［解説記事］

Using Gaussian-03 software for the computational chemistry research in Information Synergy Center of Tohoku University

Ilya D. Gridnev

COE Laboratory, Department of Chemistry, Graduate School of Science, Tohoku University

This is a practical guide that is aimed to help non-experienced users to carry out their own quantum chemical computations on the supercomputing facilities of Tohoku University. I will not try to cover numerous theoretical problems, ideas or approaches. Neither shall I touch the hardware/software requirements or technical details. Rather, I am going to follow step by step the actions necessary for submitting a job to the computer and getting back the result. According to my own experience this is one of the limiting stages in starting one's own computational research. The basics of the theoretical chemistry can be learned from the textbooks, the examples of the practical application of the computational results are numerous in the research papers. However, a chemist decided to compute some molecule often realizes that he/she just does not know how to begin, especially when it comes to submitting a computation to the super computer. In this manual we are trying explain this important step in detail assuming that a researcher can find other relevant information in the corresponding literature.

This approach sets certain limits of presentation. I will describe the procedure for a specific case of Windows PC and Gaussview-03 as a building software in hope that the corresponding procedure from any other hard/software combination can be deduced by analogy.

1. Type of computation.

The first thing to decide is what kind of computation you are going to start. I will describe here two types of computation: energy optimization and transition state search. There are numerous other possibilities, but one can easily understand how to submit them if he/she is familiar with optimization and transition state computations.

1.1 Structure optimization.

This is the most frequently used type of computation. The researcher builds a molecule (it can be charged or contain free radicals) and asks the computer to find the energy minimum for this combination of atoms. Important thing to realize is that the

computer does not know *a priori* what particular minimum you are after – it just looks for the nearest one (or the one accidentally approached during the computations). If there are two minima (*e.g.* conformational) on the PPE, it would depend on the starting configuration (that you had submitted) whether the output structure will correspond to the local minimum or to the global minimum.

The results of the energy optimizations can be used in many different ways. One can compare the energies of series of compounds, investigate the structural features of the molecules, use the optimized structures for the search of transition states (see next chapter), *etc.*

1.2 Search for transition state

Transition states are used to estimate the activation barriers of reactions. These are valuable data, but technically this type of computation is much more difficult than the energy optimization, therefore it is advised to master the energy optimizations first. Anyway, you must do the energy optimization of the starting material and the reaction product of the reaction (elementary step actually) before you can submit the transition state computation.

2. Level of Theory

This is another matter that must be decided before you start the computations. There are two parameters: theoretical approach and basis set that are specified together divided by the slash, *e.g.* B3LYP/6-31G* or MP2/LANDL2DZ. The best thing for beginner is to find in the literature some recent papers in which similar molecules were computed and use the same level of theory. Detailed information on the scope of use, theoretical approaches, and computer time requirements can be found in software manuals, Gaussian help file, *etc.*

3. Building a molecule

3.1. For the structure optimization

Any software can be used for manually building a molecule designated for the optimization. In most cases the molecular mechanics optimization (built-in feature in most of the programs) helps to create the structure resembling the expected configuration. Don’t try to start quantum mechanics optimizations from an obviously

unrealistic structure – it will only result in waste of time. While building, keep controlling the charge/multiplicity combination – if it becomes unusual (*e.g.* you can get only "doublet" multiplicity for a neutral molecule) it means that the structure is erroneous (*e.g.* extra hydrogen or another atom is attached somewhere).

3.2. For the search of transition state.

In this case you must have three structures at hand: starting compound, product, and suggested structure of the transition state. It is reasonable suggestion that the structures of the starting compound and the product are optimized before starting this job, however this is not a strict requirement.

In the case of Gaussview03 do the following:

1. Open a file with the starting configuration.
2. Click Open; in the dialog box choose "**Add all files to the active molecule group**" from the "**Target**" menu; then specify the file with the structure of the product.
3. Repeat step 2 using the file with the structure of transition state.
4. It is absolutely necessary that all three structures had the same numbers for the same atoms. Usually, when you have just finished the steps 1-3 the numeration is different in the three structures. To bring it to order use **Connection Editor** (is switched on by golden chain icon in the toolsbar or in **Edit** menu). When you start the connection editor it checks automatically if the atom numbering matches to each other. If it does – you are lucky – go to the next step. If not, try first **Enable Autofixing**, then **Z-mat. tools.**
5. Now you are ready to save the file for submission, proceed to **4.**

4. Saving a submission file.

4.1. File structure.

Only specific Gaussian format is recognized by the computing program. Therefore it is necessary to understand how the file is organized. Several software packages are capable of creating Gaussian input files from the molecular configuration created by this software. However, quite frequently these files need modification before they can be used, therefore the knowledge of the input file structure is essential even for the beginner. There are two extensions for the Gaussian files: *.gjf (Gaussian Job File) is

used on Windows machine, *.com files are required for Unix systems, including the Supercomputer at Information Synergy Centre.

```
%chk=cyach.chk
%mem=16gb
%nproc=8
# opt b3lyp/sdd

Title Card Required

1 1
```

The above shows the typical heading of the Gaussian input file. First three lines (called Link 0) specify the computer parameters to be used in computation, these lines must be started with % character followed by the parameter without blank space.

First line specifies the name of checkpoint file if you want it. Checkpoint file keeps all the information on the present computation and can be used to restart it if it has been stopped for some reason. Hence, it is better to have this file, although it has a tendency to become quite large.

Second line specifies how much memory can be used by the computation. For the Center's Supercomputer just keep this value at 16 Gb.

Third line specifies the number of processors to be used in the computation. There are two options at present: you can use either 8 processors or 16; the jobs are submitted to the queue **a8** or **a16**, accordingly. Of course, 16 processors will do the job faster, but this queue is usually used for longer computations, hence it might be necessary to wait longer before your computation will start. Hence, if your molecule is not very large, it can be better to submit the job for 8 processors.

The fourth line specifies the parameters of your computation. It starts with the sign **#** followed by a blank space. At least two parameters must be shown (as in the example above). First is the type of computation (see **1.**); parameter value **opt** is for structure optimization (see *1.1*); for the transition state search (see *1.2*) **opt=qst3** must be used.

After the fourth line there is blank line followed by the title (for your own use – computer will skip it), another blank line, and charge and multiplicity values (**1 1** in the above example means cation in singlet spin state; neutral molecule will have **0 1** in that line).

Further follow the atomic coordinates, *etc.* specifying your input geometry (skipped in the above example). In the case of the job for transition state there are three sections describing the molecule – for each of the three structures (starting compound, product, and suggested transition state) separated by *blank line-title-blank line* sequence.

4.2. Convenient procedure to make submission files in Gaussview03.

1. Build your structure.
2. Choose **Calculate**, then **Gaussian**. Pop-up menu appears.
3. In the **Job Type** section choose **Optimization**. If you are doing optimization job, choose *Optimize to a* **Minimum**. For transition state choose **TS(QST3)**.
4. In the **Method** section specify your level of theory choosing necessary values from the menus. Check if the **Charge** and **Multiplicity** values are correct.
5. Put the title in the **Title** section if you want.
6. Edit the **Link 0** section (see *4.1.*).
7. Disable **Write Connectivity** option in **General** section.
8. Press **Submit** – you will be prompted for saving the file – do it (this is actually what we are after).
9. Press **Cancel** twice – we do not want to submit computation on the local computer.
10. Proceed to *4.3.*

4.3. Problems with different OS and their solution.

If you have saved your input Gaussian file in Windows, you cannot use it for submitting your job to the supercomputer directly: the file must be modified. We will describe a simple procedure for doing this without explaining any details.

1. Prepare your file in Windows.
2. Transfer it to your directory on the supercomputer (see **5.**).
3. Edit your file ***on the supercomputer*** using **Edit** button of the WinSCP interface (see **5.**). You can change the name of the checkpoint file or number of processors – anything at all, just to make some change. When closing the editor you will be prompted for saving – do it. Now your file is ready for the submission.

5. Transferring your files to the supercomputer.

You must get the WinSCP freeware from the Web and install it on your computer. Set the Host name as follows

gen.isc.tohoku.ac.jp

and specify your account name and password.

After this is done, you'll be automatically connected to the file manager showing on the left side your computer disks, and on the right side your directory on the supercomputer. The software is convenient and straightforward, no additional explanation is necessary.

6. Starting the job.

You will need a secure telnet software to communicate with the Supercomputer. Freeware Putty is good enough. After being prompted for the login details you are automatically transfered to your home folder at the Supercomputer. You will need to know around dozen UNIX commands to navigate there. Please refer to the first couple of pages in any UNIX manual.

Move to the folder in which you have copied (and edited!) your job file (see **5.**). Type the following command

subg03_a8_filename (sign _ indicates space)

If you are submitting computation for 16 processors type a16 instead of a8.

Filename must be specified ***without extension***, thus if your file is xyz.com, type

subg03_a8_xyz

That's it – you must get a message confirming that your file has been added to the queue.

7. Checking the status of your job.

7.1. Controlling the status of submitted jobs.

By typing

qstat

you will bring on the screen the list of all your submitted jobs with their id number, status (RUN or QUE), and the amount of memory which they are currently using. The most left number in the line of a queued job shows how many jobs are in the queue before that one.

You can stop computation of a running job or remove a queued job from the queue by typing

qdel_jobid

*7.2. Checking the progress of the job. Structure of the *.log file.*

When the computations proceeds, all information on each iteration step is saved in the *.log file, which is being continuously written in the folder with your job. You can either transfer this file to your computer (see **5.**) and open it in Gaussview or open it directly in WinSCP by pressing **Edit**. After each iteration computer calculates energy in atomic units (1 a. u. = 627.51 kcal/mol) and checks four parameters for convergence. You can follow the current status of your computation by looking at the last entries in the *.log file:

```
SCF Done:  E(RB+HF-LYP) =  -1402.46599933     A.U. after   18 cycles
             Convg  =    0.6716D-08             -V/T =  2.0076
             S**2   =    0.0000
```

for energy, and

```
         Item               Value     Threshold  Converged?
 Maximum Force            0.003268     0.000450     NO
 RMS     Force            0.000241     0.000300     YES
 Maximum Displacement     0.365339     0.001800     NO
 RMS     Displacement     0.062739     0.001200     NO
 Predicted change in Energy=-3.106722D-04
```

for convergence. When all for parameters will have YES mark, the computation will be terminated, and at the end of the file you will have something like this:

```
IT'S WHAT YOU LEARN AFTER YOU KNOW IT ALL
 THAT COUNTS.
 Job cpu time:  0 days  0 hours 11 minutes  4.5 seconds.
 File lengths (MBytes):  RWF=     46 Int=      0 D2E=      0 Chk=     12 Scr=      1
 Normal termination of Gaussian 03 at Mon Aug 21 15:19:57 2006.
```

If you want to have the final value of energy, you can go several screens up and find the last energy value. Another way to have this value is to copy the *.log file on your computer, open it in Gaussview, then click **Results** and **Summary**.

Conclusion.

As mentioned in the beginning, this manual did not cover even one million's part of the rapidly developing field of computational chemistry. However, I hope that you might be able to find in it some useful hints that would allow you to start your own research, and then learn other things from books and articles. Enjoy.