

Example: H₂O (the car file)

As a practical example of DFT methods we calculate the energy and electronic properties of the water molecule. In order to carry out the DFT calculation you will need a set of starting nuclear coordinates. I will use a set of coordinates that is clearly not correct (assuming a 90° H-O-H bond angle) for purposes of illustration. The coordinate file has the following appearance.

> vi h2o.car

```
!BIOSYM archive 3
PBC=OFF
```

```
!DATE      Jan 17 16:58:17 2004
O1         0.000000000    0.000000000    0.000000000 H2O1 1      o      O      -0.820
H2         0.000000000    1.000000000    0.000000000 H2O1 1      h      H      0.410
H3         1.000000000    0.000000000    0.000000000 H2O1 1      h      H      0.410
end
end
```

Example: H₂O (the dmol file)

The dmol file is the format used in the original version of DMol3. This file format is much simpler. You may examine it by using

```
> vi h2o.dmol
```

You may convert from car to dmol using the program car_dmol. You will need to know the number of atoms in your molecule.

```
$coordinates
```

```
H      0.0000000000000000  0.0000000000000000  0.0000000000000000
O      0.0000000000000000  0.0000000000000000  1.94853623793378
H      0.0000000000000000  1.0000000000000000  3.34853623793378
```

```
$end
```

Example: H₂O (the input file)

Since we know the geometry is not correct the calculation type must be **optimize** or **optimize_frequency**. The latter performs the frequency calculation once the geometry is optimized. If the molecule is smaller than 15 atoms it probably reasonable to select **optimize_frequency** using DMol3 and running less than 4 processors. For now we can leave all other options in the input file as the default values.

> vi h2o.input

```
#Calculate          optimize
Calculate         energy
#Calculate          optimize_frequency
#Calculate          ts_search
#Calculate          frequency
#Calculate          gradient
#Calculate          Molecular_Dynamics
#Calculate          Simulated_Annealing
```

Example: H₂O (the input file)

Since we know the geometry is not correct the calculation type must be **optimize** or **optimize_frequency**. The latter performs the frequency calculation once the geometry is optimized. If the molecule is smaller than 15 atoms it probably reasonable to select **optimize_frequency** using DMol3 and running less than 4 processors. For now we can leave all other options in the input file as the default values.

> vi h2o.input

```
#Calculate          optimize
#Calculate          energy
Calculate         optimize_frequency
#Calculate          ts_search
#Calculate          frequency
#Calculate          gradient
#Calculate          Molecular_Dynamics
#Calculate          Simulated_Annealing
```

Example: H₂O (the input file)

Let's look at a few important options in the input file.

Use the search command and type

`/functional`

The functional selected is `gga` (i.e. GGA). Other common choices are `blyp` (BLYP) and `vwn` (SVWN) that were discussed earlier.

The search for `/basis`.

Here you will see that the default is `DNP`. This is the equivalent of a double-zeta quality basis set with one polarization function. The basis set in DMol3 is numerical. It does not use Gaussian functions. One aspect of DMol3 that is interesting is the ability to design and improve the basis set (but this is an advanced topic beyond our present objective).

Example: H₂O (the functional)

Functional

#Functional

#Functional

#Functional

#Functional

#Functional

#Functional

#Functional

#Functional_Post_LDA

#Functional_Post_LDA

gga

pwc

ks

jmw

vwn

vwn-bp

bp

blyp

off

BLYP

Example: H₂O (the basis set)

```
#Basis min
#Basis dn
#Basis dnd
Basis dnp
#Basis extended
#Basis all
#Basis user
# 1 7 0 0 0 0 2 2 2
# 6 11 0 0 0 0 0 0 0 2 2 2 2
# 7 11 0 0 0 0 0 0 0 2 2 2 2
# 8 11 0 0 0 0 0 0 0 2 2 2 2
# 44 15 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Example: H₂O (the job file)

To launch the job we will use the following script:

```
$ cbrc [root name] [dir name] > [root name].job
```

The script tells the computer how many processors to use (we will use 4) and how much memory to allocate. You can see the queues typing:

```
$ bqueues
```

Run the job using:

```
$ bsub < [root name].job
```

Check the job using

```
$ bjobs
```

Example: H₂O (the outmol file)

For H₂O you will not have long to wait. In about 1 minute the output (outmol) file appears. You can see it by typing

```
$ vi h2o.outmol
```

Go to the bottom by typing `<shift>G`. You will see how long the job took. Go back to top by typing `:1`. Let's check the job. Searching for the text `/electrons` (and then following that with the letter `n` to find the second occurrence) you will see the number of basis functions and electrons in the calculation.

```
Symmetry orbitals
```

```
  n  norb  representation
  1   24         a
```

```
total number of valence orbitals:      24
```

```
molecule charge=      0.0  active electron number=      10.0
including core=      10.0  (without charge=      10.0)
```

Example: H₂O (the basis set)

Just above you will see the description of the numerical basis set.

Specifications for Basis Set Selection

atomic cutoff radius 6.00 au

Hydrogen	nbas= 1	z= 1.	3 radial functions, e_ref= -0.0461310Ha	
n=1 L=0	occ= 1.00	e=	-0.222910Ha	-6.0657eV
n=1 L=0	occ= 0.00	e=	-0.844526Ha	-22.9807eV
n=2 L=1	occ= 0.00	e=	-1.999912Ha	-54.4204eV
Oxygen	nbas= 2	z= 8.	7 radial functions, e_ref= -0.0607807Ha	
n=1 L=0	occ= 2.00	e=	-18.896548Ha	-514.2014eV
n=2 L=0	occ= 2.00	e=	-0.868840Ha	-23.6424eV
n=2 L=1	occ= 4.00	e=	-0.322562Ha	-8.7774eV
n=2 L=0	occ= 0.00	e=	-2.144678Ha	-58.3597eV
n=2 L=1	occ= 0.00	e=	-1.593751Ha	-43.3682eV
n=3 L=2	occ= 0.00	e=	-2.722082Ha	-74.0717eV
n=3 L=2	occ= 0.00	e=	-1.383811Ha	-37.6554eV eliminated

Example: H₂O (the SCF cycles)

From this point search for the text **/SCF** and you will see the how the energy of the calculation is altered as the variational principle is applied to find the best coefficients. Convergence is set in the input file to 10⁻⁶ Hartrees.

~~~~~ Start Computing SCF Energy/Gradient ~~~~~

|    | Total E (au) | Binding E (au) | Convergence | Time (m) | Iter |
|----|--------------|----------------|-------------|----------|------|
| Ef | -76.4907801  | -0.4371008     | 0.16E+00    | 0.021    | 1    |
| Ef | -76.4409749  | -0.3872957     | 0.91E-01    | 0.024    | 2    |
| Ef | -76.4308187  | -0.3771395     | 0.61E-01    | 0.026    | 3    |
| Ef | -76.4309567  | -0.3772774     | 0.25E-01    | 0.028    | 4    |
| Ef | -76.4296483  | -0.3759690     | 0.32E-02    | 0.030    | 5    |
| Ef | -76.4296274  | -0.3759482     | 0.13E-02    | 0.032    | 6    |
| Ef | -76.4296135  | -0.3759343     | 0.27E-03    | 0.034    | 7    |
| Ef | -76.4296117  | -0.3759324     | 0.62E-04    | 0.036    | 8    |
| Ef | -76.4296120  | -0.3759327     | 0.36E-04    | 0.038    | 9    |
| Ef | -76.4296122  | -0.3759329     | 0.11E-05    | 0.040    | 10   |
| Ef | -76.4296122  | -0.3759329     | 0.76E-06    | 0.042    | 11   |

# Example: H<sub>2</sub>O (the basis set)

Just above you will see the description of the numerical basis set.

---

## Specifications for Basis Set Selection

---

atomic cutoff radius 6.00 au

|          |           |       |                                         |                       |
|----------|-----------|-------|-----------------------------------------|-----------------------|
| Hydrogen | nbas= 1   | z= 1. | 3 radial functions, e_ref= -0.0461310Ha |                       |
| n=1 L=0  | occ= 1.00 | e=    | -0.222910Ha                             | -6.0657eV             |
| n=1 L=0  | occ= 0.00 | e=    | -0.844526Ha                             | -22.9807eV            |
| n=2 L=1  | occ= 0.00 | e=    | -1.999912Ha                             | -54.4204eV            |
| Oxygen   | nbas= 2   | z= 8. | 7 radial functions, e_ref= -0.0607807Ha |                       |
| n=1 L=0  | occ= 2.00 | e=    | -18.896548Ha                            | -514.2014eV           |
| n=2 L=0  | occ= 2.00 | e=    | -0.868840Ha                             | -23.6424eV            |
| n=2 L=1  | occ= 4.00 | e=    | -0.322562Ha                             | -8.7774eV             |
| n=2 L=0  | occ= 0.00 | e=    | -2.144678Ha                             | -58.3597eV            |
| n=2 L=1  | occ= 0.00 | e=    | -1.593751Ha                             | -43.3682eV            |
| n=3 L=2  | occ= 0.00 | e=    | -2.722082Ha                             | -74.0717eV            |
| n=3 L=2  | occ= 0.00 | e=    | -1.383811Ha                             | -37.6554eV eliminated |

---

# Example: H<sub>2</sub>O (the energy states)

Once convergence is reached the energies of the molecular orbitals is calculated. Only one unoccupied orbital is listed here. The others are occupied. More detailed output can be requested. The MOs can be output as shown on the website.

Energy of Highest Occupied Molecular Orbital -0.24970Ha -6.795eV

| state |   |   |   | eigenvalue |          | occupation |
|-------|---|---|---|------------|----------|------------|
|       |   |   |   | (au)       | (ev)     |            |
| 1     | + | 1 | a | -18.762994 | -510.567 | 2.000      |
| 2     | + | 2 | a | -0.912899  | -24.841  | 2.000      |
| 3     | + | 3 | a | -0.434975  | -11.836  | 2.000      |
| 4     | + | 4 | a | -0.354427  | -9.644   | 2.000      |
| 5     | + | 5 | a | -0.249700  | -6.795   | 2.000      |
| 6     | + | 6 | a | 0.027694   | 0.754    | 0.000      |

# Example: H<sub>2</sub>O (the geometry optimization)

At this point the program starts the geometry optimization. Notice that the molecule has been transformed so the symmetry axis is coincident with z axis, but otherwise the geometry is not changed yet. Search on the word */Cycle* (case must right!) and track the progress of the optimization.

```
** GEOMETRY OPTIMIZATION IN CARTESIAN COORDINATES **  
Searching for a Minimum
```

```
Optimization Cycle: 1
```

```
Input Coordinates (Angstroms)
```

```
-----  
      ATOM      X      Y      Z  
1      O      0.000000  0.000000 -0.471405  
2      H      0.707107  0.000000  0.235702  
3      H     -0.707107  0.000000  0.235702  
-----
```

# Example: H<sub>2</sub>O (convergence)

It took 6 cycles in the present example. The final cycle had the appearance:

|       | Cycle | Total Energy | Energy change | Max Gradient | Max Displacement |
|-------|-------|--------------|---------------|--------------|------------------|
| opt== | 6     | -76.4352059  | -0.0000006    | 0.000588     | 0.001194         |

Again the criteria for convergence are set in the input file. Note that minimization of the energy is one criterion for convergence. Keep in mind that a set of iterations to find the SCF coefficient is carried for each geometry (i.e. 6 times in the present case). If you look at the car file you will find the updated geometry (original car file is overwritten).

```
!BIOSYM archive 3
PBC=OFF
```

```
!DATE      Jan 22 02:24:58 2004
O1         0.000000000    0.000000000   -0.400578191 H2O1 1      o      O   -0.820
H2         0.760530176    0.000000000    0.200289095 H2O1 1      h      H    0.410
H3        -0.760530176    0.000000000    0.200289095 H2O1 1      h      H    0.410
end
end
```

# Example: H<sub>2</sub>O (charges and dipole)

The properties are calculated next.

```
+++ Entering Properties Section +++
```

```
Charge partitioning by Hirshfeld method:
```

```
O      1 charge      -0.3008
H      2 charge       0.1503
H      3 charge       0.1503
```

```
symmetry unique atoms:   1   2   3
```

```
dipole moment vector (au):      0.00000      0.00000      0.76730
dipole magnitude:                0.76730 au                1.9503 debye
```

```
Mulliken Population analysis
```

```
Mulliken atomic charges:
```

```
charge  spin
O(  1)  -0.487  0.000
H(  2)   0.243  0.000
H(  3)   0.243  0.000
```

Gas phase water  
experimental value  
is 1.86 Debye.

# Example: H<sub>2</sub>O (esp fitting)

Electrostatic potential fitting is considered more accurate than Hirshfeld or Mulliken charge.

Limits of the ESP box:

```
-6.850260 -6.377828 -6.963917  
6.850260 6.377828 5.791740
```

=====

Summary of ESP fitting calculations

```
Number of points:      6288.76  
Spacing between:      0.25Ang  
Sigma:                 0.2059E-02  
RMS of V(exact):      0.1703E-01  
RRMS fit:             12.09%
```

ESP-fitted charges:

| n | Elem | chg    | vdW(in) | vdW(ex) |
|---|------|--------|---------|---------|
| 1 | O    | -0.704 | 1.72    | 3.22    |
| 2 | H    | 0.352  | 1.30    | 2.80    |
| 3 | H    | 0.352  | 1.30    | 2.80    |

=====

TIP3P water used  
in MD force fields  
has

O -0.82

H 0.41

H 0.41

However, this is  
a gas phase calculation.  
We would need to model  
a water cluster or in  
periodic boundaries to  
obtain comparable values.

# Example: H<sub>2</sub>O (frequencies)

Searching on the text `/freq` you will find the calculated vibrational frequencies.

vibrational frequencies, intensities

| mode | au_amu   | cm-1   | km/mol |
|------|----------|--------|--------|
| 7    | 0.315034 | 1619.4 | 69.54  |
| 8    | 0.721257 | 3707.6 | 6.72   |
| 9    | 0.743767 | 3823.3 | 43.56  |

| Experimental values are: | Error                    |
|--------------------------|--------------------------|
| bending mode             | 1654 cm <sup>-1</sup> 2% |
| asymmetric stretch       | 3825 cm <sup>-1</sup> 8% |
| symmetric stretch        | 3935 cm <sup>-1</sup> 3% |

## Further info: H<sub>2</sub>O (frequencies)

If the geometry is properly optimized you should have six zero frequencies. These are the translations and rotations of the molecule. Sometimes there will be a negative frequency. This may be unavoidable. For example, exocyclic amino groups have a double well potential and do not obey the harmonic approximation.

You can determine if this is the case by looking at normal mode projections. To calculate normal mode projections using your frequency outmol file and dmol file you will use the program `mpnm_dmol`.

# Program: mpnm\_dmol

This program uses the eigenvectors in the frequency outmol file to create projections along normal modes. Usually you will use 21 files. It is always a good idea to choose a number like 21, 41, 61 or 81 so that you will have a the mid-point as the equilibrium geometry and then an equal number of structures on either side of the equilibrium.

You will usually use the default value for the extent (1.0). However, if you are interested in low frequency modes then you may use larger values. The program will automatically generate the requested number of projections for all of the normal modes. This is typically a large number of files. You will likely only want to calculate the potential energy profile for certain modes.

# Generating a script for energy profiles

The output files have the following form

root#disp\_#mode.dmol

The #disp is the displacement number from 1 to 21 (if you choose 21 displacements). The mode number is obvious.

Imagine you want to calculate the profile for mode 15. Then do the following

```
$ ls root*_#mode.dmol > mkrun
```

Then edit mkrun

```
$ vi mkrun
```

Replace /root/ with /cbr root/

Replace /.dmol/ with / [dir]/ where [dir] is your working directory. Then :wq

Now activate the script

```
$ chmod +x mkrun
```

# Run the mkrun script to make a job file

Now run mkrun

```
$ mkrun > all.job
```

The job file all.job contains all 21 files you selected.

When you launch this job file it will sequentially calculate each of the energies.

```
$ bsub < all.job
```

Make sure the your INPUT\_DMOL file is an “energy” calculation. If you have an “optimize” calculation it will optimize each one of the files and they will all give you the same value!

Once this has completed you can extract the energies using the grep command.

# Extracting energies and MOs using grep

To extract the binding energy use the grep command as follows:

```
$ grep 'energy extra' root_#mode.dmol > energy#mode
```

This will extract the energies and write them to a file called energy#mode.

To extract molecular orbital energies we need to look in the Outmol file to find the number of the MO. Let's imagine that the number is 26. In that case there will be a number of spaces before and after the number 26 on the line that reports the MO energy. We use the follow:

```
$ grep ' 26 ' root_#mode.dmol > mo26_#mode
```

We can do the same for any other MOs.

# Eliminate extraneous text

Now that you have extracted the energies it would be most convenient to extract the column of numbers that reports the energies in electron volts. You would also like to have the files sorted in order in terms of the geometries from 1 to 21. The next step is to eliminate the text at the beginning of the file. You may use the program 'textelim' as follows:

1. Copy your file to the name "input"
2. Run textelim using redirect to give it a name

```
$ cp energy#mode input  
$ textelim > energy#mode_e
```

The `_e` at the end tells you that this is an intermediate file.

# Sort the file in order of the geometries

To sort the file you may use the LINUX command "sort"

```
$ sort -g energy#mode_e > energy#mode_sort
```

Finally you want to extract a column of numbers from the file. For this you use the awk command.

```
$ awk '{print $10}' energy#mode_sort > energy#mode_final
```

The syntax is telling you that the column containing the numbers of interest is column 10. I believe that is usually the case for energy output (if you want eV). But, you should check your file. For the Mos it will be different (I believe it is column 7 so you will use '{print \$7}').

# Plotting and analysis

You may copy your files over to your PC. If you use Igor you can automate the process of reading in the files.