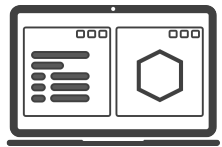
Introduction to Computational Chemistry

Part 3: Preparing structures

Eno Paenurk & Patrick Finkelstein & Felix Pultar



10.05.2023

Webinar overview

- Part 1: Takeoff
- Part 2: Reaction pathways from scratch
- Part 3: Preparing structures
- Part 4: Density Functional Theory
- Part 5: Organometallics
- Part 6: Spectroscopy
- Part 7: Chemical concepts

What you should already know

- Accessing the cluster
- Navigating in the terminal
- Generating SMILES

Goals for today

- Visualizing xyz files
- Geometry optimization: some (light) theory
- Getting decent 3D coordinates fast

.xyz file – the universal file for 3D coordinates

```
1 44
2 Comment line
3 C 5.594185 0.243820 0.071276
4 C 4.736750 1.001947 0.032354
5 N 5.445473 2.151205 0.195846
...
...
46 N 4.190242 4.817675 3.062092
```

Important: there is no information about bonds. Visualization programs use empirical distance-based criteria to show single/multiple bonds.

Visualizing coordinates

- External:

Avogadro

Chemcraft

...

- Open from the terminal:

MOLDEN

Why structure optimization?

What is needed for any QM calculation:

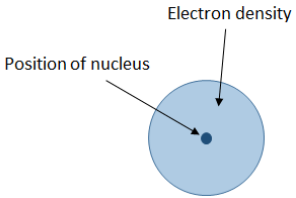
- Number, types, and **positions of the nuclei**
- Number (and location) of electrons and the spin state (multiplicity)

We want to find the positions of the nuclei that minimize the energy – i.e., the most accurate representation of the molecule.

The theoretical minimum: Atoms

$$\hat{H}\Psi = E\Psi$$

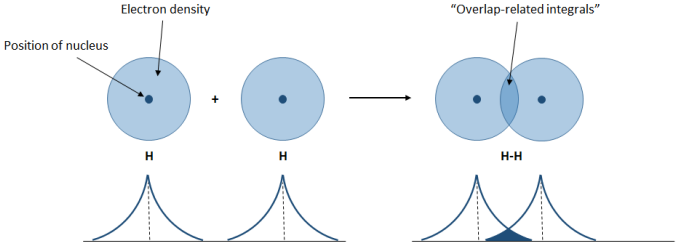
- Energy: depends on the type of nucleus (the element) and number/location of electrons (electron density)
- Electron density: depends on the orbitals (made up from basis functions)



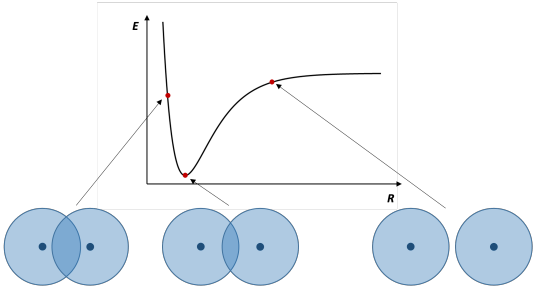
The theoretical minimum: Molecules

$$\hat{H}\Psi = E\Psi$$

- Energy: depends on the position/type of nuclei and electron density
- Position/type of nuclei: xyz coordinates
- Electron density: generated from a combination of basis functions
- **New: Overlap of the basis functions on different centers gives "new integrals"**



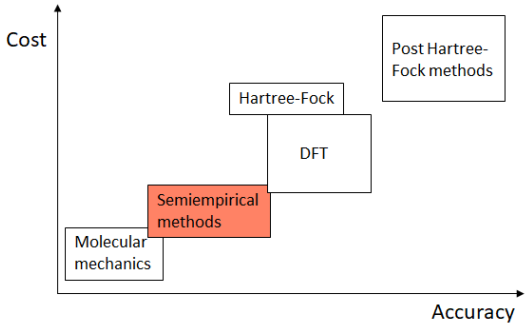
The theoretical minimum: Geometry optimizations



1. Calculate the energy of the starting structure
2. Introduce small change to the position of the nuclei
3. Calculate the energy of new structure
4. Repeat steps 2 and 3 until you reach a minimum

Different levels of theory

One big difference between the levels in the middle is how the integrals are dealt with.

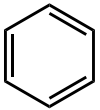


Disclaimer: This is merely a simplified representation

Semiempirical methods

- Trade-off between computational cost and accuracy
- Quantum mechanical foundation
- Simplification of **integrals**
- Use of element-specific parameters

Best known example: Hückel method



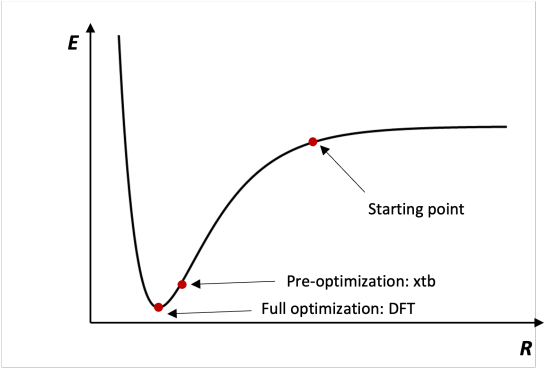
$$H = \begin{pmatrix} \alpha & \beta & 0 & 0 & 0 & \beta \\ \beta & \alpha & \beta & 0 & 0 & 0 \\ 0 & \beta & \alpha & \beta & 0 & 0 \\ 0 & 0 & \beta & \alpha & \beta & 0 \\ 0 & 0 & 0 & \beta & \alpha & \beta \\ \beta & 0 & 0 & 0 & \beta & \alpha \end{pmatrix}$$

Grimme's extended tight-binding semiempirical program package
Initially developed with a focus on

- Geometries
- Vibrational frequencies
- Non-covalent interactions

More information at: `https://xtb-docs.readthedocs.io/en/latest/contents.html`

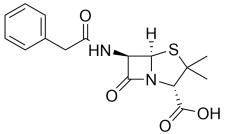
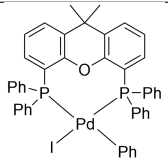
Geometry optimization – a fast strategy



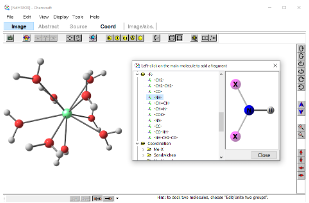
- Optimization with `xtb`: cheap and gives a good starting structure
- Use DFT for the “last mile”

Caveat: sometimes the structure you want is “unstable” in `xtb`.
Solution: skip the xtb pre-optimization.

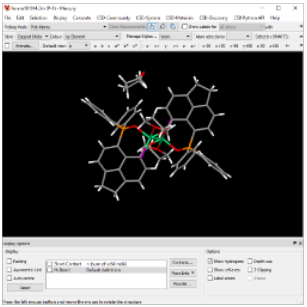
Computational cost of optimization with DFT

Molecule	Unoptimized	xtb-preoptimized
H–H	19 s	17 s
 $C_{16}H_{18}N_2O_4S$	1 h	0.5 h
 $C_{45}H_{37}IOP_2Pd$	17.5 h	3.5 h

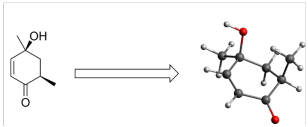
How to get coordinates



Draw in Avogadro/Chemcraft

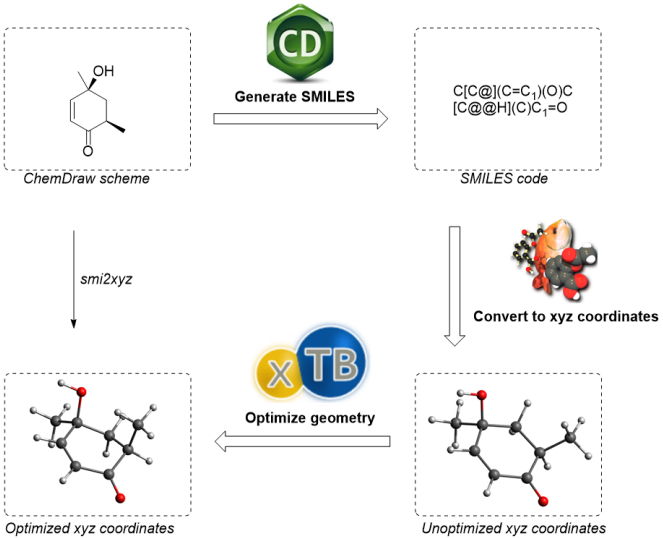


Extract from X-ray crystallography data



Generate from ChemDraw via SMILES

Convert ChemDraw scheme to coordinates



Our workflow

Known organometallic complex:
Extract xyz coordinates from the `cif` file

Unknown organometallic complex:
Either extract xyz coordinates from a similar structure and modify with Avogadro/Chemcraft or use `smi2xyz`

Organic molecule:
Combination of `smi2xyz` and Avogadro/Chemcraft

Live Demo

Exercises