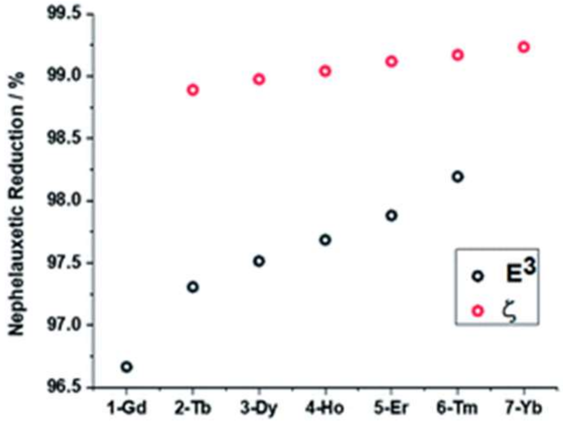
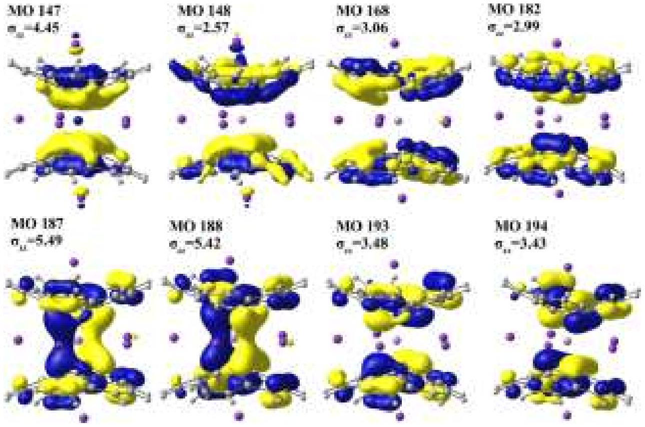
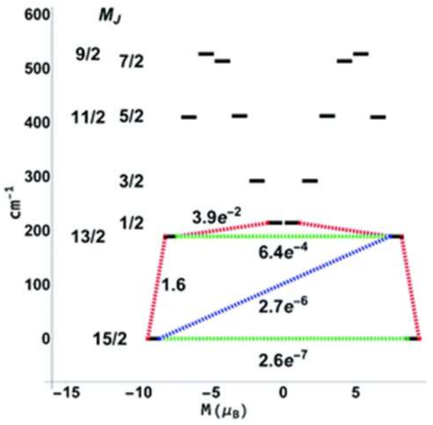
Orca

An Introduction and Running Basic Jobs







What is Orca?

The program ORCA is a modern electronic structure program package written by Prof. F. Neese, with contributions from Ute Becker, Dmytro Bykov, Dmitry Ganyushin, Andreas Hansen, Robert Izsak, Dimitrios G. Liakos, Christian Kollmar, Simone Kossmann, Dimitrios A. Pantazis, Taras Petrenko, Christoph Reimann, Christoph Riplinger, Michael Roemelt, Barbara Sandhöfer, Igor Schapiro, Kantharuban Sivalingam, Frank Wennmohs, Boris Wezisla and contributions from our collaborators: Mihály Kállay, Stefan Grimme, Edward Valeev. The binaries of ORCA are available free of charge for academic users for a variety of platforms.

ORCA is a flexible, efficient and easy-to-use general purpose tool for quantum chemistry with specific emphasis on spectroscopic properties of open-shell molecules. It features a wide variety of standard quantum chemical methods ranging from semiempirical methods to DFT to single- and multireference correlated ab initio methods. It can also treat environmental and relativistic effects.

Due to the user-friendly style, ORCA is considered to be a helpful tool not only for computational chemists, but also for chemists, physicists and biologists that are interested in developing the full information content of their experimental data with help of calculations.

The Basics

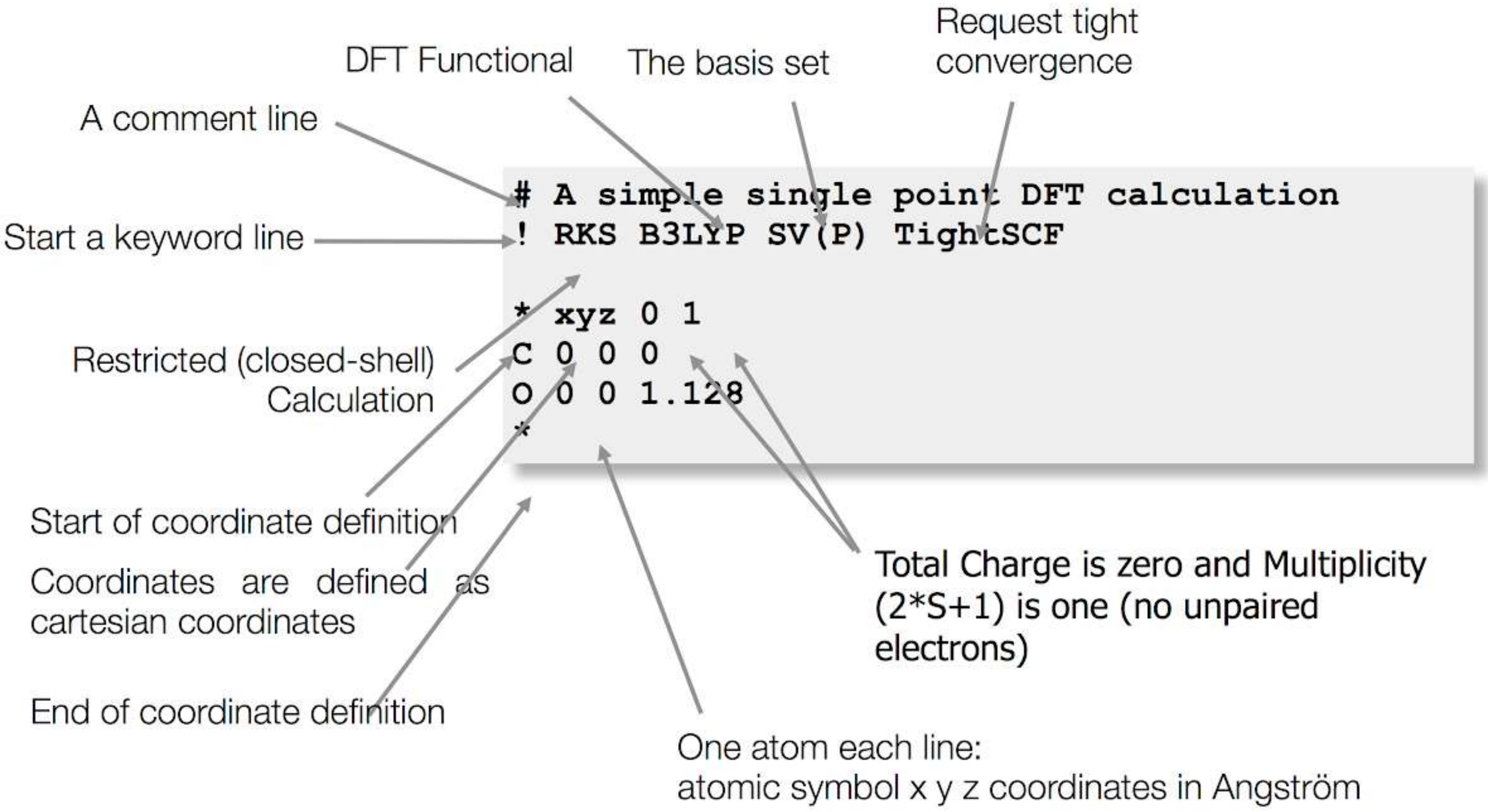
	Basic Computational Models	Basic Computations	Basic Analysis
Molecule - 3D Coords Charge Multiplicity	Density Functional Theory (DFT) Hartree-Fock (HF) + Basis Set	Energy Gradient Hessian	MO Energies MO Isosurfaces Partial Charges Bond Orders Vibrational Frequencies Checking Optimized Geometry IR Spectra

You Must Provide!!!

Your First Task When You Have a Problem You Can’t Answer

RTFM

The Input File



```
# 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

DFT Functional

The basis set

Request tight convergence

Start a keyword line

Restricted (closed-shell) Calculation

Start of coordinate definition

Coordinates are defined as cartesian coordinates

End of coordinate definition

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

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

The Basics

Molecule
 - 3D Coords

- You need to have a “good” set of coords for the program to work
- My suggestions to obtain a decent guess:
  1. Use WebMO (as discussed in the WebMO tutorial), and import coordinates from the WebMO guess
  2. PyMOL can do organics
  3. Chem3D (if you have a PC)
  4. Avogadro
  5. OpenBabel
  6. wxMacMolPlt
  7. ChemCraft
  8. Mercury

Charge

You MUST know this beforehand

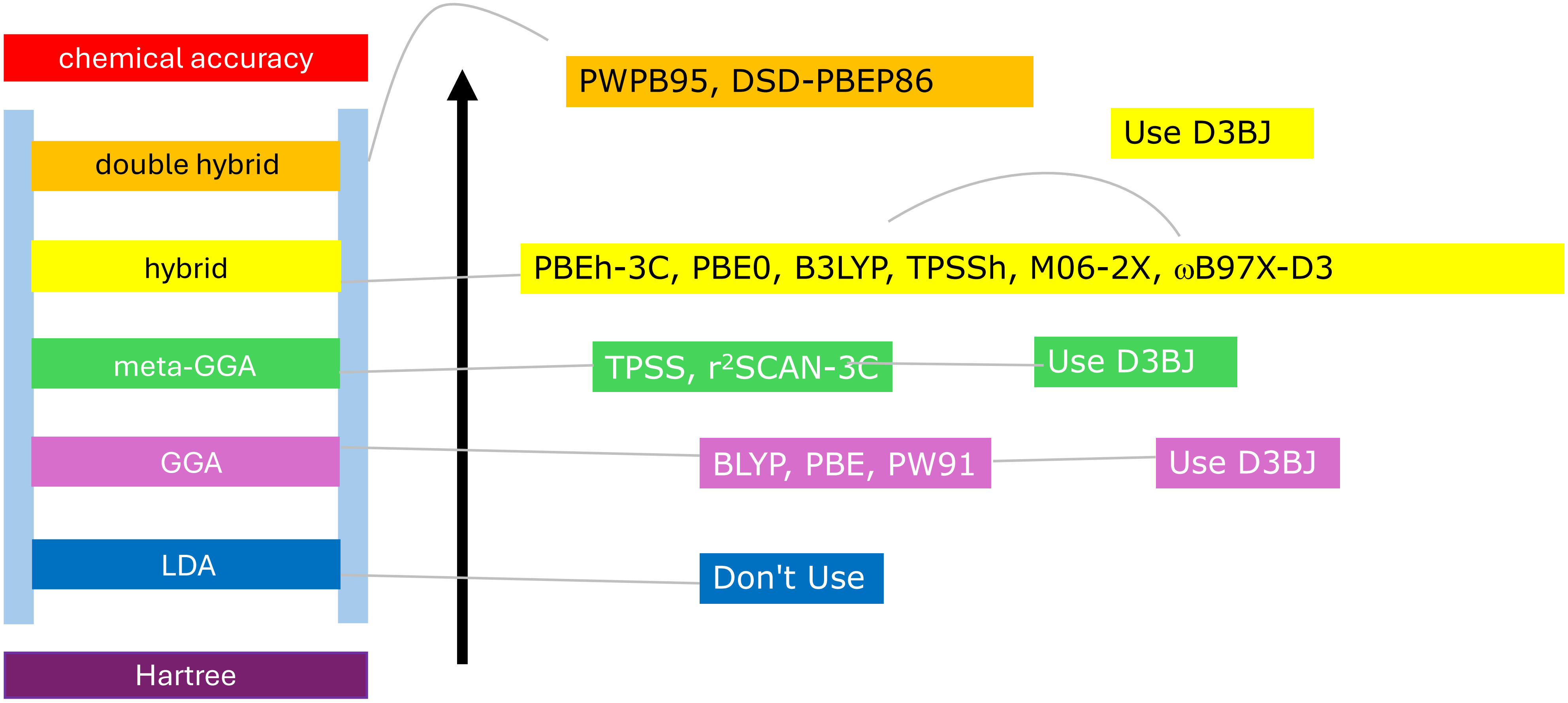
Multiplicity (2S+1)

Is your molecule magnetic?
No? 2S+1=1

Is it an open-shell singlet?
More complicated…

Transition Metal?
High/Intermediate/Low Spin?

DFT – Recommendations for GS energies



chemical accuracy

double hybrid — PWPB95, DSD-PBEP86

Use D3BJ

hybrid — PBEh-3C, PBE0, B3LYP, TPSSh, M06-2X, ωB97X-D3

meta-GGA — TPSS, r^{2}SCAN-3C — Use D3BJ

GGA — BLYP, PBE, PW91 — Use D3BJ

LDA — Don't Use

Hartree

DFT – Recommendations Optimization Basis Sets

POPLE [6-31G(d), 6-311+G(2d,p), etc...]
I don't recommend the Pople Basis sets for use with Orca

- No Auxiliary Basis Sets Available
- No "automatic" effective core potentials (ECP).
- Limited scope of periodic table (PT)

CORRELATION CONSISTENT [cc-pVDZ, cc-pVTZ, etc...]

- I don't recommend for DFT optimizations.
- This basis set is tailor-made for post-HF methods like MP2, so I definitely recommend this BS for these types of theories.

KARLSRUHE/AHLRICHS [def2-SVP, def2-TZVP, def2-QZVP, etc...]
My basis set of choice in most cases.

- Built-in "automatic" Auxiliary Basis Set (DEF2/J)
- Automatically loads ECP
- Systematic (SVP-TZVP-QZVP-QZVPP)
- Covers most of the PT
- Well-known to work well with DFT

JENSEN [pcseg-0, pcseg-1, etc...]

- Better than POPLE, but I still don't recommend for use with ORCA for "normal" calculations.
- Useful for some properties, however!

DFT – Recommendations for Optimizations

PBEh-3C, r^2SCAN-3C: Built-In Basis Set

Inorganic: def2-TZVP
PBE0, B3LYP, TPSSh, M06-2X, ωB97X-D3
Organics: def2-SV(P)

PWPB95, DSD-PBEP86

- Wouldn't optimize with these

Inorganic: def2-TZVP
TPSS, r^2SCAN
Organics: def2-SV(P)

Inorganic: def2-TZVP
BLYP, PBE, PW91
Organics: def2-SV(P)

Default convergence and DFT grid usually OK

DFT – Recommendations for SP

PBEh-3C, r^2SCAN-3C: Built-In Basis Set

Inorganic: As Big As Possible
PBE0, B3LYP, TPSSh, M06-2X, ωB97X-D3
Organics: def2-TZVP

PWPB95, DSD-PBEP86
DEF2-TZVP

Inorganic: As Big As Possible
TPSS, r^2SCAN
Organics: def2-TZVP

BLYP, PBE, PW91
Woudn't Use for SP

I Would Use TightSCF and consider DEFGRID3

Under the hood: the RIJCOSX Approximation

By default, ORCA uses the RIJCOSX Approximation

For Coulomb integrals, normally we have:

$$J_{\mu\nu\sigma\tau} = P_{\mu\nu}P_{\sigma\tau}\left\langle\mu(1)\nu(1)\middle|r_{12}^{-1}\middle|\sigma(2)\tau(2)\right\rangle$$

In the RIJ approximation, we make the following approximate substitution:

$$P_{\sigma\tau}\sigma\tau \approx \sum_k P_k\zeta_k$$

This turns a 4-center integral into a 3-center integral (which is faster to evaluate)

Requires an Auxiliary Basis Set (As mentioned previously)

Under the hood: the RIJCOSX Approximation

By default, ORCA uses the RIJCOSX Approximation

For Exchange integrals, normally we have:

$$K_{\mu\nu\sigma\tau} = P_{\mu\nu} P_{\sigma\tau} \left\langle \mu(1)\nu(1) \middle| r_{12}^{-1} \middle| \sigma(2)\tau(2) \right\rangle$$

In the COSX approximation, we make the following approximate substitution:

$$\left\langle \mu(1)\sigma(1) \middle| r_{12}^{-1} \middle| \nu(2)\tau(2) \right\rangle \approx \sum_g X_{\mu g} X_{\sigma g} A_g^{\nu\tau}$$

where with g being grid-points

$A_g^{\nu\tau} = \int \frac{\nu(r)\tau(r)}{|r-r_g|} dr$ and $X_{\mu g} = \sqrt{w_g}\,\mu(r_g)$

We thus reduce the 4-center exchange integral to a two-center integral applied over a numerical grid.

Integration Grids: the “Other” Basis Set

By default, ORCA uses the so-called DEFGRID2 which sets a numerical integration grid for both the xc-kernel as well as the COSX grid.

It might sometimes be necessary to improve the grid by moving to DEFGRID3:

1. This can sometimes happen when very tight accuracy is needed for an energy calculation.
2. Weaker grids can be susceptible to numerical problems for optimizations with shallow PESs.

If DEFGRID3 is still not good enough, you can turn off RIJCOSX by using the keyword NORI

Our First Job: A Single Point Run

Let’s do a single-point energy calculation with the r^{2}SCAN-3C model on cisplatin.

1. First we’ll build the molecule

2. Next we’ll create the input file.

3. Finally, we’ll create the runscript and submit the job.

Build of cisplatin

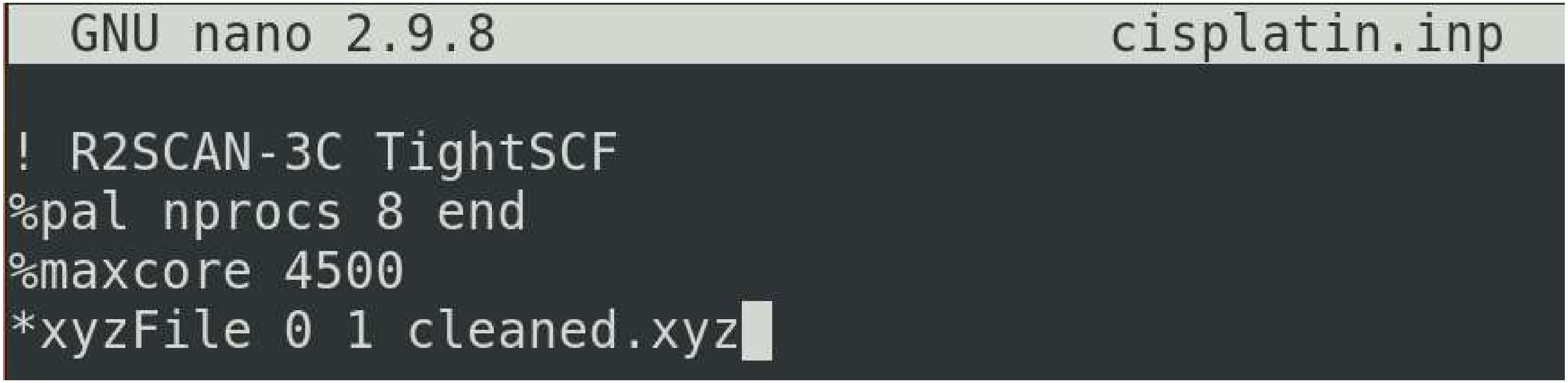
We'll use a combination of WebMO and OpenBabel

1. Use the symmetry fragments in WebMO to get a square planar structure and make the necessary modifications.

2. Go to the WebMO submission dialog to get the Cartesian coordinates, and Copy the coords

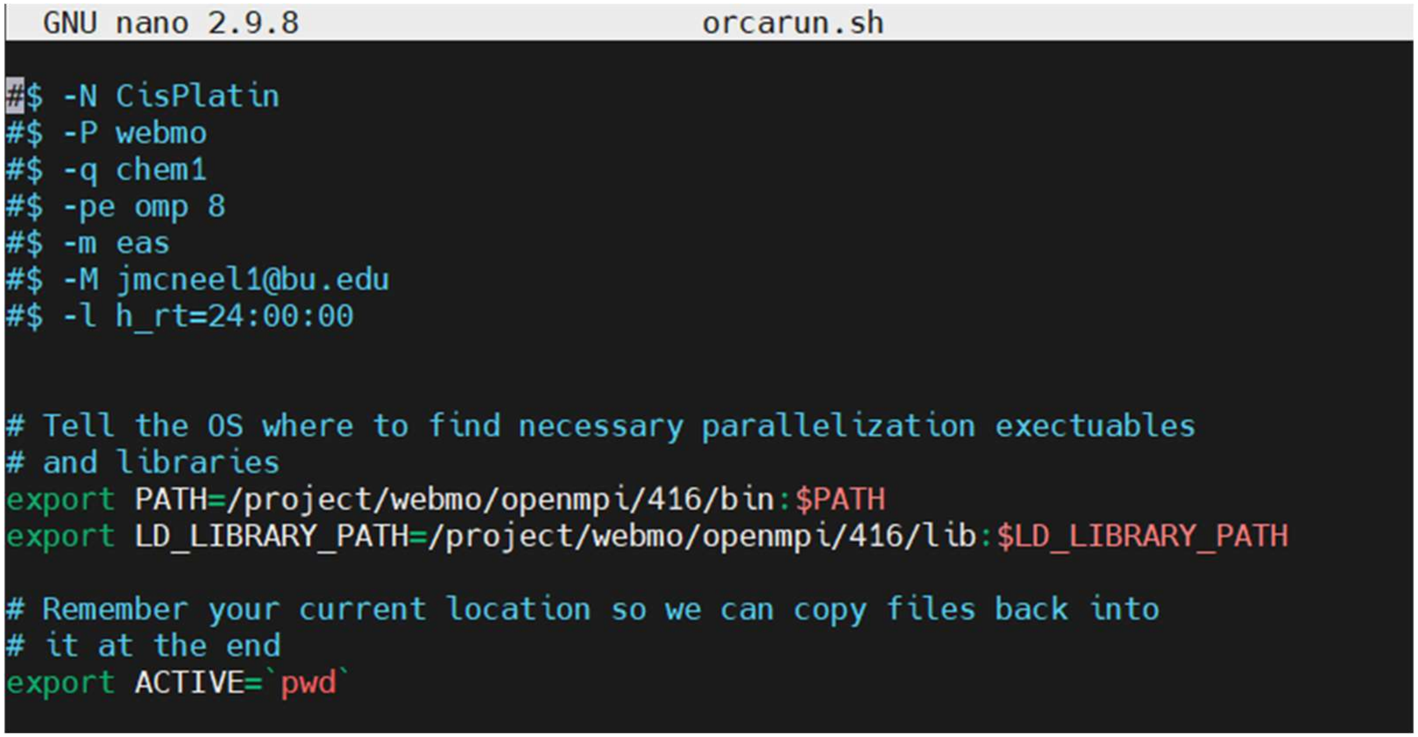
3. Use OpenBabel to clean the geometry with the UFF force field.

Input File



```
! R2SCAN-3C TightSCF
%pal nprocs 8 end
%maxcore 4500
*xyzFile 0 1 cleaned.xyz
```

Runscript



```
#$ -N CisPlatin
#$ -P webmo
#$ -q chem1
#$ -pe omp 8
#$ -m eas
#$ -M jmcneel1@bu.edu
#$ -l h_rt=24:00:00


# Tell the OS where to find necessary parallelization exectuables
# and libraries
export PATH=/project/webmo/openmpi/416/bin:$PATH
export LD_LIBRARY_PATH=/project/webmo/openmpi/416/lib:$LD_LIBRARY_PATH

# Remember your current location so we can copy files back into
# it at the end
export ACTIVE=`pwd`
```

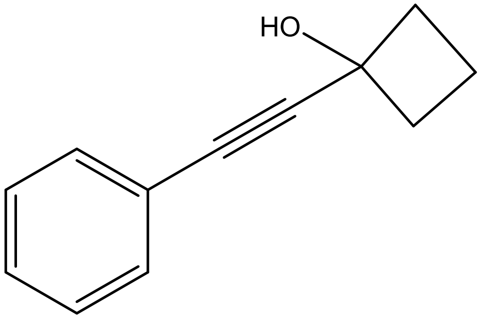
Submitting and Checking Status

- qsub *runscript_name*
- qstat –u *your_username*
- qacct –d *num_days* –o *your_username* -j

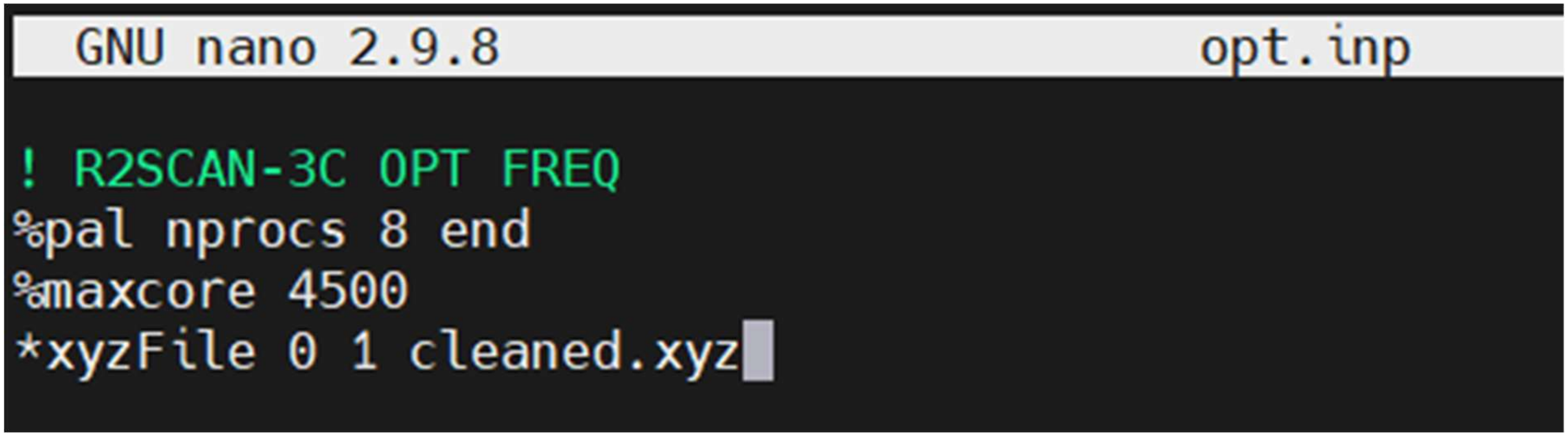
Our Second Job: A Geometry Optimization + Hessian

We'll use the same level of theory and use Chemdraw+OpenBabel to build a alkynylcyclobutanol recently listed in JOC.

https://pubs.acs.org/doi/10.1021/acs.joc.4c00505



Input File



```
  GNU nano 2.9.8                                    opt.inp

! R2SCAN-3C OPT FREQ
%pal nprocs 8 end
%maxcore 4500
*xyzFile 0 1 cleaned.xyz
```

Optimizations in Orca

	Opt	TightOpt	VeryTightOpt
TolE	5e-6	1e-6	2e-7
TolRMSG	1e-4	3e-5	8e-6
TolMAXG	3e-4	1e-4	3e-5
TolRMSD	2e-3	6e-4	1e-4
TolMAXD	4e-3	1e-3	2e-4

An important keyword to remember if you’re having convergence problems and you notice the energy is “oscillating”.

Use the TRUST option in the %geom end block. Maybe set to -0.1 and then keep going down until the oscillations stop

Checking for Imaginary Frequencies

```
grep imag output_file
```

OPT:

```
Total number of imaginary perturbations ...      0
```

TIGHTOPT:

```
Total number of imaginary perturbations ...      0
          6:      -1.50 cm**-1 ***imaginary mode***
```

VERYTIGHTOPT:

```
Total number of imaginary perturbations ...      0
```

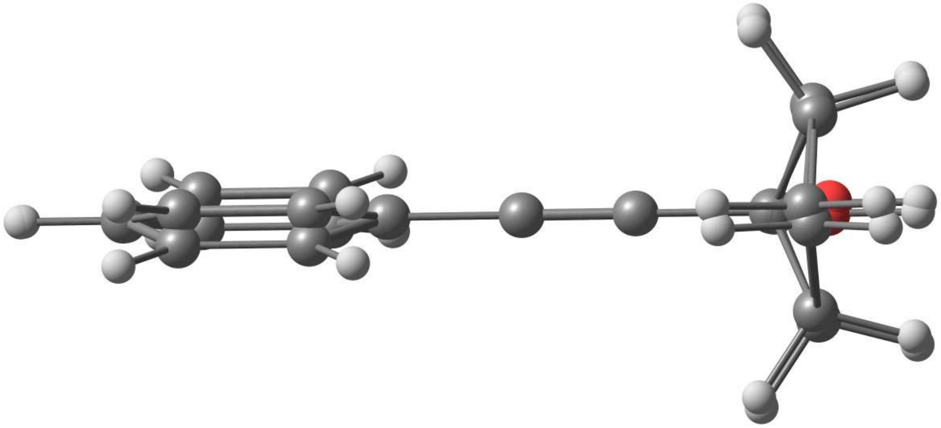
Solving Imaginary Frequencies

```
for i in `seq 5 5 20`; do
mkdir P${i}
mkdir M${i}
head –n4 tightopt.inp > P${i}/tightopt.inp
head –n4 tightopt.inp > M${i}/tightopt.inp
d=`echo “${i}/100” | bc -l`
na=`head –n1 tightopt.xyz`
sed –i “s/xyzFile/xyz/g” P${i}/tightopt.inp
sed –i “s/xyzFile/xyz/g” M${i}/tightopt.inp
sed –i “s/cleaned.xyz/\ /g” P${i}/tightopt.inp
sed –i “s/cleaned.xyz/\ /g” M${i}/tightopt.inp
/projectnb/webmo/jmcneel1/bin/displace.py tightopt.hess 6 ${d} | tail –n${na} >> P${i}/tightopt.inp
/projectnb/webmo/jmcneel1/bin/displace.py tightopt.hess 6 -${d} | tail –n${na} >> M${i}/tightopt.inp
echo “*” >> P${i}/tightopt.inp
echo “*” >> M${i}/tightopt.inp
cp orcarun.sh P${i}/
sed –i “s/verytightopt/tightopt/g” P${i}/orcarun.sh
cp P${i}/orcarun.sh M${i}/orcarun.sh
done
```

Solving Imaginary Frequencies

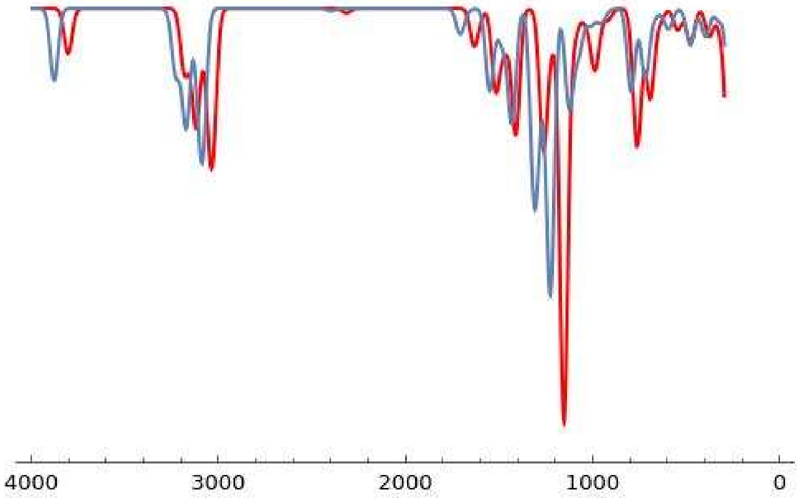
```
for i in `seq 5 5 20`; do
cd P${i}
qsub orcarun.sh
cd ../M${i}
qsub orcarun.sh
cd ../
done
```

Comparing Models



ωB97X-D3BJ shifts the plane of the ring

ωB97X-D3BJ is blueshifted



Grabbing Energies

The Electronic Energy

```
grep FINAL output_file | tail –n1
```

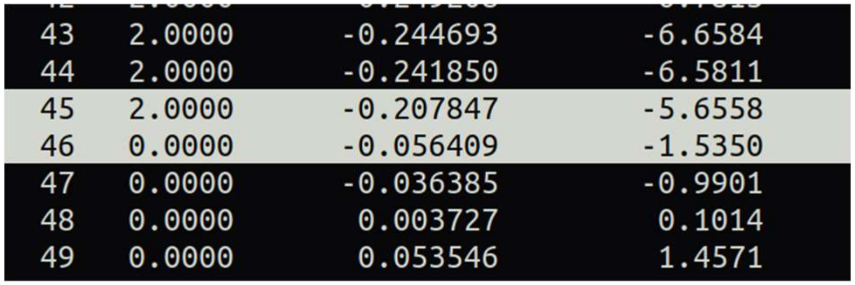
The Gibbs Free Energy

```
grep Gibbs output_file | tail –n2 | head –n1
grep FREEENERGYG property_file
```

Absolute energies are fairly useless, YOU SHOULD ONLY BE PAYING ATTENTION TO RELATIVE ENERGIES

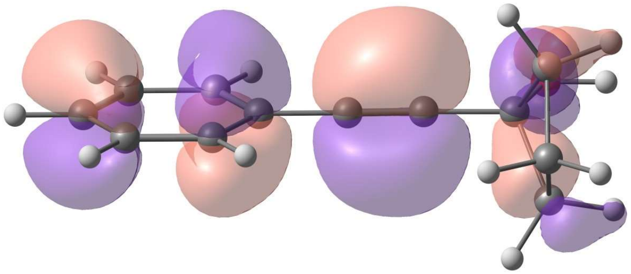
HOMO/LUMO Gaps

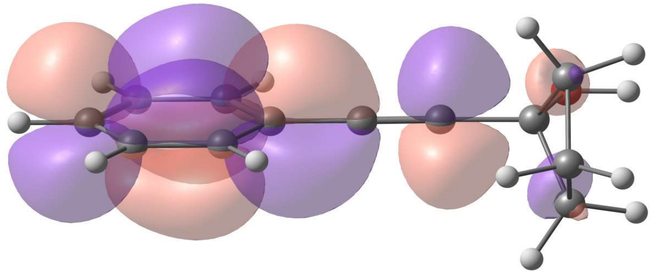
grep –A 200 ORBITAL\ EN *output_file*



```
43    2.0000        -0.244693        -6.6584
44    2.0000        -0.241850        -6.5811
45    2.0000        -0.207847        -5.6558
46    0.0000        -0.056409        -1.5350
47    0.0000        -0.036385        -0.9901
48    0.0000         0.003727         0.1014
49    0.0000         0.053546         1.4571
```

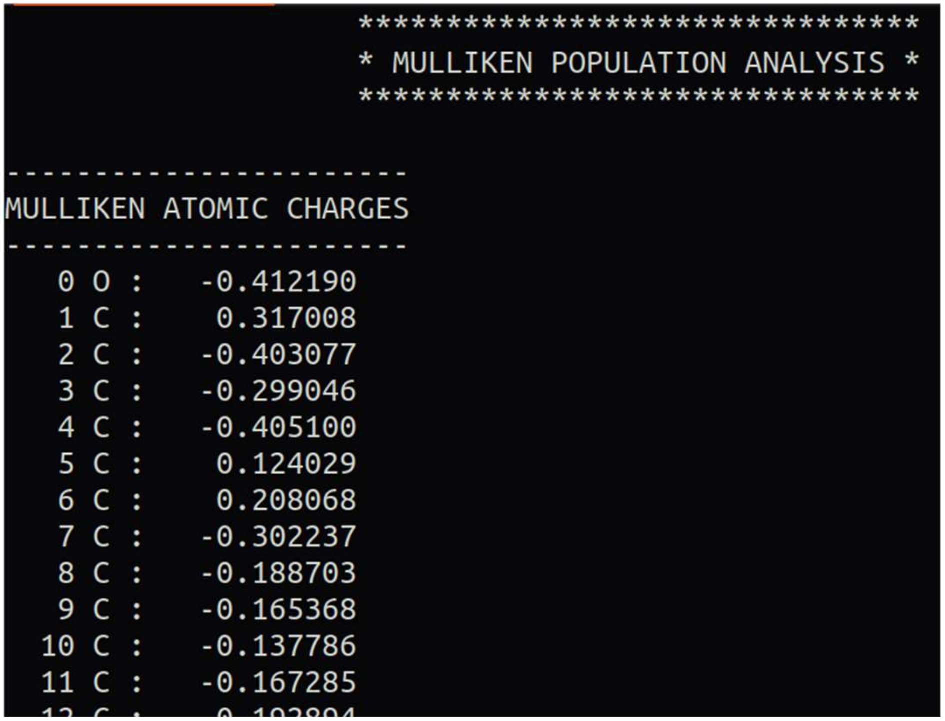
/projectnb/webmo/orca/6_0_0/orca_plot *gbw_file* -i





Population Analysis

grep -n MULL *output_file*
less +# *output_file*



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


-----------------------
MULLIKEN ATOMIC CHARGES
-----------------------
   0 O :   -0.412190
   1 C :    0.317008
   2 C :   -0.403077
   3 C :   -0.299046
   4 C :   -0.405100
   5 C :    0.124029
   6 C :    0.208068
   7 C :   -0.302237
   8 C :   -0.188703
   9 C :   -0.165368
  10 C :   -0.137786
  11 C :   -0.167285
  12 C :    0.193894
```