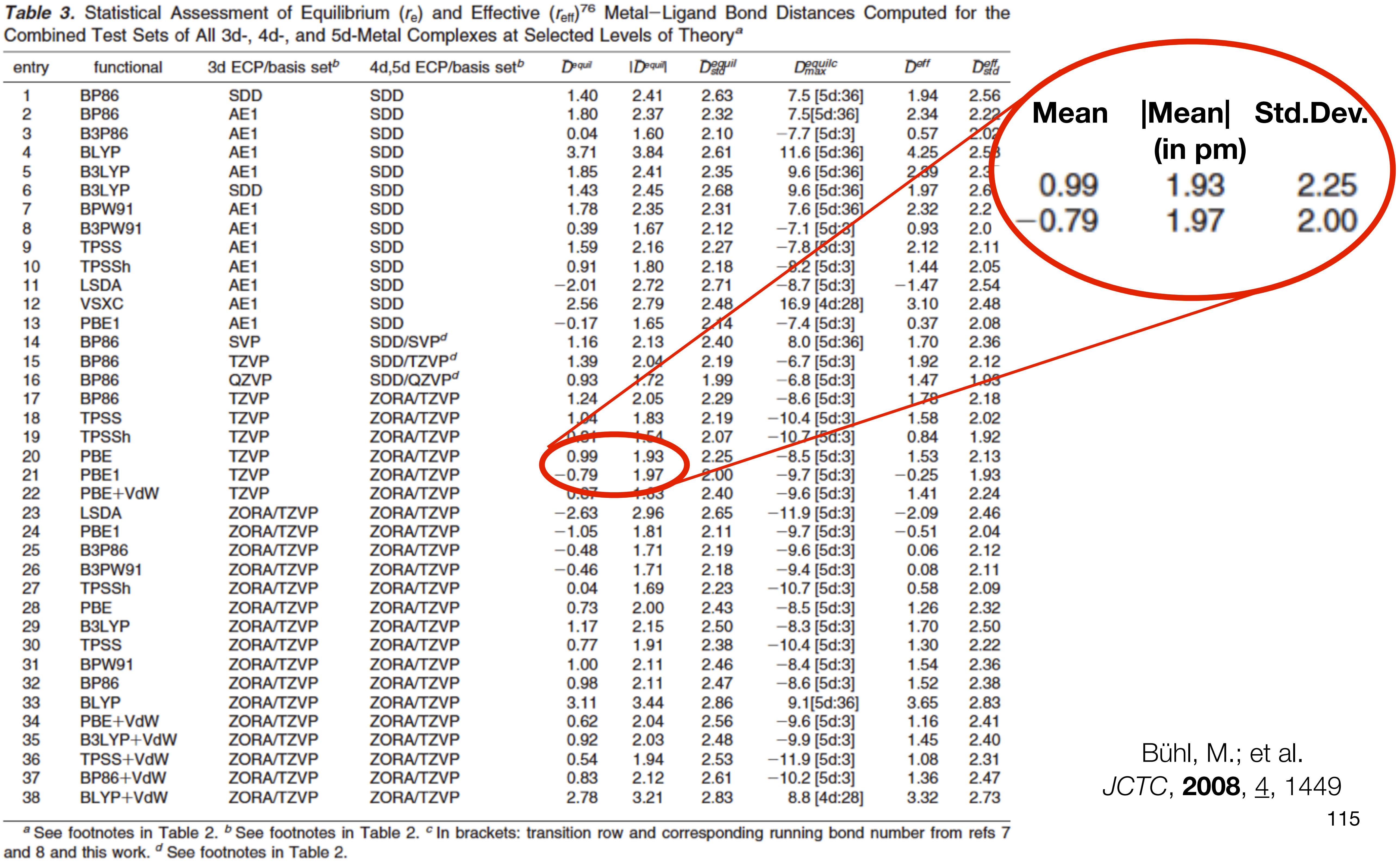


Table 3. Statistical Assessment of Equilibrium (r_e) and Effective (r_{eff})[76] Metal−Ligand Bond Distances Computed for the Combined Test Sets of All 3d-, 4d-, and 5d-Metal Complexes at Selected Levels of Theory[a]

entry	functional	3d ECP/basis set[b]	4d,5d ECP/basis set[b]	$\bar{D}^{equil}$	$\|\bar{D}^{equil}\|$	D^{equil}_{std}	$D^{equil\,c}_{max}$	$\bar{D}^{eff}$	D^{eff}_{std}
1	BP86	SDD	SDD	1.40	2.41	2.63	7.5 [5d:36]	1.94	2.56
2	BP86	AE1	SDD	1.80	2.37	2.32	7.5[5d:36]	2.34	2.22
3	B3P86	AE1	SDD	0.04	1.60	2.10	−7.7 [5d:3]	0.57	2.02
4	BLYP	AE1	SDD	3.71	3.84	2.61	11.6 [5d:36]	4.25	2.58
5	B3LYP	AE1	SDD	1.85	2.41	2.35	9.6 [5d:36]	2.39	[illegible]
6	B3LYP	SDD	SDD	1.43	2.45	2.68	9.6 [5d:36]	1.97	[illegible]
7	BPW91	AE1	SDD	1.78	2.35	2.31	7.6 [5d:36]	2.32	[illegible]
8	B3PW91	AE1	SDD	0.39	1.67	2.12	−7.1 [5d:3]	0.93	[illegible]
9	TPSS	AE1	SDD	1.59	2.16	2.27	−7.8 [5d:3]	2.12	2.11
10	TPSSh	AE1	SDD	0.91	1.80	2.18	−8.2 [5d:3]	1.44	2.05
11	LSDA	AE1	SDD	−2.01	2.72	2.71	−8.7 [5d:3]	−1.47	2.54
12	VSXC	AE1	SDD	2.56	2.79	2.48	16.9 [4d:28]	3.10	2.48
13	PBE1	AE1	SDD	−0.17	1.65	2.14	−7.4 [5d:3]	0.37	2.08
14	BP86	SVP	SDD/SVP[d]	1.16	2.13	2.40	8.0 [5d:36]	1.70	2.36
15	BP86	TZVP	SDD/TZVP[d]	1.39	2.04	2.19	−6.7 [5d:3]	1.92	2.12
16	BP86	QZVP	SDD/QZVP[d]	0.93	1.72	1.99	−6.8 [5d:3]	1.47	1.93
17	BP86	TZVP	ZORA/TZVP	1.24	2.05	2.29	−8.6 [5d:3]	1.78	2.18
18	TPSS	TZVP	ZORA/TZVP	1.04	1.83	2.19	−10.4 [5d:3]	1.58	2.02
19	TPSSh	TZVP	ZORA/TZVP	0.81	1.54	2.07	−10.7 [5d:3]	0.84	1.92
20	PBE	TZVP	ZORA/TZVP	0.99	1.93	2.25	−8.5 [5d:3]	1.53	2.13
21	PBE1	TZVP	ZORA/TZVP	−0.79	1.97	2.00	−9.7 [5d:3]	−0.25	1.93
22	PBE+VdW	TZVP	ZORA/TZVP	0.87	1.83	2.40	−9.6 [5d:3]	1.41	2.24
23	LSDA	ZORA/TZVP	ZORA/TZVP	−2.63	2.96	2.65	−11.9 [5d:3]	−2.09	2.46
24	PBE1	ZORA/TZVP	ZORA/TZVP	−1.05	1.81	2.11	−9.7 [5d:3]	−0.51	2.04
25	B3P86	ZORA/TZVP	ZORA/TZVP	−0.48	1.71	2.19	−9.6 [5d:3]	0.06	2.12
26	B3PW91	ZORA/TZVP	ZORA/TZVP	−0.46	1.71	2.18	−9.4 [5d:3]	0.08	2.11
27	TPSSh	ZORA/TZVP	ZORA/TZVP	0.04	1.69	2.23	−10.7 [5d:3]	0.58	2.09
28	PBE	ZORA/TZVP	ZORA/TZVP	0.73	2.00	2.43	−8.5 [5d:3]	1.26	2.32
29	B3LYP	ZORA/TZVP	ZORA/TZVP	1.17	2.15	2.50	−8.3 [5d:3]	1.70	2.50
30	TPSS	ZORA/TZVP	ZORA/TZVP	0.77	1.91	2.38	−10.4 [5d:3]	1.30	2.22
31	BPW91	ZORA/TZVP	ZORA/TZVP	1.00	2.11	2.46	−8.4 [5d:3]	1.54	2.36
32	BP86	ZORA/TZVP	ZORA/TZVP	0.98	2.11	2.47	−8.6 [5d:3]	1.52	2.38
33	BLYP	ZORA/TZVP	ZORA/TZVP	3.11	3.44	2.86	9.1[5d:36]	3.65	2.83
34	PBE+VdW	ZORA/TZVP	ZORA/TZVP	0.62	2.04	2.56	−9.6 [5d:3]	1.16	2.41
35	B3LYP+VdW	ZORA/TZVP	ZORA/TZVP	0.92	2.03	2.48	−9.9 [5d:3]	1.45	2.40
36	TPSS+VdW	ZORA/TZVP	ZORA/TZVP	0.54	1.94	2.53	−11.9 [5d:3]	1.08	2.31
37	BP86+VdW	ZORA/TZVP	ZORA/TZVP	0.83	2.12	2.61	−10.2 [5d:3]	1.36	2.47
38	BLYP+VdW	ZORA/TZVP	ZORA/TZVP	2.78	3.21	2.83	8.8 [4d:28]	3.32	2.73

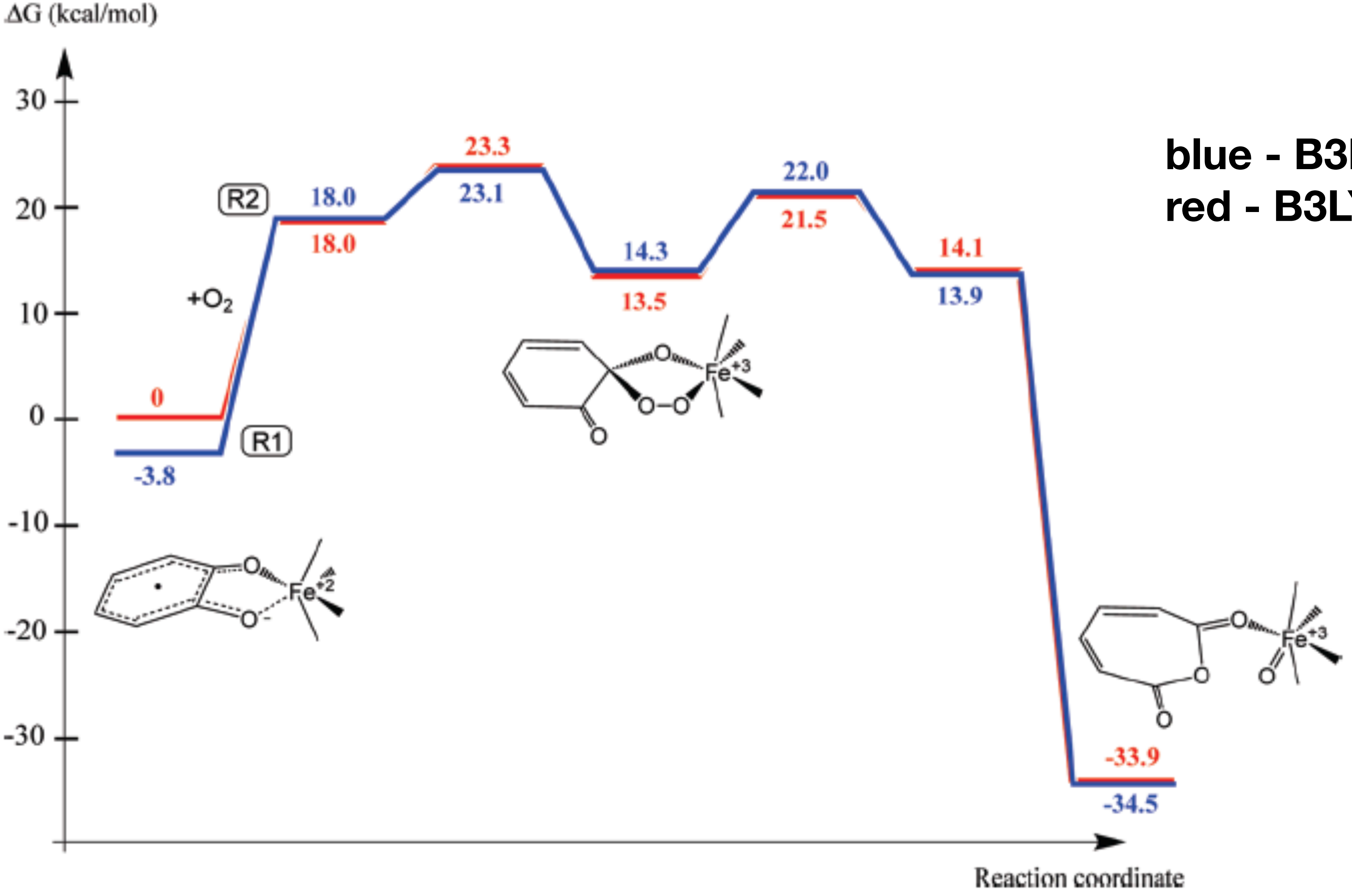
[a] See footnotes in Table 2. [b] See footnotes in Table 2. [c] In brackets: transition row and corresponding running bond number from refs 7 and 8 and this work. [d] See footnotes in Table 2.

Mean	**|Mean|** **(in pm)**	**Std.Dev.**
0.99	1.93	2.25
−0.79	1.97	2.00

Bühl, M.; et al.
JCTC, **2008**, 4, 1449

Example: Van der Waals

Normal situation

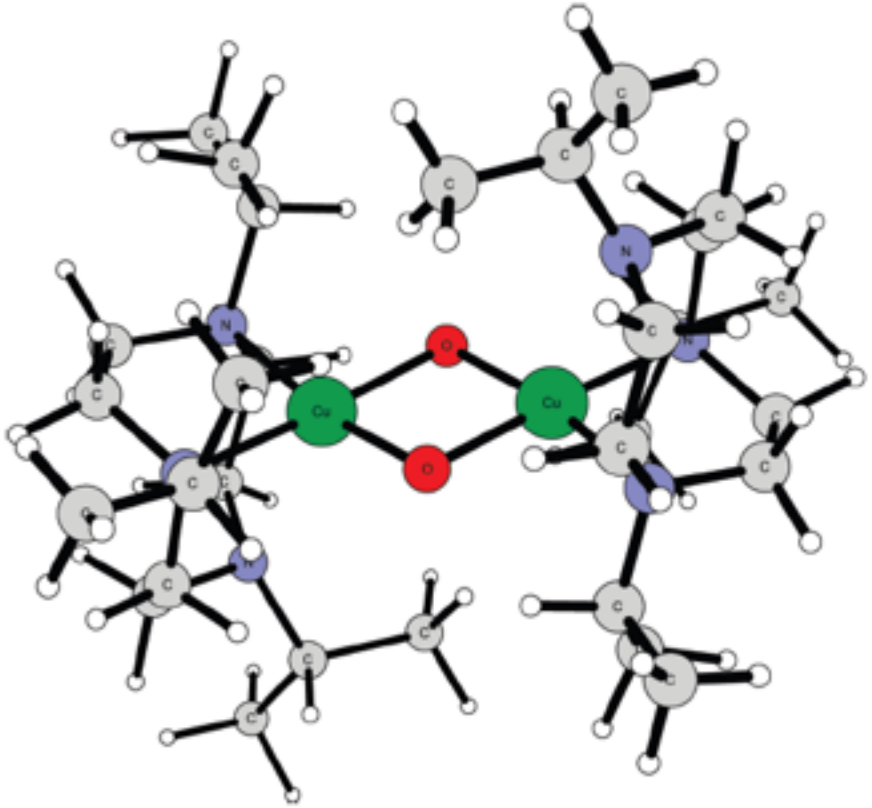


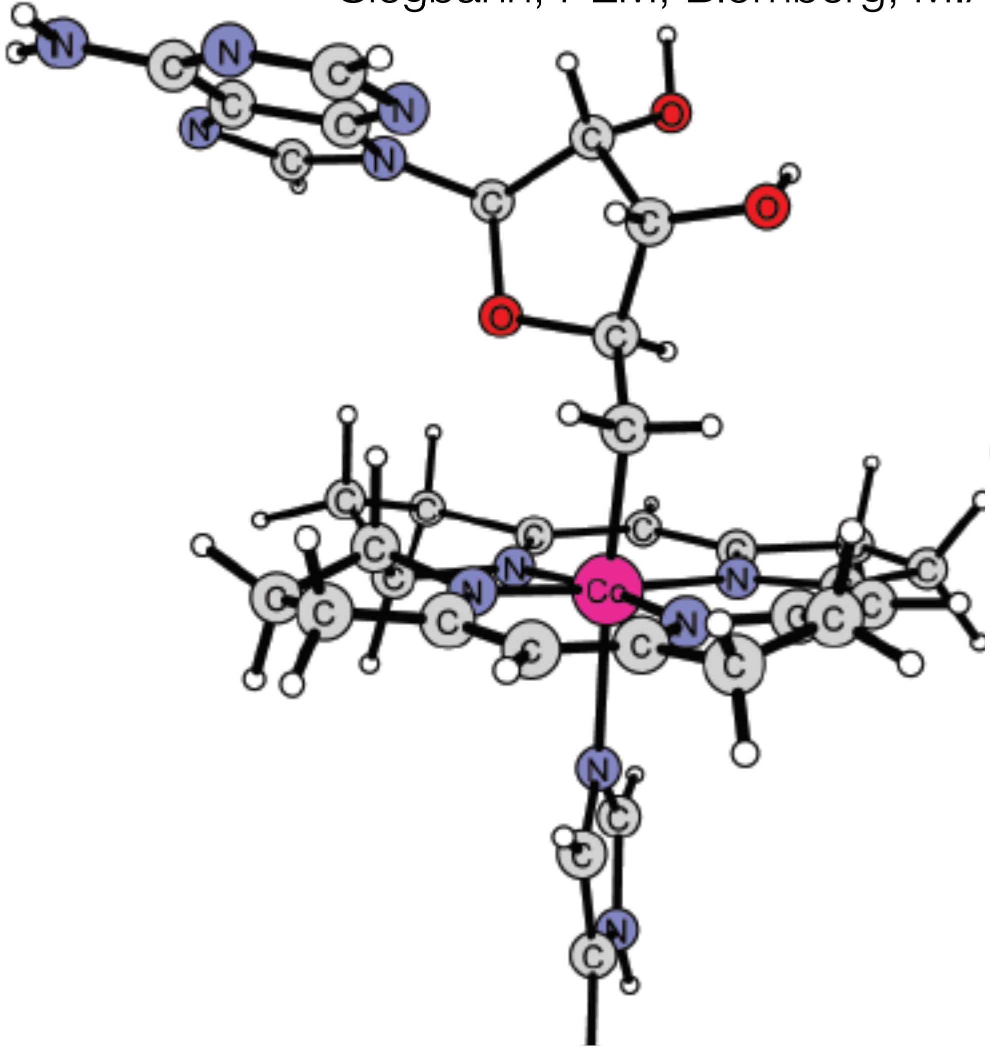
blue - B3LYP
red - B3LYP-D

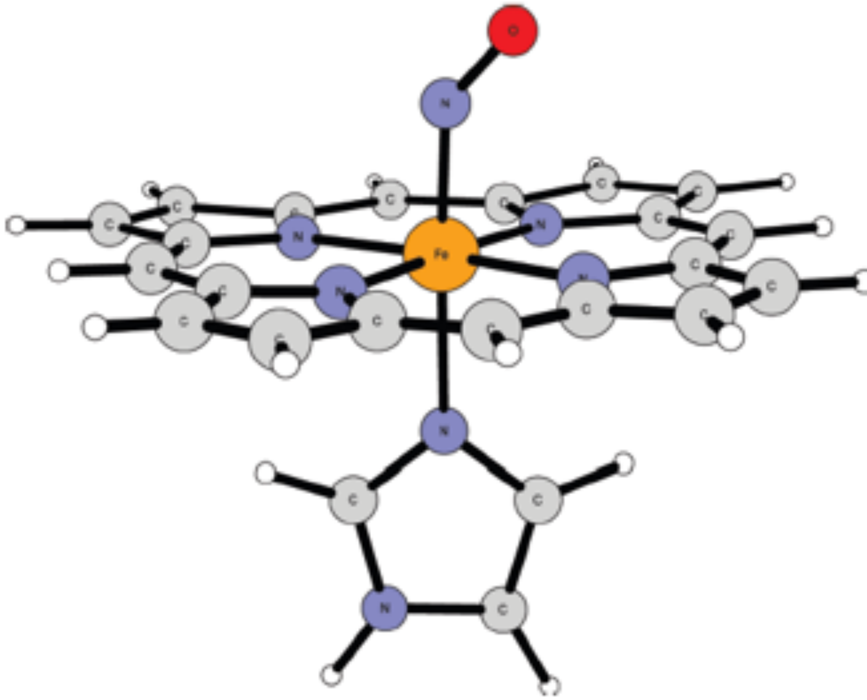
Siegbahn, PEM, Blomberg, M.A.; Chen, S.-L. *JCTC*, **2010**,6, 2040

Significant van der Waals effects

Siegbahn, PEM, Blomberg, M.A.; Chen, S.-L. *JCTC*, **2010**,6, 2040







Peroxo/bis-mue-oxo Isomerization energy

B3LYP	+15.4 kcal/mol
B3LYP*	+ 4.2 kcal/mol
B3LYP*-D	-0.6 kcal/mol
exp	**~0**

Methyl Binding

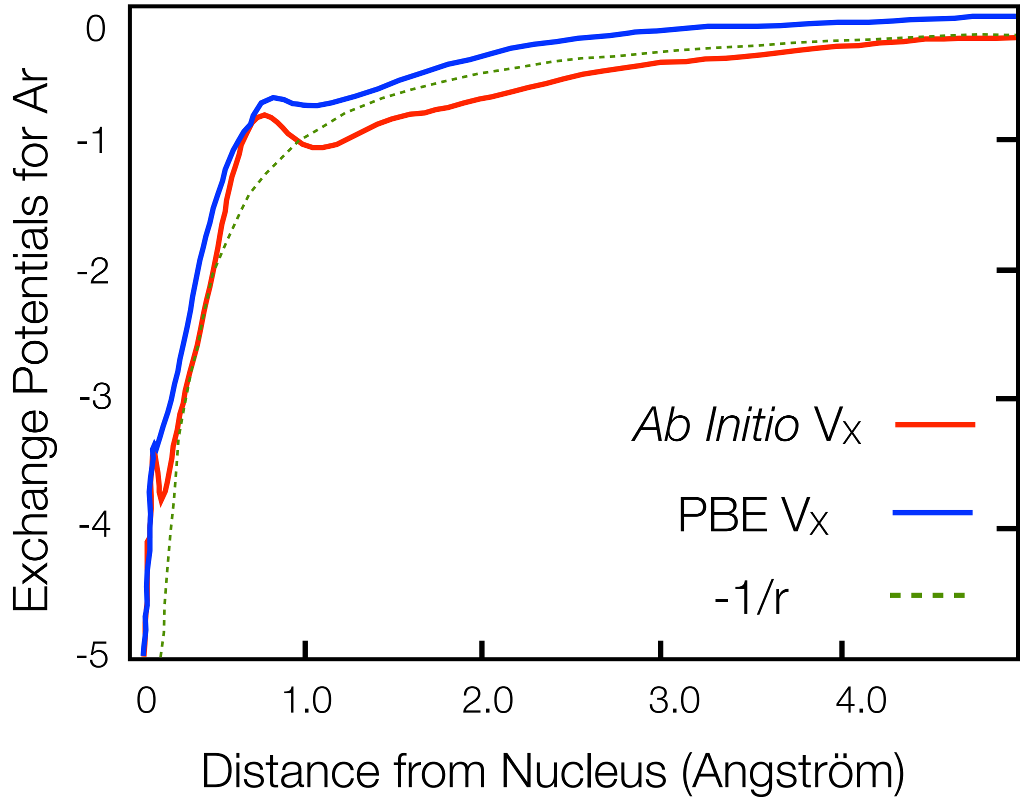
B3LYP	+16.2 kcal/mol
B3LYP*	+20.7 kcal/mol
B3LYP*-D	+32.4 kcal/mol
exp	**~37.3 kcal/mol**

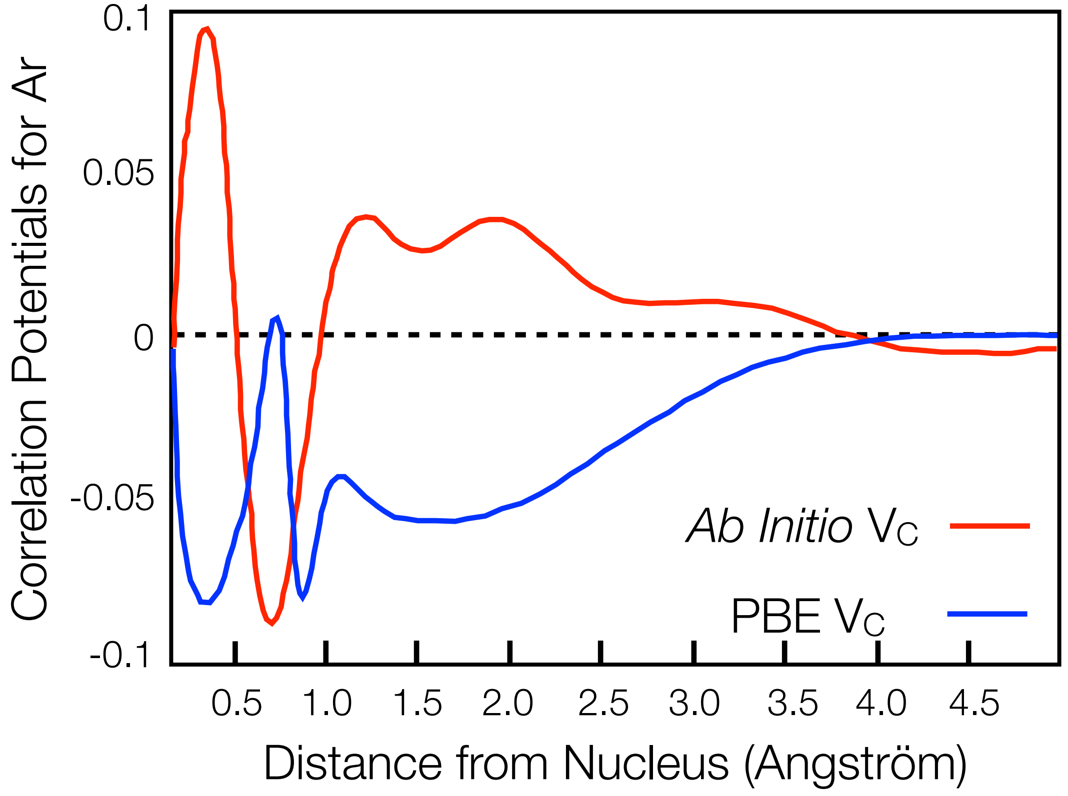
NO Binding

B3LYP	+7.6 kcal/mol
B3LYP*	+16.3 kcal/mol
B3LYP*-D	+25.6 kcal/mol
exp	**~22.8 kcal/mol**

However: *Ab Initio* vs actual DFT Potentials

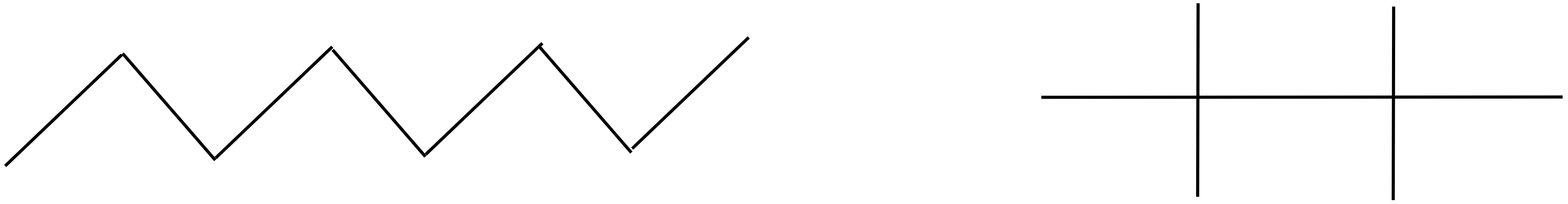
There are (expensive!) ways to construct very good KS potentials from accurate densities (red). These can be compared with “typical” present day potentials (blue).





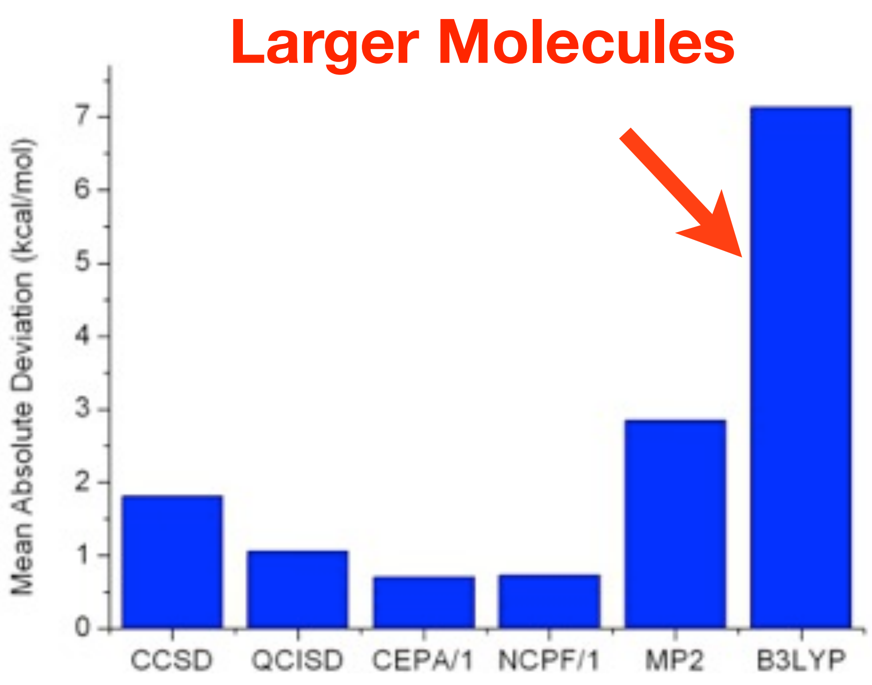
The presently used potentials are far from being correct and all present day DFT results rely on cancellation of large errors.

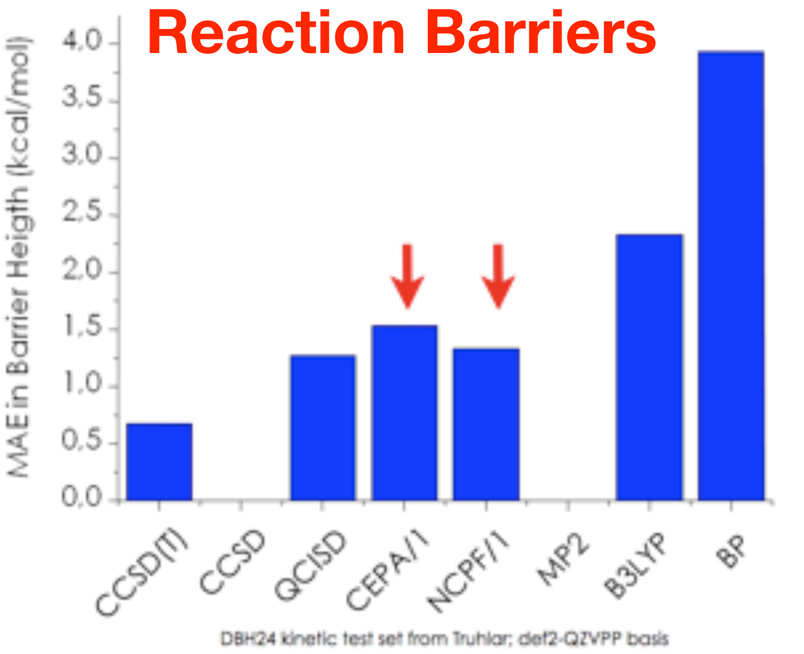
Big qualitative problems still exist ...



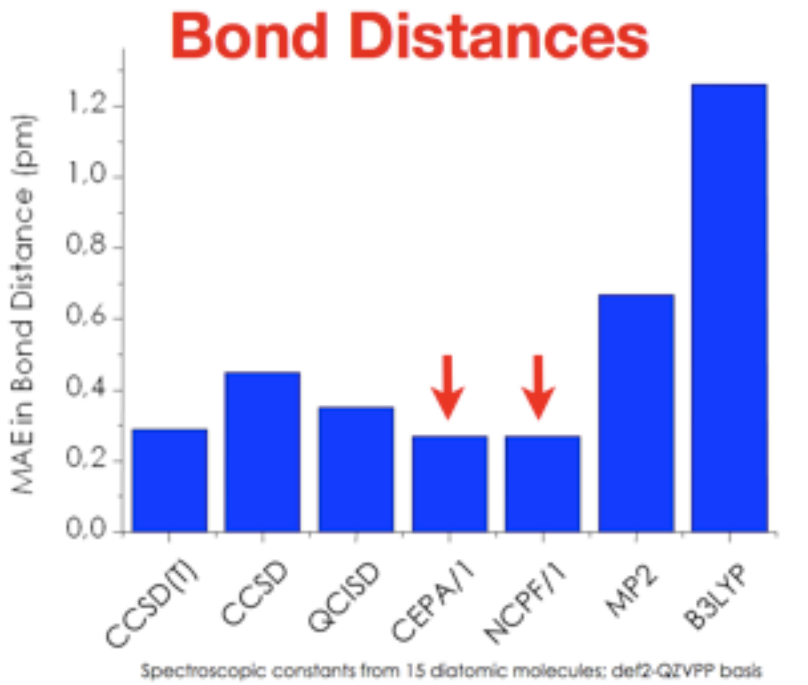
ΔE	=	+1.9±0.5	kcal/mol	Exp.
		+1.4	kcal/mol	SCS-MP2
		-11.5	kcal/mol	HF
		-8.4	kcal/mol	B3LYP
		-9.9	kcal/mol	BLYP

Energies - DFT vs *ab initio*

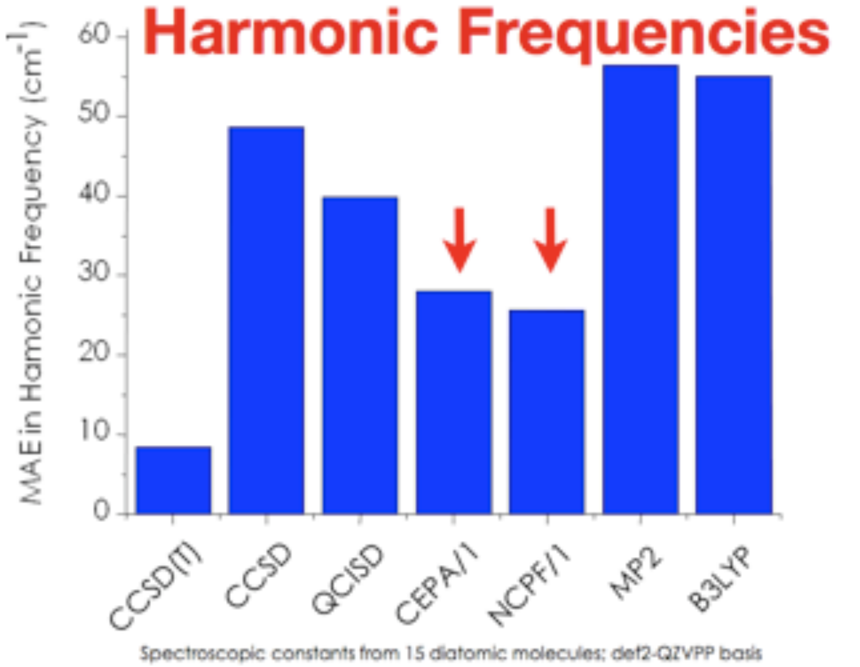

Larger Molecules


Reaction Barriers

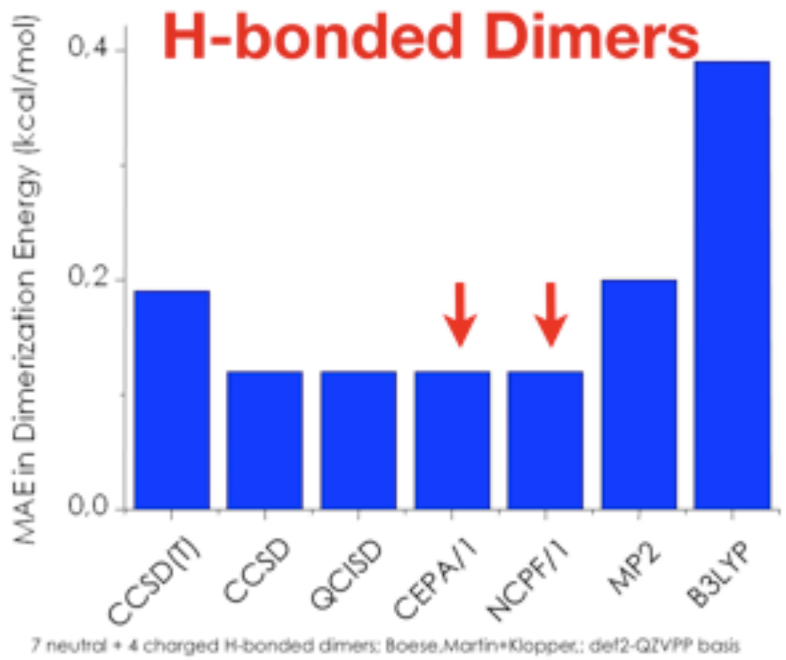
DBH24 kinetic test set from Truhlar; def2-QZVPP basis


Bond Distances

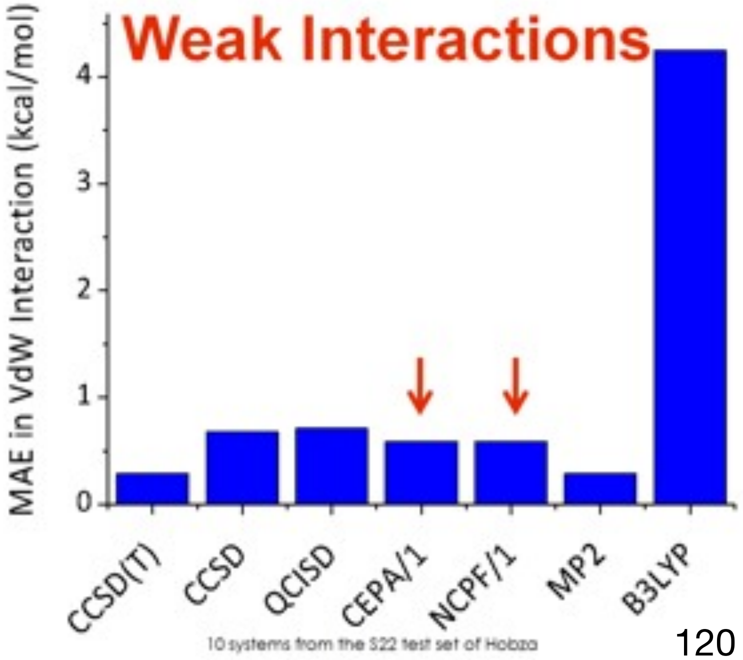
Spectroscopic constants from 15 diatomic molecules; def2-QZVPP basis


Harmonic Frequencies

Spectroscopic constants from 15 diatomic molecules; def2-QZVPP basis


H-bonded Dimers

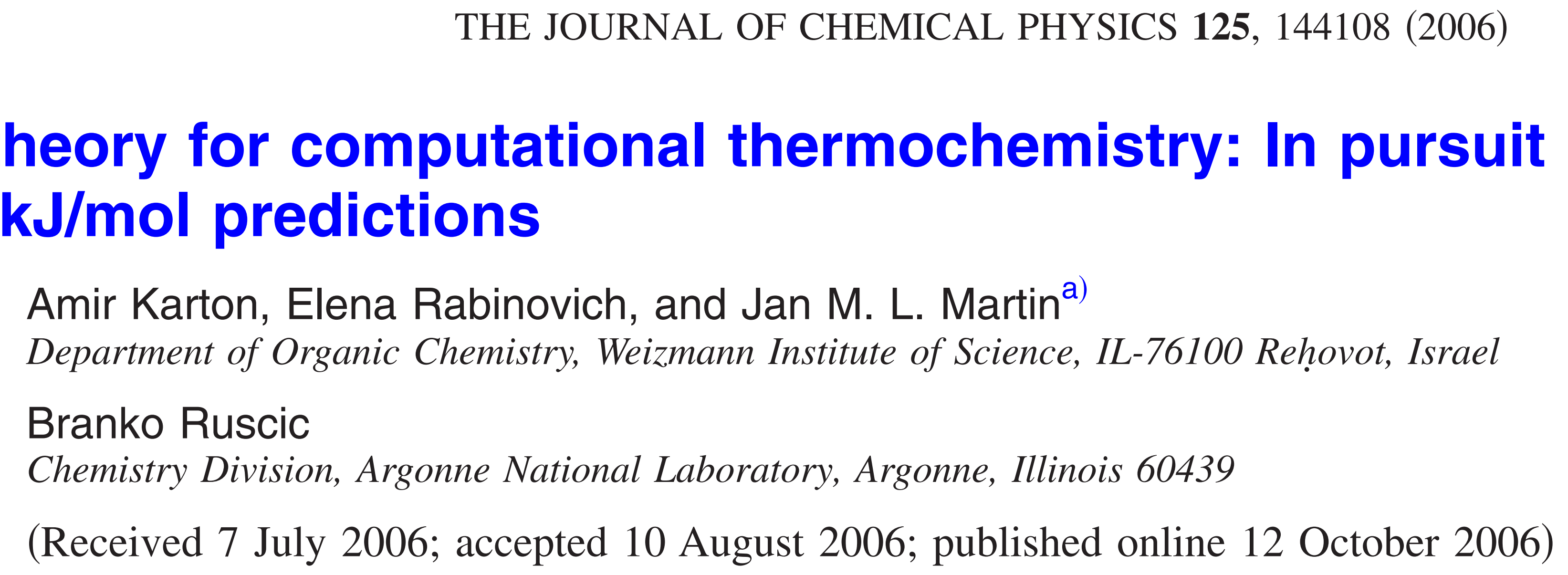
7 neutral + 4 charged H-bonded dimers; Boese,Martin+Klopper.; def2-QZVPP basis


Weak Interactions

10 systems from the S22 test set of Hobza

FN, A. Hansen, F. Wennmohs, S.- Grimme (**2009**), Acc. Chem. Res. 42, 641

Systematically approaching the solution



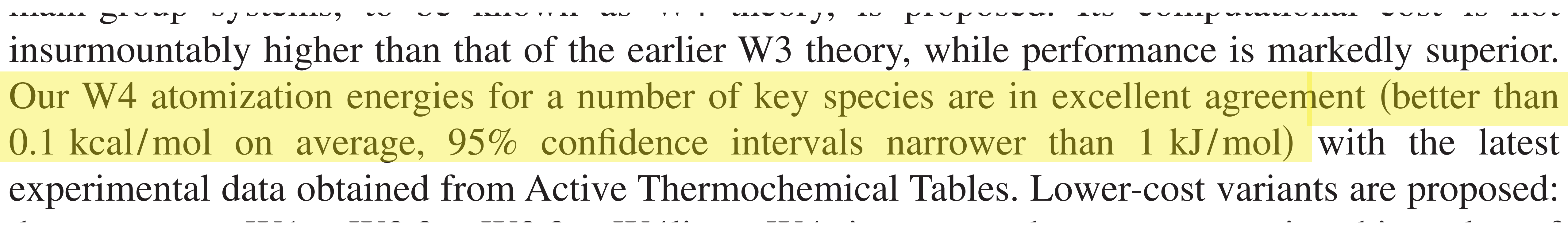
THE JOURNAL OF CHEMICAL PHYSICS **125**, 144108 (2006)

W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions

Amir Karton, Elena Rabinovich, and Jan M. L. Martin[a)]
Department of Organic Chemistry, Weizmann Institute of Science, IL-76100 Reḥovot, Israel

Branko Ruscic
Chemistry Division, Argonne National Laboratory, Argonne, Illinois 60439

(Received 7 July 2006; accepted 10 August 2006; published online 12 October 2006)



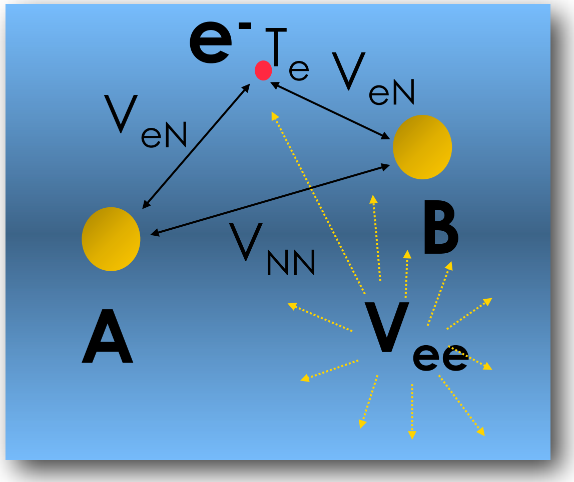
insurmountably higher than that of the earlier W3 theory, while performance is markedly superior. Our W4 atomization energies for a number of key species are in excellent agreement (better than 0.1 kcal/mol on average, 95% confidence intervals narrower than 1 kJ/mol) with the latest experimental data obtained from Active Thermochemical Tables. Lower-cost variants are proposed:

... meaning the (non-relativistic) Schrödinger equation is solved to an accuracy of 0.0001 Eh which is ~99.9999% or ~1 part in 10^6!

... For **really** small systems (1-6 *electrons*), we can today reach „crazy accuracy“, e.g. Nakatsuji calculated the H_2^- ground state energy to be **-0.597 139 063 123 405 074 834 134 096 025 974 142 a.u.** (36 significant digits!)

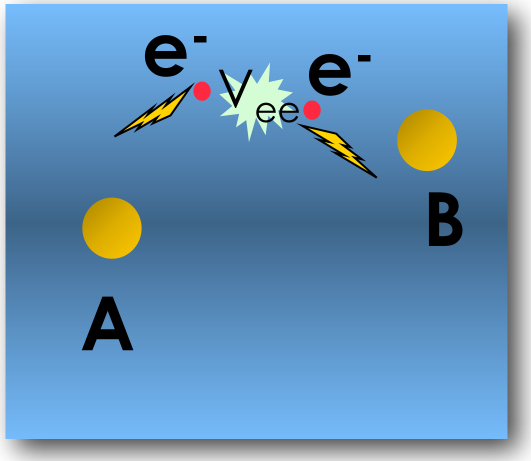
Electron Correlation: The Physical Problem

Exact Energy =



"Mean Field"
Hartree-Fock

+



Instantaneous electron-electron interaction

$$\text{Correlation energy} = \frac{1}{2}\sum_{i,j \text{ Electron pairs}} \varepsilon_{ij}(\uparrow\uparrow) + \varepsilon_{ij}(\uparrow\downarrow)$$

Fermi-Correlation ↑ *Relatively easy due to "Fermi hole" in the mean-field*

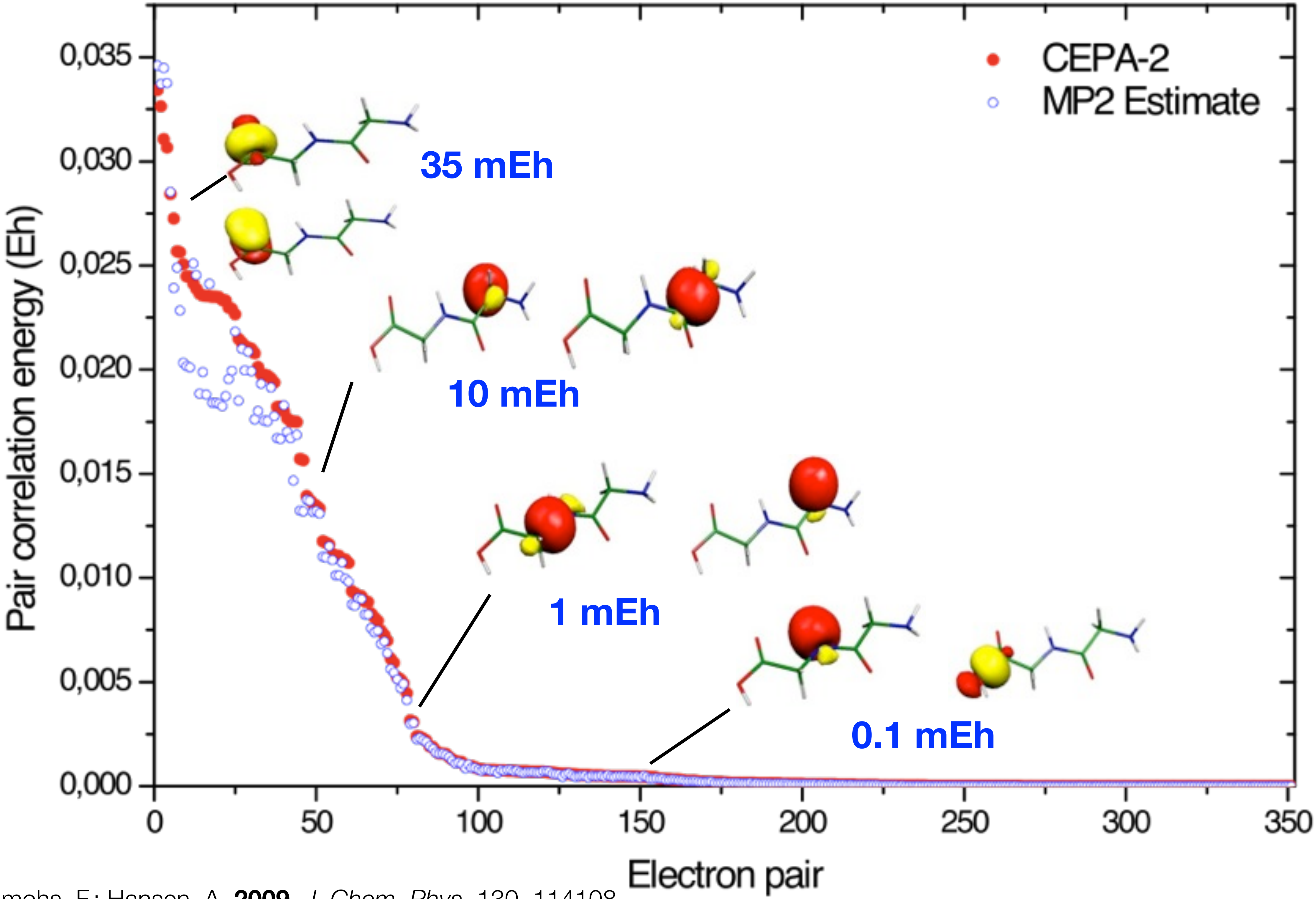
Coulomb-correlation ↑ *Hard to calculate due to interelectronic cusp at the coalescence point* $\boldsymbol{r}_1=\boldsymbol{r}_2$

F12 Methods

Electron Correlation: The Computational Problem

Local Correlation

Locality of Pair Correlation Energies



FN, Wennmohs, F.; Hansen, A. **2009**, *J. Chem. Phys.* 130, 114108

Appendix: More information on Multireference Methods

Generalizing: Detecting Multireference Character

Generally

- ✓ **Use chemical common sense!**
  - Using, e.g. ligand field theory or Lewis resonance structures you will in a large majority of cases be able to figure out off-hand that your system is multideterminantal or multiconfigurational

In DFT:

- ✓ Examine if the RHF wavefunction is stable and/or look at $\langle S^2 \rangle$
  - Note: this may well require several different initial guesses for the SCF to find alternative electronic states.
  - For example, the ***ORCA*** program let‘s you try different guesses according to BS(m,n)
  - Note: the more Hartree-Fock exchange the more likely your solution is to be „unstable“ and converge to spin-coupled states!
- ✓ Analyze the corresponding orbitals of the various solutions for overlaps significantly smaller than unity.
- ✓ Determine magnetic coupling parameters and solve the Heisenberg Hamiltonian for an estimate of pure spin-state energies

In Wavefunction Theory:

- ✓ Analyze the UHF wavefunction in the same way
- ✓ Look at the natural orbital occupation numbers of MP2 (or CCSD). Values significantly different from 0,1 or 2 indicate multireference character
  - These natural orbitals are excellent guesses for a subsequent CASSCF calculation
- ✓ Examine the largest doubles amplitudes of a CCSD calculation. Large values (approaching unity) indicate multireference character
  - Note that the frequently used T_1-diagnostic is not a good measure or MR character

Which Multireference Methods?

- **Complete Active Space self consistent field**
  - CASSCF is almost always the starting point for a MR calculation. You have to be careful about a number of points:
    - Which orbitals go in the active space? A bad choice spoils convergence. If the occupation number of an active orbital approaches 2.0 or 0.0 during the optimization this usually signals trouble. The active space should just cover the „essential physics“ (whatever that means to you ...)
    - Where do you get your initial guess orbitals from? Typically some kind of natural orbitals is a good idea. Never do a CASSCF calculation without looking at the orbitals that you put in the active space! Even then you may need to experiment with alternative active spaces
    - Over how many roots do you average?
    - Be aware of the restriction to about 14 active orbitals. Alternatives that allow larger active spaces exist (RASSCF, DMRG,...) but are not yet part of the standard arsenal.
  - CASSCF is not automatically size consistent. It depends on your choice of active space.
  - Geometry optimizations are reasonably efficient as CASSCF is fully variational
  - Do not forget that CASSCF is of the same overall quality as HF is for closed shell molecules. You cannot expect miracles and the CASSCF orbitals may be as desastrous for transition metals as RHF or ROHF orbitals.

Which Multireference Methods?

✓ **Multireference Perturbation Theory**

- Second order MR-PT has been developed into a general and powerful post-CASSCF theory. However, you should not forget that this is still second-order perturbation theory.
  - ➡ You get improved energies but not improved wavefunctions
  - ➡ The intrinsic accuracy cannot be expected to be much higher than single reference MP2 inside its valid domain *unless* you put substantial dynamic correlation in the reference.
- The most popular MR-PT2 variant is CASPT2 which is implemented in MOLCAS or MOLPRO. It has additional options:
  - ➡ The precise choice of H0. Depending on the type this may involve additional empirical parameters to be input
  - ➡ An empirical level shift to avoid intruder states
  - ➡ The option for „multistate“ treatments that make the results sensitive to the number and nature of roots to be determined
- An alternative is NEVPT2 (Dalton,MOLPRO, ORCA) or MCQDPT (Gamess, Firefly).
- For close-lying or crossing states you may have to resort to „multi-state“ treatments. These come with their own package of problems.
- Gradients are only sparsely available

Which Multireference Methods?

- ✓ **Multireference Coupled Cluster Theory**
  - From the point of view of computational chemistry this is „emerging technology“ that is barely applicable in its present form - but it is a very active field.

- ✓ **Multireference Configuration Interaction**
  - In principle, a high accuracy method that provides variational energies and wavefunctions, BUT:
    - ➡ If you pursue an „uncontracted“ MR-CI the calculations are of explosive cost with respect to the active space. (Very few programs: check COLUMBUS)
    - ➡ „Internal contracted MR-CI (e.g. MOLPRO) is much more efficient but still fairly limited in its applicability to larger molecules.
    - ➡ Approximate and more affordable MR-CI methods exist (e.g. SORCI in ORCA) but are specialist domain.
    - ➡ MR-CI is not size consistent. Popular options to approximately deal with this are MR-ACPF or MR-AQCC
    - ➡ A powerful approach for the calculation of energy differences that is less prone to size consistency problems and is much less expensive than full MR-CI is „difference dedicated CI“ (Malrieu, Caballol) as implemented in the programs of the Toulouse group or ORCA