Texas A&M HPRC Short Course Series

Introduction to Quantum Chemistry Simulations with ORCA

Xin Yang

DIVISION OF RESEARCH
TEXAS A&M UNIVERSITY

Outline

10:00 -10:20 Intro to Computational Chemistry
10:20 -10:35 Hands-on Session 1 –Set up an ORCA calculation

10:35-11:10 Basics of Quantum Mechanical Simulation
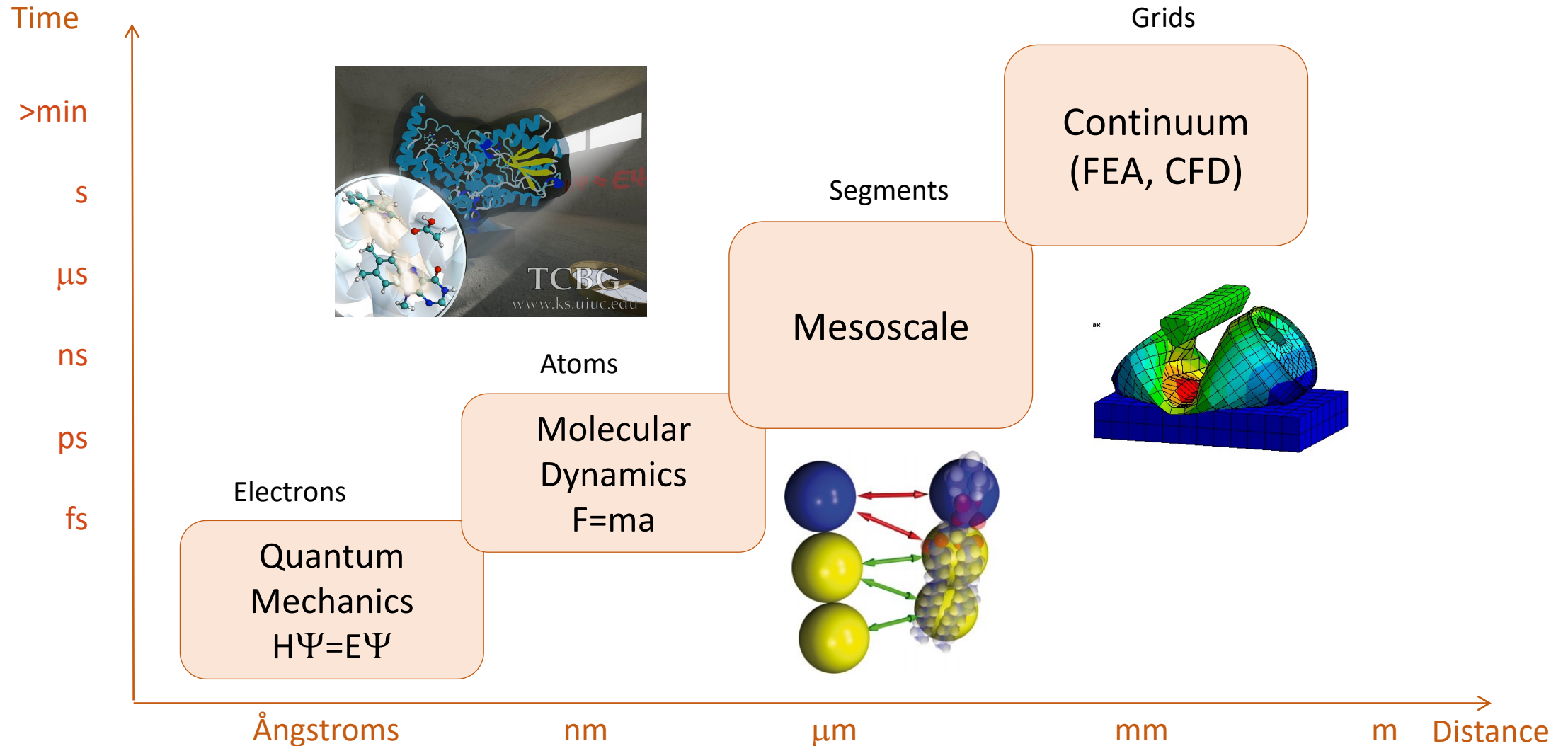11:10-11:40 Hand-on Session 2 – Geometry optimization and frequency calc.

11:40-11:55 Calculations of Molecular Properties
11:55-12:15 Hands-on session 3 – Prediction of UV/Vis Spectra

12:15-12:30 Wrap-up Lecture

Microscopic ⇔ Macroscopic

Time

>min

s

μs

ns

ps

fs

Grids

Continuum (FEA, CFD)

Segments

Mesoscale

Atoms

Molecular Dynamics F=ma

Electrons

Quantum Mechanics HΨ=EΨ

Ångstroms   nm   μm   mm   m   **Distance**

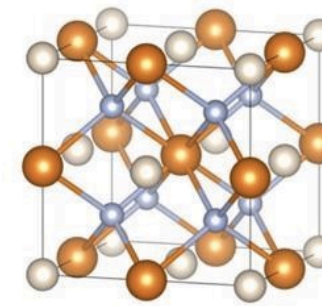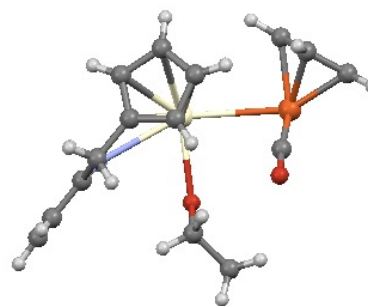# What is computational chemistry?

- Computational Chemistry: Use mathematical approximations and computer programs to obtain results relative to chemical problems.
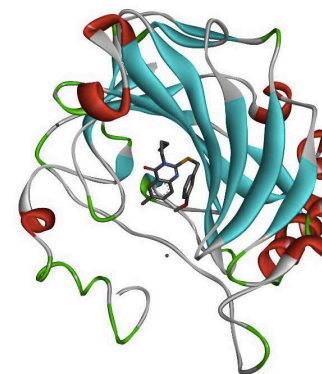
    Quantum Mechanical
    i.e., via *Schrödinger* Equation
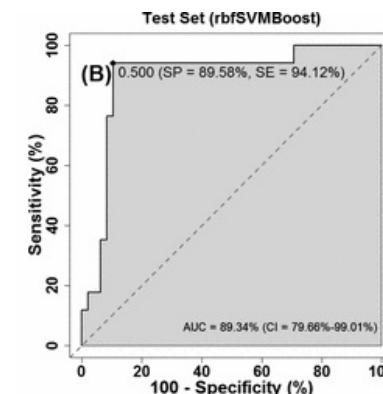   also called *Quantum Chemistry*

    Molecular Mechanical
    i.e., via *Newton*'s law F=ma
    also called *Molecular Dynamics*

    Empirical/Statistical/Machine Learning
    e.g., *QSAR*, etc., widely used in clinical and medicinal chemistry
    also called *Cheminformatics*

Computational Quantum Chemistry

- Solve the Schrödinger equation for molecular systems

$$H\Psi = E\Psi$$

$\hat{H}$ Hamiltonian Operator
ψ Wavefunction (eigenfunction)
E Energy of the system (eigenvalue)

- What is the typical size of system do we deal with?

  Assuming typical computing setup (number of CPUs, memory, disk space, etc.)
  - Semi-empirical method: *larger than DFT* (approximate solution, highly parameterized)
  - DFT method: *~100 atoms*
  - *Ab initio* method: between 1 and *~50 atoms* depending on the level of theory

Quantum Mechanical Methods

- ***ab initio* methods**

  - Hartree-Fock Methods: HF, RHF, URH

  - Post-Hartree Fock Methods: MPn, CI, CC, QCI

  - Multireference Methods: CASSCF (NEVPT2, CASPT2)

  - Applicable to any system, in principle

  - Computationally expensive

  - Typically used for small system (normally < 50 atoms)

  - Scaling: N^n n=2, 3, 4, 5, 6, …

Quantum Mechanical Methods

- **Density Functional Theory (DFT)**

 - Total energy of a system depends only on the electron density

  $$E_v[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + \bar{T}_s[\rho] + \frac{1}{2}\iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}}d\mathbf{r}_1 d\mathbf{r}_2 + E_{xc}[\rho]$$

  - Single determinant method

  - Includes electron correlation with little cost compared to ab initio methods

  - Not as dependent on the quality of the basis set as wave function methods

  - Exact functional is not known, results may vary with the choice functional

Reaction Mechanism Calculations

- Geometry optimization
- Transition state search
- Thermochemistry
- Intrinsic reaction coordinate

To understand the regioselectivity of polymer initiation step.
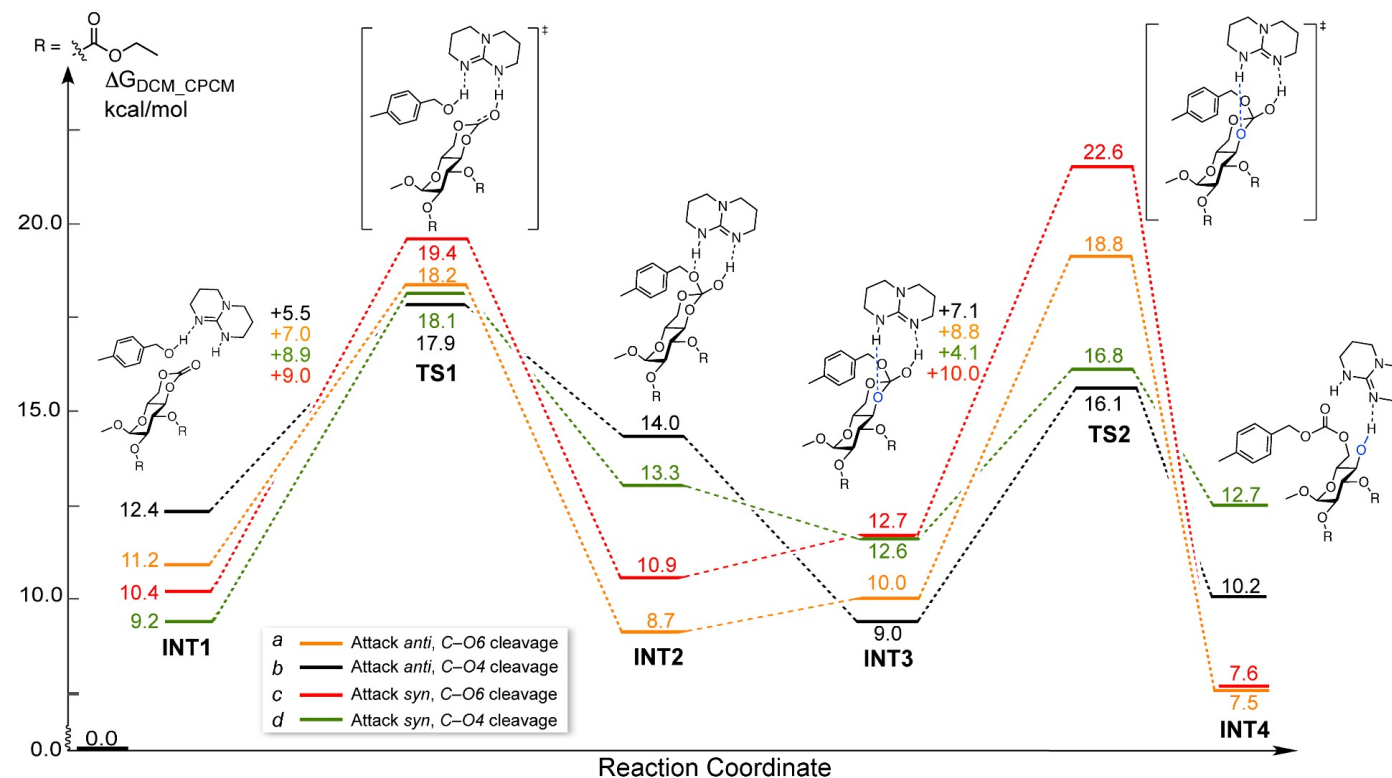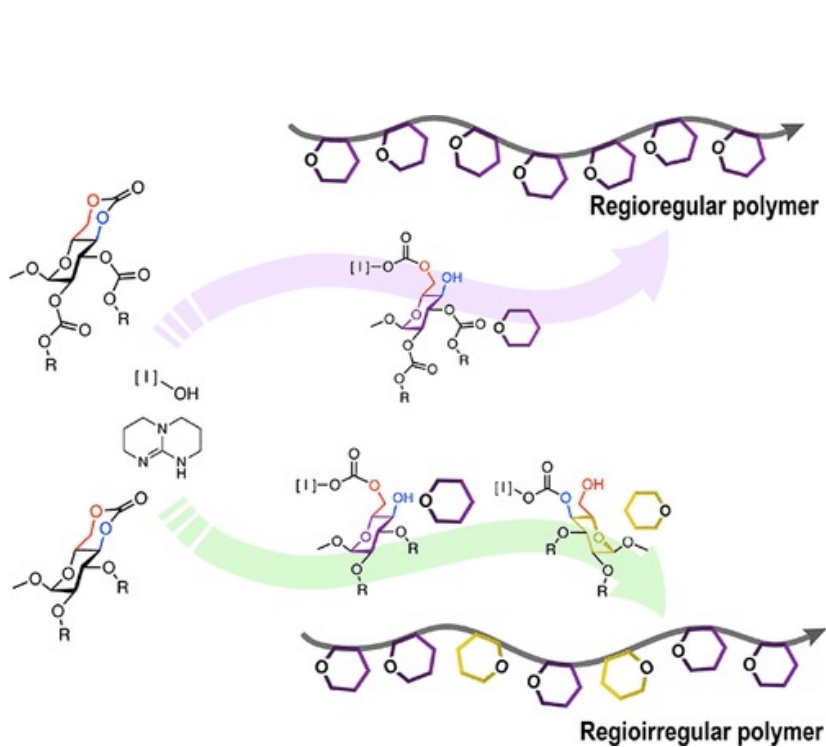
IR Spectra

- Calculated by default for frequency calculations in ORCA and most QM codes.
- Only fundamental transitions are predicted, so no overtones and combination bands are included.
- Correct frequencies with a scaling factor

Water

Infrared Spectrum

Experiment

DFT Calculation

1638.14	100.000
3786.79	8.194
3882.70	46.854
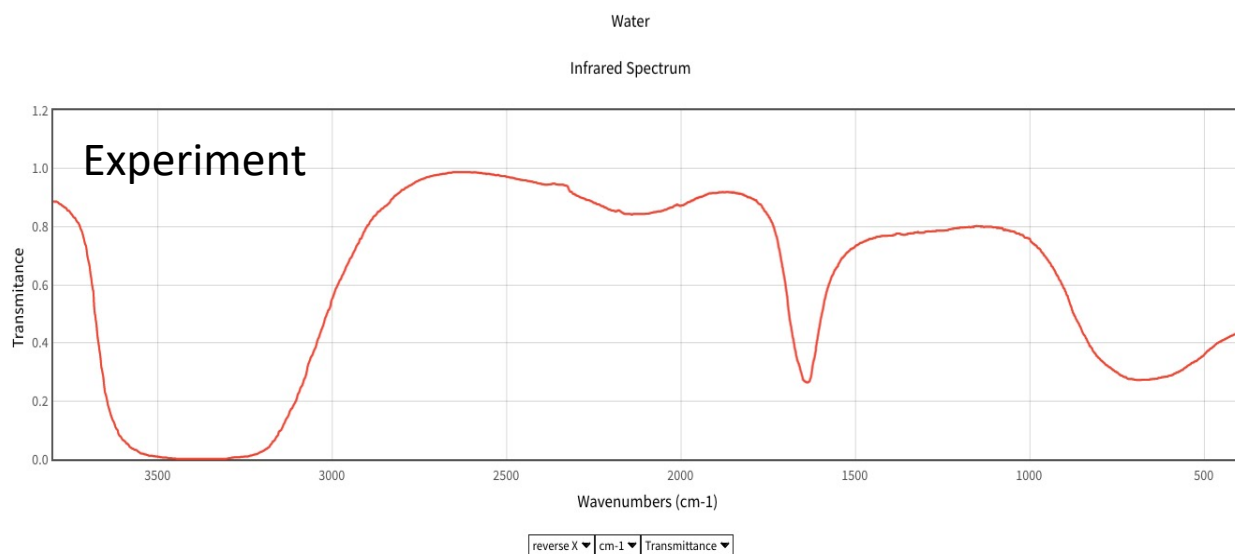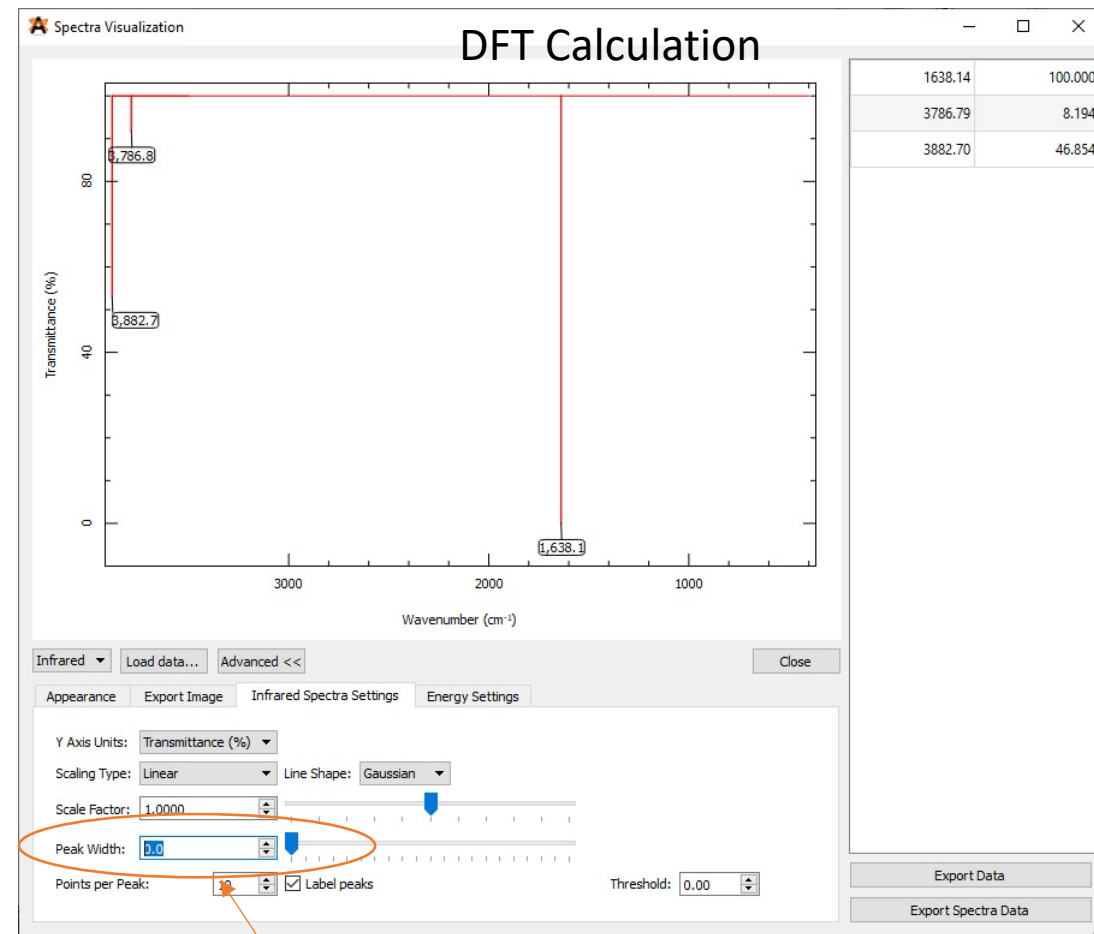
By default, only thin lines corresponding to each frequency are printed. In order to make your predicted spectra look more like an experimental one, some **line broadening** is needed.

UV/Vis Spectra via TD-DFT

a)   1Z   1E   b)   1Z   1E
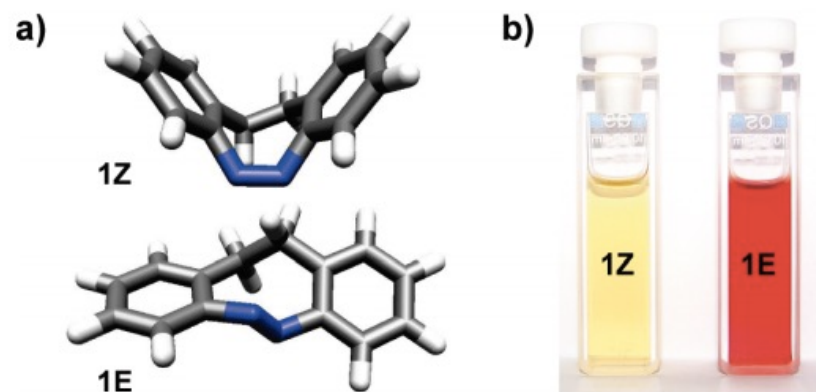

TD-DFT predicted Spectra
— Z isomer
— E isomer
exp. 490 nm
cal. 503 nm
exp. 404 nm
cal. 405 nm
ABS Intensity (arb. units)
Wavelength (nm)

TD-DFT Calculation for **1E**

STATE 1: E= 0.112459 au 3.060 eV   24682.0 cm**-1
50a -> 55a : 0.056581 (c= 0.23786860)
54a -> 55a : 0.913932 (c= -0.95599797)
54a -> 59a : 0.010570 (c= 0.10281285)

State 1 is composed of about 91% a HOMO to LUMO transition, from orbital 54 to orbital 55 (the "a" there means alpha orbital).

HOMO

LUMO

orb #55

orb #56
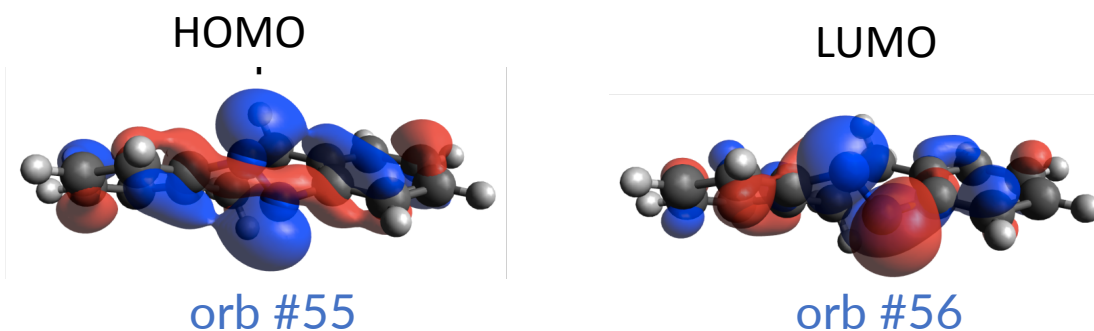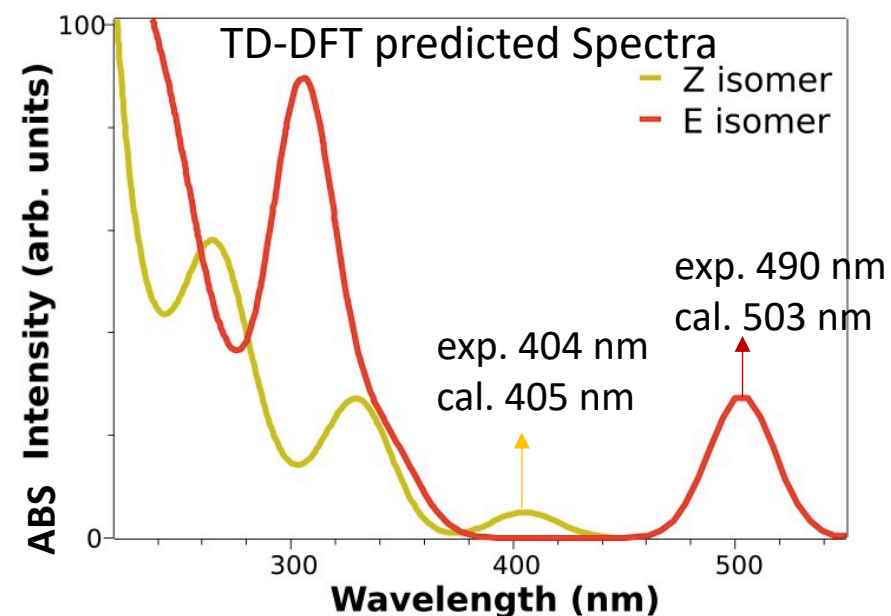
!!! Avogadro counts orbitals starting from 1, but ORCA starts counting from 0. So orbital 54 in ORCA will be orbital 55 in Avogadro.

NMR Spectra

- magnetic shielding constants (or the chemical shifts δ)
- spin–spin coupling constants

nmrdb NMR predict — theory, conformer 2 — theory, full CRE —
theory, one rotamer — theory, conformer 1 — exptl., inverted —

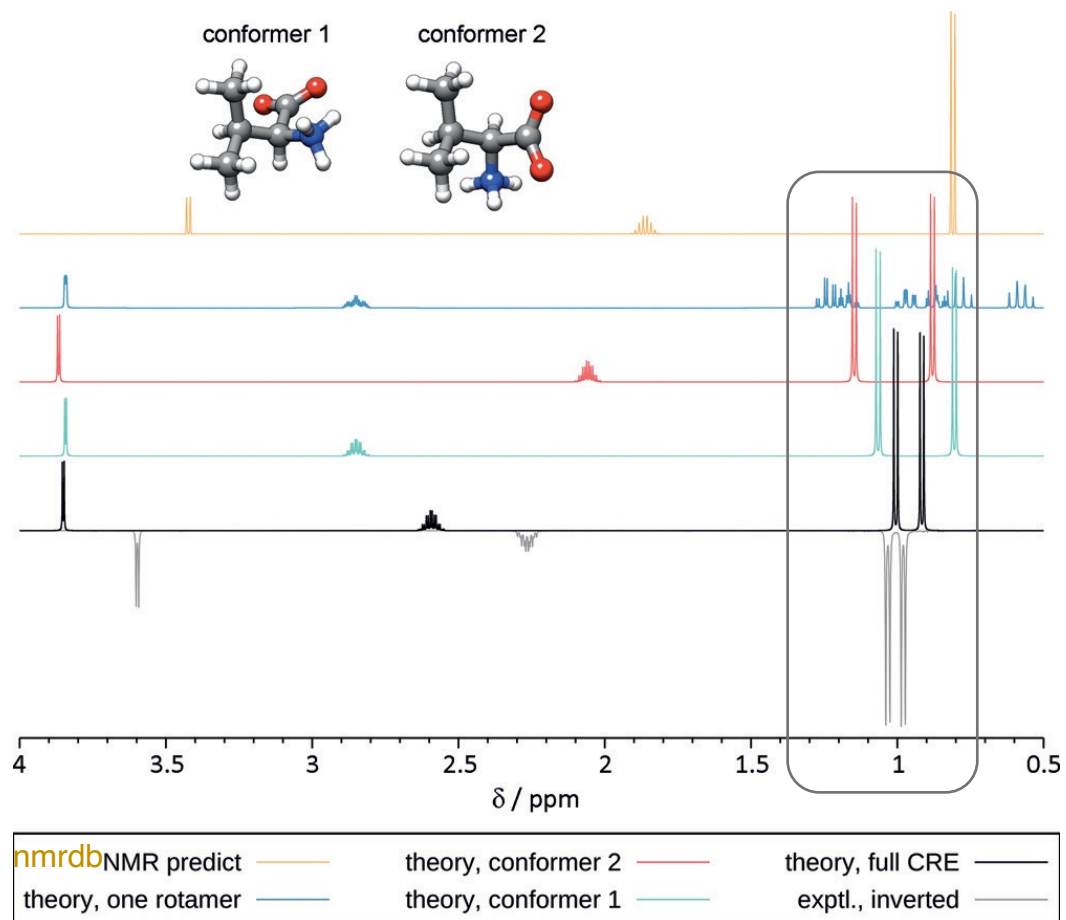! NMR shifts are quite dependent on the conformer you choose. A conformer search or even a Boltzmann weighting might be necessary!

Generate a library of conformers

(Conformational search, eg. MacroModel)

Determine optimal geometries, free energies, chemical shift δ, and J for each conformer

(DFT, eg. ORCA)

Calculate Boltzmann-weighted chemical shifts and coupling constant

Nat Protoc **9,** 643–660 (2014).
Angew.Chem. Int. Ed. **56**, 14763 −14769 (2017).

QM Simulation Software

Name	Description	Price*
ABINIT	QM (Molecular and Periodic Systems)	Free
ADF	QM (Slater orbitals)	$
AMPAC	QM (Semi-empirical)	$
GAMESS-US	QM	Free
Gaussian	QM	$
MOLPRO	QM (specializing in high-level calculations)	$$
NBO	Wavefunction analysis program	$
NWChem	QM	Free
ORCA	**QM specializing in spectroscopic properties**	Free
Quantum ESPRESSO	QM solid state and surfaces	Free
SIESTA	QM specializing in electron transport and Solids	Free
VASP	QM specializing in QMD and ultra-soft ECPs	$$

* Academic Pricing

ORCA

ORCA is an ab initio, DFT, and semi-empirical SCF-MO package developed by Frank Neese *et al.* at the Max Planck Institut für Kohlenforschung.
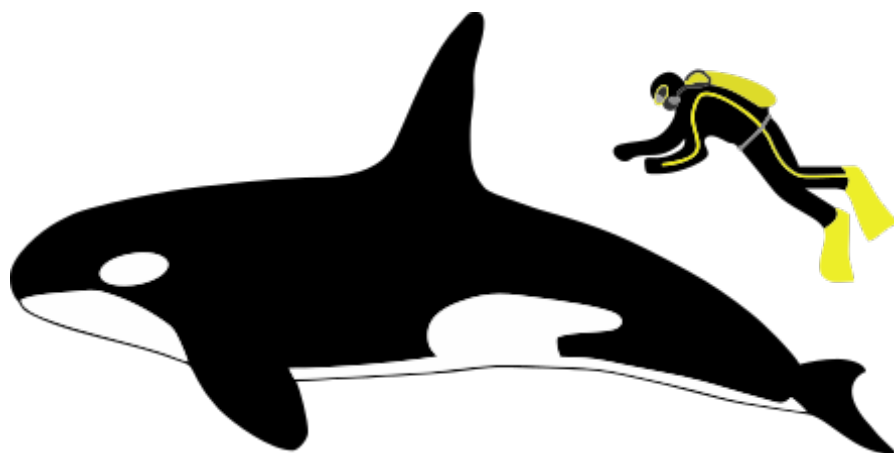
[ORCA Forum](#)
[ORCA Tutorial](#)

- Robust
- Free and open-source
- Cross platforms
- Runs parallel
- Easy to install
- Active in development
- …

- Runs completely from command line
- Needs third-party visualization tools

Neese, F. (**2012**) The ORCA program system, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2, 73-78.
Neese, F. (**2017**) Software update: the ORCA program system, version 4.0, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 8, e1327.
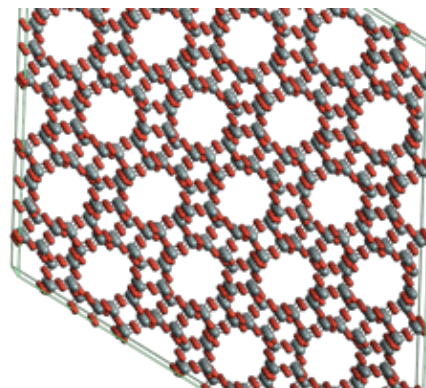
Avogadro

http://avogadro.cc/

Avogadro is a free, open source molecular editor and visualization tool, designed for use on Mac, Windows, and Linux in computational chemistry, molecular modeling, bioinformatics, materials science, and related areas. It offers flexible high quality rendering and a powerful plugin architecture.

Figures from http://avogadro.cc/
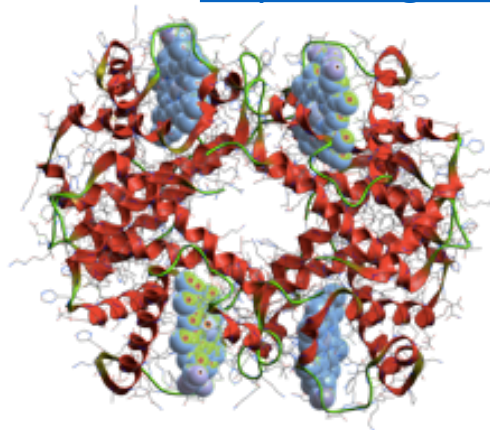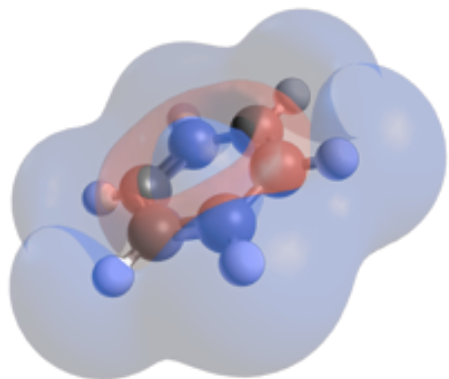
Cross-Platform
Free, Open Source
International
Intuitive
Fast
Extensible
Flexible

Marcus D Hanwell, Donald E Curtis, David C Lonie, Tim Vandermeersch, Eva Zurek and Geoffrey R Hutchison; "Avogadro: An advanced semantic chemical editor, visualization, and analysis platform" *Journal of Cheminformatics* **2012**, 4:17.

Hands-on Session #1 – 15 minutes

- In this hands-on session, you will run an ORCA calculation using pre-prepared input.

- Login to the Vidi Portal: https://vidiportal.chem.tamu.edu/

- Open a terminal and load the appropriate modules

- Take a look at the format of the h2.inp file

- Use the command line to submit an orca job

Getting Started with QM simulations

What do we need?

- Atomic coordinates
- Symmetry (Optional)
- Charge
- Multiplicity

- Level of Theory (Methods)
- Basis Set
- Job Type: single point, geometry optimization, frequency calculation

ORCA Sample Input

! B3LYP Def2-SVP OPT FREQ

*xyz 0 1
O   -5.47612   3.94470   0.00000
H   -4.50612   3.94470   0.00000
H   -5.79945   3.48041   -0.78791
*

Atomic Coordinates

- ## All atoms including hydrogens

  - ### cartesian coordinates
  - ### internal coordinates
    - z-matrix
    - Internal redundant coordinates

- ## File formats: .xyz, .mol, .pdb, *etc.*
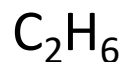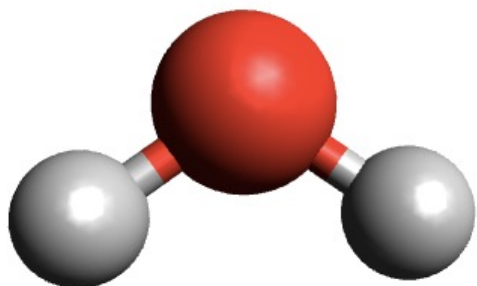
C_2H_6
D$_{3d}$
Charge = 0
Multiplicity = 1

ORCA Simple Input: XYZ coordinates in Å

```
*xyz 0 1
O    -5.47612       3.94470        0.00000
H    -4.50612       3.94470        0.00000
H    -5.79945       3.48041       -0.78791
*
```

ORCA Simple Input: Z-matrix format

```
*gzmt 0 1
O
H    1    0.97000
H    1    0.97000     2     109.47100
*
```

Atomic Coordinates

- Atomic (nuclear) coordinates are required for all atomistic calculations.

- A reasonably good starting geometry is required for electronic structure calculations.
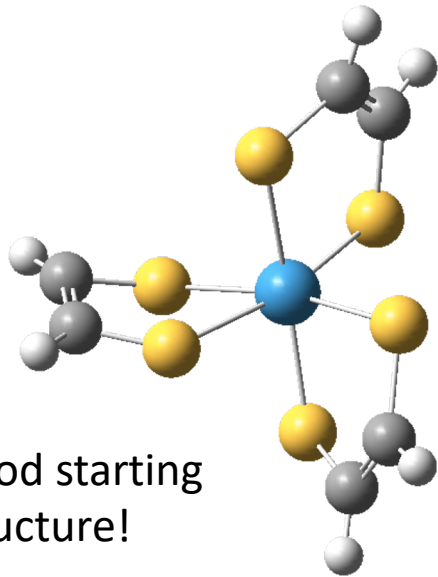
A reasonable starting geometry is required for electronic structure calculations.

Good starting structure!
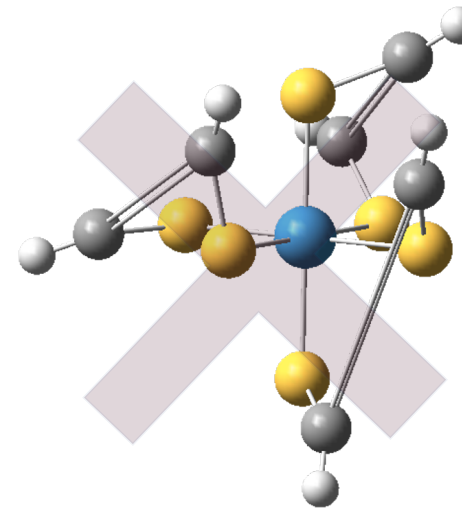
Bad starting structure!

Will most likely fail with an SCF convergence failure or possibly optimize to an excited state.

Charge, Spin, and Multiplicity

- Charge (overall charge of the molecular system)

- Spin State – How many unpaired electrons?

- Spin Multiplicity = 2S+1 where S is the total Spin

$S=\Sigma m_s$ $m_s = \frac{1}{2}$ spin up, alpha $m_s = -\frac{1}{2}$ spin down, beta

Number of unpaired e⁻	Spin	Spin Multiplicity	Quantum State
0	S=0	1	Singlet
1	S=1/2	2	Doublet
2	S=1	3	Triplet
3	S=3/2	4	Quartet
4	S=2	5	Quintet

Charge, Spin, and Multiplicity

EXAMPLE 1: H_2O

      neutral, charge = 0

      no unpaired e^-, S=0, Multiplicity = 2S+1 =1

```
*xyz 0 1
O    -5.47612     3.94470      0.00000
H    -4.50612     3.94470      0.00000
H    -5.79945     3.48041     -0.78791
*
```

EXAMPLE 2: O_2

      neutral, charge = 0

      2 unpaired e^-, S= ½ + ½ = 1, 2S+1 =3 Triplet

```
*int 0 3
O
O    1    1.31600
*
```
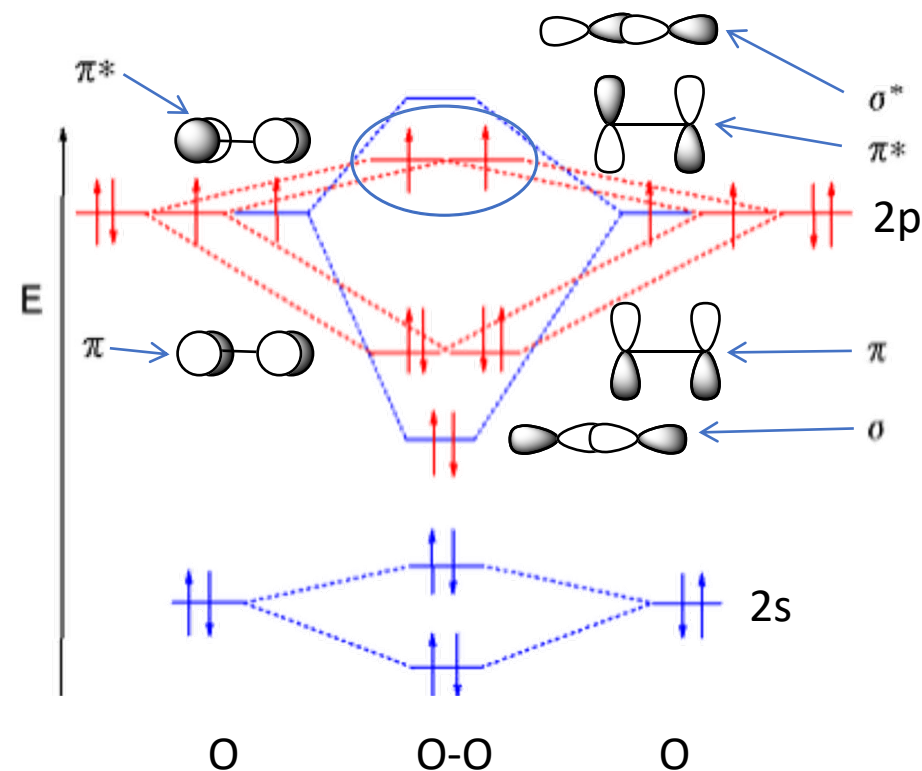

Molecular Oxygen Molecular Orbital Diagram
Triplet is the Ground State

adapted from https://commons.wikimedia.org/wiki/File:MO_diagram_dioxygen.png

Getting Started with QM simulations

- Choose a level of theory
  - The level of theory determines the Hamiltonian (H) used to solve the Schrödinger equation $H\Psi=E\Psi$
  - HF, MP2, MP3, DFT (B3LYP, etc.), CCSD, etc

- Choose a basis Set
  - Basis sets is a set of coefficients and exponents used to describe the atomic orbitals that will be using to describe the wave function
  - 6-31G(d), def2-TZVP, Effective Core Potential (ECP):SDD, LANL2DZ

- Choose a job type
  - Single point (default), geometry optimization, frequency calculations, etc.

Quantum Mechanics – Many e⁻ Hamiltonian

$$\hat{H} = -\sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{r_{AB}}$$

$$\hat{H} = \hat{T}_n + \hat{T}_e + V_{ne} + V_{ee} + V_{nn}$$

$\hat{T}_n$ = Kinetic energy operator for the nuclei

$\hat{T}_e$ = Kinetic energy operator for the electrons

V_{ne} = Coulombic attraction between the elecrons and nuclei

V_{ee} = Coulombic repulsion between the elecrons

V_{nn} = Coulombic repulsion between the nuclei

Quantum Mechanics

$$\hat{H} = \hat{T}_n + \hat{T}_e + V_{ne} + V_{ee} + V_{nn}$$

$$\hat{H} = \hat{H}_n + \hat{H}_e = \hat{H}_n + \hat{H}^o + \hat{H}^1 = \left(\hat{T}_n + V_{nn}\right) + \left(\hat{T}_e + V_{ne}\right) + \left(V_{ee}\right)$$

$$\psi^o = \text{Slater Determinant}$$

$$= \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \overline{\phi_1(1)} & \cdots & \phi_{N/2}(1) & \overline{\phi_{N/2}(1)} \\ \phi_1(2) & \phi_1(2) & \cdots & \phi_{N/2}(2) & \overline{\phi_{N/2}(2)} \\ \vdots & \vdots & & \vdots & \vdots \\ \phi_1(N) & \overline{\phi_1(N)} & \cdots & \overline{\phi_{N/2}(N)} & \overline{\phi_{N/2}(N)} \end{vmatrix}$$

Satisfies indistinguishability of the electrons and antisymmetry requirement of the wavefunction.

Hartree-Fock Theory

$$\hat{H}_e = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\frac{Z}{r_i} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}$$

$$\hat{H}_e = \underbrace{\hat{T}_e + \hat{V}_{ne}}_{\text{Sum of 1e}^- \text{ interactions}} + \hat{V}_{ee} \longleftarrow \text{2e}^- \text{ interactions}$$

N = number of electrons

$\hat{T}_e$ = Kinetic energy operator for the electrons

$\hat{V}_{ne}$ = Coulombic attraction between the elecrons and the nuclei

$\hat{V}_{ee}$ = Coulombic repulsion between the elecrons

Hartree-Fock Theory

Variational $(E_{HF} \geq E_1)$

Size-Extensive $(E_{A \ldots A} = E_A + E_A)$

    Size-extensive:

        $2E_A = E_{A \ldots A}$

        Where, E_A is the energy of molecule A and $E_{A \ldots A}$ is the energy of two molecules of A
        separated by a large distance (i.e. non-interacting)

    Size-intensive:

        $2E_A \neq E_{A \ldots A}$

Neglects instantaneous e^- correlation

Electron Correlation

- Electron Correlation – the motion of electrons are correlated
  - Coulomb hole – the probability of finding an electron of opposite spin near another electron is small ↑↓
  - Fermi hole – the probability of finding an electron with the same spin near another electron is small ↑↑
    - Hartree-Fock theory included electron correlation of electrons of the same spin (Fermi Hole) but does not include electron correlation of electrons of opposite spin.

- How to include electron correlation for electrons of the opposite spin (instantaneous e⁻ correlation)?
  - Post-HF: Møller-Plesset Perturbation theory (MPx (x=2, 3, 4, …)), Configuration Interaction (CI), Coupled-Cluster (CC), ….
  - Density Functional Theory

Ab initio Summary

- HF<MP2<CISD<MP4(SDQ)~CCSD<MP4<CCSD(T)
  - Hartree-Fock does not include dynamic e^- correlation
  - The CCSD(T) method is known as the gold-standard in quantum chemistry and widely used as a benchmark
- PROS:
  - Hierarchy which includes more e^- correlation as you move up the hierarchy
- CONS:
  - Computational cost for post-scf calculations is high
  - The accuracy of post-scf (post-HF) levels of theory are highly dependent on the quality of the basis set

Density Functional Theory

- Functional – A function whose argument is also a function is called a functional.

- 1964 Hohenberg-Kohn Theorem (Kohn (DFT) split the Nobel prize with Pople (QM Methods) in 1998)

  - For $E_o(\rho)$, where ρ is the electron density of the ground state then ρ determines the number of electrons and the potential. Therefore, it also describes the ground state properties.

- Kohn-Sham formalism

Density Functional Theory

$$E_{DFT}\left[\rho(r)\right] = T\left[\rho(r)\right] + V_{ne}\left[\rho(r)\right] + V_{ee}\left[\rho(r)\right] + \Delta T\left[\rho(r)\right] + \Delta V_{ee}\left[\rho(r)\right]$$

$T\left[\rho(r)\right] =$ Kinetic Energy of the non-interacting electrons

$V_{ne}\left[\rho(r)\right] =$ nuclei-electron coulombic attraction

$V_{ee}\left[\rho(r)\right] =$ electron-electron coulombic repulsion

$\Delta T\left[\rho(r)\right] =$ correction to the Kinetic Energy due to the interacting nature of all electrons

$\Delta V_{ee}\left[\rho(r)\right] =$ all non-classical corrections to the electron-electron interactions

$$E_{DFT}\left[\rho(r)\right] =$$

$$\sum_i^N \left(\langle \chi_i | -\frac{1}{2}\nabla^2 | \chi_i \rangle - \langle \chi_i | \sum_k^A \frac{Z_A}{|r_i - r_k|} | \chi_i \rangle + \langle \chi_i | \frac{1}{2}\int \frac{\rho(r')}{|r_i - r'|}dr' | \chi_i \rangle \right) + E_{xc}\left[\rho(r)\right]$$

$$E_{xc} = \Delta T\left[\rho(r)\right] + \Delta V_{ee}\left[\rho(r)\right]$$

note : E_{xc} is commonly called the exchange-correlation energy, but remember that it contains a correction to the non-interacting electron kinetic energy term.

Implementations of DFT

Which functional should I choose? It depends!

https://www.chem.uci.edu/~kieron/dft/pubs/RCFB08.pdf
https://pubs.acs.org/doi/10.1021/ar700111a

QUANTUM MECHANICAL HEAVEN

$H\Psi = E\Psi$

Dispersion Corrected Functionals

- • Double-Hybrid functionals: B2PLYP, B2K-PLYP
- • Range-Separated and Dispersion Corrected: wB97X-D3, CAM-B3LYP

$E(\rho, \nabla\rho, \nabla^2\rho) + HF$ Hyper-GGA

- • Hybrid meta-GGA: TPSSh, PW6B95, M06, M06-2X

$E(\rho, \nabla\rho, \nabla^2\rho)$ Meta-GGA

- • Meta-Generalized Gradient Approximation: TPSS, M06L
- • Hybrid GGA: B3LYP, BHLYP, PBE0

$E(\rho, \nabla\rho)$ GGA

- • Generalized Gradient Approximation: PBE, BP86, BLYP

$E(\rho)$ LDA

- • Local Density Approximation: SVWN

Dispersion Correction

A large vdW complex as an example: the C_{60} dimer
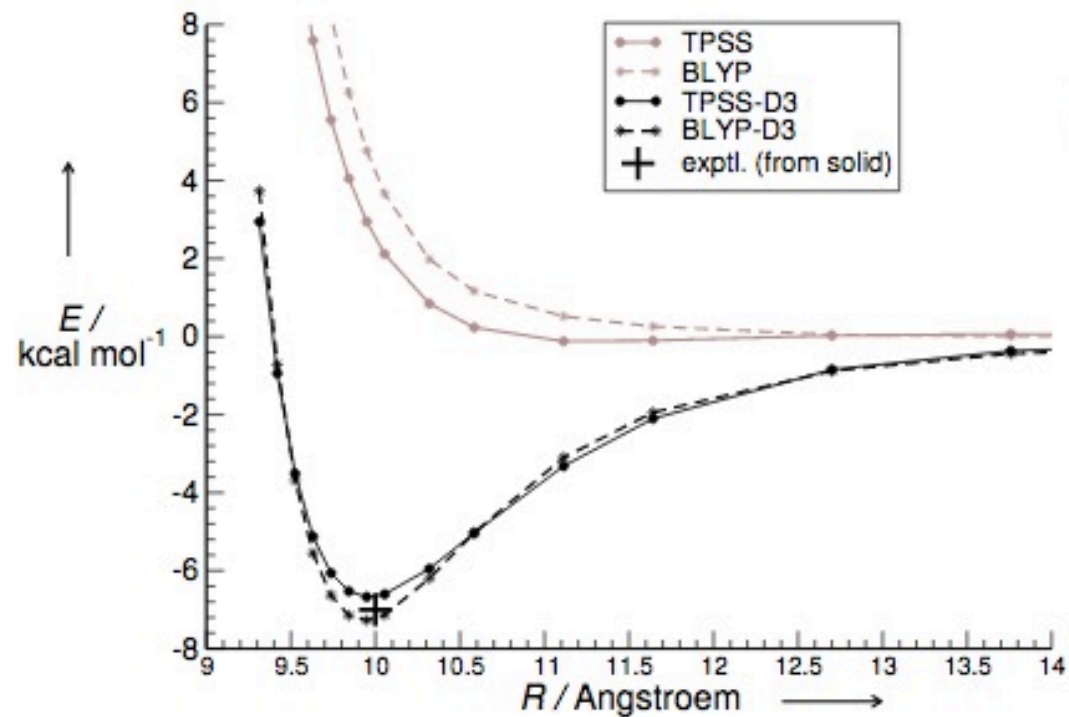
Dispersion Correction

- DFT calculations with atom-wise dispersion correction
  - **D3BJ** Atom-pairwise dispersion correction to the DFT energy with Becke-Johnson damping (recommended)
  - **D3ZERO** Atom-pairwise dispersion correction with zero damping
  - **D2** Empirical dispersion correction from 2006 (not recommended)
  - applicable to many established functionals (e.g. BLYP, BP86, PBE, TPSS, B3LYP, B2PLYP, …)
- Range-separated functionals(some with dispersion)
  - LC-BLYP
  - wB97
  - wB97X
  - wB97X-D3
  - wB97X-V, wB97X-D3BJ
  - wB97M-V, wB97M-D3BJ
  - CAM-B3LYP
- DFT calculations with the non-local, density dependent dispersion correction: DFT-NL
  - compute the Non-Local (NL) dispersion contribution only the electron density, eg. **VV10**

We recommend the DFT-D3 dispersion correction in general and the DFT-NL method
for checking purposes by single-point calculations. ---ORCA Manual

Density Functional Theory

- In theory, DFT is:
    - *ab initio*
    - variational
    - size-extensive
- The exact functional form is not known, therefore, an approximate functional that has been parameterized to fit experimental data is used. The functionals used today for DFT calculations are approximate and parameterized.
- Currently implemented DFT is:
    - not strictly *ab initio*
    - not variational
    - size-extensive

Basis Set Terminology

- MO's built from AO's:

Linear Combination of Atomic Orbitals (LCAO)

- 3 GTO's are needed to adequately describe 1 STO

  - Single-ζ (SZ) Each AO is described by 3 GTO

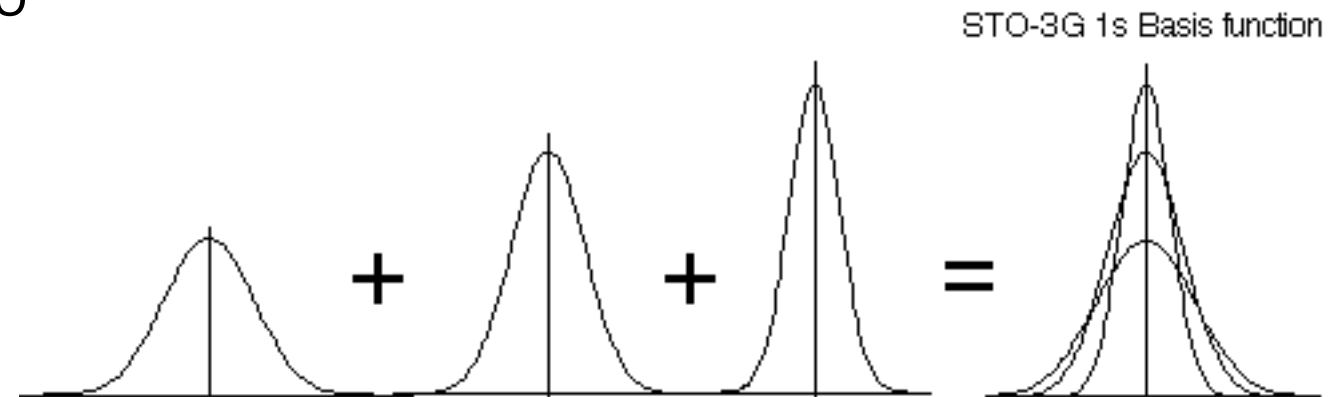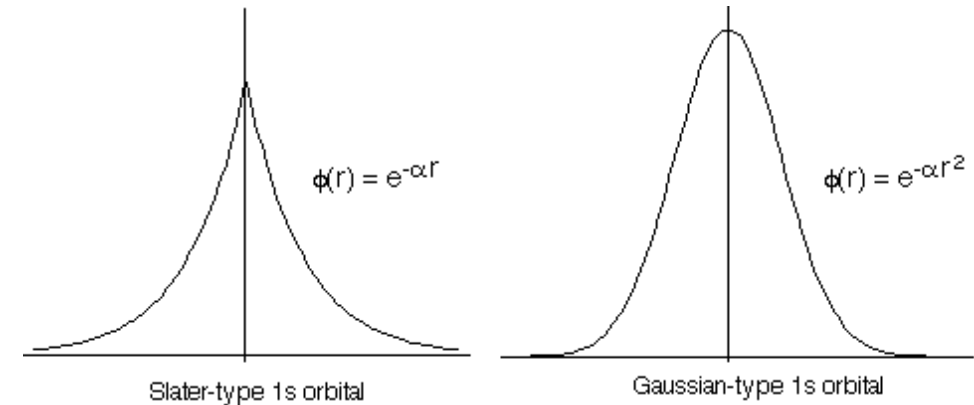   - EXAMPLE: STO-3G (minimal basis set, rarely used)

  - Double-ζ (DZ) Each AO is described by 6 GTO

    - EXAMPLE: 6-31G, cc-pVDZ, …

  - Triple-ζ (TZ) Each AO is described by 9 GTO

    - EXAMPLE: 6-311G, def2-TZVP, cc-pVTZ, …

  - etc

$\phi(r) = e^{-\alpha r}$

$\phi(r) = e^{-\alpha r^2}$

Slater-type 1s orbital

Gaussian-type 1s orbital

STO-3G 1s Basis function

\+ \+ =

Minimal Basis Set: STO-3G

$$\chi_i = \sum_{s=1}^{b} C_{si}\phi_s \quad where,$$

$$\phi_s = Nr^{(2n-2-l)}e^{-\varsigma r^2}Y_l^m(\theta,\phi)$$

\# Basis set for element : H
NewGTO H

S 3	**Exponent - ζ**		**Coefficient C**
1	3.4252509100	ζ_{a1}	0.1543289707 c_{a1}
2	0.6239137300	ζ_{a2}	0.5353281424 c_{a2}
3	0.1688554000	ζ_{a3}	0.4446345420 c_{a3}

end;

\# Basis set for element : C
NewGTO C

S 3	**Exponent - ζ**		**Coefficient C**
A	1	71.6168370000 ζ_{a1}	0.1543289700 c_{a1}
	2	13.0450960000 ζ_{a2}	0.5353281400 c_{a2}
	3	3.5305122000 ζ_{a3}	0.4446345400 c_{a3}

S 3

B	1	2.9412494000 ζ_{b1}	-0.0999672301 c_{b1}
	2	0.6834831000 ζ_{b2}	0.3995128303 c_{b2}
	3	0.2222899000 ζ_{b3}	0.7001154705 c_{b3}

P 3

C D E	1	2.9412494000 ζ_{c1} ζ_{d1} ζ_{e1}	0.1559162721 c_{c1} c_{d1} c_{e1}
	2	0.6834831000 ζ_{c2} ζ_{d2} ζ_{e2}	0.6076837282 c_{c2} c_{d2} c_{e2}
	3	0.2222899000 ζ_{c3} ζ_{d3} ζ_{e3}	0.3919573953 c_{c3} c_{d3} c_{e3}

end; (P_x, P_y, P_z) (P_x, P_y, P_z)

H atom, 1s orbital

1 basis functions
3 primitive gaussians

$$\sum_{i=1}^{3} c_{ai}e^{-\varsigma_{ai}r^2}$$

0.1543289707 e$^{-3.4252509100r^2}$ +
0.5353281424 e$^{-0.6239137300r^2}$ +
0.4446345422 e$^{-0.1688554000r^2}$

C atom 1s orbital
2s orbital
$2p_x, 2p_y, 2p_z$

5 basis functions
15 primitive gaussians

$$A\sum_{i=1}^{3} c_{ai}e^{-\varsigma_{ai}r^2} + B\sum_{i=1}^{3} c_{bi}e^{-\varsigma_{bi}r^2}$$

$$+C\sum_{i=1}^{3} c_{ci}e^{-\varsigma_{ci}r^2} + D\sum_{i=1}^{3} c_{di}e^{-\varsigma_{di}r^2} + E\sum_{i=1}^{3} c_{ei}e^{-\varsigma_{ei}r^2}$$

Quiz: Basis Set

What is the minimal number of basis functions (STO-3G) that you need to describe a methane (CH_4) molecule? How many primitive Gaussian functions?

C atom	1s orbital	
	2s orbital	5 basis functions
	$2p_x$, $2p_y$, $2p_z$	15 primitive gaussians

H atom, 1s orbital	1 basis functions	
	3 primitive gaussians	$\times 4$

CH_4	9 basis functions
	27 primitive gaussians

Polarization and Diffuse Functions

- Polarization Function – Higher angular momentum function than occupied. For example, a d function added to the C atom basis set would be a polarization function since C only has S and P orbitals occupied. e.g. 6-31G(**d**)
  - Important to allow the molecular orbitals to change shape.
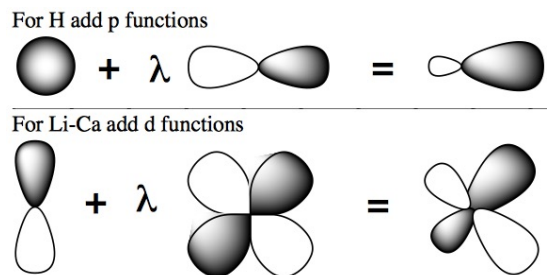
- Diffuse Function – Functions with small exponents (ζ) e.g. 6-31**+**G
  - Important for anions, NMR, excited states, and any calculation where a good description of the wavefunction far from the nucleus is important.

Effective Core Potentials

Assume that the core e⁻ are not important for reactivity and chemical properties. Core e⁻ are not treated explicitly:

$$U_{ECP}(r) = \sum_i a_i r^{ni} e^{-\alpha_i r^2}$$

ECP's normally include the contraction of the inner orbitals due to relativistic effects

Large Core ECP for Pb has 78 e⁻ in the core and 4 e⁻ in the valence
   Core: 1s 2s 3s 4s 5s 2p 3p 4p 5p 3d 4d 5d and 4f
   Valence: 6s and 6p

Medium Core ECP for Pb has 68 e⁻ in the core and 14 e⁻ in the valence
   Core: 1s 2s 3s 4s 5s 2p 3p 4p 5p 3d 4d and 4f
   Valence: 6s 6p and 5d

Small Core ECP for Pb has 60 e⁻ in the core and 22 e⁻ in the valence
   Core: 1s 2s 3s 4s 2p 3p 4p 3d 4d and 4f
   Valence: 5s 6s 5p 6p and 5d

Common Job Types

Approximately 90% of all Calculations!

- Single-Point Energy Calculations (default)

- Geometry Optimizations (Opt)
  - Equilibrium geometries
  - Transition states (OptTS) and reaction rates

- Frequency Calculations (Freq, NumFreq)
  - Characterization of stationary points
  - Thermodynamic properties – Vibrational spectra

ORCA Sample Input

Job Type

```
! B3LYP Def2-SVP OPT

*xyz 0 1
O     -5.47612      3.94470       0.00000
H     -4.50612      3.94470       0.00000
H     -5.79945      3.48041      -0.78791
*
```

Hands-on Session #2 – 30 minutes come back

Prepare an ORCA input file with Avogadro and run the calculation with ORCA

Visualize calculation results with Molden and Avogadro

- View geometry convergence and bond distance change

- Molecular orbitals

- Normal modes

- IR spectra
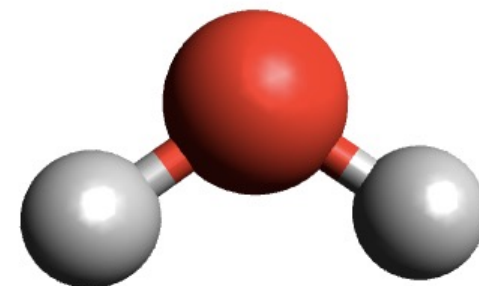
COME BACK AT 11:40

Geometry Optimization

A geometry optimizations minimizes the overall energy of the system by changing the atomic coordinates.
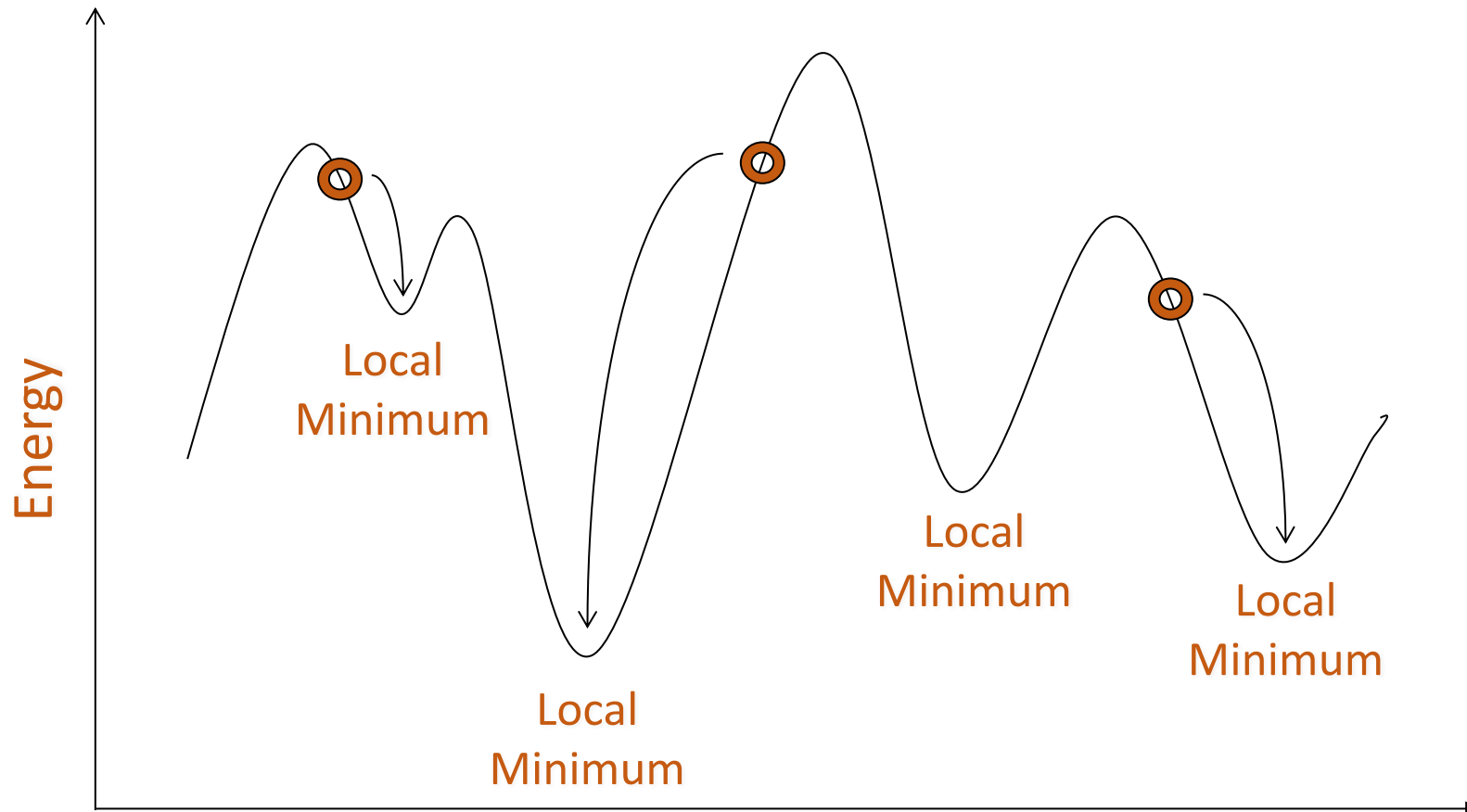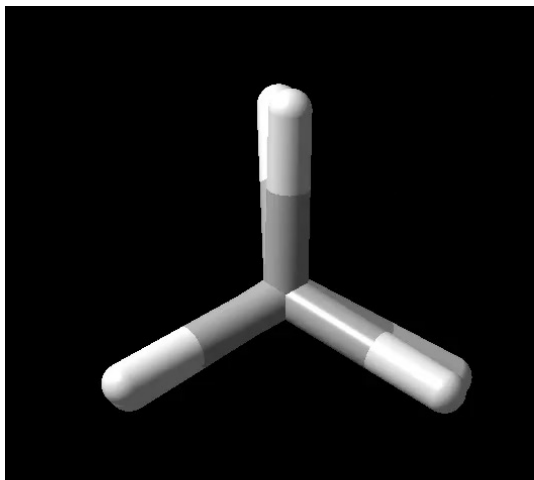

Minimum: All gradients have to vanish

$$f_i = \left. \frac{\partial E(\mathbf{R})}{\partial R_i} \right|_{\mathbf{R}=\mathbf{R}^{(k)}}$$

Problem: Leads to the **nearest** local minimum, not the **global** minimum of the system!

Geometry Optimization

Stationary Points

Imaginary mode 298.1i cm^{-1}
1st order saddle point
transition state
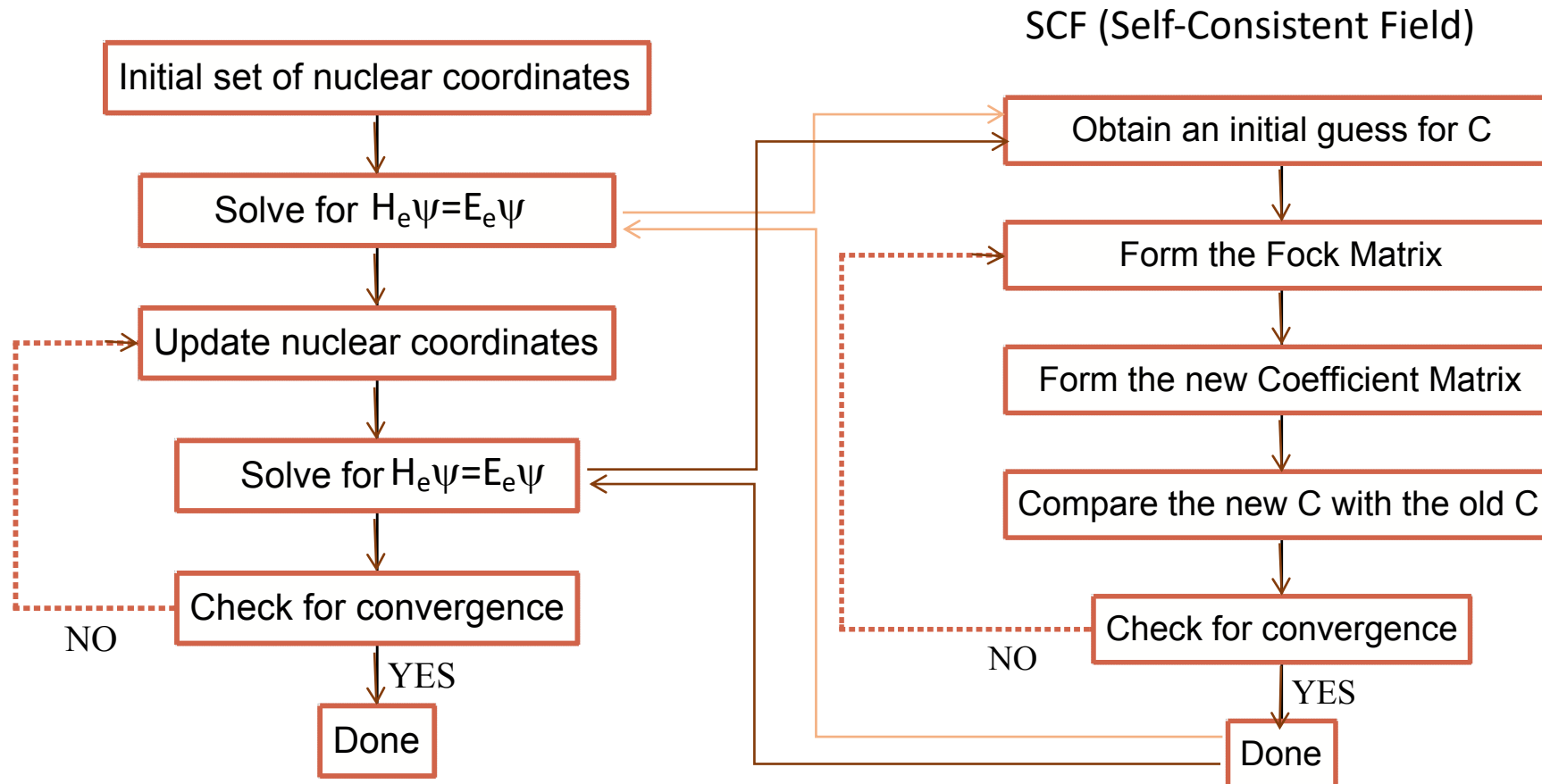
Stationary Point: Gradients with respect to all internal coordinates are zero.

Vibrational modes

All real = minimum
1 imaginary = 1st order saddle point (transition state)
2 imaginary = 2nd order saddle point
3 imaginary = 3rd order saddle point
etc
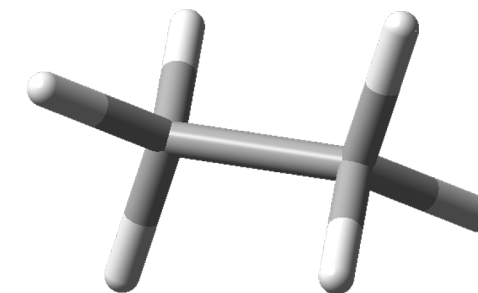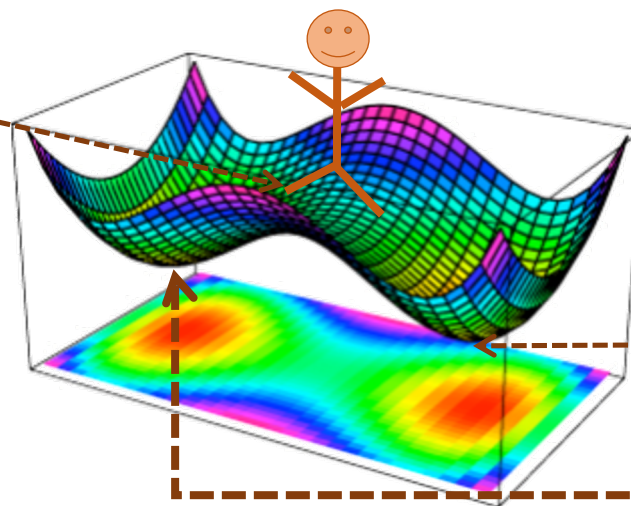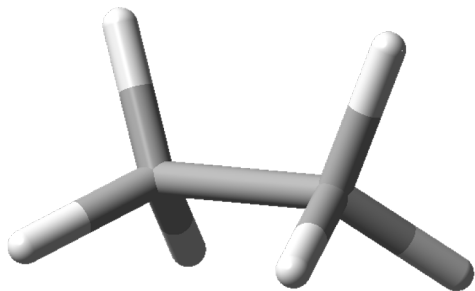
staggered
Minimum (D$_{3d}$)

Frequency Calculations

Why do we need frequency calculations?

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

ORCA Sample Input

analytical (default) or Numfreq

```
! B3LYP def2-TZVP     Opt     Freq

* xyz 0 1
  C      -7.89820      1.60729      0.00000
  C      -6.38418      1.60729     -0.00000
  H      -8.28648      2.56918      0.34866
  H      -8.28648      0.82440      0.65869
  H      -8.28648      1.42830     -1.00735
  H      -5.99591      2.56917      0.34866
  H      -5.99591      1.42830     -1.00735
  H      -5.99590      0.82440      0.65868
*
```

Calculation of Vibrational Frequencies

- The vibrations of a molecule are given by its normal mode.
- Linear molecules have 3N-5 normal modes.
- Non-linear molecules have 3N-6 normal modes.

Classical treatment of Normal Mode: Harmonic Oscillator Approximation

- Each normal modes acts like a simple harmonic oscillator
- The center of mass does not move
- Normal modes are independent, they do not interact

Calculation of the mass weighted Hessian matrix, i. e. the second derivative of the energy with respect to all coordinates

$$H_{AB} = \frac{\partial^2 E^{(I)}}{\partial X_A \partial X_B}\bigg|_{X_A = X_B = \ldots \bar{X}}$$

Upon diagonalization we get the frequencies of the normal modes:

$$\omega = \frac{1}{2\pi c}\sqrt{\frac{k}{m}}$$

ω wavenumber, k force constant, m reduced mass

Calculation of Vibrational Frequencies
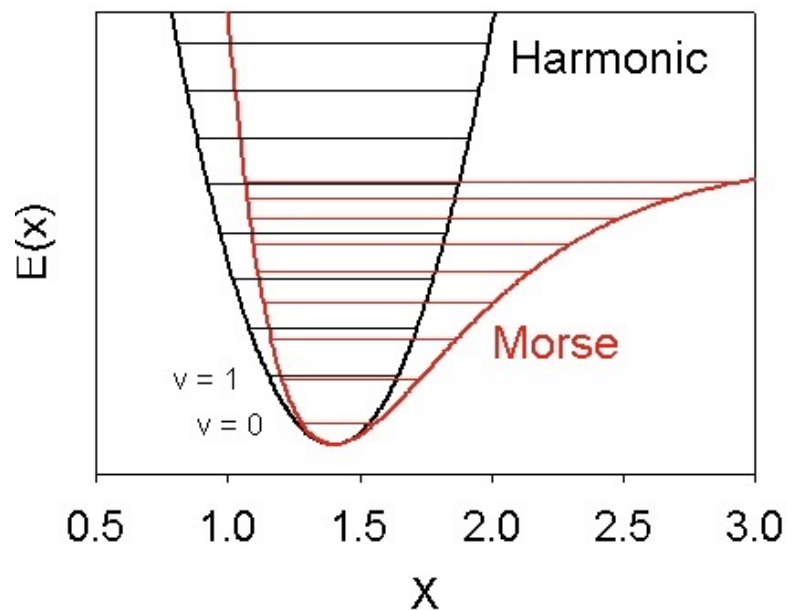
Since the vibrational partition function depends on the frequencies, you must use a structure that is either a **minimum** or a **saddle point**.

Transition State

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

Scaling factor for frequencies =
1.000000000  (already applied!)

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:      -298.14 cm**-1
***imaginary mode***
7:       896.63 cm**-1
8:       897.91 cm**-1
...

...
23:      3105.25 cm**-1
```

Vibrational modes
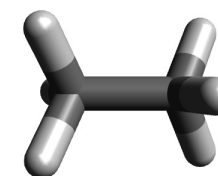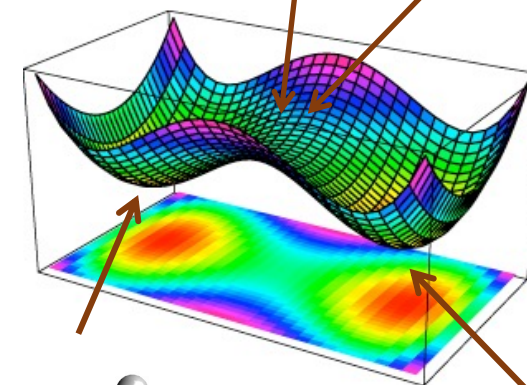All real = minimum
1 Imaginary = 1st order saddle point (TS)
2 Imaginary = 2nd order saddle point
3 Imaginary = 3rd order saddle point
etc
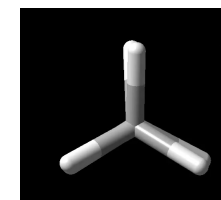
1st order saddle point
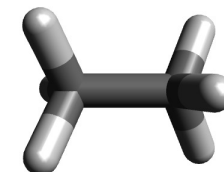
Thermodynamics

Sample ORCA Output

```
-----------
INNER ENERGY
-----------

The inner energy is: U= E(el) + E(ZPE) + E(vib) + E(rot) + E(trans)
E(el) = E(kin-el) + E(nuc-el) + E(el-el) + E(nuc-nuc)
E(ZPE)  - the the zero temperature vibrational energy from the frequency calculation
E(vib)  - correction to E(ZPE) due to population of excited vibrational states
E(rot)  - is the rotational thermal energy
E(trans)- is the translational thermal energy


Summary of contributions to the inner energy U:
Electronic energy                ...     -79.79450251 Eh
Zero point energy  E(0K)...      0.07393626 Eh     46.40 kcal/mol
Thermal vibration            ...       0.00021706 Eh     0.14 kcal/mol
Thermal rotation  E(298K) ...      0.00141627 Eh     0.89 kcal/mol
Thermal translational correction ...    0.00141627 Eh      0.89 kcal/mol
-----------------------------------------------------------------------
Total thermal energy             -79.71751665 Eh
```

```
--------
ENTHALPY
--------

The enthalpy is H = U + kB*T
Thermal Enthalpy correction   0.00094421 Eh
-----------------------------------------------------------
Total Enthalpy                ...     -79.71657244 Eh
-------
ENTROPY
-------
The entropy contributions are T*S =
T*(S(el)+S(vib)+S(rot)+S(trans))
-----------------------------------------------------------
Final entropy term  0.02686916 Eh
-----------------
GIBBS FREE ENERGY
-----------------
The Gibbs free energy is G = H - T*S
-----------------------------------------------------------
Final Gibbs free energy        -79.74344160 Eh
```

Energies

- Absolute energies are not accurate due to the approximations that are made to solve the Schrödinger equation.

- Relative energies have a cancellation of errors to give more accurate results.

- To calculate relative energies, you must use the same Hamiltonian and mathematical description of ψ (basis set):

  - $\Delta H^o_{rxn} = H^o_{prod} - H^o_{react}$, H^o_{prod} and H^o_{react} must be calculated at the same level of theory with the same basis set.
  - Similarly, $\Delta G^o_{rxn} = G^o_{prod} - G^o_{react}$, G^o_{prod} and G^o_{react} must be calculated at the same level of theory with the same basis set.
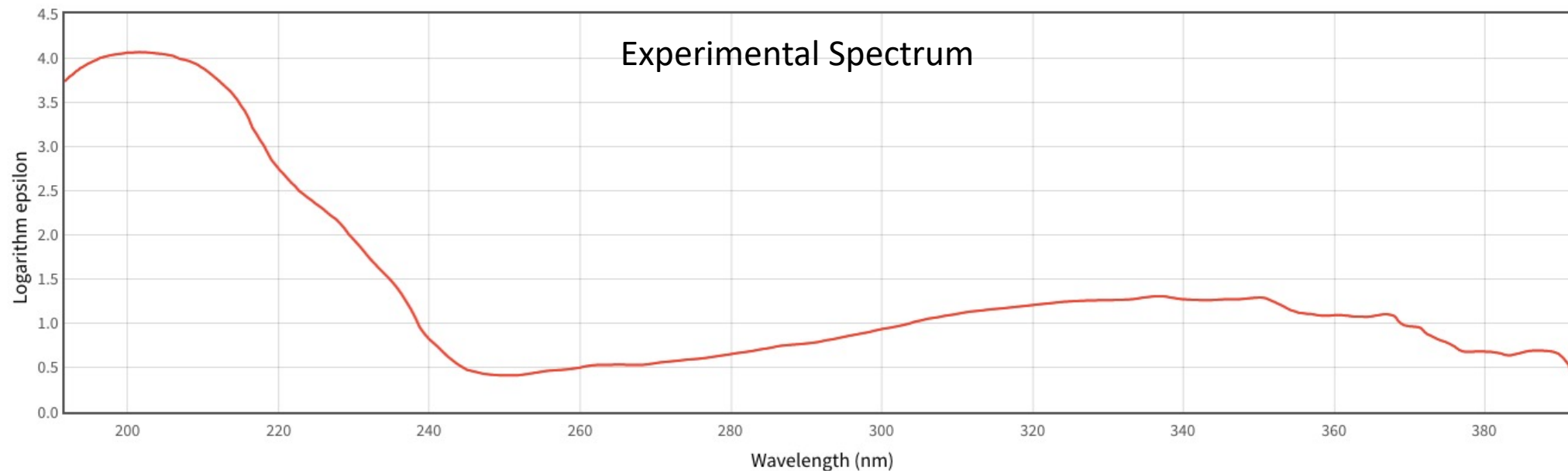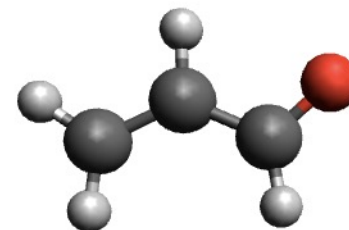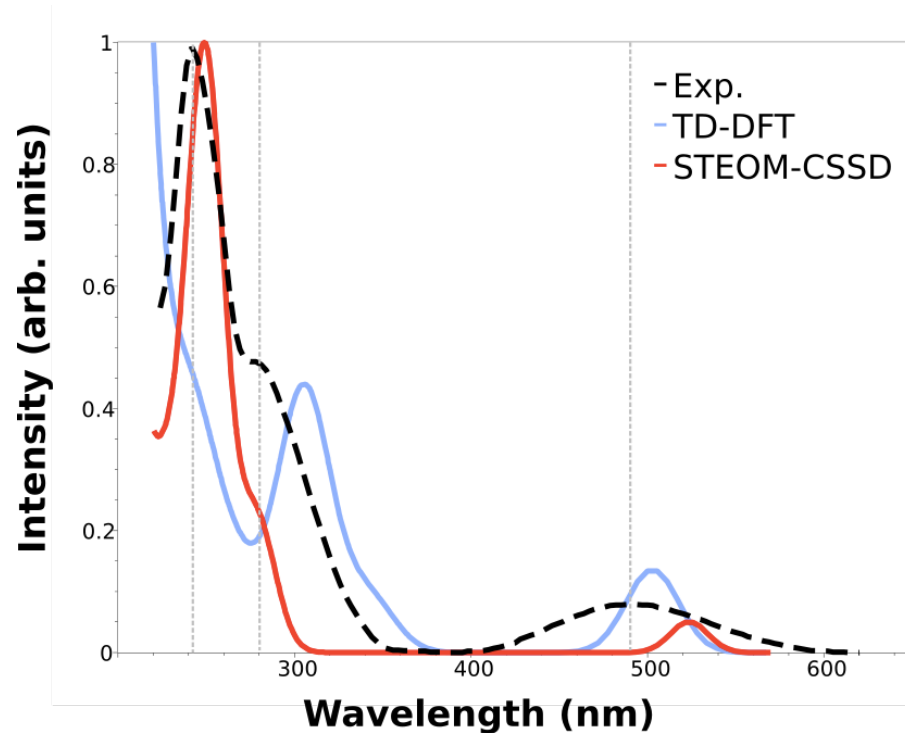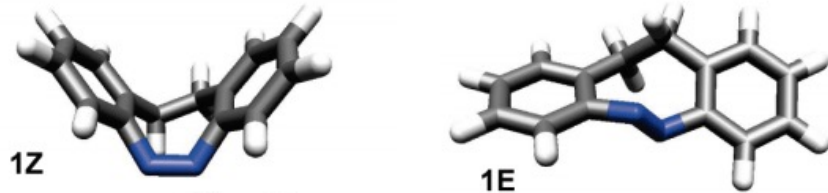
Excited States Calculations

In ORCA, there are several methods that can compute excited state properties with higher or lower accuracy:

- **TD-DFT**
  simpler and widely used
  presents a good speed to accuracy trade-off

- **STEOM-CCSD**
  Similarity Transformed Equation of Motion CCSD
  includes correlation energy into the calculation of the excited states
  At the high-end of excited state methods, expensive

- **STEOM-DLPNO-CCSD**
  Similar cost as TD-DFT with higher accuracy

https://www.orcasoftware.de/tutorials/spec/UVVis.html
J. Am. Chem. Soc. 2009, 131, 43, 15594–15595

Solvation

The Problem:

- Traditional quantum calculations are carried out in the gas-phase

- Most reaction chemistry and/or spectroscopy is performed in solution

- Gas-phase calculations fail to properly describe the energetics of charged and/or neutral polar molecules involved in a reaction that occurs in solution.

Possible Solutions:
- Explicit Solvation - include explicit solvent molecules.
    - Computationally expensive
    - Convergence problems
    - Can overestimate solvent interaction if an inadequate number of solvent molecules are included

- Implicit Solvation - estimate the effect of the solvent molecules using an implicit solvation model.
    - Gives reasonable results for bulk polar solvent effects.
    - Fails if the solvent molecule is directly involved in the reaction chemistry (ie strong H-bonding, coordinated, etc).

Implicit Solvation Model

The solvation energy is then decomposed in two main terms, electrostatic (ΔG_{ENP}) and cavity-dispersion (ΔG_{CDS}):

$$\Delta G^o_{solv} = \Delta G_{ENP} + \Delta G_{CDS}$$

and the methods differ on how to compute these terms.

Conductor-like Polarizable Continuum Model (CPCM)

- The CPCM method [Cossi1998], the bulk solvent is treated as a conductor-like polarizable continuum and the main parameters to define the method are the **refractive index** and the **dielectric constant** of the medium.

- ΔG_{ENP} that arises from the interaction of the medium and the molecular surface charges is included in the SCF calculation - so that you even get "solvated" orbitals - and the cavity term (ΔG_{CDS}) is obtained form a simple linear equation with respect to the cavity area.

`!CPCM(solvent)`

Universal Solvation Model (SMD)

- The SMD method [Truhlar2009] can be thought as an improvement over the CPCM, since it uses the full solute electron density (not just the surface) to compute the cavity-dispersion contribution instead of the area only.

- ORCA has currently 179 solvents available
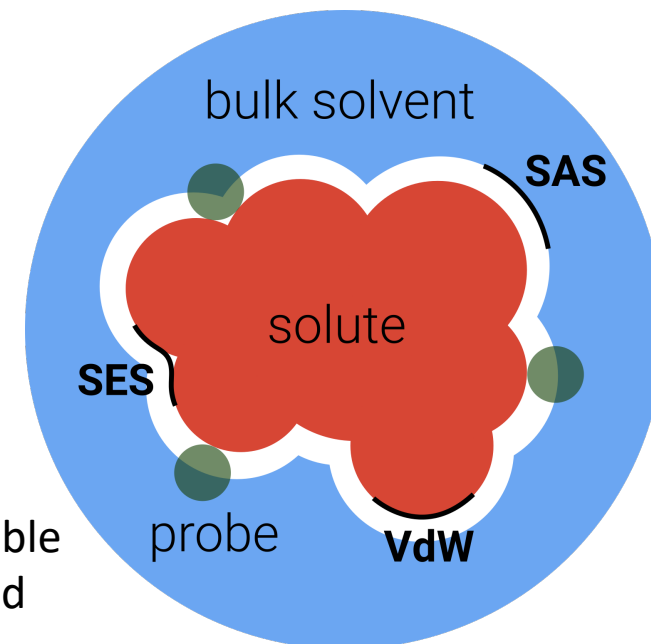
```
%CPCM SMD TRUE
SMDSOLVENT "SOLVENT"
END
```
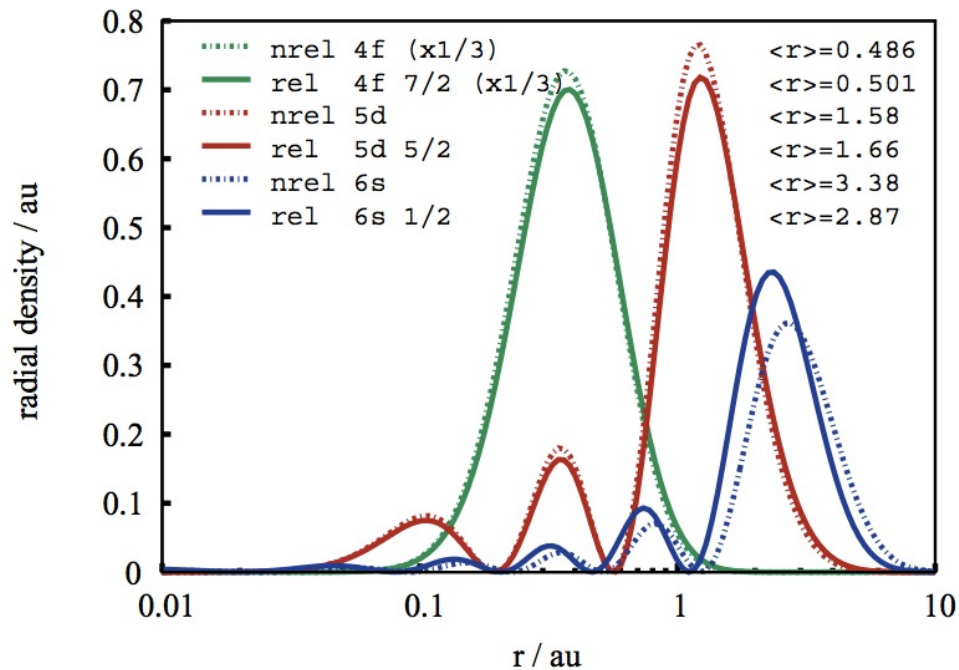
https://www.orcasoftware.de/tutorials/prop/CPCM.html

Relativistic Corrections

When calculating properties or geometries of systems that contain heavy elements (fourth row and beyond), relativistic effects can have a big impact and should not be ignored.

EXAMPLE: Relativistic Effects on selected valence orbitals of gold atom. 6s contracts relativistically.

1) The mass-velocity correction
2) The Darwin term   } scalar relativistic
3) Indirect relativistic effect
4) The spin-orbit coupling

Inclusion of relativistic correction in ORCA

- Relativistic Effective Core Potential (RECPs)
- All-electron Hamiltonian

  (Note that here we are using specific basis for each method, named ZORA- or DKH-DEF2-TZVP.)

  - Zero-Order Regular Approximation (**ZORA**, [Baerends1996])

    `!B3LYP DKH DKH-DEF2-TZVP`

  - Douglas-Kroll-Hess (**DKH**, [Kroll1974] [Hess1985]) Hamiltonian.

    `!B3LYP ZORA ZORA-DEF2-TZVP`

Basis Sets in ORCA

Pople-style basis sets : 6-31G, 6-3111++G**, …

Karlsruhe def2 family: def2-SVP, def2-DZP, …

Ahlrichs basis sets: DZP, TZP, …

Correlation-consistent basis sets: cc-pVDZ, cc-pVTZ, …

Auxiliary basis sets
    for Coulomb fitting: Def2/J, SARC/J
    for simultaneously fitting Coulomb and exchange:
    Def2/JK, cc-pVnZ/JK, aug-cc-pVnZ/JK, …
    for correlation calculations: Def2-TZVP/C, cc-pVnZ/C, …

Relativistically recontracted Karlsruhe basis sets
    DKH-def2-SVP and ZORA-def2-SVP

ORCA allows to treat different atom types or even individual groups in a molecule according to the desired accuracy

```
# CuCl4
! UHF HF
%basis basis "SV"
newGTO Cl "DUNNING-DZP" end
end
* xyz -2 2
Cu   0   0   0   newGTO "TZVPP" end
Cl   2.25   0   0
Cl  -2.25   0   0
Cl   0   2.25   0
Cl   0  -2.25   0
*
```

Blocks start with **%** and end with **end**. In general the input is not case sensitive.

Other Tasks can be performed/interesting features

- Calculation of ground state properties

  – Looking at charge distributions and orbitals

  – 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, ...

Wrapping Up

What do we need?

- Atomic coordinates
- Symmetry (Optional)
- Charge
- Multiplicity

- Level of Theory (Methods)
- Basis Set
- Job Type: single point, geometry optimization, frequency calculation

What can we get?
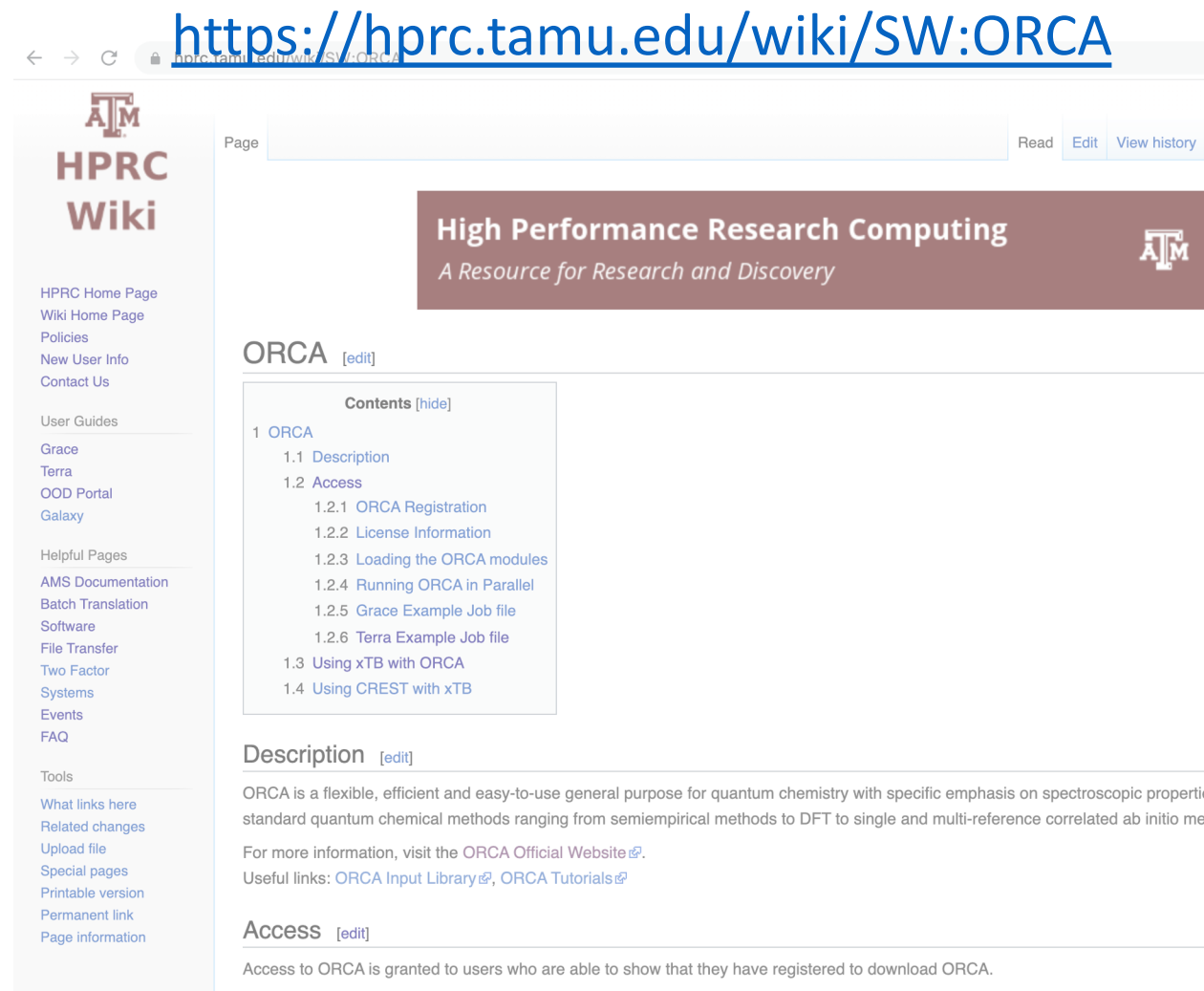
- Molecular Geometry:
  Equilibrium and Transition State
- Dipole and Quadrupole Moments
- Electron Density Map
- Potential Energy surfaces
- Thermochemical data:
  Enthalpy, Entropy, Free Energy
- Vibrational Frequencies
- Spectroscopies:
  IR, Raman, UV/Vis, NMR, CD
- …

Running ORCA on Grace and Terra

General Suggestions

1. Refer to the user guide

2. Create a directory for each calculation

3. Build job files for batch systems

4. Specify the number of cores in input when running ORCA in parallel

Need Help? Contact the HPRC Helpdesk

Website: hprc.tamu.edu
Email: help@hprc.tamu.edu
Telephone: (979) 845-0219

Help us, help you -- we need more info
- Which Cluster (Terra, Grace)
- NetID (NOT your UIN)
- Job id(s) if any
- Location of your jobfile, input/output files
- Application used if any
- Module(s) loaded if any
- Error messages
- Steps you have taken, so we can reproduce the problem