

# CHEM-596 Practice Session 3: Vibrational Frequency Analysis and Modeling of Chemical Reactions

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

Due: March 14, 2023 (11:59 pm)

In this practice, you will use the DFT methods available in Q-Chem to (i) compute harmonic frequencies and IR spectrum of the CO<sub>2</sub> molecule; (ii) locate the transition state and calculate the reaction (free) energy barrier for a Diels-Alder reaction.

Besides running Q-Chem calculations through IQmol, you will also get a chance to practice running Q-Chem on a Linux command line (the last set of calculations in **Exercise 2**). 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 this practice session only). You are encouraged to run Q-Chem calculations with 4-thread OpenMP parallelization. The command:

```
qchem -nt 4 [input] [output] &
```

where the “&” at the end allows you to do other things on the terminal when the job is running. For jobs involved in this practice, please have *only one* of them running on the server at a time.

You will need to **complete the tables and short-answer questions on this handout** and also prepare the requested files for submission. 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: Vibrational frequencies of CO<sub>2</sub>

In this practice, you will compute the harmonic vibrational frequencies of the CO<sub>2</sub> molecule using two different “model chemistries” (B3LYP/6-31+G(d) and r2SCAN/def2-TZVP) and compare the results to theoretical references and experimental fundamental frequencies. You will then generate a computed IR spectrum using one of these methods.

While the *theoretical best geometry* (“TBG”) of CO<sub>2</sub> is provided, vibrational frequency analyses are most commonly performed using the geometry optimized at the same level of theory as for frequency analysis (“OptG”). Let’s start with obtaining the “OptG” frequencies using B3LYP/6-31+G(d):

- Open “CO2\_TBG.xyz”. Set up a geometry optimization calculation (“Calculate → Geometry”) using the B3LYP functional and the 6-31+G(d) basis (equivalent to 6-31+G\*).
- Click the green “+” button to add a 2nd job. The \$molecule section should just contain “read”. Now set “Calculate → Frequencies”. Then under “SCF Control”, change “Guess” to “READ”, which will read in the converged SCF solution for the last (optimized) geometry from the first job.
- Submit the job. Once finished, you can find the computed harmonic frequency for each mode under “Frequencies” on the left column of the IQmol window (“Model View”). Report the frequencies under “OptG” for the corresponding method in **Table 1. (0.25 pt)**

Repeat the same procedure with “METHOD = r2SCAN” and “BASIS = def2-TZVP” and report the obtained frequencies in **Table 1. (0.25 pt)** Now, visualize the vibrational mode corresponding to each frequency based on the r2SCAN/def2-TZVP results (you just need to double-click each frequency under Model View). **Describe the motions in words: (0.25 pt)**

The definition of *root-mean-square error* (RMSE) is as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (1)$$

where each  $i$  corresponds to an individual data point and  $N$  is the total number of data points;  $y_i$  and  $\hat{y}_i$  are the observed and true values for each data point, respectively. Using this formula, calculate the RMSEs of the “OptG” frequencies of those two methods, with each measured against the *theoretical reference values for harmonic frequencies* (the last column of **Table 1**). Report the two “OptG” RMSEs in the last row of **Table 1**. (0.25 pt)

**Note:** (i) consider  $\omega_1$  and  $\omega_2$  as one data point as they are degenerate modes; (ii) you should try to do the RMSE calculation using Excel.

**Table 1:** Harmonic frequencies of  $\text{CO}_2$  (in  $\text{cm}^{-1}$ ) calculated using B3LYP/6-31+G(d) and r2SCAN/def2-TZVP at the optimized geometry of each theory (“OptG”) and the theoretical best geometry (“TBG”). The DFT frequencies are *unscaled* and compared to the high-level reference values for the harmonic frequencies

	B3LYP/6-31+G(d)		r2SCAN/def2-TZVP		Theor. Ref.
	OptG	TBG	OptG	TBG	
$\omega_{1,2}$					672.9
$\omega_3$					1353.8
$\omega_4$					2393.5
RMSE					—

In a DFT benchmark paper for vibrational frequencies discussed in class ([Liang and Head-Gordon et al.](#)), it was suggested that the errors in DFT’s prediction of harmonic frequencies may be reduced by using the theoretical best geometries (TBGs). Let’s apply the idea here. Since the TBG for  $\text{CO}_2$  has been given (C02\_TB.G.xyz), to get the results you only need to set up frequency calculations with the two methods *without* optimizing the geometry first. Report the TBG frequencies calculated using B3LYP/6-31+G(d) and r2SCAN/def2-TZVP in the corresponding columns in **Table 1** and calculate the RMSEs. (0.5 pt)

Based on the RMSEs, did using the TBG of  $\text{CO}_2$  help improve the accuracy of the frequency predictions? **Explain your answer.** (0.25 pt)

As introduced in the lecture, the calculated harmonic frequencies need to be *scaled* to make them comparable to experimental frequencies. The scaling factor for B3LYP/6-31+G(d) can be found on the [CCCBDB website](#) (note: 6-31+G(d) is identical to 6-31+G(d,p) for C and O atoms), while the factor for r2SCAN/def2-TZVP has been provided in **Table 2**. Report the scaling factor for B3LYP as well as the *scaled* frequencies of the two methods evaluated at “OptG” then calculate the RMSEs relative to the experimental reference values. The results should all be collected in **Table 2**. (0.5 pt)

Based on the RMSEs, which method combined with a scaling factor works better in predicting the experimental frequencies? Use data to support your answer. (0.25 pt)

As the final task in **Exercise 1**, you are supposed to generate an IR spectrum using the superior method you identified above. Steps to follow:

- Under **Model View** for the corresponding job (with “OptG”), double-click on “Frequencies”, a new window with title “Vibrational Frequencies” should pop up.

**Table 2:** Scaled DFT frequencies (in  $\text{cm}^{-1}$ , evaluated at OptG) compared to the experimental values

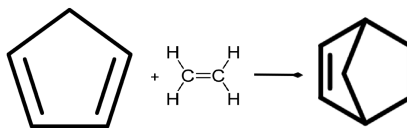
	B3LYP/6-31+G(d)	r2SCAN/def2-TZVP	Expt. Ref.
scale factor used		0.967	—
$\omega_{1,2}$			672.9
$\omega_3$			1332.9
$\omega_4$			2349.1
RMSE			—

- On the left bottom of the window, set the **Scale Factor** to be the same as in **Table 2**.
- Under the generated spectra on the right of the window, select the “**Lorentzian**” line shape. Now you should see the spectral lines are broadened.
- Save the spectrum as an image, which needs to be submitted with your report. **(0.25 pt)**

Looking at calculated IR spectrum for  $\text{CO}_2$ , how many peaks are visible? Which vibrational mode shows no IR signal? Explain why that is the case. **(0.5 pt)**

## Exercise 2: Predicting the reaction energy and barrier heights of a Diels-Alder reaction

In this exercise, you will be calculating the reaction (free) energy and the (free) energy barrier height for the Diels-Alder (D-A) reaction between ethylene and cyclopentadiene, which produces *norbornene*: Before



**Figure 1:** The D-A reaction to be investigated

the last set of questions (for **Table 4**), *all* the calculations involved in this exercise should be performed at the B3LYP-D3(BJ)/6-31G(d) level. Please *do not* forget to add the -D3(BJ) correction (on IQmol it is added by “**Advanced** → **DFT** → **Dispersion Correction** → **DFT-D3(BJ)**”). Here a very small basis set (6-31G\*) is employed to allow the jobs to finish within the 10 minutes limit on the IQmol cloud server. In real production calculations a larger basis set should be used.

First let’s determine the reaction free energy:  $\Delta G_{\text{rxn}} = G(\text{P}) - G(\text{R})$ . The equation that we employ to calculate the Gibbs free energy of each species is as follows:

$$G = E_e + H_{\text{t,r,v}} - TS_{\text{t,r,v}} \quad (2)$$

where  $E_e$  is the electronic internal energy of the reactant (R) or product (P),  $H_{\text{t,r,v}}$  and  $S_{\text{t,r,v}}$  are translation, rotation, and vibration’s contributions to *enthalpy* and *entropy*, respectively, and  $T$  is the temperature (use

**Table 3:** B3LYP/6-31G(d) results for the reactant, product, and transition state (TS) of the Diels-Alder reaction

	ethylene	cyclopentadiene	norbornene	Transition state
$E_e$ (a.u.)				
$H_{t,r,v}$ (kcal/mol)				
$S_{t,r,v}$ [cal/(mol·K)]				
$G$ (kcal/mol)				
$\Delta E_e$ (kcal/mol)	—	—		
$\Delta G$ (kcal/mol)	—	—		

298.15 K here). To obtain these quantities, we need to perform harmonic frequency calculations at *optimized* reactant and product structures:

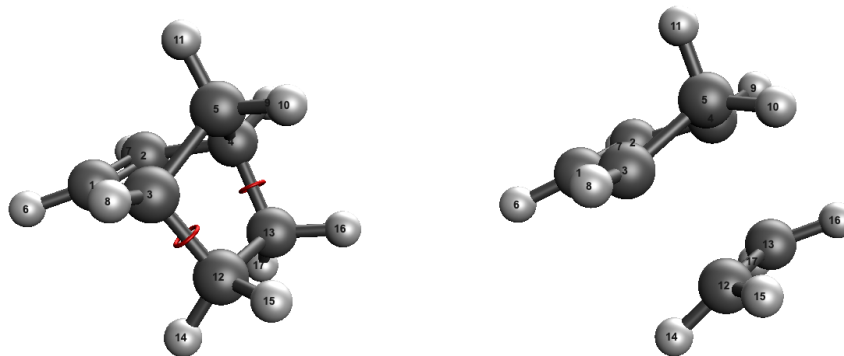
- Build the structure of *cyclopentadiene* using IQmol. Remember to check the number of hydrogens and use IQmol’s energy minimizer (“E↓”) to pre-optimize the structure.
- Set up geometry optimization followed by vibrational frequency calculation at the B3LYP-D3(BJ)/6-31G(d) level. The procedure is the same as how you set up the “OptG” frequency calculations in **Exercise 1**. Again, don’t forget to add the “-D3(BJ)” correction.
- You need to open the output text file to get all the numbers you need:  $E_e$  can be found by searching for “Final energy” (keep **6 digits** after decimal point), and the values of  $H_{t,r,v}$  and  $S_{t,r,v}$  can be found near the very end of the output after “Total Enthalpy” and “Total Entropy”.
- Repeat the same procedure for *ethylene* and *norbornene*. While the calculation for ethylene is fast, the “Opt + Freq” for norbornene can take 8–9 minutes on the IQmol server simply due to its relatively large number of atoms. To get the results faster, you can also choose to do the norbornene calculation on Linux server.

After collecting the data for the reactant and product species (**0.75 pt**), calculate the total free energy (in kcal/mol) for each compound at 298.15 K using Eq. (2) and then evaluate the values of  $\Delta E_{e, \text{rxn}}$  and  $\Delta G_{\text{rxn}}$  for this D-A reaction. Report the numbers in **Table 3**. (**0.5 pt**)

**Note:** (i) 1 a.u. = 627.5095 kcal/mol; (ii) the unit for entropy reported by Q-Chem is in cal/(mol·K) rather than kcal/(mol·K).

Up to this point, you should have already completed columns 2–4 in **Table 3**. Let’s now locate the transition state (TS). The D-A reaction involves the formation of two new bonds. As introduced in lecture, at the TS the length of the newly formed bonds should be elongated by 40–50% compared to equilibrium bond lengths. If we use the lower bound (40%), then the estimated C···C distance at the TS should be close to 2.2 Å. Based on this estimation, let’s perform a *constrained geometry optimization* to generate a good guess structure for TS search:

- We should start from the optimized norbornene geometry. Identify the two C–C bonds that are formed in the D-A reaction (whose bond lengths should be close to 1.56 Å).
- Set up constraint on the lengths of these two bonds. Select the two C atoms for one bond, then click “Build → Set Geometric Constraint”. On the pop-up window, you should choose “Constrain”, and then enter the value “2.2 Å”. Repeat the same for the other C–C bond. After setting up the constraints, your molecule should look similar to that on the left panel of **Figure 2**.



**Figure 2:** IQmol view of (i) the norbornene structure after geometry constraints are set (left) and (ii) resultant structure of the constrained optimization.

- Go to “Q-Chem Setup”. The first thing you should do is to hit “Reset”, which will load the optimized structure of norbornene (instead of the initial structure of the Opt job). Then set “Calculate → Geometry”. Once you did that, the following lines should appear in the generated input file:

```
$opt
CONSTRAINT
stre 3 12 2.2
stre 4 13 2.2
ENDCONSTRAINT
$end
```

This indicates that the geometry constraints are ready in place. **Note:** the indexes of the bonding atoms may be different in your molecule depending on how the molecules were built initially.

- Now you just need to set up the correct **Method** and **Basis** to use and then submit the calculation. On the IQmol server, this calculation may take 6–8 minutes to complete.

After the constrained optimization is done, you should confirm that the lengths of the two C–C bonds are 2.2 Å. **Save the XYZ coordinates obtained from constrained optimization and submit it with your report. (0.25 pt)** You should expect to see something similar to the right panel of **Figure 2**.

This constrained optimization should give us a good guess to the TS structure. Starting from that, we are going to set up a frequency calculation first to generate the Hessian at the initial structure, and then perform a TS search:

- At the output of the constrained optimization job that was copied back, go to “Q-Chem Setup”. Click “Reset” to load the final structure of the constrained Opt.
- Set up a vibrational frequency calculation using B3LYP-D3(BJ)/6-31G(d) first (don’t forget the D3). Then add a second job, set the job type to “TS” (Calculate → Transition State). You also need to have the 2nd job read in the Hessian computed by the 1st job. To set that, go to “Advanced → Geometry Optimization”. Set “Initial Hessian” to “None” first and then set it back to “READ” (this way to get it work looks pretty weird, which might be a small glitch of IQmol). In the end, you should make sure that “GEOM\_OPT\_HESSIAN = READ” appears in the generated input.
- Submit the TS search job. This will take 5–6 minutes to finish on the IQmol server. Report the electronic energy ( $E_e$ ) of the optimized TS structure and evaluate the reaction energy barrier in **Table 3**:  $\Delta E_e^\ddagger = E_e(\text{TS}) - E_e(\text{reactants})$  **(0.5 pt)**
- To verify that you successfully located a TS, you need to do another vibrational frequency calculation at the optimized TS structure. Make sure that you are currently at the TS optimization job under Model

View, then go to “Q-Chem Setup” again. Click “Reset” first (to load the optimized TS structure) and then set up a simple frequency calculation using B3LYP-D3(BJ)/6-31G(d) *without* doing any optimization. Once finished, check out the frequencies first. Based on the number of imaginary frequencies, is this structure a valid TS? **(0.25 pt)**

Then double-click on the first imaginary mode to visualize the vibration. **Describe the motion using text below.** Based on the visualization, is this imaginary mode relevant to the D-A reaction? **(0.25 pt)**

- From the output of the frequency calculation at the optimized TS, retrieve the enthalpy and entropy corrections and report them in the last column of Table 3. **(0.25 pt)** Calculate the Gibbs free energy of the TS and the activation free energy of this reaction at 298.15 K:  $\Delta G^\ddagger = G(\text{TS}) - G(\text{reactants})$ . Fill your results in **Table 3**. **(0.25 pt)**

Based on your computational results in **Table 3**, sketch a *reaction free energy diagram* below, in which you should use 3 *horizontal lines* to represent reactant (R), transition state (TS), and product (P), and 2 *double-headed arrows* to indicate  $\Delta G_{\text{rxn}}$  and  $\Delta G^\ddagger$ . **(0.5 pt)**

Answer the following questions based on your DFT calculations:

1. Is this reaction endothermic or exothermic? **(0.25 pt)**

**Hint:** To answer this question you should check the sign of  $\Delta H_{\text{rxn}}$ , which includes contributions from  $E_e$  and  $H_{\text{t,r,v}}$

2. Describe the effect of *entropy* on the thermodynamics (more/less spontaneous) and kinetics (faster/slower) of this reaction. **(0.25 pt)**
3. The theoretical reference for the energy barrier height ( $\Delta E_e$ ) of this reaction is 18.3 kcal/mol (obtained using a highly accurate composite electronic structure method). It’s known that the B3LYP functional tends to underestimate the reaction energy barrier. Do you observe the same trend here? Based on our discussion of the remaining challenges of DFT in lecture, which issue is related to this underestimation? **(0.25 pt)**

It is very common practice to refine the energetic results using a more sophisticated electronic method at the structures optimized using a lower-tier but less costly method. We are now going to recalculate

the reaction energy barrier ( $\Delta E_e$ ) at the  $\omega$ B97M-V/def2-TZVPD level of theory (*without* reoptimizing the structures):

- Find the “Opt + Freq” job for *cyclopentadiene* (if you have closed IQmol in between you can simply use IQmol to open the output file). Click “Reset” to load the optimized geometry.
- Set up a single-point energy calculation with “METHOD =  $\omega$ B97M-V”, “BASIS = def2-TZVPD”. Set SCF\_CONVERGENCE = 8 and “GUI = 0”.
- This job may not be able to finish on the IQmol server within 10 minutes so you should **use the Linux server** instead. You can simply copy the input generated by IQmol and paste that into a new text file called “Cp\_ $\omega$ B97MV\_tzvpd.in”. You can directly generate the input on the remote server using the vi editor on the terminal, or create a text file on your local computer first and then upload that to the server. Then use the command provided on the first page to execute Q-Chem.
- After the job finishes, you can copy the output file back to your local computer and open it up using IQmol. You can also directly find the data (DFT energy) you need by executing the following command on terminal:

```
grep "Total energy" [output]
```

which will return

```
Total energy in the final basis set = [...]
```

**Table 4:** Recalculating the reaction energy barrier of the D-A reaction using  $\omega$ B97M-V/def2-TZVPD

	cyclopentadiene	ethylene	transition state
$E_e$ (a.u.)			
$\Delta E_e$ (kcal/mol)	—	—	

Repeating this for the optimized ethylene (another reactant) and transition state structures as well. Report the DFT energies in **Table 4** and calculate the  $\Delta E_e$ . Compared to the theoretical reference value (18.3 kcal/mol), is the accuracy for  $\Delta E_e$  significantly improved when  $\omega$ B97M-V/def2-TZVPD is employed? **(0.5 pt)**