

# CP2K Tutorial

## Exercises

### Core Level Spectroscopy

Problem: Calculation of the x-ray absorption and emission spectra of a molecule in gas-phase before and after dimerization.

For this tutorial some input and output files are given in order to present a complete procedure to solve the given problem. Some hints are also given to help in the analysis of the results. In order to be able to run these examples, some paths need to be correctly set in the input files (i.e. set the variables LIBPATH, XYZPATH, RUNPATH).

The tasks to be completed are:

- Optimize the geometry of the molecule in gas-phase choosing an appropriate level of theory. Different isomers of the same molecule can be stable. We take into account one CIS isomer and one TRANS isomer.
- Compute the C 1s absorption and emission spectra for the optimized geometries by applying the Slater transition potential approximation.
- Optimize the structure of the dimer in gas phase. Also in this case we can consider different isomers.
- Calculate the C 1s absorption and emission spectra of the different optimized isomers.

## 1 First task: geometry optimization of two isomers of the monomer

It is expected that the monomer in gas phase can be stable in either the CIS or the TRANS isomer, therefore the optimization of the structure is performed two times starting from two different initial configurations. Two input files are prepared for the two different topologies:

`gop_monomer_cis.inp` and `gop_monomer_trans.inp`

The initial coordinates are described in the xyz files :

`monomer_cis_init.xyz` and `monomer_trans_init.xyz`

In the SUBSYS section the information on the structure and the simulation box are:

```

&TOPOLOGY
  &CENTER_COORDINATES T
  &END CENTER_COORDINATES
  COORD_FILE_NAME ../monomer_trans.xyz
  COORDINATE XYZ
  CONNECTIVITY OFF
&END TOPOLOGY
&CELL
  ABC [angstrom] 22.0 22.0 22.0
&END CELL

```

We run the optimization at the DFT level of theory using revPBE as XC functional and the Grimme correction for the dispersion forces.

```

&XC
  &XC_FUNCTIONAL
    &PBE
      PARAMETRIZATION REVPBE
    &END PBE
  &END XC_FUNCTIONAL
  &XC_GRID
    XC_SMOOTH_RHO NN50
    XC_DERIV NN50_SMOOTH
  &END XC_GRID
  &VDW_POTENTIAL
    POTENTIAL_TYPE PAIR_POTENTIAL
    &PAIR_POTENTIAL
      SCALING 1.0
      TYPE DFTD2
      R_CUTOFF [angstrom] 16
    &END PAIR_POTENTIAL
  &END VDW_POTENTIAL
&END XC

```

For the optimization we could use the GPW method and treat the core electron through the use of pseudopotentials. However, since later on for the calculation of the spectra we have to switch to all-electron calculations and the systems in this case are not too big, we work from the beginning with GAPW and all-electrons. For the optimization of the structure we use all-electron basis sets of moderate size, 6-311Gxx. Later for the calculation of the spectra it is recommendable to select more extended basis sets (see next task). You can try to change the input file to employ GPW and pseudopotentials and compare the results with the GAPW optimization.

The output (.out), the trajectory (.xyz), and the CP2K restart (.restart) files obtained by running CP2K with the given inputs, can be found in the subdirectories :

GOP\_MONO\_CIS, GOP\_MONO\_TRANS

The `.wfn` files storing the information on the electronic structure have not been reported here because of the size. They are not essential, since a calculation can always restart from an atomic guess of the initial density matrix. Restarting from a previously calculated density matrix is obviously a better guess, thus requiring fewer iterations of the optimization scheme to converge. The behavior of the gradients and of the total energy along the optimization can be monitored by grepping "Max. grad" and "Total Energy " from the output file. Once the optimization is concluded the final configurations are copied into new xyz files:

`mono_cis_pbe_311gxx_opt.xyz` and `mono_trans_pbe_311gxx_opt.xyz`

and stored in the XYZ directory. The following command can be used:

```
tail -n 24 ${xray_path}/DIMERIZATION/GOP_MONO_CIS/gop_mono_cis_pbe_311gxx-pos-1.xyz >
${xray_path}/XYZ/mono_cis_pbe_311gxx_opt.xyz
```

These are the structures used for the calculation of the C-1s spectra.

## 2 Second task: C-1s spectra calculation for the monomer

The input file for the calculation of the spectra are given:

`hch_C1s_mono_cis.inp` and `hch_C1s_mono_trans.inp`.

The run-type is a simple ENERGY calculation, but in the DTF section an additional XAS section is added to invoke the calculation of the spectra. In order to compute the spectra, first a ground state calculation is performed. For this calculation, it is possible to restart the wavefunction from the last restart file written during the geometry optimization. See in the DFT section

RESTART\_FILE\_NAME

This is possible even if we change basis set, as indeed it is done, since the spectra are quite sensitive to the basis set, in particular in the vicinity of the excited core orbital. In the example the  $6 - 311 + +G2d2p$  basis set is used.

In order to reproduce the x-ray absorption and emission spectra, transition energies and intensities related to the promotion process of a core electron to an unoccupied orbital or the radiative decay process from a valence state to the emptied core state must be computed. This is done by assuming that initial and final state of the transition can be directly approximated by selected one-electron orbitals, solutions of the KS equations with a properly modified core-hole potential on the absorbing atom. Once the ground state wavefunction is available, a loop is started over all the atoms for which the spectrum has to be calculated, in the present case over all the C atoms. The core-hole potential is set by adjusting the occupation number of the absorbing core-state, and the KS equations are solved under the new conditions. In the present set of examples, the Slater transition potential approximation is adopted, i.e. the occupation of the absorbing core state is set to 0.5.

The section that needs to be added in the input reads:

&XAS

CONVERGENCE 5.E-6

```

EPS_DIIS    0.05
SCF_OPTIMIZER GENERAL
&MIXING
    ALPHA 0.2
    METHOD DIRECT_P_MIXING
&END
METHOD      TP_HH
DIPOLE_FORM VELOCITY
MAXSTEP     200
STATE_TYPE  1s
STATE_SEARCH 22
ATOMS_LIST   1 2 3 4 5 6 12 13 14 19
ADDED_MOS    500
&LOCALIZE
&END
&PRINT
    &PROGRAM_RUN_INFO
    &END
    &RESTART
        FILENAME ./mono_cis_hch
        &EACH
            XAS_SCF 15
        &END
        ADD_LAST NUMERIC
    &END
    &XAS_SPECTRUM
        FILENAME ./mono_cis_hch
    &END
    &XES_SPECTRUM
        FILENAME ./mono_cis_hch
    &END
&END
&END

```

Whatever is the optimization scheme (OT, diagonalization,...) used for the ground state calculation, the SCF with the core-hole potential is done using the diagonalization scheme plus mixing of the charge density. The keyword **METHOD** indicates which type of core-hole potential is used. **STATE.TYPE** indicates which orbital in the core is excited. **ATOMS LIST** is the list of atoms that are going to be excited to compute the respective spectra. The localization of the ground state orbitals is necessary to facilitate the assignment of the core orbitals to the different atoms. To this purpose a unitary rotation is applied to a subset of orbitals, in order to transform them into the description corresponding to maximally localized wavefunctions. **ADDED\_MOS** in this section indicates how many states are to be added a posteriori to compute the spectrum. These states are not included during the optimization of the wavefunction.

Hence, the meaning of this keyword must not be confused with the `ADDED_MOS` keyword in the subsection SCF. The output files are collected in the directories

`HCH_MONO_CIS_311++G2d2p` and `HCH_MONO_TRANS_311++G2d2p`

The files with extension `xas_at#.spectrum` and `xes_at#.spectrum` are the actual absorption and emission spectra for the atom with index `#`. These are the raw spectra that need to be convoluted before plotting and analyzing them. The files with extension `.rst` are the restart files for the x-ray spectroscopy calculations, one file for each excited core orbital (one for each atom in the list `ATOMS_LIST`). By adding in the section XAS the keyword `RESTART`, the code looks for these files and restart the calculation from the wavefunctions contained in these files. This is useful when the SCF has not converged, or after changing some parameter, as for example the convergence criteria, the mixing, the number of `ADDED_MOS`. It is not possible restarting with a different `METHOD` or a different `STATE_TYPE`.

For the analysis of the results, it is first necessary to convolute the data. One possibility is to sum up all the data, i.e. copying all the computed spectra for all the carbon atoms in one single file, and then applying the convolution, thus obtaining one single C-1s spectrum including all the contributions. The different contribution to the spectrum are instead analyses by generating one spectrum for each atom. In the directory `LIB_TOOLS` there is a simple routine in Fortran that can be used to convolute the data. It needs a few input arguments as the number of lines of the raw spectrum, the minimum width of the gaussians used for the convolution and energy, in eV, from where the Gaussian width starts to be linearly increased to better mimic the broadening of the experimental signals (life-time of the excited states).

```
./my_convolute.x -pnt $ntot -e1 $e1 -e2 $e2 -smin $sigma
```

By default this program reads an input file named `spectrum.inp`, with seven columns, where the first column is the state index, the second column is the transition energy, third, fourth, and fifth columns are the transition intensities projected in *x*, *y*, and *z* respectively (polarized), the sixth column is the total intensity of the transition (transition probability), and the seventh column is the occupation number of the final state. The file has no header (delete the header of the output of `cp2k`). The output of the convolution is written in `spectrum.out`.

In order to have a more accurate reference for the transition energy, it is a good practice to compute the first excitation energy by a  $\Delta$  SCF approach. The energy of the first excited state is computed by promoting one electron into the first unoccupied state and solving the KS equation with this charge redistribution. The difference of the total energy obtained for this excited state and the energy obtained in the ground state is then the reference energy. Hence, the spectra can be rigidly shifted in order to align the on-set energy with the corresponding first transition energy attained by  $\Delta$ SCF.

The  $\Delta$ SCF calculations for the two optimized monomers are performed using the input files : `dscf_C1s_mono_cis.inp`, `dscf_C1s_trans_cis.inp`

For the  $\Delta$ SCF calculations, some unoccupied states are needed also by the optimization of the electronic structure, because one electron has to be physically promoted to the lowest unoccupied state. Therefore states need to be added also at the level of ground state calculation

and the OT optimization scheme cannot be used. Diagonalization and direct mixing of the density matrix are going to be used (which is the default method). The SCF subsection within the DFT section reads:

```
&SCF
  MAX_SCF 100
  EPS_SCF 5.0E-6
  SCF_GUESS RESTART
  ADDED_MOS 10 10
  &PRINT
    &RESTART
      ADD_LAST NUMERIC
      &EACH
        MD 10
      &END EACH
    &END RESTART
  &END PRINT
&END SCF
```

The XAS section in the input looks very similar to the what seen for the half-core-hole potential calculations. The value of the keyword **METHOD** has been changed to **DSCF**, no additional states and no print keys are needed since no spectra are going to be calculated in this case.

```
&XAS
  CONVERGENCE 5.E-6
  EPS_DIIS 0.01
  SCF_OPTIMIZER GENERAL
  &MIXING
    METHOD DIRECT_P_MIXING
    ALPHA 0.1
  &END
  METHOD DSCF
  MAXSTEP 200
  STATE_TYPE 1s
  STATE_SEARCH 22
  ATOMS_LIST 1 2 3 4 5 6 12 13 14 19
  &LOCALIZE
  &END
  &PRINT
    &PROGRAM_RUN_INFO
  &END
  &RESTART
    FILENAME ./dehydro_xas_dscf
  &EACH
```

```

XAS_SCF 15
&END
ADD_LAST NUMERIC
&END
&END
&END

```

The output of these calculations are collected in the directories:

DSCF\_MONO\_CIS\_311++G2d2p and DSCF\_MONO\_TRANS\_311++G2d2p

The first transition energy can be computed for each individual carbon from the energy difference  $E_{\Delta\text{SCF}} - E_{\text{GS}}$ . These energies have to be compared with the first transition energy computed by the core-hole potential approximation and stored in the files with extension `xas_at#.spectrum`. Calculate the  $\Delta\text{SCF}$  shift for each C atom individually.

### 3 Third task: geometry optimization of four isomers of the dimer

The molecule studied in the previous tasks can dimerize by cycloaddition. In the XYZ directory, four possible structures of the dimer are prepared:

`dimer_p13_updown.xyz`, `dimer_p12_updown.xyz`, `dimer_p13_upup.xyz`, `dimer_p12_upup.xyz`

The structure optimization of the dimer can be carried out using the following input files:

`gop_dimer_p13_upup.inp`, `gop_dimer_p12_upup.inp`,  
`gop_dimer_p13_updown.inp`, `gop_dimer_p12_updown.inp`

Beware that more isomers might be possible. The set-up for the structure optimization is the same that has been used for the monomer. It might be interesting checking whether by changing something in the set-up, basis set, functionals, potentials (from all electrons to PP) something changes in the optimized structures and in particular in the relative stability among the isomers. Output and trajectory files of the optimizations are collected in the directories:

`GOP_DIM_P13_UPUP`, `GOP_DIM_P12_UPUP`, `GOP_DIM_P13_UPDOWN`, `GOP_DIM_P12_UPDOWN`

The optimized structures are copied in the XYZ directory.

### 4 Fourth task: C-1s spectra calculation for the dimer

Once the optimized structures are available, the C-1s spectra following the same procedure already used for the monomers. The proposed input files are:

`hch_C1s_dim_p13_upup.inp`, `hch_C1s_dim_p12_upup.inp`,  
`hch_C1s_dim_p13_updown.inp`, `hch_C1s_dim_p12_updown.inp`

In `ATOMS_LIST` all the C atoms of the dimers are given. The results of these simulations are stored in the following directories:

`HCH_DIM_P13_UPUP_311++G2d2p`, `HCH_DIM_P12_UPUP_311++G2d2p`  
`HCH_DIM_P13_UPDOWN_311++G2d2p`, `HCH_DIM_P12_UPDOWN_311++G2d2p`

Possible changes in the spectral profile due to the dimerization can be investigated, in particular checking the spectra of those C atoms directly involved in the dimerization process. Is it possible to distinguish among the different isomers from the spectra? Possible effects of functional and basis-set can be also studied by changing the simulation set-up.