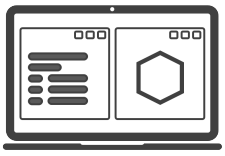


Introduction to Computational Chemistry
Exercises Part 6

Eno Paenurk & Patrick Finkelstein & Felix Pultar

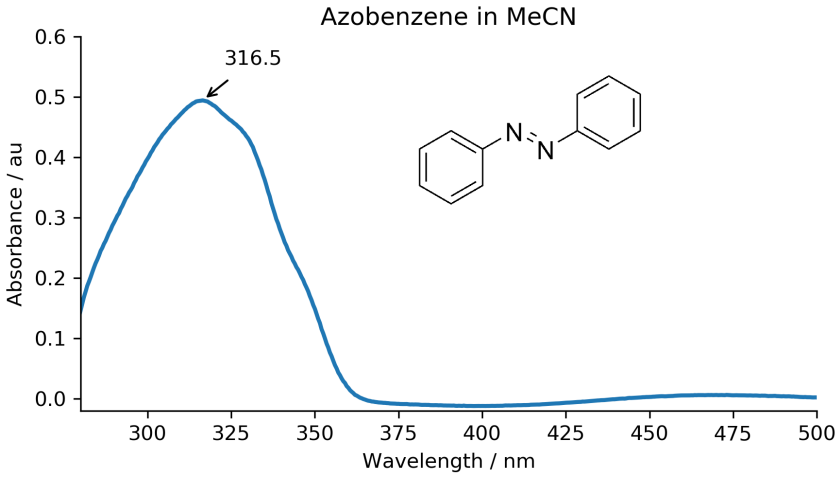
May 31, 2023

Goals:

- Calculate and visualize spectroscopic data
  - UV-Vis
  - NMR
  - IR

1 UV-Vis

We start with UV-Vis, as it involves some longer calculations, which you can have running in the background while you do the IR and NMR exercises.



The experimental value for the maximum absorption azobenzene in acetonitrile was (professionally, by us) measured to be approximately 317 nm. Your task is to use and compare three different methods to see how close they get to the experimental value.

- BP86 functional and def2-TZVP basis set
- B3LYP functional and def2-TZVP basis set
- Extra: STEOM-DLPNO-CCSD and def2-TZVP basis set

The first two calculations should be straightforward and be done in a few minutes (see Section 1.1), while the third requires some extra explanation and we will walk you through it in more detail (see Section 1.2). Compare the two/three different methods by plotting the spectra they predict (see Section 1.3). Additionally, check if the two DFT methods predict the same type of lowest-energy transition (see Section 1.4).

1.1 BP86 and B3LYP

By now, you should be familiar with the input file structure for the first two calculations. We have provided you with the input for the BP86 calculation below, you can adapt this for the B3LYP calculation by exchanging `BP86` with `B3LYP`.

```
! BP86 def2-TZVP cpcm(acetonitrile) tightscf
%pal nprocs 4 end
%maxcore 2048
%tddft
  Nroots 5
end
*xyzfile 0 1 azobenzene.xyz
```

With `Nroots` you can change the number of excited states that are calculated. In general, you might want to ask for more than 5 states. Note that even if you only care about the first excited state, it is recommended to still ask for at least `nroots 3`. As we are only interested in the first major band, calculating the first 5 states should be fine.[1] We have provided the optimized coordinates `azobenzene.xyz` in the `git` repository for this week.
Submit the calculation with:

```
suborca.sh bp86.inp
```

Repeat the same steps for `B3LYP`.

1.2 Extra: STEOM-DLPNO-CCSD

So far we have only performed semiempirical (`xtb`) and DFT calculations. Today, we introduce you to the "next level" as a bonus. One of the drawbacks of this method is that it takes a while to run and it generates a ton of temporary data. For this reason, **do not run this job in your home directory**. We will instead use your scratch directory (we briefly introduced this in the first lecture). To appreciate how much more space you have in your scratch directory, type `lquota` in your terminal.

1.2.1 Step 1: Move to your scratch directory

To access your scratch directory on Euler, type

```
cd /cluster/scratch/$USER/
```

and on Grace:[2]

```
cd /home/$USER/palmer_scratch
```

You can confirm that you are in the scratch directory by using the `pwd` command. Just as in your home directory, the same rules/recommendations apply. A key difference is that after 2 weeks of not modifying a file (60 days on Grace), it will be deleted. So, after the calculations have finished, make sure to transfer any important data to your home directory.

1.2.2 Step 2: Writing the input and submitting

Below is the input for the calculation.

```
! STEOM-DLPNO-CCSD def2-TZVP def2-TZVP/C RIJCOSX def2/J cpcm(acetonitrile) tightscf
%pal nprocs 8 end
%maxcore 4096
%mdci
Nroots 5
dorootwise true
dosolv true
end
*xyzfile 0 1 azobenzene.xyz
```

[1] This is a technical detail related to how the states are calculated.
[2] On Grace, there is also a project directory you can use. Both scratch and project are linked to your home directory, so you can access them easily.

There are several keywords that are likely unfamiliar to you – some of them are required for these calculations, others help speed it up. You can read more about them in the DLPNO-related sections in the ORCA manual. More information on specifically STEOM-DLPNO-CCSD can be found in the ORCA 5.0.3 manual Sections 8.5.6 and 8.5.9 (and links therein). Submit the calculation with:

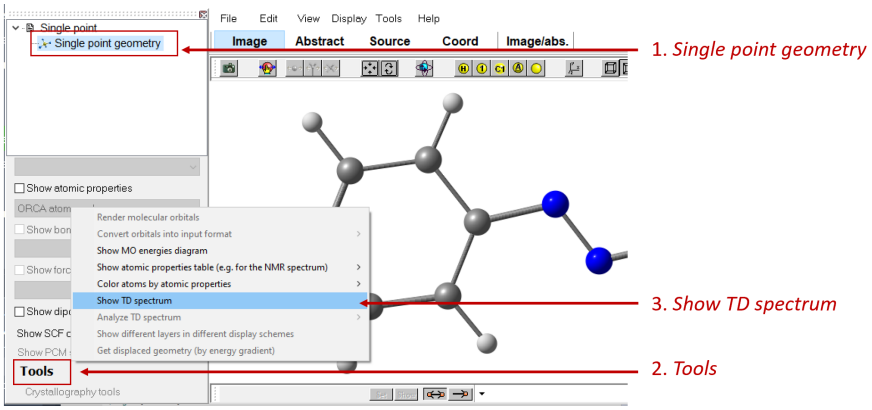
```
suborca.sh ccsd.inp
```

1.2.3 Step 3: Evaluating the results

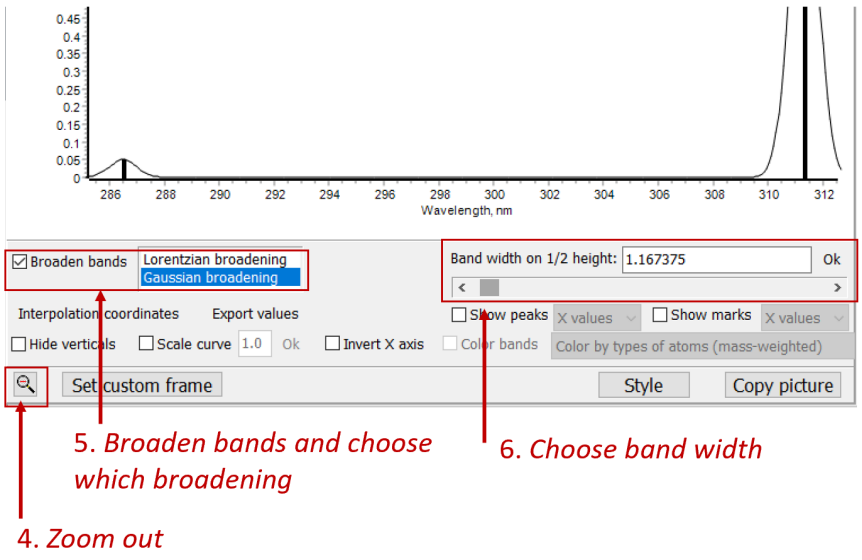
Once the job has finished, don't forget to copy the output file and whatever else you want to keep to your home directory. In the following sections we will show you how you can visualize your results in order to compare the different methods to the experimental value. We have a separate section for Chemcraft (Section 1.3.1), Avogadro (Section 1.3.2), and how to plot the data by using `orca_mapspc` (Section 1.3.3).

1.3 Visualizing UV-Vis data

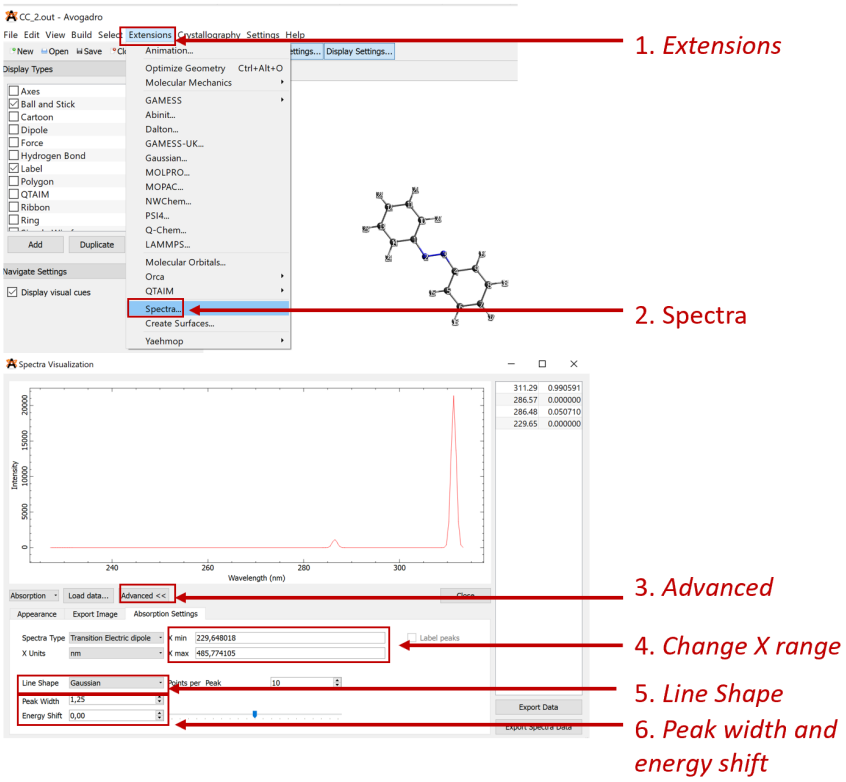
1.3.1 Chemcraft



Open the `.out` file and click in the top left on *Single point geometry* (**1**). You can then select *Tools* (**2**) in the bottom left and *Show TD spectrum* (**3**) so that a new window appears. In the top left of this window you have the option of switching the units of the x-axis (eV, cm^{-1}, or nm). Sometimes, only a part of the spectrum is shown. To view the whole, press on the little magnifying glass in the bottom left (**4**). You have the option to broaden the bands either by Lorentzian or by Gaussian broadening (**5**) and you can choose the band widths with the slider on the right (**6**). There are plenty other options to modify the data treatment and visualization, and we encourage you to play around and explore the options.

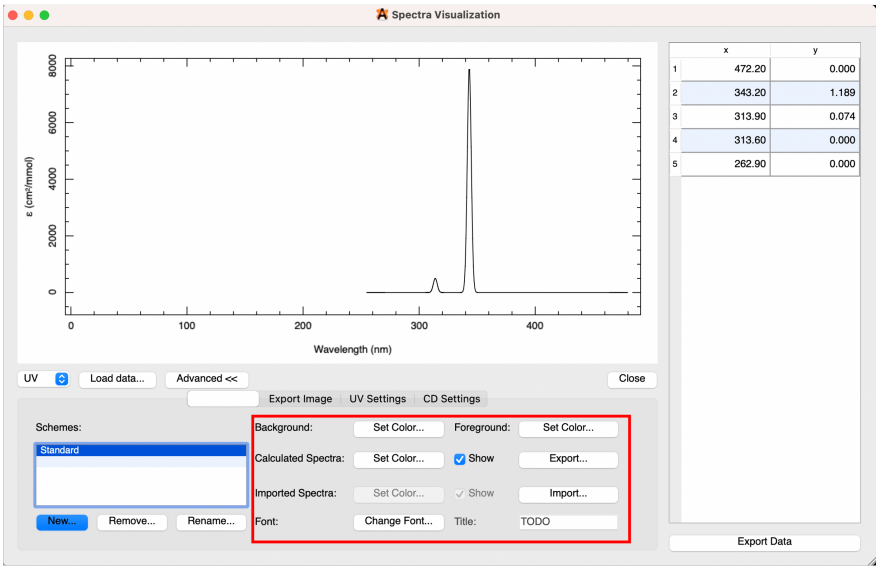


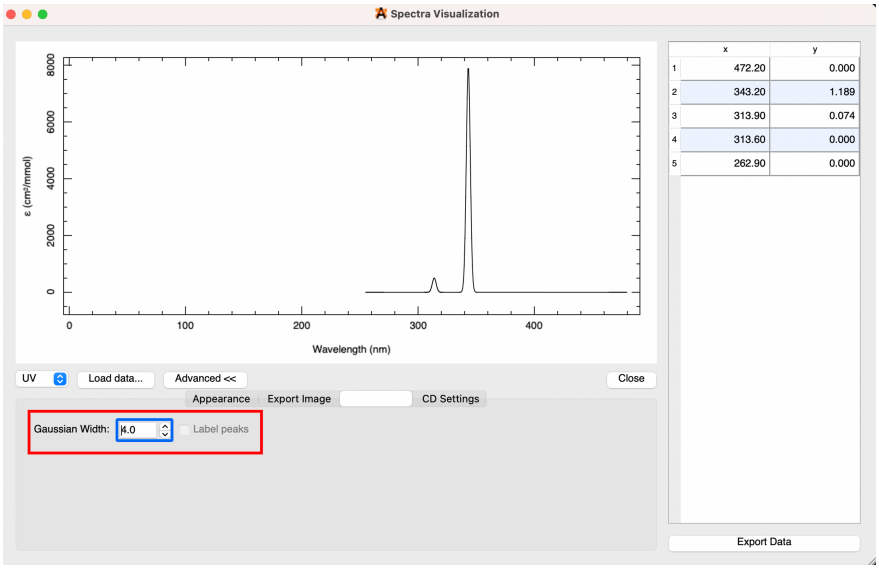
1.3.2 Avogadro



Open the `.out` file and in the top menu click on *Extensions* (**1**) and *Spectra* (**2**). In the bottom left you can change from *CD* to *Absoprtion*. Click on *Advanced* (**3**) to access more options. You can change the range of the X values (usually in [nm]) (**4**), modify the line shape (**5**), peak width (**6**), and even introduce an energy shift.

Note: Avogadro 1 for Mac looks different than the version for Windows and generally seems to have fewer features (and Avogadro 2 still completely lacks support for UV/Vis spectrum visualization). It is also possible that you only see a black screen before adjusting the color scheme. We recommend using `orca_mapspc` (next section) and visualizing your data with Excel or another software package. In case you still want to give Avogadro a shot, change the color scheme under the "Apearance" tab of the Advanced section and find the peak broadening under the "UV Settings" tab of the Advanced section.





1.3.3 `orca_mapspc`

`ORCA` has a tool to extract the data from a spectroscopy calculation and save it as a simple text file for you to plot with another program, *e.g.*, Excel or Python. To call it type

```
orca_mapspc
```

This will show all the general options you have at your disposal. For the UV-Vis spectrum of azobenzene you would need to ask for an absortpion spectrum (with the `ABS` option) and you may want to plot the data in eV (with the `-eV` flag) as there is no option for wavelength (to get wavelength in nm, you can divide 1239.84 by the eV values). Additionally, you can choose to only show the spectrum between 2.5 eV and 6 eV (with the `-x0` and `-x1` flag). By default, Gaussian line broadening is applied and you can specify the full width at half maximum (with the `-w` flag). As only two significant digits are printed, make sure to only plot 100 points per 1 eV, *i.e.*, take th6e spectral width (in eV), multiply it by 100 and take this as the number of points to be plotted (with the -n flag). In our example below it would thus be 400 points. The full command would thus be

```
orca_mapspc dummy.out ABS -eV -x02 -x16 -w0.05 -n400
```

This will give you two text files with the spectrum data in the region of 2 eV – 6 eV. The `.abs.dat` file contains the broadened spectrum with five columns. The first column is energy and the second column the total intensity (columns 3–5 represent intensities w.r.t. individual polarizations along the x, y, and z axis). The second file (`.abs.stk`) only contains the (stick) peaks with intensities.

1.4 Excited states

You can check the details of the excitations by opening the output file in an external text editor or in `vim`, and searching for the "TD-DFT/TDA EXCITED STATES" section. An example output is shown below.

```
-------------------------------------
TD-DFT/TDA EXCITED STATES (SINGLETS)
-------------------------------------

the weight of the individual excitations are printed if larger than 1.0e-02

STATE  1:  E=   0.084842 au      2.309 eV    18620.6 cm**-1 <S**2> =   0.000000
    47a ->  48a  :     0.993617 (c=  0.99680356)

STATE  2:  E=   0.123959 au      3.373 eV    27205.9 cm**-1 <S**2> =   0.000000
    45a ->  48a  :     0.360134 (c=  0.60011127)
    46a ->  48a  :     0.597956 (c=  0.77327607)
```

In this case, the lowest-energy excitation is from the orbital 47a to the orbital 48a and has a weight of approximately 1.0. We will learn next week how to visualize orbitals, so you can interpret the nature of

this transition. In contrast, the second state is mainly made up of two orbital excitations (but it is still a one-electron excited state). *Advanced:* If you want to interpret such "mixed" excited states, you can try calculating natural transition orbitals (ORCA 5.0.3 Manual, Section 9.26.9).

You can find various spectra with the wavelengths and intensities printed in the "TD-DFT/TDA-EXCITATION SPECTRA" section:

```
----------------------------
TD-DFT/TDA-EXCITATION SPECTRA
----------------------------
```

where the typical UV/Vis spectrum is listed as:

```
-------------------------------------------------------------------------------
         ABSORPTION SPECTRUM VIA TRANSITION ELECTRIC DIPOLE MOMENTS
-------------------------------------------------------------------------------
State   Energy    Wavelength  fosc         T2        TX        TY        TZ
        (cm-1)      (nm)                (au**2)    (au)      (au)      (au)
-------------------------------------------------------------------------------
   1   18620.6    537.0   0.000000000   0.00000   0.00000  -0.00000   0.00000
   2   27205.9    367.6   0.597202410   7.22659   2.52214  -0.93028  -0.00000
   3   27685.2    361.2   0.000000000   0.00000   0.00000  -0.00000   0.00000
   4   27990.5    357.3   0.633228097   7.44775   2.72007  -0.22133  -0.00000
   5   32464.9    308.0   0.000000000   0.00000   0.00000  -0.00000   0.00000
```

where the fosc column lists the oscillator strengths of the transitions (which are dimensionless quantities that give you the relative intensities of the transitions).

For STEOM-DLPNO-CCSD, the UV/Vis spectrum is given in the same format, but the excited states are listed under:

```
------------------
STEOM-CCSD RESULTS
------------------
```

2 IR

For the later NMR exercise we will be looking at everyone's favorite NMR spectra impurity next to water – ethyl acetate. To get decent data we need to run the NMR calculation on an optimized structure; to make sure it is a minimum, we will also calculate the frequencies (see Week 4). We can take the data from the frequency calculation to get familiar with plotting IR spectra.

Generate the coordinates of ethyl acetate with your method of choice (drawing the molecule in a visualization software, smi2xyz, *etc.*). Run a cheap DFT geometry optimization and frequency calculation to obtain the IR data.

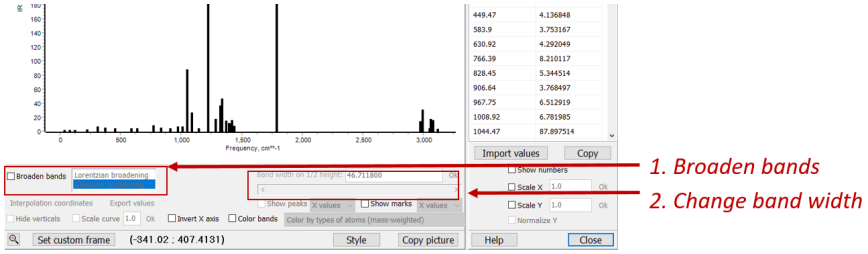
```
! BP86 def2-SVP D3BJ tightopt freq
%pal nprocs 4 end
%maxcore 1024
*xyzfile 0 1 ethyl_acetate.xyz
```

Check that you have a minimum energy structure by looking at the obtained frequencies (you should have 0 imaginary ones), either with a text editor (notepad, vim, *etc.*) or in a visualization software. In the next section we will show you how to visualize the data, so that you can compare it to the experimental one (for instance, from NIST). Note that we are calculating the IR spectrum in the gas phase.

2.1 Visualizing IR data

2.1.1 Chemcraft

Open the .out file. The frequencies are shown on the left. Remember, the first 6 are translations and rotations which have been projected out and are thereby 0. To view the animation of a specific vibration, simply click on the frequency value. In the bottom of the left sidebar you can change a few settings, such as the displacement scaling and the cycle time.



On the bottom of the left sidebar you have the option to *Show spectrum*. Click on it and select *IR intensities* – you now have the vibrational spectrum. You have the option to broaden the bands either by Lorentzian or by Gaussian broadening (**1**) and you can choose the band widths with the slider on the right (**2**). You can zoom into a specific area of the spectrum by holding down the left mouse button and selecting the area. As already mentioned above, there are many other options you can try out for treating and visualizing the data.

2.1.2 Avogadro

Unfortunately, neither Avogadro 1 nor Avogadro 2 have support for finding information on vibrational spectra in output files generated by `ORCA` 5.0.3. If you don't want to or cannot use Chemcraft, you will have to resort to `orca_mapspc` (next section) and plotting the data yourself in your software of choice. For identifying specific vibrations, you could use the `orca_pltvib` tool we introduced in Part 4 or explore using Molden.

2.1.3 `orca_mapspc`

The same guidelines apply as in Section 1.3.3, but this time use the `IR` option. As cm^{-1} is the default unit we don't have to specify it explicitly.

```
orca_mapspc dummy.out IR -x0300 -x14000 -w50
```

The example above will save two text files for the IR spectrum data in the region of 300 to 4000 cm^{-1} and with a full-width at half-maximum height of 50 cm^{-1} (the `-w`). One with the full broadened spectral data `dummy.out.ir.dat` and one where only the peaks are listed with the intensities `dummy.ir.stk`.

3 NMR

As mentioned in the the IR section, we will be first looking at ethyl acetate. We will run the calculation in chloroform and see if we can get the same values as you can find in the literature (or in your own spectra). You can use the coordinates from the optimization in the IR exercise.
To calculate the NMR data, we will switch to a larger basis set, and use the keyword `!NMR`. An example input is shown below, where we add the **`eprnmr`** block **after** the coordinates.

```
! BP86 def2-TZVPP cpcm(chloroform) NMR
%pal nprocs 4 end
%maxcore 1024

*xyzfile 0 1 coord.xyz

%eprnmr
Nuclei = All H {Shift}
Nuclei = All C {Shift}
end
```

There are many settings you can tune for NMR calculations, and we encourage you to explore the `ORCA` 5.0.3 manual Sections 8.9.7 and 9.42.3.

Open the output file either with `vim` or an external text editor. After the results from the single point calculations (SCF, orbital energies, *etc.*) you should find a new section:

```
ORCA EPR/NMR CALCULATION
```

Scroll further down to the CHEMICAL SHIFTS section. For every nucleus that you asked for (the default is every, in our example we only asked for H and C), you will get the tensor values for the dia- and paramagnetic contributions. Below this, you will finally get a summary of the chemical shielding of all nuclei in ppm (we will be using the isotropic values). You can find the summary section when you seach for

```
CHEMICAL SHIELDING SUMMARY (ppm)
```

Comparing these to literature values (^{13}C-NMR: 14.2, 21.0, 60.5, 171.4 and ^{1}H-NMR: 1.26, 2.05, 4.12) you will notice some severe differences. This is because we calculated absolute values and these are not directly comparable to experimental data. We need to calculate the reference shielding and subtract the absolute values from this reference. Another difference is that equivalent nuclei have different shieldings. To compare them to the literature values, you can simply average them.

We will be using tetramethylsilane as the reference for the C and the H shifts. Just as with ethyl acetate, run a quick geometry optimization and use the same input file (but different coordinates) for the NMR calculation. Again, use the averages of the C and H shifts as your reference value.
To now properly compare the average calculated with the experimental values, use

$$\sigma_{\text{calculated}} = \sigma_{\text{TMS}} - \sigma_{\text{EtOAc}}$$

Note: You do not always have to reference to an external molecule. A nice alternative is to reference to a known shift within the compound. Another popular approach is to use linear regression to address functional and basis set dependent errors from NMR spectrum calculation (see M. W. Lodewyk, M. R. Siebert and D. J. Tantillo, Chem. Rev. **2012**, *112*, 1839–1862 for an overview).

> Similarly to `autodE` for reaction mechanisms, there are tools that automate the process of NMR spectra calculation. If you want to stay in the Grimme software suite, consider `enso` and `anmr` (`https://xtb-docs.readthedocs.io/en/latest/enso_doc/enso.html`).