

# CHEM-596 Practice Session 4: Chemical reactions in solvation environment; excited-state calculation and analysis using TDDFT

Name: \_\_\_\_\_

Due: April 5, 2023 (11:59 pm)

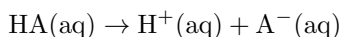
In this practice, you will first get some further practice on studying the thermodynamics of chemical reactions by predicting the pKa's of  $\text{CH}_3\text{COOH}$  and  $\text{CF}_3\text{COOH}$ , in which a solvent model is essential. You will then practice TDDFT calculations and excited-state analysis with two molecules: (i) pyridine; (ii) dimethylaminobenzonitrile (DMABN).

In the last practice session you have gained some experience in using the Linux servers. In this practice, you should use the servers for **Exercise 1 & 3**, while for **Exercise 2** it is up to your choice. You can still use IQmol to create the input files for you though. Just as last time, in order to evenly distribute the workload on the two servers, **please use the ITCOTCK server if your last name initial is A–J and use the DGX server otherwise** for the calculations involved in this practice. You are encouraged to run Q-Chem calculations with 4-thread OpenMP parallelization using `"qchem -nt 4 [input] [output] & "`. Some regulations regarding the cluster usage have been posted on Canvas (under "Pages").

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.

## Exercise 1: Predicting the pKa values of acetic acid and trifluoroacetic acid

Acetic acid ( $\text{CH}_3\text{COOH}$ ) and trifluoroacetic acid ( $\text{CF}_3\text{COOH}$ ) are two common organic acids. Their acidity can be measured using their pKa values, which is defined as the negative logarithm of the equilibrium constant ( $K^\circ$ ) of the dissociation reaction:



where "aq" stands for aqueous solution. Based on

$$\Delta G^\circ = -RT \ln K^\circ = -2.303RT \log_{10} K^\circ, \quad (1)$$

we can derive that

$$\text{pKa} = \Delta G^\circ / 2.303RT \quad (2)$$

If we convert  $\Delta G^\circ$  to be in kcal/mol, then at 298.15 K we have  $RT = 0.592$  kcal/mol. Therefore, the problem of pKa prediction is essentially evaluation of the Gibbs free energy change of the acid dissociation reaction in aqueous solution. Here we employ the direct method, in which the electronic energy and thermodynamic corrections will all be calculated under a solvation environment. The theoretical model we employ here is M06-2X/6-31+G(d)/CPCM. For the C-PCM model, we set the dielectric constant  $\epsilon$  to 78.39 for water.

We will start with calculating the pKa of  $\text{CH}_3\text{COOH}$ . Note that the Gibbs free energy of proton ( $\text{H}^+$ ) in water is available as a constant:  $G(\text{H}^+, \text{aq}) = -270.273$  kcal/mol. Therefore, we only need to calculate the Gibbs free energy of  $\text{CH}_3\text{COOH}$  and its conjugate base  $\text{CH}_3\text{COO}^-$ :

**Table 1:** Worksheet to determine the pKa values of CH<sub>3</sub>COOH and CF<sub>3</sub>COOH at the M06-2X/6-31+G(d)/CPCM level of theory

	CH <sub>3</sub> COOH(aq)	CH <sub>3</sub> COO <sup>-</sup> (aq)	CF <sub>3</sub> COOH(aq)	CF <sub>3</sub> COO <sup>-</sup> (aq)
$E_e + E_{\text{solv}}$ (a.u.)				
$H_{\text{vib}}$ (kcal/mol)				
$S_{\text{vib}}$ [cal/(mol·K)]				
$G^\circ$ (kcal/mol)				
$\Delta G^\circ$ (kcal/mol)		—		—
pKa		—		—

- For the starting geometry of CH<sub>3</sub>COOH, we can load the built-in geometry through “Add Fragment → Molecules → Carboxylic Acids → Acetic Acid”. Set up a geometry optimization job at the M06-2X/6-31+G(d) level, and then apply the C-PCM solvation model by selecting “Advanced → Solvent Model” and setting the “Solvent Method” to “PCM”. **Note:** you may need to select “PCM” *twice* to fully apply the model. After doing that, you should see lines

```
$pcm
  THEORY  CPCM
$end

$solvent
  DIELECTRIC  78.39
$end
```

appear at the bottom of your input. The dielectric constant for water has been loaded by default so you don’t need to change any other parameters.

- Now add a second job to perform a vibrational frequency calculation at the optimized geometry (remember to change the `jobtype`). Typically the keywords for the setup of the solvent model in the first job should be carried over to the second by IQmol, but if not, just set up the C-PCM model for water again for the frequency calculation.
- Now you can paste the generated input to an input file on the server and execute your calculation there. Once successfully finished, by searching for “**Final energy**” in the output as last time you will obtain the  $E_e + E_{\text{solv}}$  value for CH<sub>3</sub>COOH(aq). For the thermodynamic contribution to enthalpy and entropy, you need to do something slightly different from last time: instead of reporting the “**Total Enthalpy**” and “**Total Entropy**”, you should only include the *vibrational* contribution here as the translational and rotational contributions as derived in the gas phase do not apply here. Evaluate

$$G^\circ = E_e + E_{\text{solv}} + H_{\text{vib}} - TS_{\text{vib}} \quad (3)$$

at  $T = 298.15$  K.

- For CH<sub>3</sub>COO<sup>-</sup>, you can just delete the proton from the either optimized or unoptimized CH<sub>3</sub>COOH and repeat the entire procedure as documented above. Make sure that you set charge to  $-1$ , i.e., the line in the input for charge and multiplicity should be “ $-1$  1”.

**Report the computational results for CH<sub>3</sub>COOH and CH<sub>3</sub>COO<sup>-</sup> in Table 1 and calculate the pKa value of acetic acid using Eq. (2). (1 pt)**

For  $\text{CF}_3\text{COOH}$ , you should start from the provided `TfAcOH.xyz` downloaded from the CCCBDB database rather than the built-in trifluoroacetic acid structure in IQmol since the latter turns out to be an energetically slightly unfavorable rotamer. The rest of the procedure should be pretty much *the same* as for the acetic acid. **Complete the last two columns of Table 1 (0.5 pt) and then answer the following questions:**

- The experimental pKa values for acetic and trifluoroacetic acids are 4.76 and 0.52, respectively. Did your calculations predict the correct qualitative trend for their acidity (ordering)? **(0.25 pt)** Quantitatively, how many pKa units do your computational results differ from the experimental values? Quantify the error for each species. **(0.25 pt)**
- One possible source of error is the employed electronic structure method. Just like what we did for the reaction barrier of the Diels-Alder reaction in our last practice, here we are going to refine the  $E_e + E_{\text{solv}}$  term at the  $\omega\text{B97M-V}/\text{def2-TZVPD}/\text{SMD}$  level of theory, which comprises more accurate density functional, larger basis set, and more sophisticated solvent model. Using **Table 2** below as a worksheet, calculate the pKa values using the refined  $E_e + E_{\text{solv}}$  values and discuss whether this has improved the accuracy for pKa prediction or not. **(0.5 pt)**

**Note:** (i) you only need to do a *single-point* calculation at the *optimized* structure of each species and the  $H_{\text{vib}}$  and  $S_{\text{vib}}$  contributions should stay unchanged; (ii) to set up the SMD model, go to “Advanced → Solvent Method” and select “SMD” under “Solvent Method”: you may also need to do it twice to see these lines appear at the bottom of your input:

```
$smx
  SOLVENT  water
$end
```

- **(0.5 bonus point)** Discuss what might be the major source of error in pKa prediction and what might be the solutions?

**Table 2:** Worksheet to determine the pKa values of  $\text{CH}_3\text{COOH}$  and  $\text{CF}_3\text{COOH}$  with the refinement at the  $\omega\text{B97M-V}/\text{def2-TZVPD}/\text{SMD}$  level

	$\text{CH}_3\text{COOH}(\text{aq})$	$\text{CH}_3\text{COO}^-(\text{aq})$	$\text{CF}_3\text{COOH}(\text{aq})$	$\text{CF}_3\text{COO}^-(\text{aq})$
$E_e + E_{\text{solv}}$ (a.u.)				
$G^\circ$ (kcal/mol)				
$\Delta G^\circ$ (kcal/mol)		—		—
pKa		—		—

## Exercise 2: Singlet and triplet excited states of pyridine via TDDFT calculations

In this practice, we will use time-dependent density functional theory (TDDFT) to investigate the singlet and triplet excited states of pyridine. Note that the Tamm-Dancoff approximation (TDA) will be used throughout this exercise, which is the default option in Q-Chem. All the excited state calculations should be done based on the given structure `pyridine.xyz`, i.e., no geometry optimization should be done in this exercise.

We will start with calculating **8 singlet states** using  $\omega$ B97X-D/aug-cc-pVDZ. Using IQmol, the job can be set up as follows:

- Open up the XYZ file and go straight to “Q-Chem Setup”. Under Method, find and select “TD-DFT”. Once you did that, some additional job control options should appear under “CIS/TD-DFT” at the left bottom of the control panel. Now you should select “omegaB97X-D” under “Exchange” and change the basis set to aug-cc-pVDZ. (**Note:** “Exchange = omegaB97X-D” + “Correlation = None” is the old way to invoke the  $\omega$ B97X-D functional in Q-Chem inputs, which has been deprecated but still works though.)
- Under the “CIS/TD-DFT” options, you should unselect “Triplets” and set “Number of Roots” to 8. We will need to use natural transition orbitals to identify the characters of these excited states later, so you should set “NTO PAIRS” to 2. Lastly, set “Convergence” to 8 under SCF control. The generated input file should look like below:

```
$molecule
0 1
C    0.0000000    0.0000000   -1.4099987
C    0.0000000    1.1932681   -0.6988838
C    0.0000000   -1.1932681   -0.6988838
C    0.0000000    1.1398403    0.6914739
C    0.0000000   -1.1398403    0.6914739
N    0.0000000    0.0000000    1.3905662
H    0.0000000    0.0000000   -2.4905277
H    0.0000000    2.1472345   -1.2045420
H    0.0000000   -2.1472345   -1.2045420
H    0.0000000    2.0535202    1.2718329
H    0.0000000   -2.0535202    1.2718329
$end

$rem
BASIS = aug-cc-pVDZ
CIS_N_ROOTS = 8
CIS_TRIPLETS = 0
EXCHANGE = omegaB97X-D
GUI = 2
NTO_PAIRS = 2
SCF_CONVERGENCE = 8
$end
```

- You can submit this job to the Q-Chem server (which will make the following analysis slightly easier), which will take  $\sim 5$  minutes to finish.

Once the TDDFT calculation finishes, you should check out the excited state information in the output file (some of the information can also be viewed by double-clicking “Excited States” under Model View (the left column)). The character of each state can be visualized by plotting the HONTO and LUNTO of each excited state, which can be found under “Surfaces”. By visualizing the NTOs, you should be able to identify the character of each excited state as  $n \rightarrow \pi^*$ ,  $\pi \rightarrow \pi^*$ , or Rydberg (typically featured by very diffuse virtual orbitals involved in the excitation). **Collect the information of the  $n \rightarrow \pi^*$  and  $\pi \rightarrow \pi^*$  excited states in Table 3 (1 pt).**

**Note:** if Table 3 has too many rows, just leave the extra ones blank; if you need more rows, add more rows at the bottom.

Now **Complete the following:**

- Which type of states are of lower excitation energy? which type of states are brighter (i.e., with larger oscillator strength)? **(0.25 pt)**

- Save the plots of the dominant NTO pairs (HONTO and LUNTO) for the  $n \rightarrow \pi^*$  states and **submit them with your report. (0.25 pt)**
- Generated an absorption spectrum **in nm** using IQmol. To do that, first make sure that your IQmol version is 3.1.2 (the older version has a compatibility issue for this feature) and that you have your output file opened. Double-click on “Excited States” and change the unit to nm. Make sure that the spectral lines are roughly centered on the spectrum (you can adjust by moving left/right or zooming in/out on the plotting area). Finally, set broadening method to “Gaussian”. **Save the spectrum plot and submit that with your report (0.25 pt).**

**Table 3:** Information about the singlet valence excitations of pyridine calculated using  $\omega$ B97X-D/aug-cc-pVDZ. Keep 4 digits after decimal point for both excitation energies and oscillator strengths. Under “State character” you should write “ $n \rightarrow \pi^*$ ” or “ $\pi \rightarrow \pi^*$ ”, and under “Dominant Orbitals” you should write something like “H $\rightarrow$ L”, “H-1 $\rightarrow$ L+1”, etc.

State index	Excitation energy (eV)	Oscillator strength	State character	Dominant Orbitals

Let’s now set up another Q-Chem job to calculate **6 triplet states**. To do that, you need need to change the line “CIS\_TRIPLETS = 0” above to “CIS\_SINGLETS = 0” and change the value of CIS\_NROOTS to 6 (of course, you can use IQmol to set this up from opening up the XYZ file). Run this calculation with two different density functionals: (1) B3LYP; (2)  $\omega$ B97X-D. For each obtained excited state (all should be valence excitations), identify their characters ( $n \rightarrow \pi^*$  or  $\pi \rightarrow \pi^*$ ) and then report the excitation energies based on the order in **Table 4: two  $n \rightarrow \pi^*$  states first, followed by four  $\pi \rightarrow \pi^*$  states. (1 pt)** Then complete the following:

1. Based on the reference values, calculate the root-mean-square errors ( $\text{RMSE} = \sqrt{\sum_{i=1}^N (y_i - \hat{y}_i)^2 / N}$ ) of B3LYP and  $\omega$ B97X-D **for these 6 triplet states** and report the values in **Table 4. (0.25 pt)** Based on the RMSEs, which functional is more accurate for pyridine’s first 6 triplet excited states? **(0.25 pt)**
2. Compared to the reference data, discuss whether the two functionals manage to predict the correct ordering of excited states. **(0.25 pt)**
3. Comparing the singlet and triplet excitation energies for the two  $n \rightarrow \pi^*$  states, are the triplet states of higher or lower energies? Does that agree your expectation? **(0.25 pt)**

**Table 4:** Calculation of 6 triplet excited states of pyridine ( $\omega$  stands for excitation energy)

State character	$\omega$ (B3LYP) (eV)	$\omega$ ( $\omega$ B97X-D) (eV)	$\omega$ (ref) (eV)
$n \rightarrow \pi^*(1)$			4.46
$n \rightarrow \pi^*(2)$			5.36
$\pi \rightarrow \pi^*(1)$			4.30
$\pi \rightarrow \pi^*(2)$			4.79
$\pi \rightarrow \pi^*(3)$			5.04
$\pi \rightarrow \pi^*(4)$			6.24
RMSE			—

### Exercise 3: Intramolecular charge-transfer states for twisted DMABN

Dimethylaminobenzonitrile(DMABN) is a commonly used model system for studying twisted intramolecular charge transfer (TICT) states. In this practice, you will use three different functionals to calculate the **two lowest singlet excited states** of *twisted* DMABN (of  $C_{2v}$  symmetry). You will also practice using Q-Chem's `libwfa` module to perform wavefunction analysis and visualization of excited states.

Since you have gained some experience in setting up TDDFT calculations from **Exercise 2**, I will provide less instructions here. First, with the given XYZ file "`twisted_DMABN.xyz`", use B3LYP/aug-cc-pVDZ to calculate 4 singlet excited states and report the excitation energies for the first two states in Table 5. **(0.25 pt)** Keep `NTO_PAIRS = 2` on in your input so that you can visualize the excited states you obtain. **IMPORTANT:** make sure that you put a keyword `symmetry = false` in your input file to turn off integral symmetry otherwise you may get unpleasant oscillatory behavior near the end of your SCF calculation making it fail to converge.

Based on the NTOs you just obtained, using your words to describe the charge transfer character associated with the  $S_1$  and  $S_2$  states (just imagine what you will write if you are discussing these excited states in a paper, e.g., electrons going from moiety X to functional group Y). **(0.25 pt)**

Based on the class materials, you should be aware that B3LYP is a global functional of 20% exact exchange. Now perform the same calculation with BLYP (semi-local GGA) and CAM-B3LYP (range-separated hybrid) and report the  $S_1$  and  $S_2$  excitation energies in the same table. **(0.25 pt)**

**Table 5:** The excitation energies (in eV) of the  $S_1$  and  $S_2$  states of DMABN

	BLYP	B3LYP	CAM-B3LYP	Reference
$S_1$				4.11
$S_2$				4.74

Compared to the reference values, did the results get better or worse with these two functionals (relative to B3LYP)? **(0.25 pt)** Rationalize the trend you observe based on the discussion of TDDFT for charge-transfer excited states in **Lecture 13**. **(0.25 pt)**

To wrap up this exercise we will perform some excited-state wavefunction analyses based on the CAM-B3LYP results. You can make a copy of the input you have for the CAM-B3LYP calculation and then do the following:

- Delete `NT0_PAIRS = 2` and `GUI = 2`
- Add the following 4 keywords to the `$rem` section:
  - `STATE_ANALYSIS = TRUE`
  - `MAKE_CUBE_FILES = TRUE`
  - `PLOTS = TRUE`
  - `CIS_RELAXED_DENSITY = TRUE`
- Add the following section to the bottom of your input to set up a plotting grid:

```
$plots
    grid_points 80 80 80
$end
```

You can then run the calculation on the server. On the job is done, a folder “[jobname].plots” will be generated with the output, under which you can find a big number of .cube files organized as S\_1, S\_2, ...

**Report the dipole moments (in Debye) for the S<sub>1</sub> and S<sub>2</sub> states calculated using their *relaxed* densities in Table 6.** To find that, first search for the line “Analysis of Relaxed Density Matrices” in your output; then under “Singlet 1 (relaxed)” and “Singlet 2 (relaxed)”, find “Dipole moment [D]:”. The number after corresponds to the total magnitude of dipole moment which you should report. In the same table also **report the ground state dipole** (search for “Ground State (Reference)” and scroll down). **(0.25 pt)**

Note: the difference dipole between ground and excited states for charge-transfer excitations are usually quite substantial.

**Table 6:** Dipole moments for the ground and S<sub>1</sub> and S<sub>2</sub> excited states of DMABN

	Ground State	S <sub>1</sub>	S <sub>2</sub>
Dipole			

Finally, visualize the attachment and detachment densities for the S<sub>1</sub> and S<sub>2</sub> states and **submit the plots with your report (0.25 pt)**. Make sure that

1. The files you visualize are `S_1.rlx.detach.cube`, `S_1.rlx.attach.cube`, `S_2.rlx.detach.cube`, and `S_2.rlx.attach.cube`
2. Use a 10× smaller isosurface value (0.002) to generate these plots
3. The detachment densities are in blue (ideally somewhat transparent) and attachment densities in red so that they can be distinguished. You can use the “Swap Colors” feature (double-click “Cube Data” under Model View) in IQmol to change the colors if IQmol didn’t give you the correct color in the first place.
4. After you are done with plotting, delete the “[jobname].plots” folder on the server using the “`rm -r [jobname].plots`” command because those cube files take a lot of disk space.