CHEM-596 Practice Session 5: Correlated wavefunction methods; energy decomposition analysis for non-covalent interactions

Name: _____

Due: April 18, 2023 (11:59 pm)

In this practice, you will first practice using a combination of correlated wavefunction methods to generate high-level quantum chemistry reference data. You will then employ the ALMO-EDA method and its variant to investigate (i) the $H_2O\cdots Cl^-$ interaction in the gas phase and implicit solvent; (ii) vibrational frequency shift of CO in the OBe$\cdots$CO complex. It is recommended to work on **Exercise 3** after next Tuesday's lecture, in which we will introduce the adiabatic ALMO-EDA method.

It is recommended that you use the Linux servers for *all* exercises this time, although most of the calculations involved in **Exercise 1** can be done through the IQmol server (except for the HF/aug-cc-pV5Z calculation). You can still use IQmol to conveniently generate the input files for you though. During the practice session today, **please use the ITCOTCK server if your last name initial is A–F and use the DGX server otherwise**. You are encouraged to run Q-Chem calculations with 4-thread OpenMP parallelization using "qchem -nt 4 [input] [output] & ".

You will need to **complete the tables and short-answer questions on this handout** and also submit the files that are requested. Please refer to the **Assignment page** on Canvas for detailed submission guidelines. **8 points** are associated with the completion of this handout and the rest **2 points** are for attendance and completion of the in-class quiz.

Exercise 1: Generating the quantum chemistry reference value for the HCN → HNC isomerization energy

In this exercise, we will calculate the energy change upon the HCN → HNC isomerization reaction at the MP2/CBS and ΔCCSD(T)/CBS levels. The isomerization energy of HCN is defined as

$$\Delta E_{\mathrm{iso}} = E(\mathrm{HNC}) - E(\mathrm{HCN}) \tag{1}$$

The MP2/CBS energy in this example is defined as the sum of HF/aug-cc-pV5Z energy and MP2 correlation energy extrapolated to the CBS limit, i.e.,

$$\begin{aligned} E_{\mathrm{MP2}}(\mathrm{CBS}) &\approx E_{\mathrm{HF}}(\mathrm{a5Z}) + E_{\mathrm{MP2}}^{\mathrm{corr}}(\mathrm{CBS}) \\ &\approx E_{\mathrm{HF}}(\mathrm{a5Z}) + \left[X^3 E_{\mathrm{MP2}}^{\mathrm{corr}}(X) - Y^3 E_{\mathrm{MP2}}^{\mathrm{corr}}(Y)\right]/(X^3 - Y^3) \end{aligned} \tag{2}$$

Here we use RIMP2 results calculated using aug-cc-pVTZ and aug-cc-pVQZ to extrapolate to the CBS limit, for which $X = 3$ and 4, respectively. Let's start with calculating these with the given `HCN.xyz` structure. To set up an RIMP2 calculation using IQmol:

- Set `Method` to "RI-MP2" and `Basis` to "aug-cc-pVTZ" or "aug-cc-pVQZ"

- Set the "RI Fitting Basis" at the bottom to "RIMP2-aug-cc-pVTZ" or "RIMP2-aug-cc-pVQZ"

- Set `GUI = 0` and the SCF convergence criterion to 8 (under `SCF Control`)

Through the output of an RI-MP2 calculation, you will be able to find:

- The Hartree-Fock (HF) energy (the number after "SCF energy in the final basis set")

- The RI-MP2 total energy ("`RIMP2    total energy`")

- The RI-MP2 correlation energy ("`RIMP2    correlation energy`")

The section showing these numbers can be easily seen by searching for "`RIMP2 correlation energy`" in the output.

You can now report these values [HF, RIMP2, and RIMP2(corr)] based on your RIMP2/aug-cc-pVTZ and aug-cc-pVQZ calculations in **Table 1**. In addition, you should run a Hartree-Fock calculation with the aug-cc-pV5Z basis (remember to set `SCF_CONVERGENCE` to 8) and fill in the blank "HF/a5Z". **(0.75 pt)**

Now use Eq. (2) to obtain the RIMP2/CBS result for HCN. **Note:** the extrapolation should only apply to the RIMP2 correlation energy rather than the RIMP2 total energy. **(0.25 pt)**

Hint: use Excel to do the extrapolation so that you can use the equation later for HNC.

Table 1: Getting the MP2/CBS and ΔCCSD(T)/CBS values for the HCN $\rightarrow$ HNC isomerization energy. aTZ, aQZ, and a5Z stand for aug-cc-pVTZ, aug-cc-pVQZ, and aug-cc-pV5Z, respectively.

Item	E(HCN) (a.u.)	E(HNC) (a.u.)	ΔE_{iso} (kcal/mol)
HF/aTZ			
HF/aQZ			
HF/a5Z			
RIMP2/aTZ			
RIMP2/aQZ			
RIMP2(corr)/aTZ			—
RIMP2(corr)/aQZ			—
RIMP2/CBS			
CCSD/aTZ			
CCSD(T)/aTZ			
ΔCCSD(T)/CBS			

Now repeat the procedure above for the HNC molecule (using the given geometry "`HNC.xyz`") and report the ΔE_{iso} values. **(1 pt)** At this point, you should have filled up **Table 1** up to the row "RIMP2/CBS".

The ΔCCSD(T)/CBS values in our case can be calculated using the following formula:

$$E_{\Delta\text{CCSD(T)}}(\text{CBS}) = E_{\text{MP2}}(\text{CBS}) + \left[E_{\text{CCSD(T)}}(\text{aTZ}) - E_{\text{MP2}}(\text{aTZ})\right] \qquad (3)$$

Since we have already done the RIMP2/aTZ calculation, we only need to perform a CCSD(T) calculation with the aug-cc-pVTZ basis for each of the species. You can use IQmol to set up this calculation (setting `Method` to "CCSD(T)"; also remember to set `GUI = 0` and `SCF_CONVERGENCE = 8`). Search for "`CCSD total energy`" and "`CCSD(T) total energy`" in the output and report the numbers for HCN and HNC in **Table 1** and calculate the corresponding ΔE_{iso} values. **(0.5 pt)**

You can then use Eq. (3) to calculate the ΔCCSD(T)/CBS energies for both species. Report the numbers (including the final ΔE_{iso} value) in **Table 1**. **(0.25 pt)**

The ΔE_{iso} value evaluated at the ΔCCSD(T)/CBS level can be regarded as the *quantum chemistry reference*. **Now answer the following questions:**

1. Comment on the accuracy of HF/aQZ, RIMP2/aQZ, and CCSD/aTZ for the prediction of the isomerization energy (e.g., overestimate or underestimate? by how much?) **(0.25 pt)**

2. With the reference value we can assess the performance of DFT functionals. Now calculate the isomerization energy using B97M-V (a semi-local meta-GGA) and M06-2X (a hybrid meta-GGA) with the def2-TZVPPD basis. Report your results in **Table 2** below. **(0.25 pt)**

3. Which functional gives a smaller error compared to the ΔCCSD(T)/CBS reference? In terms of the absolute error, are they doing better or worse than RIMP2/CBS? **(0.25 pt)**

Table 2: The HCN $\rightarrow$ HNC isomerization energy calculated using B97M-V and M06-2X with the def2-TZVPPD basis set

Functional	E(HCN) (a.u.)	E(HNC) (a.u.)	ΔE_{iso} (kcal/mol)
B97M-V			
M06-2X			

Exercise 2: Energy decomposition analysis for the $H_2O \cdots Cl^-$ complex

In this exercise, we will use the second-generation ALMO-EDA to investigate the interaction between a water molecule and a Cl^- anion in the gas phase and in implicit solvent. All calculations will be performed using the given XYZ structure file "`water_cl.xyz`" and at the ωB97X-V/def2-TZVPD level of theory.

Let's first set up the gas-phase ALMO-EDA calculation with the following steps:

- Open "`Q-Chem Setup`". First change the charge of the system to -1 and make sure that the multiplicty is 1. Then you need to specify the charge and multiplicity of each fragment *manually*: fragment 1 is the chloride anion so should be "-1 1", while fragment 2 is the water molecule which should be "0 1". You also need two dashes "--" to indicate the beginning of each fragment. The completed `$molecule` section should look like:

```
$molecule
-1 1
--
-1 1
 Cl  -1.1242249     0.0179827     0.0000000
--
 0 1
 O    1.9940516    -0.0909265     0.0000000
 H    1.0163933    -0.2184602     0.0000000
 H    2.0588108     0.8652887     0.0000000
$end
```

- Under "`Calculate`" select "`Energy Decompose`" (which gives you "`JOB_TYPE = EDA`"), then manually add "`EDA2 = 1`" which stands for "decomposition of frozen energy + FERF-DQ polarization"

- Set the functional to "wB97X-V" and basis to "def2-TZVPD"; finally, set the SCF_CONVERGENCE to 8 and set GUI = 0.

You can now run the gas-phase EDA calculation on the server. Once finished, from the output you should be able to find results for the total interaction energy (INT) and its 5 components: permanent electrostatics (ELEC), Pauli repulsion (PAULI), dispersion (DISP), polarization (POL), and charge transfer (CT). Here we focus on the results of the "classical" version of frozen decomposition so the keywords we need to search for in the output file are:

INT: E_int   ELEC: E_cls_elec   PAULI: E_mod_pauli   DISP: E_disp   POL: E_pol   CT: E_vct

Report the numbers under the "Gas phase" column of **Table 3 (0.5 pt)** and answer the following questions:

1. Among the 5 energy components, which ones are attractive (< 0) and which ones are repulsive (> 0)? **(0.25 pt)**

2. It was proposed that hydrogen bonding interactions (like the interaction between H_2O and Cl^-) are dominated by charge-transfer interactions. Does our result support this? If not, which attractive force makes the largest contribution to the $H_2O \cdots Cl^-$ interaction? **(0.25 pt)**

3. Based on the ALMO-EDA results, estimate by how much (in %) the interaction energy would be underestimated if the employed density functional were not able to describe any dispersion interaction. **(0.25 pt)**

Table 3: ALMO-EDA results for the $H_2O \cdots Cl^-$ complex using ωB97X-V/def2-TZVPD (energy components in kJ/mol)

Energy components	Gas phase	SMD water
ELEC		
PAULI		
DISP		
POL		
CT		
INT		

Let's now set up the ALMO-EDA-II calculation with implicit solvent (SMD water). The procedure is the same as how you add the SMD solvation model for your pKa calculations last time: go to "Advanced $\rightarrow$ Solvent Model", and select "SMD" as the "Solvent Method". You may need to click that twice to make sure that "SOLVENT water" is specified under the $smx input section. Execute this calculation on the server and complete the "SMD water" column in **Table 3** based on the output. **(0.25 pt)** For permanent electrostatics and Pauli repulsion, you should look for E_cls_elec(solv) and E_mod_pauli(solv) this time.

Comparing the EDA results for $H_2O\cdots Cl^-$ in the gas phase and in the solvation environment, discuss what you observe from the data. E.g., how much does the total interaction energy change, which energy components are significantly impacted by the solvation environment and how do they change, etc. **(0.5 pt)**

The final task of **Exercise 2** is to generate the complementary occupied-virtual pairs (COVPs) that describe the charge transfer interaction between H_2O and Cl^-. Starting from your input for the *gas-phase* EDA, make the following modifications:

- Change the `EDA2` option to "2", which indicates the use of the original ALMO scheme (instead of FERF) for the decomposition of POL and CT

- Add the following keywords to the `$rem` section:

  - `EDA_PCT_A = 1` (doing perturbative charge-transfer analysis)

  - `EDA_COVP = TRUE` (turn on the COVP analysis)

  - `EDA_PRINT_COVP = AUTOMATED` (this enables the automated generation and selection of the most significant COVPs)

  - `PLOTS = TRUE`

- Add a `$plots` section to the input to set up a plotting grid for the cube files:

      $plots
       grid_points 80 80 80
      $end

The calculation is ready to go. Complete the following once it is finished:

- The POL and CT terms from this calculation should be different from those under the "Gas phase" column in **Table 3**. Based on the comparison, did the switch from FERF (`eda2 = 1`) to the original ALMO scheme (`eda2 = 2`) make a significant difference? **(0.25 pt)**

- Based on the results under "`Perturbative CT Analysis`", is the charge transfer primarily from Cl^- to H_2O or the opposite? **(0.25 pt)**

  Hint: (i) You should look at the 2×2 matrices under "`PCT Energy lowering`" and "`PCT Charge displacement`", whose off-diagonal entries (1-2 and 2-1) correspond to the charge transfer between fragments; (ii) in our calculation fragment 1 is Cl^- and fragment 2 is H_2O

- From the genereated `.cube` files in the `[jobname].plots` folder, plot the most significant COVPs for the $Cl^-\rightarrow H_2O$ and $H_2O\rightarrow Cl^-$ charge transfer. The indexes of the two orbitals in a COVP can be found under the section "`COVP Analysis for Charge Transfer`" in the output file, which should be the same as the number `x` in the names of the cube files (`covp_a_x.cube`). Specific requirements on the graphics:

  - Use isovalue 0.05 to generate the plots

– Each pair of orbitals should be plotted in the same figure. For the acceptor orbitals, use "`Lines`" instead of "`Fill`" which can be set by double-clicking "`Cube Data`" under Model View.

Submit the genereated COVP plots (2 pairs, 4 orbitals) with your report **(0.25 pt)**. Based on the visualization, describe the character of the orbital of Cl^- when it serves as the donor and acceptor in the two pairs. **(0.25 pt)**

Exercise 3: Decomposing the vibrational frequency shift of CO in the OBe$\cdots$CO complex

In this exercise, we are going to use the adiabatic ALMO-EDA method to investigate the change in the vibrational frequency of CO when it is bound to BeO. All the calculations will be performed at the ωB97X-D/def2-TZVPP level of theory.

Let's start from an "Opt + Freq" calculation for the CO molecule using ωB97X-D/def2-TZVPP. To get a good initial structure, you can draw a CO molecule in IQmol with 3 bonds connecting C and O, and then hit the "E↓" button to minimize the energy. Here I will no longer repeat the procedure to set up an "Opt + Freq" calculation: you can refer to the pKa exercise in Practice 4 for the details. Report the optimized bond length and frequency of CO on the right of "Free CO" in **Table 4**. **(0.25 pt)**

Table 4: Adiabatic ALMO-EDA results for the OC$\cdots$BeO complex using ωB97X-D/def2-TZVPP

Stage	CO bond length (Å)	CO frequency (cm^{-1})
Free CO		
FRZ		
POL		
Full		

The input file for the adiabatic EDA calculation has been provided (`OC_BeO_adiabatic_eda.in`), which contains 6 consecutive Q-Chem jobs:

- Jobs 1 & 2: Opt + Freq calculation for the complex on the full SCF surface

- Jobs 3 & 4: Opt + Freq on the polarized surface (with `POL_GEOM = TRUE`)

- Jobs 5 & 6: Opt + Freq on the frozen surface (with `FRZ_GEOM = TRUE`)

Take a look at the input file and make sure you understand the above. Execute the input on the server. Once finished, open up the output using IQmol. You will find three sets of "`Frequencies`" under the Model View, which, from top to bottom, correspond to the frequencies calculated on the FULL, POL, and FRZ surfaces. Now find the vibrational mode that correspond to the CO stretch (which should not be too far away from the free CO stretch frequency) and report them in the right places in **Table 4**. **(0.5 pt)**

To complete **Table 4**, you now need to fill in the optimized C−O bond length on each of these surfaces. To find those, you can search for "`CONVERGED`" in the output (there are 3 occurrences in total), under which you can find the Cartesian coordinates of the optimized structure followed by the `Z-matrix print`. From the Z-matrices you can easily find out the C−O distance in the optimized structures. You may also save the Cartesian coordinates as XYZ files and use IQmol to open them and measure the bond lengths. **(0.25 pt)**

Answer the following questions based on the results you obtained in **Table 4**:

1. Is the CO frequency in the fully relaxed complex red- or blue-shifted compared to the free CO stretch frequency? Then describe the respective contributions from frozen interaction, polarization, and charge transfer to the net frequency shift relative to free CO. **(0.25 pt)**

2. Are CO frequencies correlated to the bond lengths at each stage? If so, describe the correlation/trend you observe. **(0.25 pt)**