DFT reaction energy calculations

Introduction

In this practical we are going to use density functional theory (DFT) to calculate the activation and reaction free energies ($\Delta G^{\ddagger}$ and ΔG^0, respectively) for the uncatalyzed transformation of chorismate to prephenate:
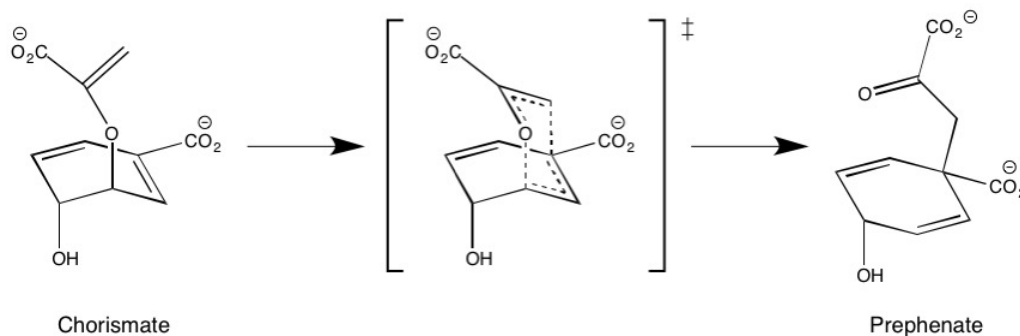

Figur 1: The unimolecular transformation of chorismate to prephenate.

These energies will later be used in the EVB reference reaction to calibrate the EVB Hamiltonian.

The reaction energies for the uncatalyzed transformation of chorismate to prephenate (Figure 1) in water are going to be calculated with DFT using Gaussian 16[1]. Structure optimization and frequency calculations at the reactant (chorismate), transition and product (prephenate) states are to be computed with the B3LYP functional[2] and the 6-31G(d,p) basis set. Dispersion effects must be included in all calculations using Grimme's B3LYP-D3 method[3,4]. Intrinsic reaction coordinate (IRC) calculations are to be performed in both directions from the TS to verify that the correct minima are connected (reactant and product state). The electronic energies will be obtained from single-point calculations on the optimized geometries (RS, TS and PS) at the b3lyp/6-311G+(2d,2p) level of theory. The final energies to be reported and used further for calibrating the EVB Hamiltonian must be corrected for zero-point energy (ZPE), dispersion and solvation effects that are calculated with the SMD model[5] using water as solvent (eps=80).

Recommended literature
- Siegbahn & Himo (2011) "The quantum chemical approach for modeling enzyme reactions"[6]
- Himo (2017) "Recent trend in quantum chemical modeling of enzymatic reactions"[7]

## Practical step by step

Setting up DFT calculations on Saga.
1. Log in to your user account on saga (alternative with TigerVNC)
   a. ssh -Y userXX@saga.sigma.no
   b. desktop.saga.sigma2.no:5901 (with tigervnc)
2. Clone the git repo:
   a. git clone https://github.com/isaksengeir/biomolmod

3. cd to biomolmod/practicals/DFT
4. Type: *module load GaussView/6.0.16*
5. Type: *gv6* (opens gaussView)
6. Normally you would now use you chemical intuition to generate a guess for the transition state. We have already done this for you, so you can open "CM_TS_guess.com" in gaussView (File → Open).
7. Set up the transition state calculation by going to "Calculate → Gaussian Calculation Setup" and make the following edits:
   a. Job type
      i. Job type: Optimization
      ii. Optimize to a: TS (Berny)
      iii. Calculate Force Constants: Once
         (Specifies that the force constants be computed at the first point using the current method)
   b. Method:
      i. Method: Ground state – DFT – Restricted – B3LYP
      ii. Basis Set: 6-31G (d,p)
      iii. Charge: -2   Spin: Singlet
   c. General:
      i. Tick the "Additional Print" box. All the other options that are clickable should here be unchecked.
   d. Additional keywords: empiricaldispersion=gd3 opt=noeigentest
      (opt=noeigentest suppresses testing the curvature in Berny optimizations. The test is on by default only for transition states in internal (Z-matrix) or Cartesian coordinates, for which it is recommended. Occasionally, transition state optimizations converge even if the test is not passed, but **NoEigenTest** is only recommended for those with large computing budgets, which is not a problem in our case.)

   The window should now look something like this:
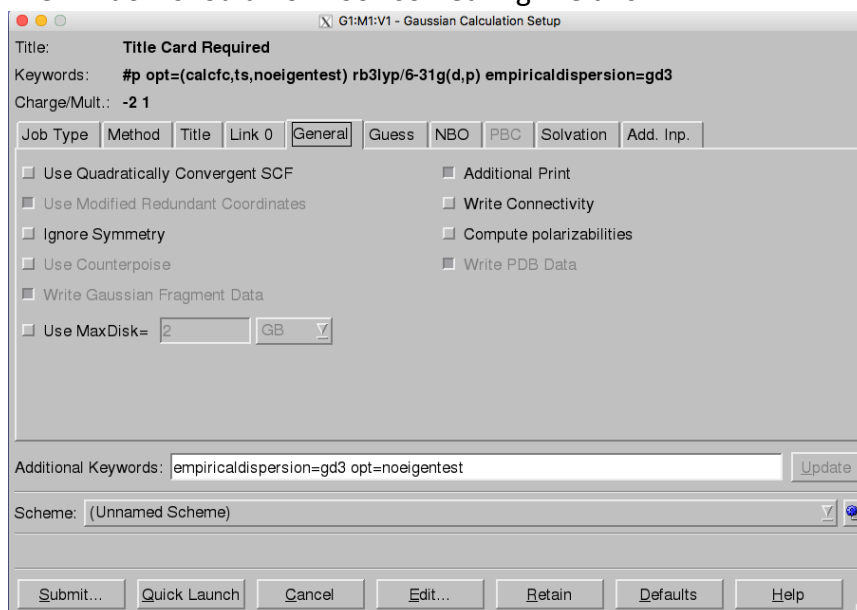
8. Press "Edit" and then "Save" as for example "CM_TS_opt.com"
9. Close GaussView

10. Submit the DFT calculation with gaussion on Saga with the g16runs.sh script by typing:
    a. g16run.sh filename nodes hours ← NOTE that <mark>filename is without extension</mark>!
        i. For example: *g16run.sh CM_TS_opt 1 1*
        ii. Note that the .com file extension is not to be used here.
        iii. The job takes approximately 15 minutes to finish,
11. You can view your jobs in the queue on stallo with the following command:
   a. squeue -u "you_username"

Analyzing the DFT geometry optimzation
1. When the calculation has finished, you will need make sure that it has converged. For this you will need to check the output file "CM_TS_opt.out". You can simply use "grep" for this:
   a. *grep -A 4 "Converged" CM_TS_opt.out*
      The job is converged if you see <u>4 lines with "YES"</u>.

	Item	Value	Threshold	Converged?
Maximum	Force	0.000013	0.000450	YES
RMS	Force	0.000002	0.000300	YES
Maximum	Displacement	0.000524	0.001800	YES
RMS	Displacement	0.000124	0.001200	YES

      If the job has not converged, you can continue the job by opening the output file in GaussView and starting from the one of the last frames (same as described above). One tips here is to standardize the Z-matrix to update coordinate list to avoid angles coming close to 0/180 before saving the new inputfile:
      Edit → Atom List Editor → Edit → Z-Matrix → Standardize

2. It is always important <u>to visually inspect your geometries</u>. You can inspect the geometry optimizations with Molden (or gaussView).
   a. Type "module load Molden/6.1-GCCcore" to use molden.
   b. Launch molden by typing molden
   c. Press "Read" and open CM_TS_opt.out
   d. Under draw mode – untick "shade" for better visualization.
   e. Under "Convergence" press "Geom.conv" – Here you can see the iterative optimization cycles with the corresponding energies. This is also a useful tool if you struggle to optimize your structure. Try pressing the "Movie" button to see the iterative optimization.
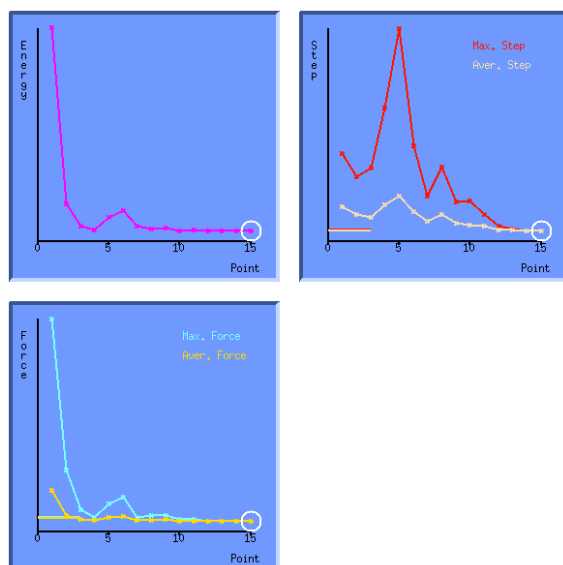
Calculating energies and corrections; frequency (ZPE), big basis and solvation

From the converged geometry optimization of the transition state, we first of all need to do a frequency calculation to check that the optimized geometry corresponds to a actual transition state (imaginary frequency). Normally you would do this in a similar fashion as described above with gaussView, starting from the geometry optimized structure – one for opt=frq and then one for the big basis set and one for solvation corrections. To save some time, we have prepared a script to prepare this 3 calculations for you.

1. Make sure that you are in the correct directory (/$HOME/biomolmo/practicals/DFT/)
2. Type: **fh_water_corrections.py** CM_TS_opt.out (output file for converged geometry calculation).
3. You now have 3 new inputfiles in your directory (*_frq.inp, *_solv.inp, *_bbs.inp).
4. Submit the frequency calculation first to verify that you have the correct transition state:
   a. **g16run.sh** CM_TS_opt_frq 1 1
5. Analyze CM_TS_opt_frq.out with molden. Press the "Norm. Mode" button in the upper right corner to visualize the normal modes. You should see one negative frequency (transition state) with a value around -293.530 cm^{-1}. You can you can set

the "Scale Factor" to 0.5 to enhance the vibrations.
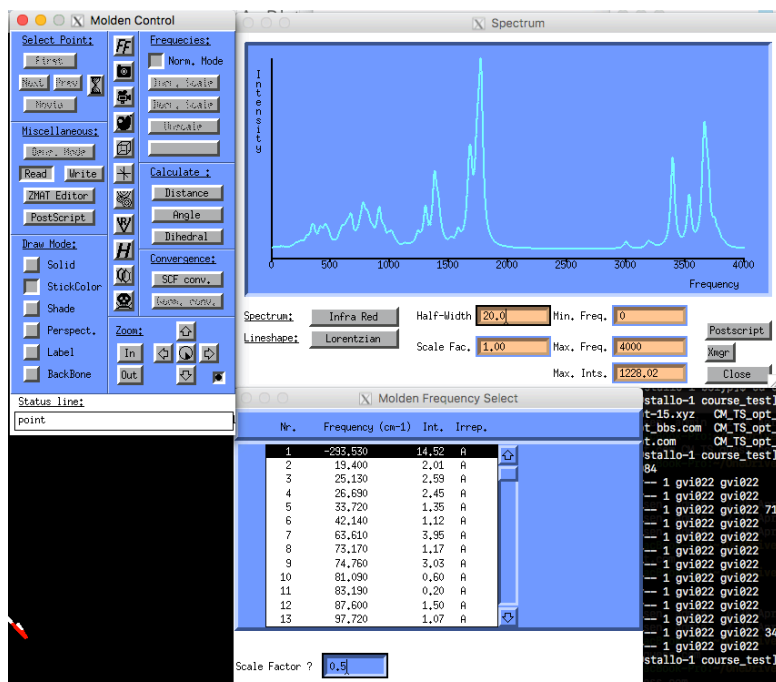
6. With the correct transition state located, submit the big basis set and the solvation single point calculations:
    a. g16run.sh CM_TS_opt_bbs 1 1
    b. g16run.sh CM_TS_opt_solv 1 1

Intrinsic reaction coordinate (IRC) calculation

You are now going to do a intrinsic reaction coordinate calculation from the transition state to locate and verify the connected reactant (chorismate) and product (prephenate). You will later geometry optimize the forward and reverse IRC endpoints. Setting up the IRC is usually done by opening the converged TS structure in GaussView similar to the above. To save time loading graphics on the cluster, we have prepared a script for this as well. To create inputfiles for the forward and reversed IRC, simply type (in the correct directory):

1. irc_frwd_rev.py CM_opt_TS.out

This will produce two new files: CM_opt_TS_rev.com and CM_opt_TS_frwd.com. Submit the IRC calculations:

2. g16run.sh CM_TS_opt_rev 1 3
3. g16run.sh CM_TS_opt_frwd 1 3

Visualize the resulting IRCs in GaussView or Molden and decide which directions (forward and reversed) that connects the reactant state (chorismate) and the product state (prephenate).

Geometry optimization of reactant and product states

Open the IRC outputfile (forward and later the reverse) in GaussView. From the final IRC structure you set up a gaussian calculation in the same way as you did for the TS calculation, only that this time you select "Optimize to a: Minimum" under job type. Your inputfile should end up with this job specification:

#p opt rb3lyp/6-31g(d,p) empiricaldispersion=gd3

Save the inputfiles as CM_RS_opt.com and CM_PS_opt.com and submit the g16 jobs as before.

When your RS and PS have converged, do the <u>frequency, big basis set and solvation calculations</u> as described above (**fh_water_corrections.py**). If your calculation does not converge, restart from the last frame through GaussView, and do Z-matrix diagonalization (Edit → Atom List Editor → Edit → Z-Matrix → Standardize).

Putting it all together
You should now have a *opt.out (small basis set in gas phase), *opt_frq.out, *opt_bbs.out and a *opt_solv.out for the RS (chorismate), TS and PS (prephenate). We calculate all energies relative to RS. All energies from Gaussian are in hartree, so we need to convert:

1 unit =	hartree	kJ mol^{-1}	kcal mol^{-1}	eV	cm^{-1}
hartree	1	2625.50	627.51	27.212	2.1947×10^5
kJ mol^{-1}	3.8088×10^{-4}	1	0.23901	1.0364×10^{-2}	83.593
kcal mol^{-1}	1.5936×10^{-3}	4.1840	1	4.3363×10^{-2}	349.75
eV	3.6749×10^{-2}	96.485	23.061	1	8065.5
cm^{-1}	4.5563×10^{-6}	1.1963×10^{-2}	2.8591×10^{-3}	1.2398×10^{-4}	1

1. Calculate the <u>electronic reaction energies</u> from the big basis set single point calculations:
   a. Get the electronic energies from the output file with grep:
      i. grep "SCF Done" filename.out
         The last line is the final energy.
   b. $\Delta E^{\ddagger} = 627.51 * (E^{TS} - E^{RS})$
   c. $\Delta E^0 = 627.51 * (E^{PS} - E^{RS})$

2. Calculate the <u>correction for solvation</u>:
   a. First, calculate the solvation energies for the RS, TS and PS:
      i. $\Delta E_{solv}^{RS} = 627.51 * (E_{solv}^{RS} - E_{gas}^{RS})$
      ii. $\Delta E_{solv}^{TS} = 627.51 * (E_{solv}^{TS} - E_{gas}^{TS})$
      iii. $\Delta E_{solv}^{PS} = 627.51 * (E_{solv}^{PS} - E_{gas}^{PS})$
         1. Get the E_{gas} and E_{solv} using grep:
            a. grep "SCF Done" *_opt.out
            b. grep "SCF Done" *_opt_solv.out
               The last line is the final energy.
   b. Then, the solvation correction to the reaction energies become:
      i. $\Delta\Delta E_{solv}^{\ddagger} = \Delta E_{solv}^{TS} - \Delta E_{solv}^{RS}$
      ii. $\Delta\Delta E_{solv}^{0} = \Delta E_{solv}^{PS} - \Delta E_{solv}^{RS}$

3. Calculate the <u>correction for zero point energies (ZPE)</u> to ΔG:
   a. First, extract the **Thermal correction to Gibbs free Energy** (∂G_{ZPE}) for each stationary point (RS, TS, PS) from the frequency calculations (*_frq.out) with grep:
      i. grep -A 10 "Zero-point" *_frq.out   (* = filename)
   b. The ZPE correction to the reaction energies are then:

i. $\Delta \partial G_{ZPE}^{\ddagger} = 627.51 * (\partial G_{ZPE}^{TS} - \partial G_{ZPE}^{RS})$

ii. $\Delta \partial G_{ZPE}^{0} = 627.51 * (\partial G_{ZPE}^{PS} - \partial G_{ZPE}^{RS})$

4. Finally, the <u>reaction energies</u> corrected for zero-point energies (ZPE) and solvation effects (dispersion included) become:

a. $\Delta G^{\ddagger} = \Delta E^{\ddagger} + \Delta\Delta E_{solv}^{\ddagger} + \Delta\partial G_{ZPE}^{\ddagger}$

b. $\Delta G^{0} = \Delta E^{0} + \Delta\Delta E_{solv}^{0} + \Delta\partial G_{ZPE}^{0}$

For the rapport?

- Figures of optimized structures with key distances of the 3 stationary points (RS, TS, PS)
- Construct a DFT energy diagram
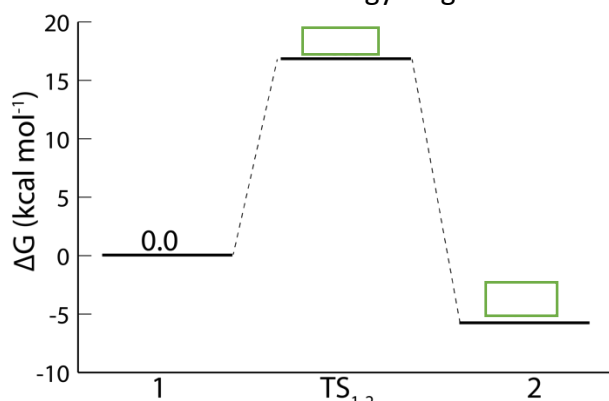
References

1   Frisch, M. J. *et al.*   (Wallingford CT, 2009).

2   Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J Chem Phys* **98**, 5648-5652, doi:10.1063/1.464913 (1993).

3   Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J Chem Phys* **132**, 154104, doi:10.1063/1.3382344 (2010).

4   Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J Comput Chem* **32**, 1456-1465, doi:10.1002/jcc.21759 (2011).

5   Marenich, A. V., Cramer, C. J. & Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J Phys Chem B* **113**, 6378-6396, doi:10.1021/jp810292n (2009).

6   Siegbahn, P. E. M. & Himo, F. The quantum chemical cluster approach for modeling enzyme reactions. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **1**, 323-336, doi:10.1002/wcms.13 (2011).

7   Himo, F. Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions. *J Am Chem Soc* **139**, 6780-6786, doi:10.1021/jacs.7b02671 (2017).