How to run a quantum chemistry program: The components of the input file

Sunday, August 17, 2008
7:16 PM

Coordinates
You must specify a geometry of the molecule you wish to study. You have two input choices:

- **Cartesian.** Specify the atomic symbol and the XYZ coordinates of the atom.
- **Z-matrix.** Specify the internal bond distances and angles between atoms. A ball and stick model is useful.

A simple way to perform transformation from Cartesian coordinates to intra-molecular is to use xmol program.

Sources of coordinates:

- **X-ray crystallography, NMR.** X-ray is good for solid phase structures and crystallized organics and proteins. NMR is useful for liquid phase structures.
- **Books.** These will provide you with bond distances between different elements. However, you will still need to determine angles on your own.
- **Previous calculations.** One can use a various set of software packages to get molecular geometries. Insight (tm) is one such useful package. You can build a molecule from fragments then optimize the geometry with a classical force field algorithm. Also, you can use the semi-empirical method for structures such as AM1. Finally, you can use any number of *ab initio* packages such Q-Chem, GAMESS, Titan, etc. to optimize a guess structure into a valid geometry. Note that when you are performing accurate methods such as coupled-cluster theory, you should use the best geometry available, i.e., one from an SCF or MP2 geometry optimization using a medium sized basis set.

Sample Z-matrix inputs
A Z-matrix is used to define connectivity between atoms in a molecule. The parameters one needs are distances, angles and dihedral angles. We will show a few simple examples of how to make Z-matrices in this text.

Sometimes it is a good idea to think before attempting to write a Z-matrix. What is it you are planning on doing with the molecule? If you are going to do a geometry optimization for the ground state, then it would be a good idea to enforce symmetry. Looking at the benzene example below, one can see that the $D_{6h}$ symmetry will never be broken. When optimizing, only the bond distances have a chance of changing, since the angles are forced to 120 degrees.

However, if one is going to do a transition state search, then the Z-matrix should be as flexible as possible, to allow for any symmetry-breaking geometry changes. Taking the time to plan what one is going to do can save time hunting for why the desired output was not achieved.

**Sample: Water**

H
O 1 OH
H 2 OH 1 OHO

OH = 1.08
OHO = 107.5

For water, all we need is a bond distance and an angle. We start with the first atom, hydrogen, on a line of its own. The next line begins with the second atom, oxygen, and
then states with which atom to measure the bond distance OH from, in this case, atom one. On the next line, the third atom, hydrogen, is OH distance away from atom two and has a bond angle of OHO in relation to atom one.

**Charge**
This determines the net charge of the system of interest and implicitly tells the program how many electrons there will be.

**Spin multiplicity**
This field determines the net difference between up and down spins in the molecule. Note how you have to tell the program beforehand whether your molecule in the ground state is a singlet, triplet, etc. If you are not sure of the multiplicity, then run some calculations with different settings and determine which ones look reasonable. Sometimes, ground state multiplicity is self-evident. Other times, it is nearly impossible to determine the proper multiplicity because either there are near-degeneracies in the ground state, or correlation and basis set effects are crucial.

**Basis set**
The importance of selecting a good basis set is a key to getting reasonable results. This topic will covered in more detail in the future. For now, it is important to know that the simplest basis set is a minimal one but it typically provides poor answers. Medium sized basis sets typically are 3 or 4 times larger and are useful for geometry optimizations and properties of well-behaved molecules. Finally large basis sets can be on the order of 10+ times larger than a minimal set and they are useful for "ultra-accurate" results.

**Methods**
One can choose from a multitude of methods which include:

- **SCF**—Self Consistent Field theory (Ground State). Hartree-Fock theory: A starting point for correlated and excited state methods.
- **DFT** - Density functional Theory (Ground State). A quick and reasonably accurate method.
- **MP2**—Møller-Plesset perturbation theory (Ground State). This is the simplest correlation correction to the SCF energy and derivatives.
- **CCD, CCSD, CCSD(T)**—Coupled-cluster theories. These are more advanced forms of correlation corrections which correspondingly take longer to perform.
- **CI**—Configuration interaction (Ground State and Excited State). CIS and CIS(D) are useful first approximations to excited states of molecules. Higher order CI such as CISDTQ can be useful a correlation correction to the ground state and excited states.
- **Geometry optimizations**. A molecule of interest can be optimized with respect to the electronic potential energy surface generated by SCF and MP2 calculations.

**Properties**
Different programs offer a wide range of calculated properties of molecules. They include energies, dipoles, polarizabilities, vibrational frequencies, NMR shifts, etc.