Introduction to Gaussian

Computational chemistry using the Arctic Region Supercomputing Center installation of *Gaussian 03*

John Keller
Department of Chemistry & Biochemistry
University of Alaska Fairbanks

1. Computational chemistry background
2. Using Gaussian
    (a) input
    (b) output

3-18-2008

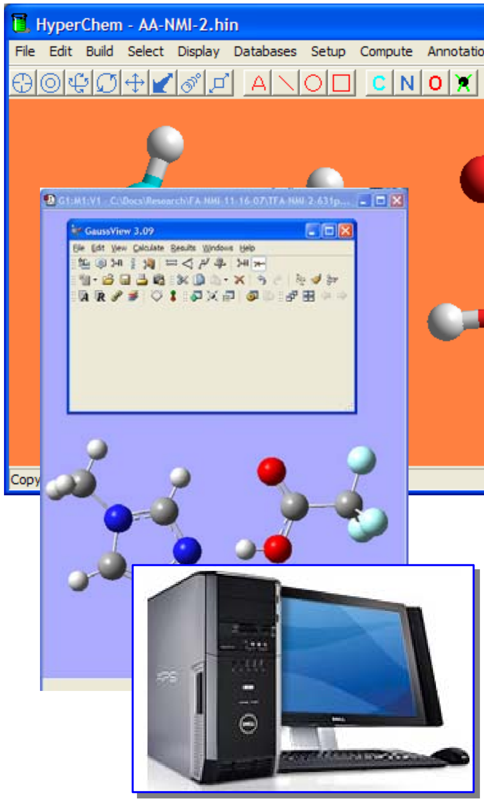
Back in the day, it was not clear that mathematical treatments would ever be useful in chemistry!

> "Every attempt to employ mathematical methods in the study of chemical questions must be considered profoundly irrational and contrary to the spirit of chemistry."
>
> – Auguste Compte, French philosopher, 1830

Today computational chemistry is a standard tool of science. For example, there are now >20 journals devoted to the subject.

Chemometrics and Intelligent Laboratory Systems
Computational and Theoretical Polymer Science
Electronic Journal of Theoretical Chemistry
International Journal of Quantum Chemistry
Journal of Chemical Information and Computer Sciences
Journal of Chemical Theory and Computation
Journal of Chemometrics
Journal of Computational Chemistry
Journal of Computer Aided Chemistry
Journal of Mathematical Chemistry
Journal of Molecular Graphics and Modelling
Journal of Molecular Modeling
Journal of Molecular Structure: THEOCHEM
Journal of the Chemical Computing Group
Journal of Theoretical and Computational Chemistry
Macromolecular Theory and Simulations
Perspectives in Drug Discovery and Design
Proteins Structure, Function, and Genetics
Structural Chemistry
Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretical Chimica Acta)

Univ of Alaska Fairbanks/ARSC Resources



Kerberos
FTP
Telnet


Midnight


Iceberg

GaussView 03 (UA site license)
HyperChem v 8

Gaussian 03
Gaussian utilities (e.g., fmchk)
GaussView 04 (UNIX)

Capabilities of Gaussian (similar to other large-scale computational chemistry packages):

*JWK has used.

- Determine most stable (optimum) molecular geometry and energy. (Used to calculate ΔE and K_{eq}.) *
- Define a **potential energy surface** by stepping through a range of values for a geometry coordinate, such as bond distance or torsion angle
- Predict IR, Raman, UV, NMR, and other spectra
- Optimize transition states
- Solvate molecules using the "polarized continuum (PCM) " or other models.
- Special tools for optimizing transition metal complexes, and other molecules containing large atoms.
- "ONIOM" technique for defining layers within one molecule where higher and lower accuracy methods can be applied.
- Model surfaces using a 2D periodic boundary condition (PBC) method, or crystals using 3D PBC.

Gaussian and HyperChem Resources

The Gaussian manual is on-line at
http://www.gaussian.com/g_ur/g03mantop.htm

ARSC issues relating to user accounts, connections, Kerberos etc are best answered by ARSC consultants. Email at consult@arsc.edu or telephone 450-8602

Technical questions on running Gaussian are best directed to Gaussian consultants at help@gaussian.com

Current users including J. Keller (ffjwk@uaf.edu) may also be able to make constructive suggestions.

The full manual of *HyperChem* methods and theory in pdf format is installed on each computer in the chemistry computer lab.

Computational chemistry background

Basically two ways to calculate molecular structure:

1. Molecular mechanics

2. Quantum mechanics

    Basis sets
    Four different methods
    Geometry optimization
    Comparing different methods & basis sets with an example

Molecular mechanics

Each bond and bond angle is assigned a potential function based on experimental bond distances and angles. These are added up and then minimized as the geometry is adjusted.

The set of potential functions is called a “force field”.

There are 3 force fields available in Gaussian: AMBER, DREIDING, UFF.

Fast. Used mainly for large molecules. AMBER is the standard force field used in protein optimization and molecular dynamics.

Major limitations: No electron density, vibrations, or other spectroscopic results.

Quantum mechanics

Based on the Schrödinger Equation.

$$H\Psi = E\Psi$$

H is energy operator which describes the kinetic and potential energy of an electron in field of nuclei and other electrons.

Usually the **nuclei** are assumed to be **stationary**.

(Electrons move at ~c/3.
Nuclei move at ~c/25,000.)

Ψ is f(x,y,z)
A set of spatial distributions describing the probability of finding electrons (orbitals).

Ψ is the product of **n** spatial functions (*ψ's*), one for each of **n** electrons in the atom or molecule.

A single atom has AOs.
Molecules have MOs.

E = sum of the energies of the orbitals, which may contain 1, or at most 2, electrons.

Gaussian approximates orbital shapes and orbital energies of a given molecular geometry using a **model chemistry** consisting of two parts: a **basis set** and a **method**.

Basis Sets

- Molecular orbitals (MOs) are approximated as linear combinations of basis functions (the **basis set**), which mathematically look like s, p, or d atomic orbitals.

- Each atom is assigned several (or many) basis functions.

- In turn, each basis set function is comprised of a number of gaussian functions called **primitives**.

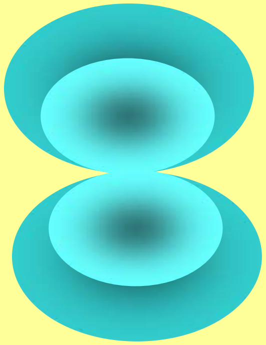
Calculations using **large basis sets** are **more accurate** because they are less restrictive on the location of the electrons.

Such calculations are also **more expensive** because they require computing more integrals.

Basis Sets, continued

Split-valence basis sets were developed by John Pople, 1998 Nobel Prize winner and founder of Gaussian, Inc.

The **basis set nomenclature** describes how the basis functions are constructed from the gaussian primitives (the “contraction scheme”.)



A p-orbital in the 6-31G basis set is ***split*** into **inner** and **outer** p-type sub-orbitals.

Example: 6-31G basis set

The number of digits (2 here) indicates how many **sub-orbitals** each **valence atomic orbital** is split into.

The first number indicates the # of gaussian primitives used for **core (inner)** electrons.

The values tell how many gaussian primitives make up each sub-orbital.

Basis Sets: adding polarization and diffuse basis functions

 # and type of basis functions

Basis set	Each H atom gets:	Each C, O, or F atom gets:
6-31G	**2** (1s,1s’)	**9** (1s,2s,2s’,3p,3p’)
6-31G(d) or 6-31G*	**2** (1s,1s’)	**15** (1s,2s,2s’,3p,3p’,6d)
6-31G(d,p) or 6-31G**	**5** (1s,1s’,3p)	**15** (1s,2s,2s’,3p,3p’,6d)
6-31G+(d,p) or 6-31+G**	**5** (1s,1s’,3p)	**19** (1s,2s,2s’,2s+,3p,3p’,3p+,6d)
6-31G++(d,p) or 6-31++G**	**6** (1s,1s’,1s+,3p)	**19** (1s,2s,2s’,2s+,3p,3p’,3p+,6d)

Diffuse functions extend further out from the nucleus, and are necessary for modeling non-bonding or weakly-bonding systems.

Methods

These differ mainly according to how, or if, "electron correlation" is treated.

Electron correlation is the tendency of electrons to **avoid each other**, even within the same orbital.

Four common quantum methods:

Semi-empirical

Uses pre-calculated orbital functions for different kinds of atoms. Fast, but less accurate.

Hartree-Fock (HF),
Møller-Plesset *n* (MP2),
Density Functional Theory (DFT)

"ab initio" methods that require significant computing power for accurate results.

Hartree-Fock (HF) method

Electron correlation is **ignored**. The many-electron wavefunction (Ψ) is estimated by the **self-consistent field (SCF)** method.

This starts with an **initial guess** of Ψ ("Ψ_o"), which is the product of the initial estimates of the molecular orbitals.

$$\Psi_o = \psi_0(1)\psi_0(2)\psi_0(3)\cdots$$

Then, while all the other MO functions are held constant (the "field"), each ψ is varied so as to minimize the total energy. This is the "**variational method**." This gives a new Ψ ("Ψ_1"):

$$\Psi_1 = \psi_1(1)\psi_1(2)\psi_1(3)\cdots$$

This process is repeated until the change in energy is close to zero. These orbitals then comprise the self-consistent field.

Møller-Plesset (MP)

Electron correlation is accounted for allowing one or more electrons to occupy higher-energy, unoccupied ("anti-bonding") MO's.

This results in an energy correction that lowers the total energy because it **lowers the electron-electron repulsion energy.**

The commonly-used MP**2** method actually considers the effect of **2 electrons** occupying anti-bonding orbitals.

MP2 is more expensive than HF because more configurations must be calculated.

Density Functional Theory

Developed by Walter Kohn of UC-Santa Barbara and others.

Kohn shared the 1998 Nobel Prize with John Pople.

Gives nearly the same accuracy as MP2, but with the (lower) computational cost of Hartree-Fock SCF.

A **function** transforms a *number* into another number.

$$2 \xrightarrow{x^3} 8$$

$$f(x) = x^3$$

A **functional** transforms a *function* into a number.

$$f(x) = x^3 \xrightarrow{\int_0^2 f(x)dx} \left.\frac{x^4}{4}\right|_0^2 = 4$$

$$F[f(x)] = \int_0^2 f(x)\,dx$$

Computational Chemistry, E. Lewars, 2003, Kluwers, p. 388.

Density Functional Theory

Hohenberg-Kohn Theorem:

Energy of a molecule
= *F* [electron density] , where electron density = f(x,y,z))

This says that there **exists** a functional that will calculate molecular energy from electron density.
But it does not say what the functional is!

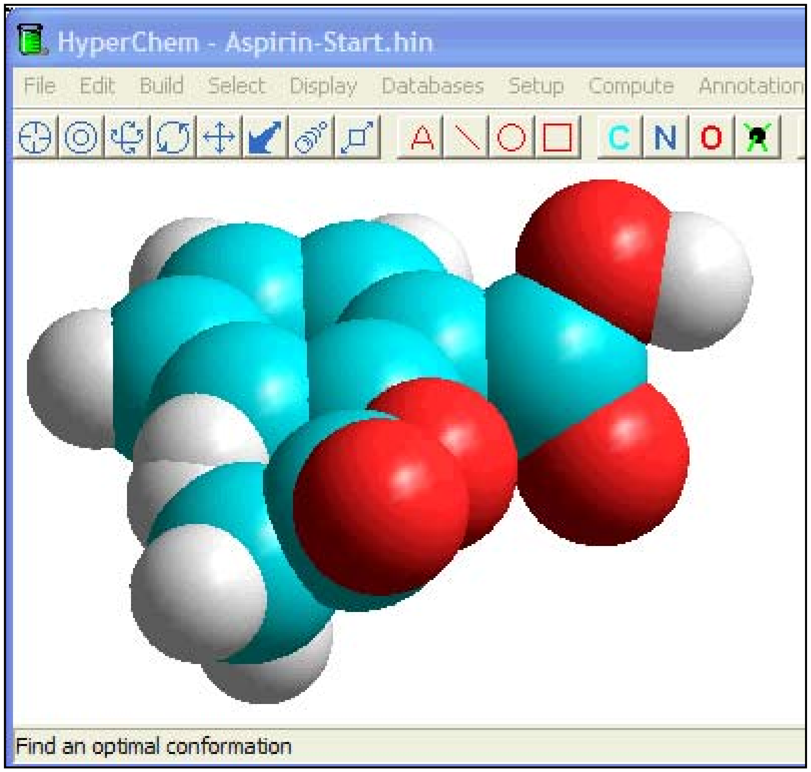
DFT methods account for electron correlation by estimating the interaction of an electron with the total electron density.

DFT orbitals are formed from **basis functions** like those used in SCF or MP2.

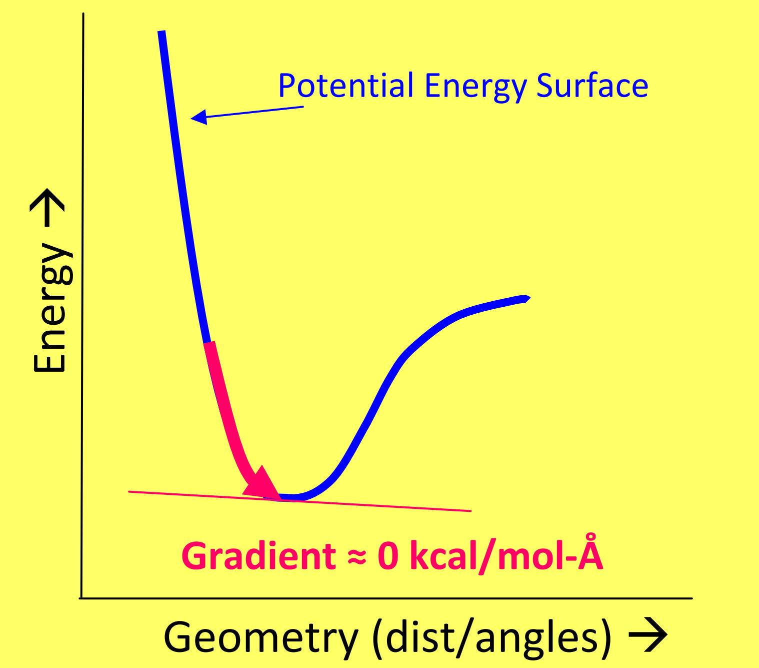
Most popular DFT method is **B3LYP**. (**B**ecke **3**-Parameter method for calculating that part of the molecular energy due to overlapping orbitals, plus the **L**ee-**Y**ang-**P**arr method of accounting for correlation.)

Geometry Optimization

AMBER optimization starting with a non-optimal conformation of the aspirin molecule. Visualized in real time with *HyperChem*.



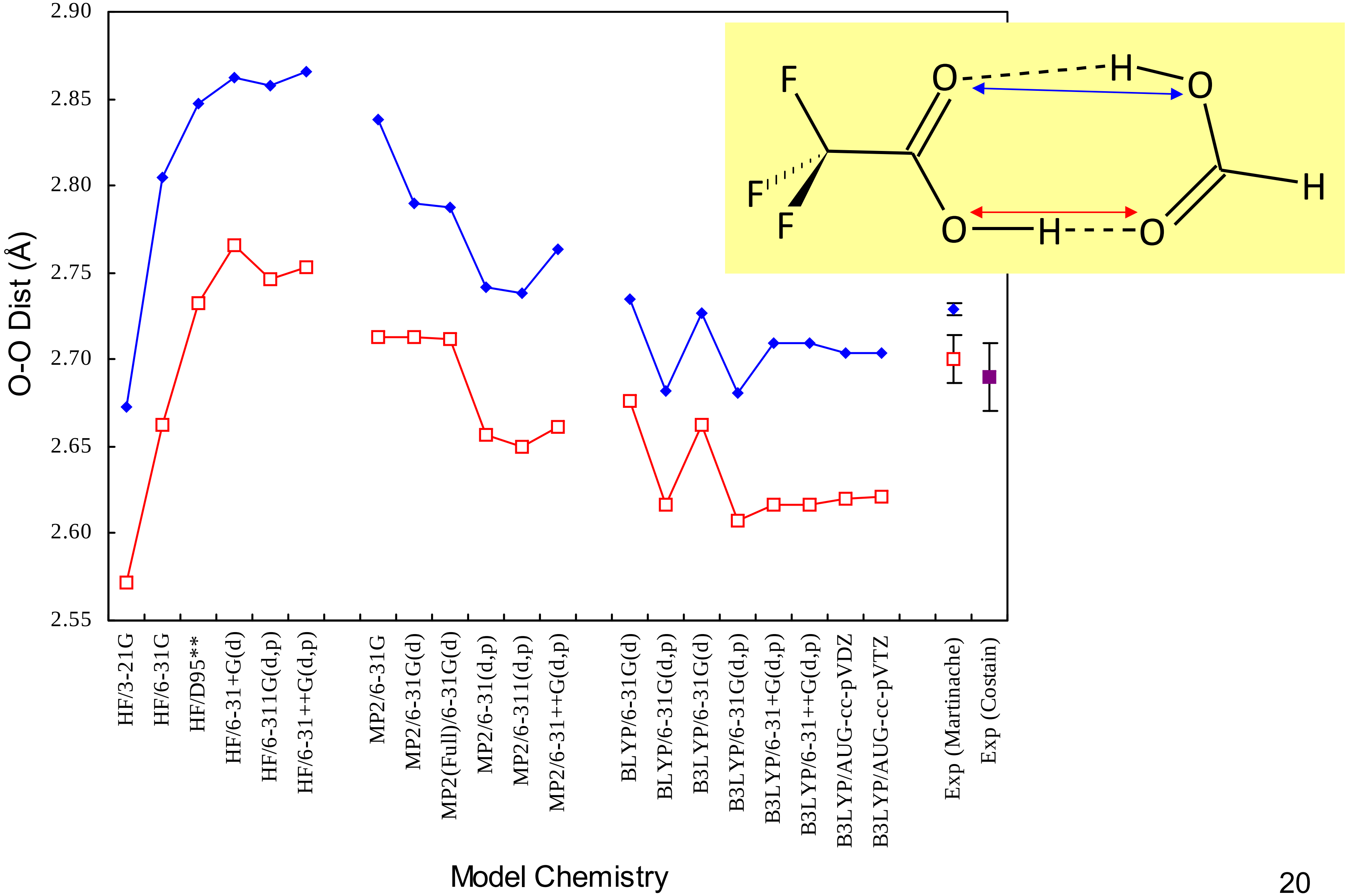
Find an optimal conformation



Optimization adjusts geometry to decrease the total energy until it reaches the nearest minimum.

In *Gaussian/GaussView,* optimization steps are visualized by querying the .log output file.

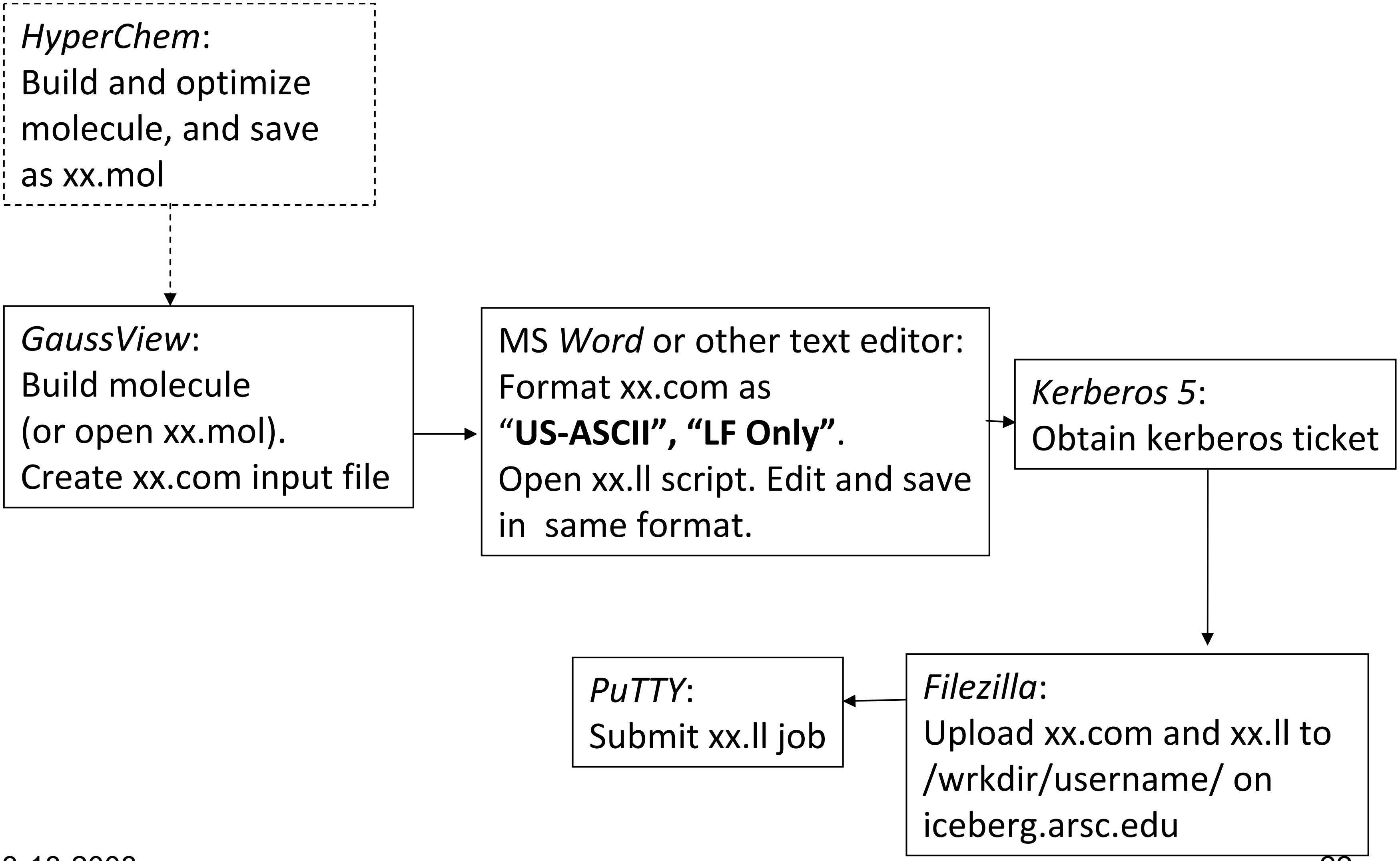
Using different **model chemistries** to estimate O-O distances in the formic acid-trifluoroacetic acid bimolecule.*



*J. W. Keller, *J. Physical Chemistry A* **108**, 4610-4618(2004)

Using Gaussian and GaussView

Create and submit a Gaussian input file



HyperChem:
Build and optimize molecule, and save as xx.mol

GaussView:
Build molecule
(or open xx.mol).
Create xx.com input file

MS *Word* or other text editor:
Format xx.com as **“US-ASCII”, “LF Only”**.
Open xx.ll script. Edit and save in same format.

Kerberos 5:
Obtain kerberos ticket

Filezilla:
Upload xx.com and xx.ll to /wrkdir/username/ on iceberg.arsc.edu

PuTTY:
Submit xx.ll job

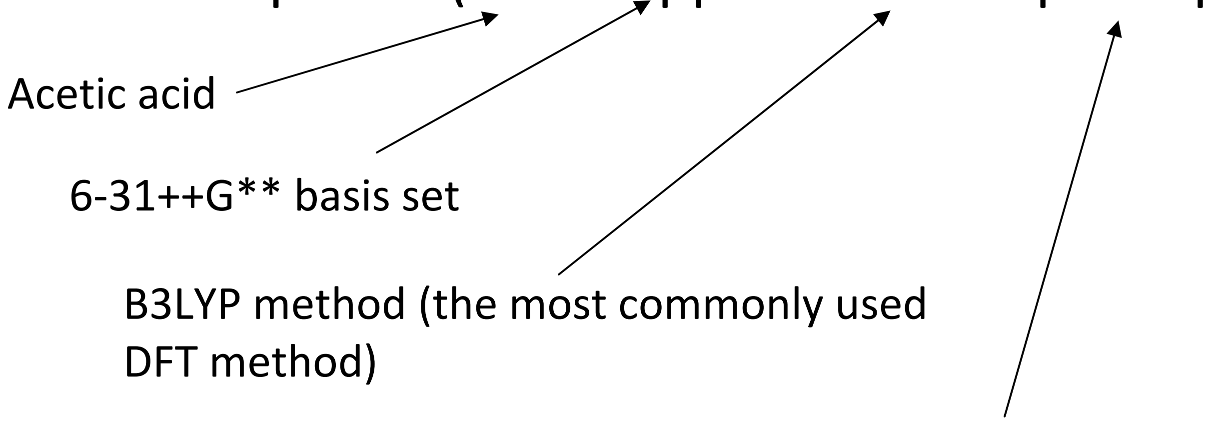
Suggestion:

Use long filenames to keep track of different computational jobs.

Molecule-BasissetMethod-Jobtype.com

Gaussian input file (AA-631ppGssB3LYP-OptFreq.com)

PBS script file (AA-631ppGssB3LYP-OptFreq.cmd)



Acetic acid

6-31++G** basis set

B3LYP method (the most commonly used DFT method)

Optimization and Vibrational frequency calculation

Gaussian input file (xx.com) can be typed in manually, or created by *GaussView*

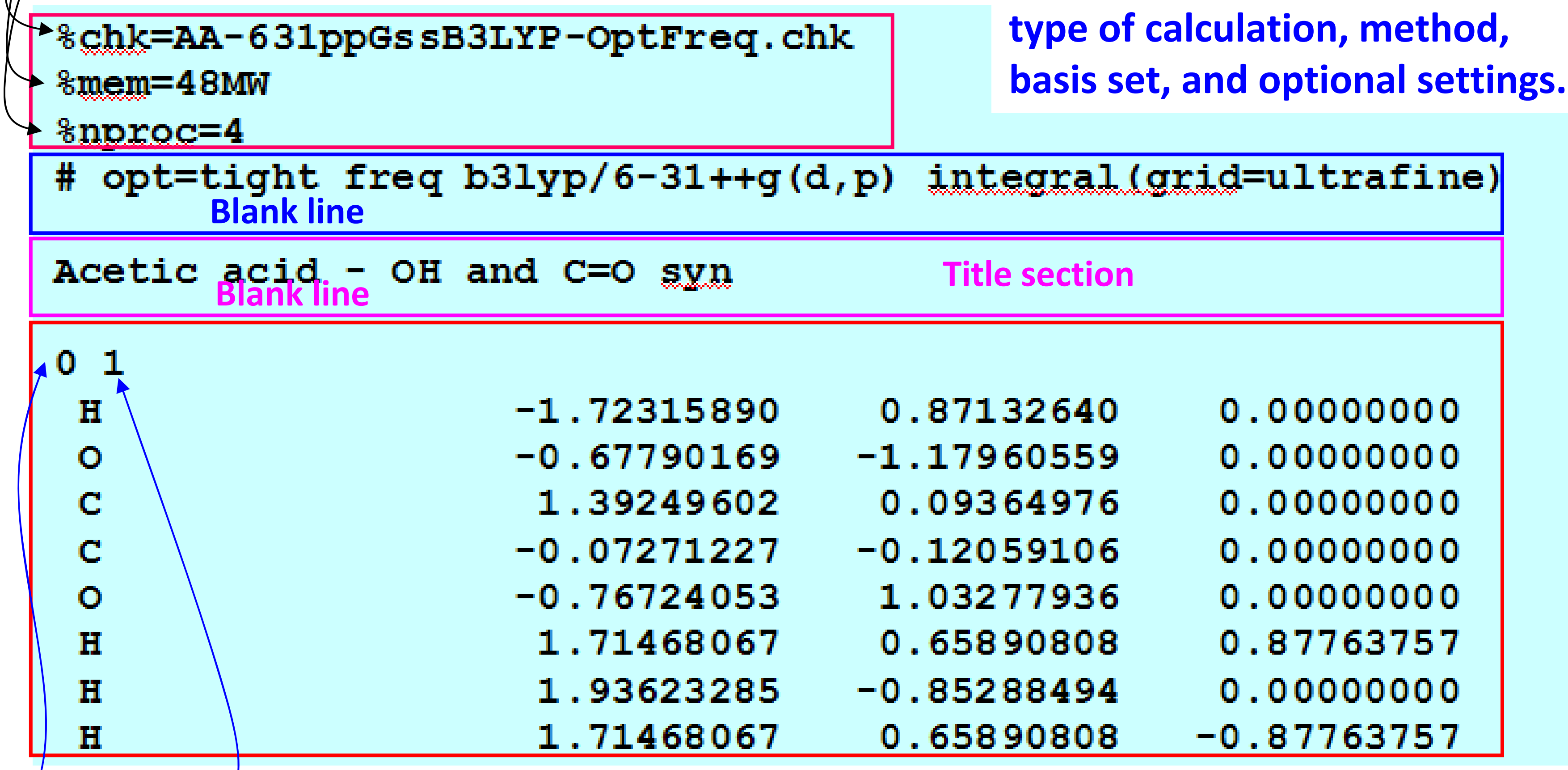
Optional “Link 0” lines for job control:

******Names checkpoint file (saves output)*******

Sets amount of shared memory

Sets number of processors (no blank line terminator)

Route section: keywords define type of calculation, method, basis set, and optional settings.



```
%chk=AA-631ppGssB3LYP-OptFreq.chk
%mem=48MW
%nproc=4
# opt=tight freq b3lyp/6-31++g(d,p) integral(grid=ultrafine)

Acetic acid - OH and C=O syn

0 1
 H                  -1.72315890    0.87132640    0.00000000
 O                  -0.67790169   -1.17960559    0.00000000
 C                   1.39249602    0.09364976    0.00000000
 C                  -0.07271227   -0.12059106    0.00000000
 O                  -0.76724053    1.03277936    0.00000000
 H                   1.71468067    0.65890808    0.87763757
 H                   1.93623285   -0.85288494    0.00000000
 H                   1.71468067    0.65890808   -0.87763757
```

Blank line

Title section

Blank line

Molecule specification

Net charge

Multiplicity (“1” for even-electron species such as molecules; “2” for odd-electron species such as NO molecule)

A **script file** is used to specify job details.

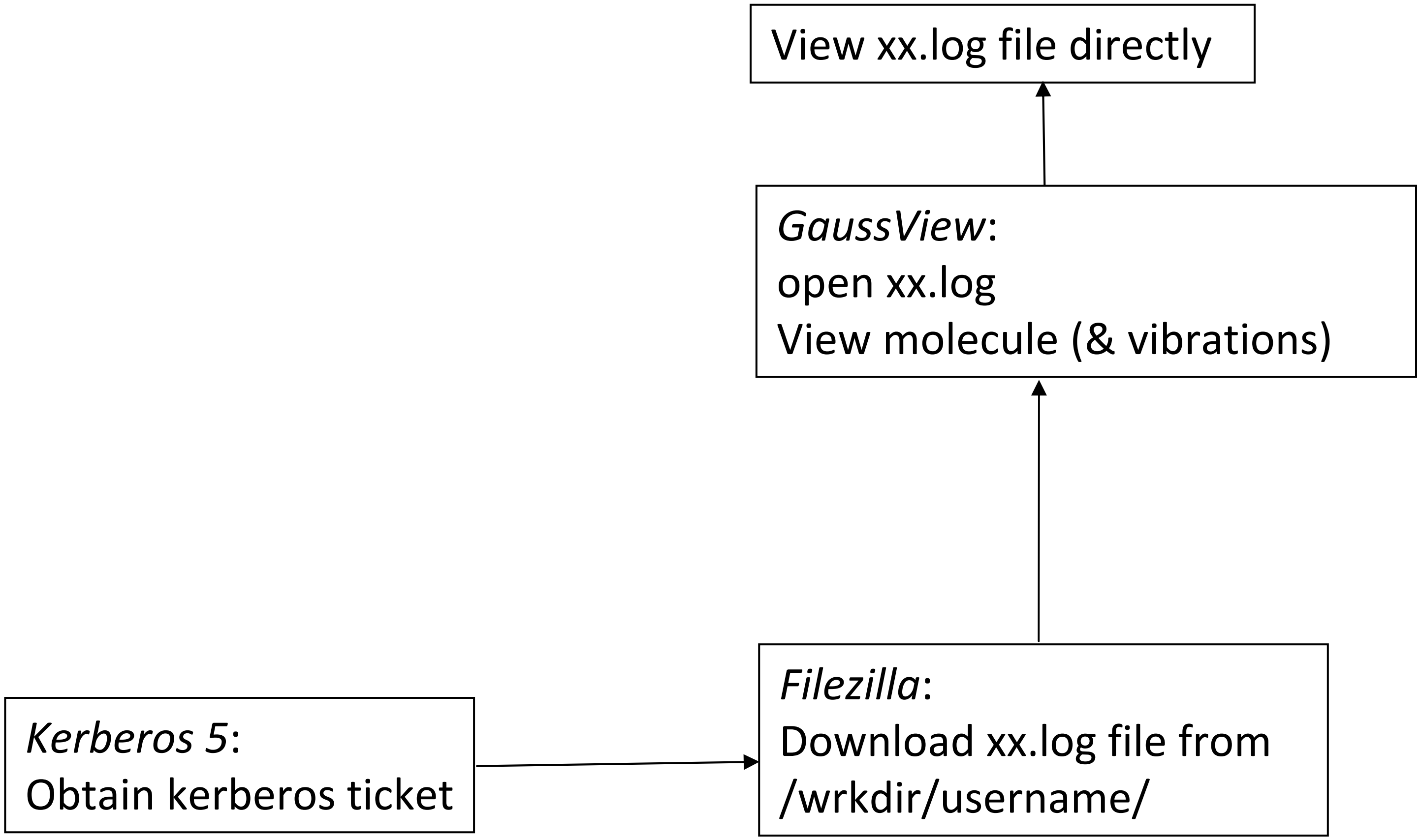
Iceberg uses the Loadleveler job queue software (xx.ll).
Midnight uses the PBS job queue software (xx.cmd).

A script must be used for each job.

```
#!/bin/ksh
#PBS -q standard
#PBS -l walltime=1:00:00
#PBS -l select=4:ncpus=4:node_type=4way
#PBS -j oe

# for midnight
# standard class

# Change directory to where the script is submitted
cd $PBS_O_WORKDIR

# Set Location of Gaussian
export GAUSS_SCRDIR=$WRKDIR
export g03root=/usr/local/pkg/gaussian/current
. $g03root/g03/bsd/g03.profile

# run the calculation
g03 AA-631ppGssB3LYP-OptFreq.com
```

This should match the Link0 “nproc” line

Some *Gaussian* output is contained in the xx.log file.



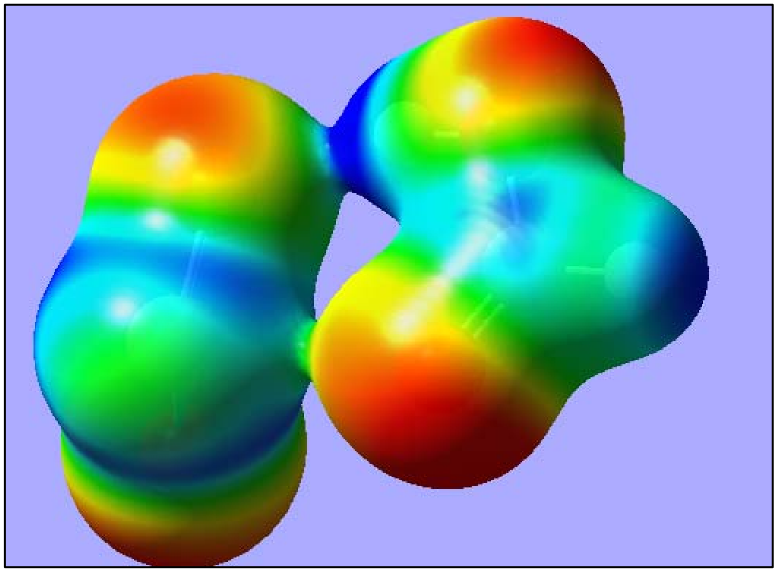
View xx.log file directly

GaussView:
open xx.log
View molecule (& vibrations)

Kerberos 5:
Obtain kerberos ticket

Filezilla:
Download xx.log file from
/wrkdir/username/

Electrostatic potential output is contained in the xx.chk file.



These surfaces are calculated by GaussView (on local or remote machine) using a “formatted checkpoint file” as input. This file is created from the standard checkpoint file (xx.chk) by the Gaussian **formchk utility** program, which runs on Iceberg or Midnight.

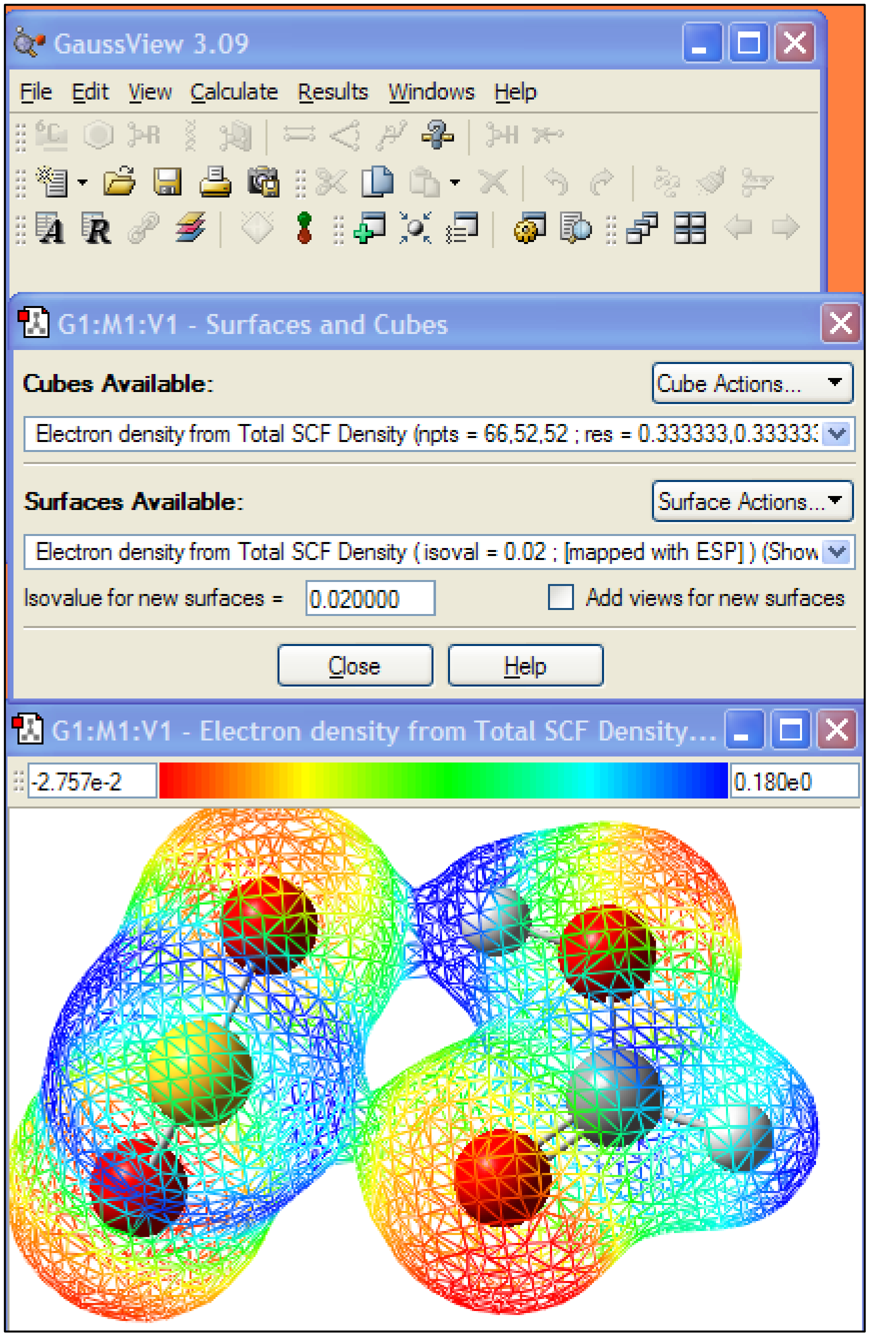
Log on to Iceberg

```
cd /wrkdir/username/
ls                              (to get filename list)
setenv g03root /usr/local/pkg/gaussian/current
$g03root/g03/formchk filename.chk    (paste it)
```

This should create the new file   filename.fchk.
Transfer filename.fchk to your machine by FTP.

Calculating and Displaying an Electrostatic Potential Surface

- In GaussView, open the xx.fchk file.
- Do Results, Surface, Cube Actions, New Cube, Total Density, Coarse, OK.
- Set Isovalue for new surface to about 0.02 (units of electrons/Å^3).
- Under Surface Actions, choose New Mapped Surface, ESP.
- Adjust color scale Min/Max



Cubes Available:

Electron density from Total SCF Density (npts = 66,52,52 ; res = 0.333333,0.33333

Surfaces Available:

Electron density from Total SCF Density (isoval = 0.02 ; [mapped with ESP]) (Show

Isovalue for new surfaces = 0.020000

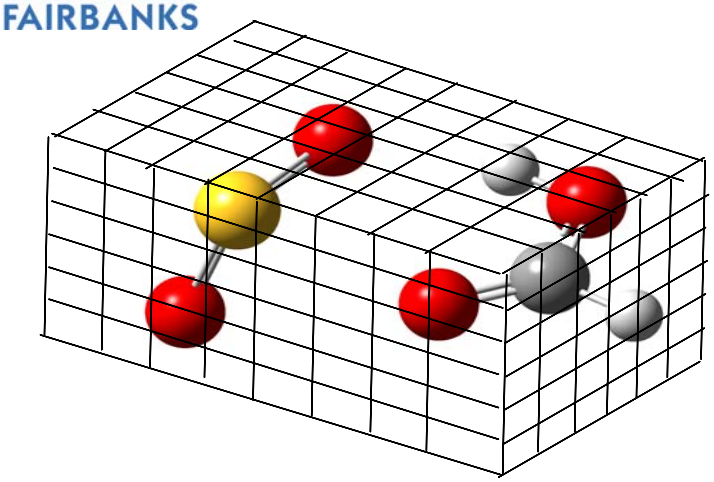
☐ Add views for new surfaces

-2.757e-2

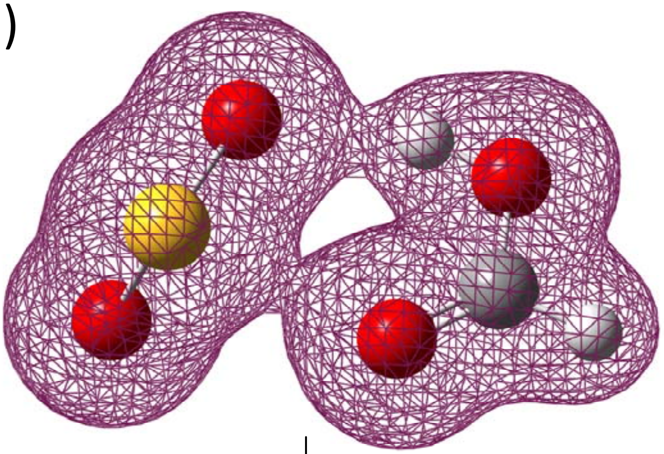
0.180e0

3-18-2008

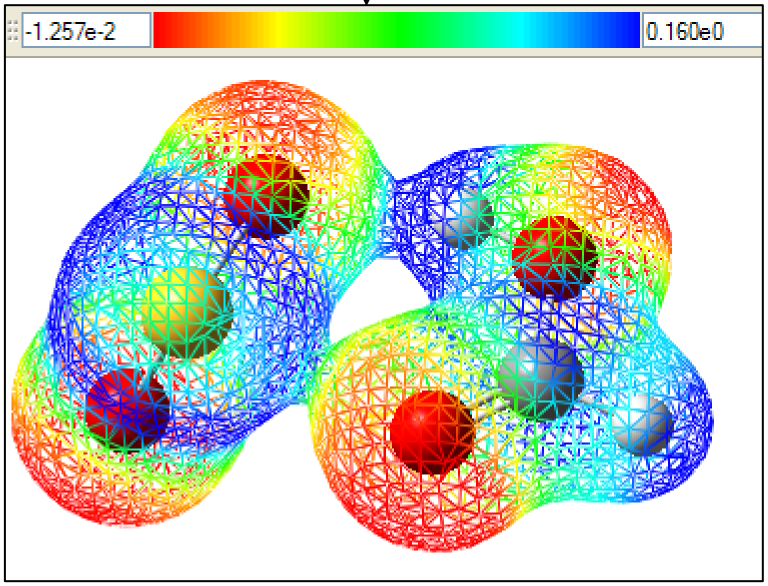
(1) A “cube” of electron density is created.



(2) The isosurface is created by connecting boxes of equal electron density ($0.02\ e/\text{Å}^3$ here)

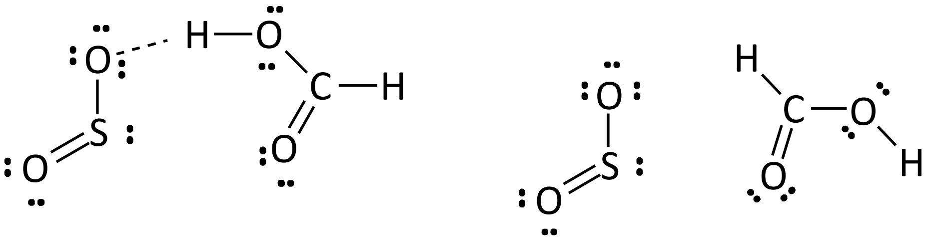


(3) The electrostatic potential is calculated at each point on the electron density surface. The potential values are **painted** on the surface using a red-to-blue scale.



One Current Project

investigates the stability and vibrational spectra of SO_2-formic acid complexes.



1) Are the complexes stable enough to exist in natural SO_2-acid gas mixtures such as volcano gases?
2) Can we observe such complexes in the laboratory?
3) How important are the S-to-O dipolar interactions?

Students

- Gaussian calculations on these complexes were carried out by **Bronwyn Harrod**, then a freshman chemistry major. She presented a poster at the 2007 Mercury Conference on Computational Chemistry .

- **Sifat Chowdhury**, a West Valley H.S. student, worked on this project as a part of his Alaska High School Science Symposium project.

Acknowledgements

National Science Foundation for purchase of *HyperChem* and *GaussView* licenses.

University of Alaska Foundation

Arctic Region Supercomputing Center for installing *Gaussian*, and providing ample amounts of computer time.

The End

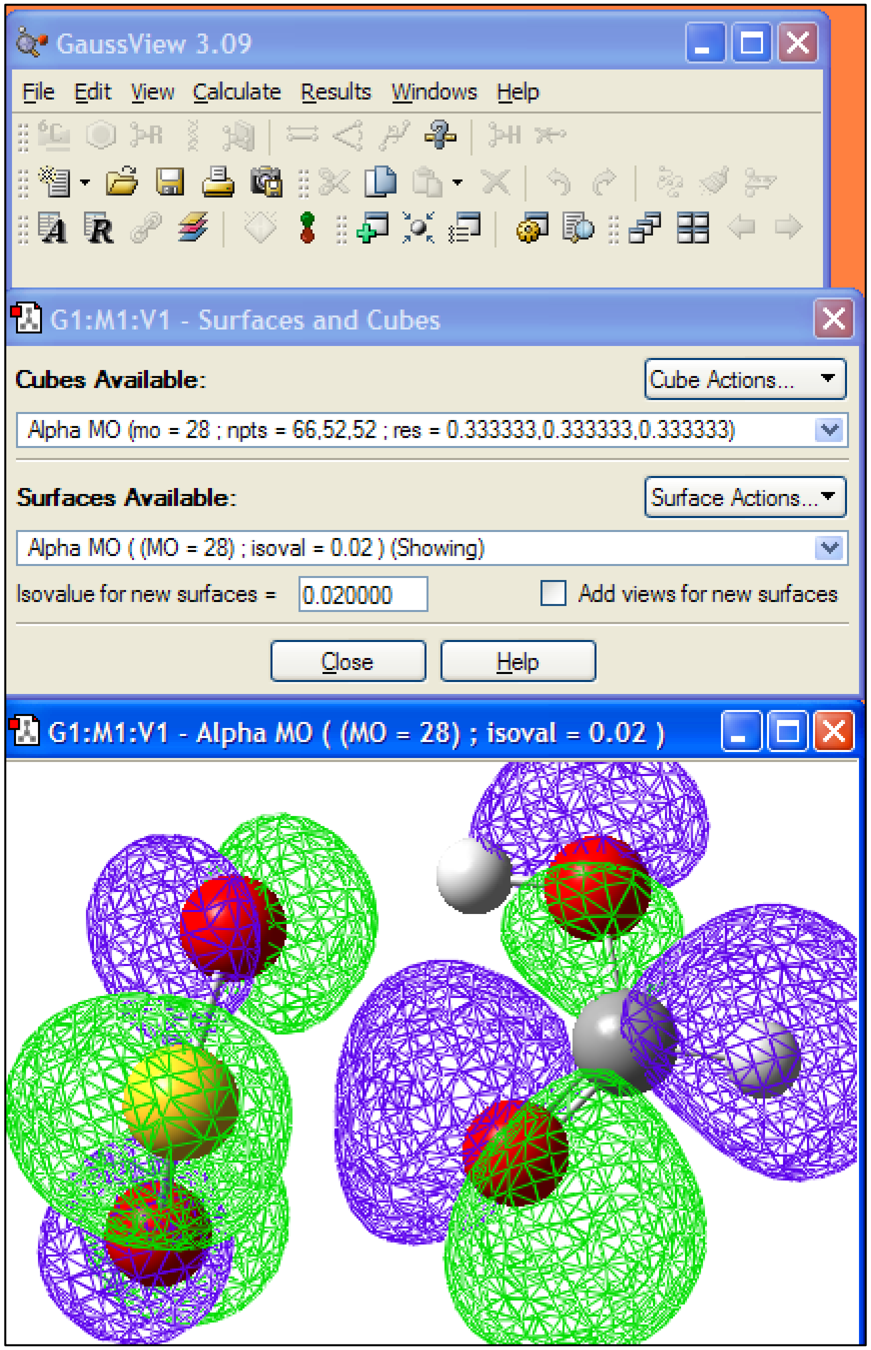
Geometry optimization of **aspirin**

	HyperChem v 8	***Gaussian 03***
	WinXP/Dual Core CPU	**Iceberg/2 Processors**
Method	PM3 semi-empirical	PM3 semi-empirical
Final gradient (kcal/mol-Å)	0.00029	0.0000092
# Iterations	596	33
Heat of formation (kcal/mole)	-145.2472	-145.1500
Dipole moment (Debyes)	4.742	4.7423
Wall clock time	1 min 22 sec	23 sec

~4x faster

Calculating and Displaying Molecular Orbitals

- In GaussView, open the xx.fchk file.
- Do Results, Surface, Cube Actions, New Cube, Molecular Orbital, HOMO, Coarse, OK.
- Set Isovalue for new surface to about 0.02 (units of electrons/Å³).
- Under Surface Actions, choose New Surface, MO
- (Surface style is set using File, Preferences, Surface, Format.)



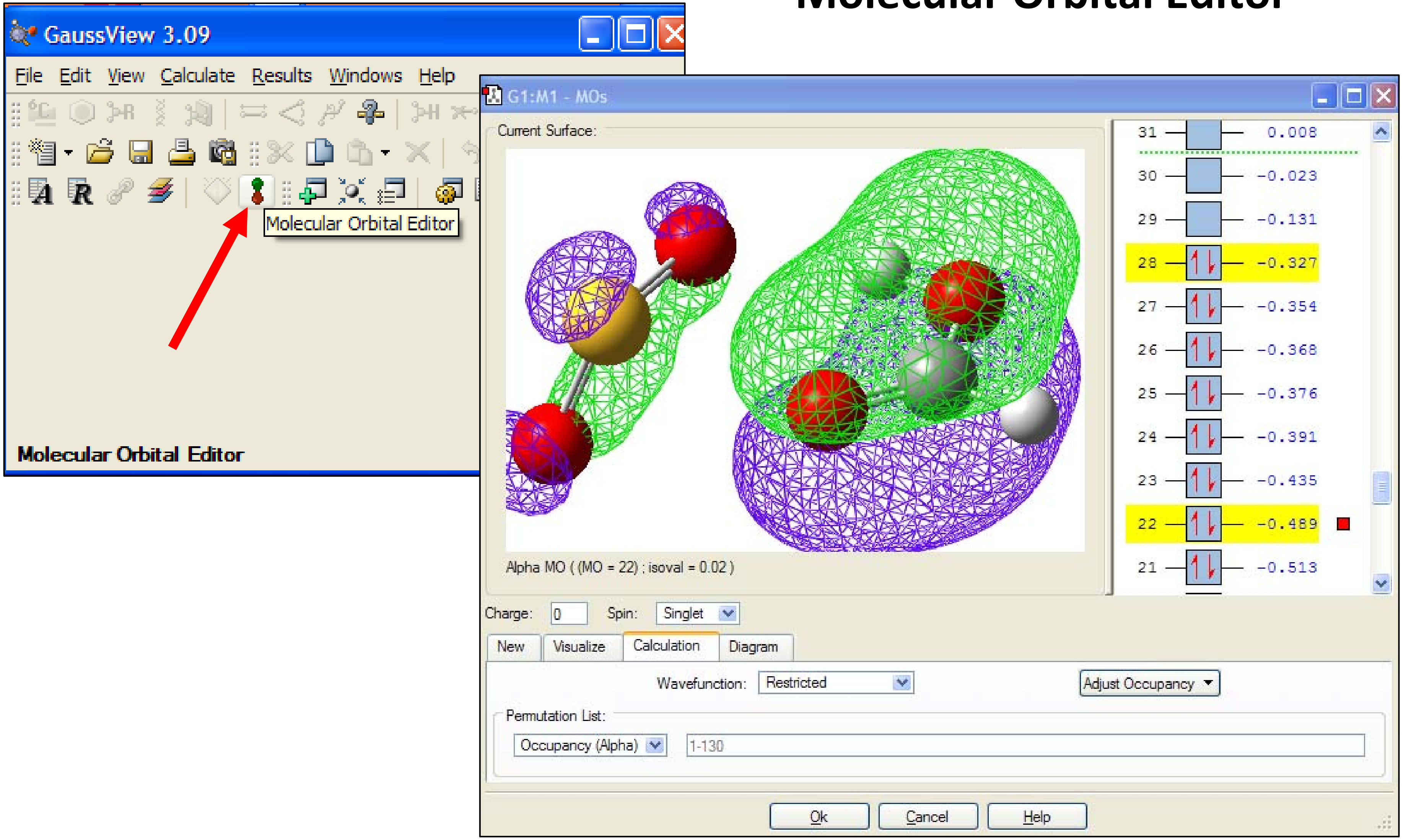
Cubes Available:
Alpha MO (mo = 28 ; npts = 66,52,52 ; res = 0.333333,0.333333,0.333333)

Surfaces Available:
Alpha MO ((MO = 28) ; isoval = 0.02) (Showing)

Isovalue for new surfaces = 0.020000

☐ Add views for new surfaces

Molecular orbital scheme can be viewed using the Molecular Orbital Editor



Molecular Orbital Editor

Alpha MO ((MO = 22) : isoval = 0.02)

31	0.008
30	-0.023
29	-0.131
28	-0.327
27	-0.354
26	-0.368
25	-0.376
24	-0.391
23	-0.435
22	-0.489
21	-0.513

Charge: 0

Spin: Singlet

Wavefunction: Restricted

Permutation List: Occupancy (Alpha) 1-130

Rotational constants (in GHz) are measured by microwave spectroscopy