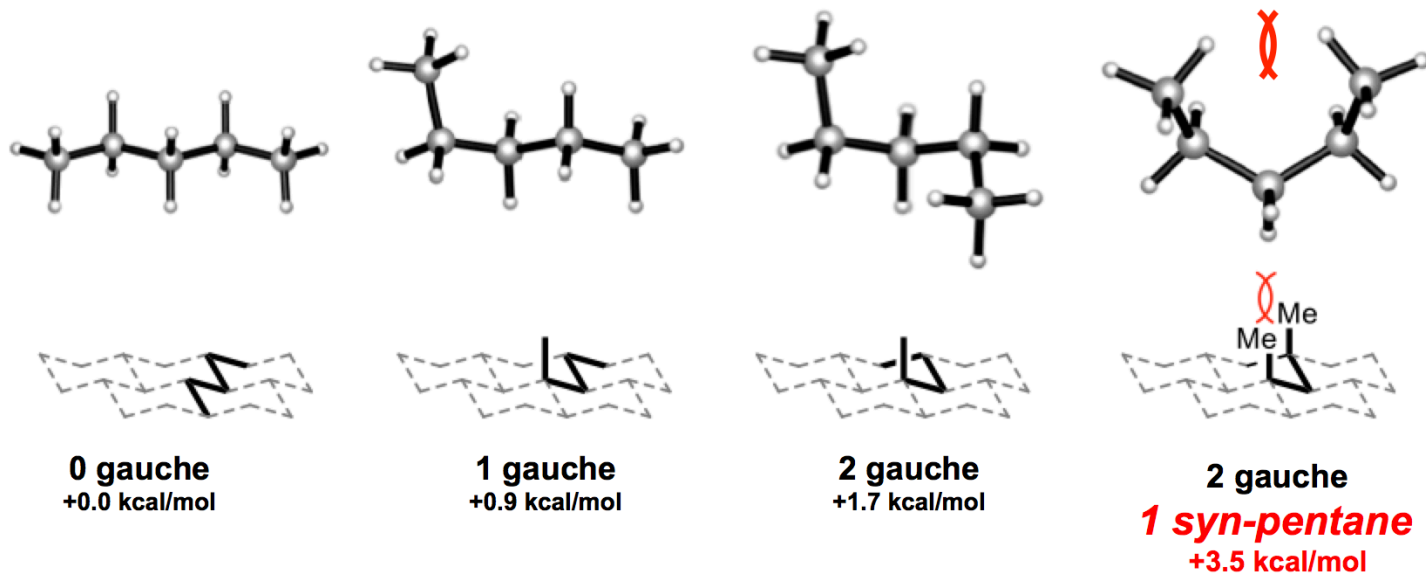Chem 106: Computational Handout 1
Conformers of Pentane

Objectives: draw structures, run your first Gaussian job, extract energies, thermochemistry

In This Exercise: we will examine the conformational landscape of *n*-pentane:

0 gauche	**1 gauche**	**2 gauche**	**2 gauche**
+0.0 kcal/mol	+0.9 kcal/mol	+1.7 kcal/mol	*1 syn-pentane*
			+3.5 kcal/mol

We will number these conformers 1–4 (from left to right). The free energies listed were calculated with B3LYP, a density functional method. Here, we will calculate the energies with a semi-empirical method (PDDG), which offers rough accuracy at very low computational cost.

Mac Setup

1. Open iTerm and connect to Odyssey: `ssh -Y -C -o ServerAliveInterval=30 -fN ody`
2. Open an SSH window: `ssh ody`

PC Setup

1. Connect to the RC VPN (`johnsmith@odyssey`).
2. Open PuTTY and connect to `rclogin.rc.fas.harvard.edu`.
3. Mount your home directory using "map network drive." (user name: `rc\johnsmith`, directory: `\\rcstore.rc.fas.harvard.edu\homes\home08\johnsmith`)

Odyssey Setup

Update your Chem 106 git repository on Odyssey:

```
cd ~/chem106_calculations
git pull
```

Draw a Structure

1. Open GaussView.
2. Click on the 6C button in the upper left of the toolbar. Select the Builder Fragment by clicking on the button on the bottom middle of the toolbar (it might say "carbon tetrahedral" or something similar).
3. Select four-coordinate carbon. Click anywhere in the empty blue window to create a methane molecule.
4. Click on a hydrogen to extend methane to ethane, propane, etc. Draw the extended conformation.
5. File…Save…`pentane_conformation_1.gjf`. Select "Write Cartesians" and press Save. You should create a directory (`mkdir dirname`) in the home directory (`~`) on your laptop to work in. That means using the Cygwin prompt (PC) or a local iTerm session (Mac).

Anatomy of a Gaussian Input File:

1. Open the file in vi: `vi pentane_conformation_1.gjf`
2. Edit the file to make it look like this:

```
%mem=1GB
#p pddg opt freq=noraman

pentane conformation 1

0 1
 C                  -2.88903966     0.76606259    -0.11351068
 H                  -2.53238523    -0.24274741    -0.11351068
 H                  -2.53236682     1.27046078    -0.98716219
 H                  -3.95903966     0.76607578    -0.11351068
 C                  -2.37569744     1.49201887     1.14389428
 H                  -1.30569744     1.49198877     1.14390371
 H                  -2.73233626     2.50083438     1.14388473
 C                  -2.88906221     0.76607801     2.40129895
 H                  -2.53242634    -0.24273855     2.40130635
 H                  -3.95906221     0.76611064     2.40129127
 C                  -2.37571583     1.49203212     3.65870347
 H                  -2.73076928     2.50140623     3.65772710
 H                  -1.30571764     1.49032236     3.65968045
 C                  -2.89135547     0.76770263     4.91610657
 H                  -2.53670050    -0.24181177     4.91683768
 H                  -2.53428182     1.27181687     5.78975824
 H                  -3.96135310     0.76983511     4.91537361
```

(your atom ordering/coordinates could vary from this)

~   File must end in at least one blank line. ~ is not
~   actually in the file (just denotes the end of the file).

Use `i` to enter insert mode, where you can type. Type `ESC` `dd` to delete a line. Press `ESC` `:wq` to save and quit (or `ESC` `:q!` to quit and discard all changes).

Route Card:
`%mem=1GB` = how much memory to use
`#p` = verbose output, `pddg` = the PDDG/PM3 semi-empirical method
`opt` = optimize this geometry to the nearest minimum
`freq=noraman` = calculate vibrational frequencies, don't calculate Raman intensities

Title Card:
`pentane conformation 1` (can be anything, separate from route card with a blank line)

Charge and Multiplicity:
`0 1` (neutral species, singlet; separate from title card with a blank line)

Geometry Specification: (no blank line before this)

```
C               -2.88903966    0.76606259   -0.11351068 (atom 1 is a carbon at these coordinates
H               -2.53238523   -0.24274741   -0.11351068  in Å, atom 2 is a hydrogen, etc.)
```

Setup Jobs Folder on Odyssey

On Odyssey: `cp -r ~/chem106_calculations/jobs_clean ~/jobs1`

In English: copy the whole `jobs_clean` folder (and everything inside it recursively, `-r`) to `~/jobs1`, where `~` is your home directory.

Transfer the Input File to Odyssey

Mac: `scp pentane_conformation_1.gjf johnsmith@ody:~/jobs1/`

PC: `cp pentane_conformation_1.gjf /cygdrive/z/jobs1/`

Run the Job on Odyssey

1. Change to the jobs1 folder: `cd ~/jobs1`

2. Submit job: `./submit.sh 1 serial_requeue 5 2000`

In English: submit all `.gjf` files in this folder (there's only one in this test). Send the jobs to the `serial_requeue` partition (see Appendix) with a maximum runtime of 5 minutes (job will be killed if it exceeds this limit) and a maximum memory usage of 2000 MB.

This creates a folder called `pentane_conformation_1`. Let's have a look:

```
[ekwan@rclogin01 jobs1]$ ls
output  pentane_conformation_1  scripts  analyze.sh  eek.sh  submit.sh  template.sh
[ekwan@rclogin01 jobs1]$ cd pentane_conformation_1
[ekwan@rclogin01 pentane_conformation_1]$ ls
analyze.sh pentane_conformation_1.gjf  pentane_conformation_1.sh

[ekwan@rclogin01 pentane_conformation_1]$ more pentane_conformation_1.sh

#!/bin/bash

#SBATCH -N 1        # number of nodes

#SBATCH -n 1    # number of cores

#SBATCH -p serial_requeue    # partition to submit to

#SBATCH --mem=2000  # memory per node in MB (see also --mem-per-cpu)

#SBATCH -t 5 # expected runtime in minutes

#SBATCH -J g09_pentane_conformation_1 # name of this job
```

The `more` command displays the contents of a text file. Press *SPACEBAR* to page down or `q` to quit. As you can see, the `submit.sh` script (a script is a short program for automating routine tasks) has

copied your request into this "SLURM submission script." SLURM is the name of the job scheduling system on Odyssey. When SLURM runs this job, it will execute this BASH script. Alternatively, you can construct SLURM scripts manually and submit them one at a time: `sbatch my_slurm_script.sh`

3. Wait for a moment for the job to start and type `ls` again:

```
[ekwan@rclogin01 pentane_conformation_1]$ ls
analyze.sh   log.txt  pentane_conformation_1.gjf   pentane_conformation_1.out
pentane_conformation_1.sh   slurm-83472424.out
```

The output of the job will be written to a `.out` file. We can see the beginning of the file with `head`:

```
[ekwan@rclogin01 pentane_conformation_1]$ head ../output/pentane_conformation_1.out
 Entering Gaussian System, Link 0=g09
 Input=pentane_conformation_1.gjf
 Output=pentane_conformation_1.out
 Initial command:
 /n/sw/g09_D.01/g09/l1.exe      "/scratch/ekwan_pentane_conformation_1/Gau-22997.inp"      -
scrdir="/scratch/ekwan_pentane_conformation_1/"
 Entering Link 1 = /n/sw/g09_D.01/g09/l1.exe PID=      22998.

 Copyright (c) 1988,1990,1992,1993,1995,1998,2003,2009,2013,
           Gaussian, Inc.  All Rights Reserved.
```

4. When the job finishes, it will be moved to `jobs1/output`:

```
[ekwan@rclogin01 pentane_conformation_1]$ cd ../output
[ekwan@rclogin01 output]$ ls
output.txt   pentane_conformation_1.out   placeholder
```

Verify that it terminated normally:

```
[ekwan@rclogin01 output]$ tail pentane_conformation_1.out


 IF AT FIRST YOU DON'T SUCCEED, TRY, TRY AGAIN.
 THEN GIVE UP;
 THERE'S NO USE BEING A DAMN FOOL ABOUT IT.

                         -- W. C. FIELDS
 Job cpu time:       0 days  0 hours  0 minutes  2.0 seconds.
 File lengths (MBytes):  RWF=      10 Int=       0 D2E=       0 Chk=       2 Scr=       1
 Normal termination of Gaussian 09 at Wed Mar 15 14:23:49 2017.
```

(You might see different flavor text.) If it terminated abnormally, check that you typed in everything correctly and resubmit the job by repeating the above procedure. `submit.sh` will ask you what to do when you try to submit a job that has the same name as one that has already been submitted. Choose `o` for "overwrite."

5. Some key information will be summarized in `output.txt`:

```
[ekwan@rclogin01 output]$ more output.txt
************************************
Running on host:
holy2a13208.rc.fas.harvard.edu
Job pentane_conformation_1 started at...
Wed Mar 15 14:23:37 EDT 2017
```

```
Job finished at:
Wed Mar 15 14:23:49 EDT 2017

Job terminated normally...

Final energy:
SCF Done: E(RPDDG) = -0.552410743198E-01 A.U. after 2 cycles

No imaginary frequencies found.
 Sum of electronic and thermal Free Energies=          0.068028
*************************************
```

As you run more jobs, output will be appended to this file. You can delete this file to "clean your slate" periodically to stay organized. If an error occurred in the job, more information may sometimes be found in the `slurm-xyz.out` file in the job directory. This can help you troubleshoot your problem.

The "final energy" is the last electronic energy in hartree (1 hartree = 627.509469 kcal/mol), which is the energy of bringing all the electrons and nuclei from isolation at infinity into their present geometry. (That is why the number is enormous and negative.) This energy does not include zero-point energy—it is strictly the potential energy at this level of theory.

Alert! Electronic energies are not directly comparable between different levels of theory or basis sets.

The lack of imaginary frequencies means that the curvature of the potential energy surface (PES) at this point is purely positive. That is, the surface looks like a normal parabola in all directions here. Because the gradient is zero here (see below), this means we are at a true local minimum. However, there is no guarantee this is a global minimum.

If we were at a true transition state (see later exercises), we expect to see one imaginary frequency (and the rest would be real, positive numbers), corresponding to a first-order saddle point.

The "sum of electronic and thermal free energies" is the Gibbs free energy (by default at 298.15 K). It is sometimes worth computing them at other temperatures (use `temperature=xyz` in the route card, where `xyz` is a temperature in K).

Transfer the Output File from Odyssey

Make sure you are in the directory you want to download the files to before completing this step.

Mac: `scp -C johnsmith@ody:~/jobs1/output/*.out .`

PC: `cp /cygdrive/z/jobs1/output/*.out .`

* is a wildcard that can stand for any filename. If multiple files match the wildcard, they will all be copied. The `-C` flag turns on compression, which is helpful when transferring large text files. The period at the end represents the current directory.

Alert! If a file of the same name already exists at the destination, it will be overwritten without warning. There is no way to undo this.

Note that your home directory has a 100 GB quota. If you exceed this, all your jobs will crash without warning or even meaningful error messages. Your home directory is also not appropriate for high throughput input/output calculations, such as high-level *ab initio* calculations (advanced topic). Please see https://rc.fas.harvard.edu/resources/odyssey-storage/ for more details.

Examine Results Locally (GaussView)

1. Open the .out file in GaussView. Ensure that the File Type is set to Gaussian Output Files and "read intermediate geometries" is selected.

2. Open Results…Summary. This lists the electronic energy. This is the simplest method for getting the energy (so that you can compare the energies of conformers, for example). We will examine more efficient ways of doing so later in this exercise.

3. In the upper left, you will see a green circle and some numbers like "1 of 6." This means that the optimization required six steps and you are currently looking at the first geometry (i.e., the input geometry.) Go to Results…Optimization to see a graph of energy vs. iteration number. Expand the window to view the RMS gradient (in Cartesian coordinates). This is the "average slope"—when it drops below approximately 1E-5 hartree/bohr, the optimization is considered to be finished.

4. Requesting `freq` in the route card computes the local curvature of the PES (i.e., the second derivative or Hessian matrix) at the final geometry. This also produces the normal modes, which can be viewed in Results…Vibrations. Try animating the various motions. You will see that similar motions are coupled in the normal mode basis.

Submit the Other Jobs

1. Draw the other conformations (see the diagram on the first page) and save them as appropriately labeled `.gjf` files.

2. Use vi to edit their headers as above.

3. Transfer the files to the cluster as before.

4. Submit all three jobs at once. You can check on the status of your jobs with the `squeue1` command. (The results from this command take about 30 seconds to refresh, so be patient.) These jobs will be extremely quick, but if you have longer jobs and wish to cancel them type `scancel jobid`, where jobid is given by the `squeue1` command.

Analysis

What are the relative energies of the various conformations?

1. Open each output file in GaussView and click on Results…Summary. The energy listed here is the "electronic energy," which you can think of as potential energy. The unit is hartree, which is 627.509469 kcal/mol.

2. Copy the results to a spreadsheet. Calculate the energy of all the conformers relative to the lowest one and convert to kcal/mol.

I get the following energies:

```
pentane_conformation_1.out 0
pentane_conformation_2.out 0.23
pentane_conformation_3.out 0.67
pentane_conformation_4.out 1.4
```

A high-level literature estimate for these energies is (*JPCA* **2009**, *113*, 11974) 0.61, 0.96, and 2.8 kcal/mol. Evidently, the PDDG electronic energies underestimate the *syn*-pentane interaction.

Thermochemical Corrections

Note: if your jobs did not complete, you can still retrieve and analyze my output files from `chem106_calculations/pentane`.

We can try to do better with thermochemical energies, which require a frequency analysis (i.e., the `freq` keyword). A frequency analysis involves calculating the Hessian at a stationary point, which you can think of as the curvature of the potential energy surface. (It is actually a matrix of second derivatives of energy. Because the potential energy surface has many dimensions, we need the second derivatives in many directions to describe the overall curvature.) The thermochemical corrections to the energy are then calculated by applying the rigid-rotor-harmonic-oscillator (RRHO) approximation.

Alert! The RRHO approximation is only valid at a stationary point. In general, one cannot calculate the free energy at an arbitrary point on the potential energy surface by this method. Practically, if you want to calculate a free energy, you must make sure that the gradient is sufficiently small (below 1E-5 hartree/bohr) before calculating the frequencies. This happens naturally when an optimization converges under normal conditions.

1. Open `output.txt` in vi, which is located in your `output/` folder on Odyssey: `vi output.txt`

2. If you did a frequency analysis, you will find an entry like:

```
No imaginary frequencies found.
 Sum of electronic and thermal Free Energies=            0.068028
```

This is the Gibbs free energy at 298 K. Collect all these energies to get the relative free energies of the conformations. I get:

```
pentane_conformation_1.out 0
pentane_conformation_2.out 0.913654
pentane_conformation_3.out 1.72189
pentane_conformation_4.out 2.97565
```

Note: a proper treatment of free energy is complicated by many factors (see *JCTC* **2016** *12* 2679). The simplest of these would include a statistical correction for the symmetry of the conformations ($R \ln n$, where n = 1, 4, 2, and 1, in order).

3. Experimental enthalpies have been measured (*JPCA* **2009**, *113*, 1012) by Raman spectroscopy for conformations 1–3 (conformation 4 is hard to observe): 0.0, 0.62, 0.94 kcal/mol. What are the PDDG enthalpies?

Neither `output.txt` nor GaussView give you the enthalpies, but they are printed out on a line like this:

```
Sum of electronic and thermal Enthalpies=            0.106416
```

Open the file in vi and type *ESC* :/Enthalpies to locate this line. Press n to cycle to the next result and N to cycle back to the previous one. What are the relative energies? I get:

```
pentane_conformation_1.out 0
pentane_conformation_2.out 0.63
pentane_conformation_3.out 1.3
pentane_conformation_4.out 2.1
```
PDDG did pretty well!

4. It can be tedious to go through each file to extract these energies. Here is a way that can speed up the process:

```
~/chem106_calculations/pentane $ grep Enthalpies *.out
pentane_conformation_1.out: Sum of electronic and thermal Enthalpies=          0.106416
pentane_conformation_2.out: Sum of electronic and thermal Enthalpies=          0.107415
pentane_conformation_3.out: Sum of electronic and thermal Enthalpies=          0.108530
pentane_conformation_4.out: Sum of electronic and thermal Enthalpies=          0.109708
```

grep is a standard tool for searching files for text. In this case, we are searching for "Enthalpies" in all of the .out files in this folder. In later exercises, we will learn more about how to use command-line tools to speed up routine tasks such as this one.

Appendix: Partitions on Odyssey

Partition	Maximum Runtime	Resources Per Node	Notes
general	7 days	64 cores, 256 GB	8 floating points and 16 integer units / core, 4 sockets, infiniband
serial_requeue	7 days (24 h realistic)	varied	very fast queuing, jobs often pre-empted
unrestricted	unlimited	same as general	100% uptime not guaranteed, slow queueing
jacobsen	unlimited	12 cores, 24 GB	Jacobsen group only, request twice as many cores as requested in %nprocshared directive

Eugene Kwan
March 2017