Contents

Software required:

Connection Software (Mac, Win, Linux):

1) Client SSH: PuTTY (Win)

2) Client SFTP: WinSCP (Win)

3) Client SFTP: gftp (Linux)

4) Client SFTP: Classic FTP for Mac 2.13 (Mac)

Molecular Editor (Mac, Win, Linux):

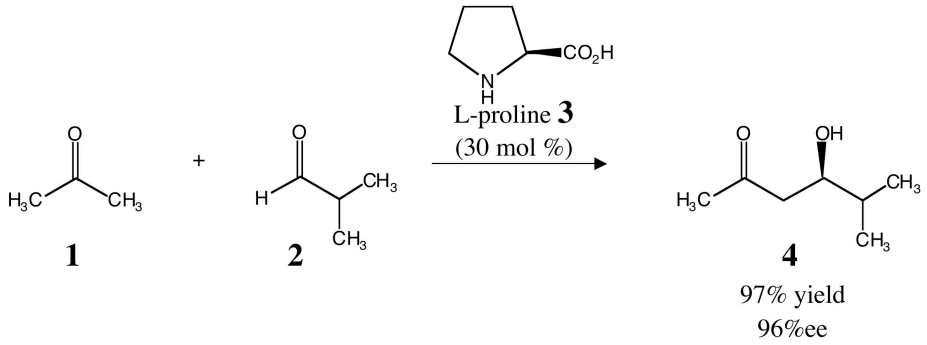
1) Avogadro

Part 4. Exercise

4.a Introduction

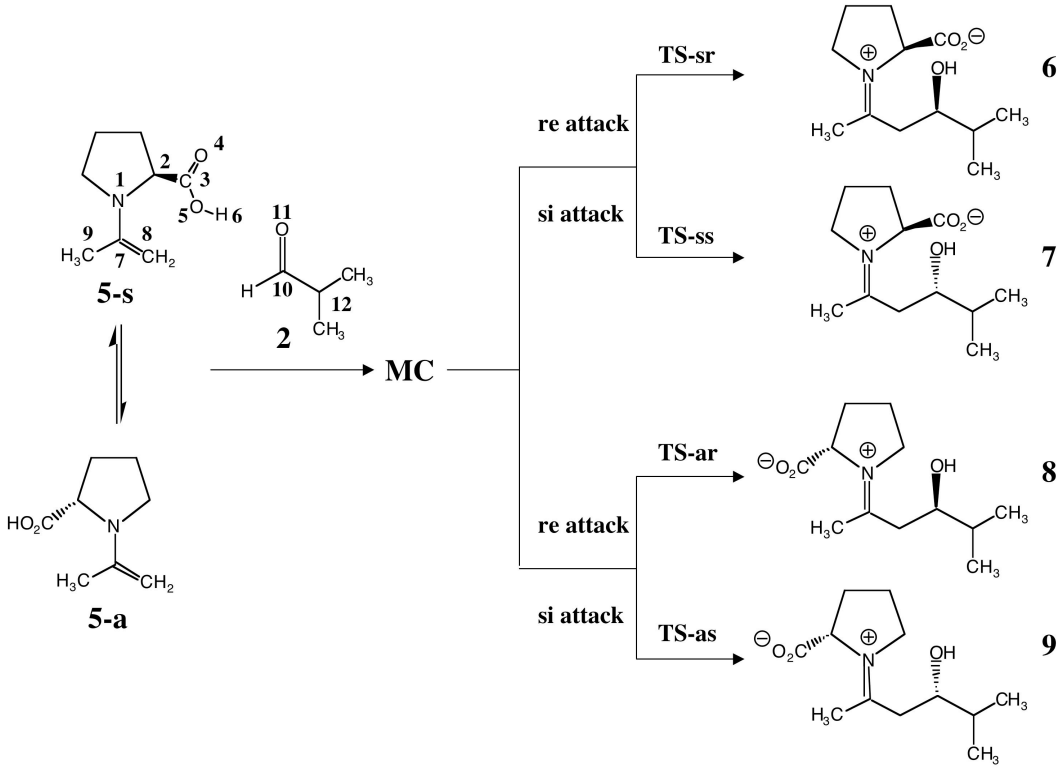
Origin of the stereoselectivity of the Direct Aldol Reaction between Acetone and Isobutyraldehyde (Arnó et al., Theor. Chem. Acc. (2002) 108:232-239)

We will take as an example of direct aldol reaction the reaction of acetone and isobutyraldehyde (Scheme 1) that illustrates the catalytic potential of simple organic molecules in asymmetric synthesis.



Scheme 1

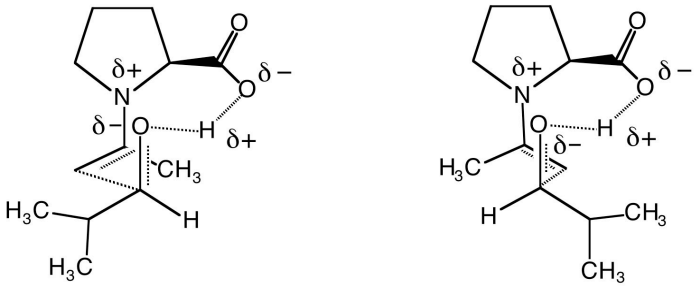
Because, as stated in section x, the computed transition states associated with the C-C bond-formation step of the proline-catalyzed intermolecular aldol reaction provide models for understanding the stereoselectivities of this reaction, we will focus our attention on these structures using density functional theory methods at the B3LYP/3-21G computational level. For this step, which is the stereocontrolling and rate-determining step, four reactive channels corresponding to the syn and anti arrangement of the active methylene of the enamine relative to the carboxylic acid group of L-proline (enamines **5-s** and **5-a** in Scheme 2, respectively) and the re and si attack modes to both faces of the aldehyde carbonyl group can be determined. In particular, they are related to the nucleophilic attack of the C8 carbon atom of the active methylene group of the enamines **5-s** and **5-a** to the re and si faces of the carbonyl group of isobutyraldehyde, to give the zwitterionic iminium intermediates **6–9** (Scheme 2). After the hydrolysis, the iminium intermediates **6** and **8**, and **7** and **9** give (R)-cetol (**4**) and its enantiomer (S)-cetol, respectively.



Scheme 2

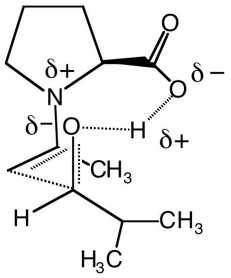
Thus, four TSs, **TS-sr**, **TS-ss**, **TS-ar** and **TS-as**, can be localized and characterized for this reaction. They are related to the attack of active methylene of the enamines **5-s** and **5-a**, named as s and a, to the re and si faces of the carbonyl group of isobutyraldehyde, named as r and s. MC in Scheme 2 stands for molecular complexes associated with an early stage of the C-C bond-formation process.

Schematic representations of the structures of the four TSs are shown in Scheme 3.

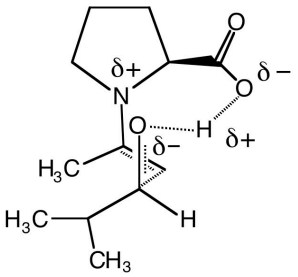


TS-ar

TS-ss



TS-as



TS-sr

Scheme 3

The formation of an intermolecular hydrogen-bond between the acidic hydrogen of L-proline and the carbonyl oxygen atom of the aldehyde in an early stage of the process catalyzes very effectively the *C-C* bond-formation by a large stabilization of the negative charge that is developing at the carbonyl oxygen atom along the nucleophilic attack, thus decreasing the activation energy of the primary enamine-catalyzed aldol reactions. Furthermore, favourable hydrogen bonding interaction between the partial positive hydrogens of the carbon adiacent to the proline nitrogen to the forming alkoxide ($^{\delta+}$NCH--O$^{\delta-}$) contribute to the electrostatic stabilization in TSs associated with the anti arrangements (see, for example, the two structures for **TS-ar** and **TS-ss** in Figure 1; $^{\delta+}$NCH--O$^{\delta-}$ distance is 2.164 *Å* in **TS-ar** while in **TS-ss** is 3.007 *Å*). Thus, as a consequence of the hydrogen-bond network formation, the reactive channels associated with the anti arrangement of the enamine are favored over the channels associated with the syn arrangement.

In addition, along the anti channels the attack of the active methylene on the re face of the aldehyde is favored over the attack on the si face, because of the steric repulsion involving the substituents, in agreement with experiment.

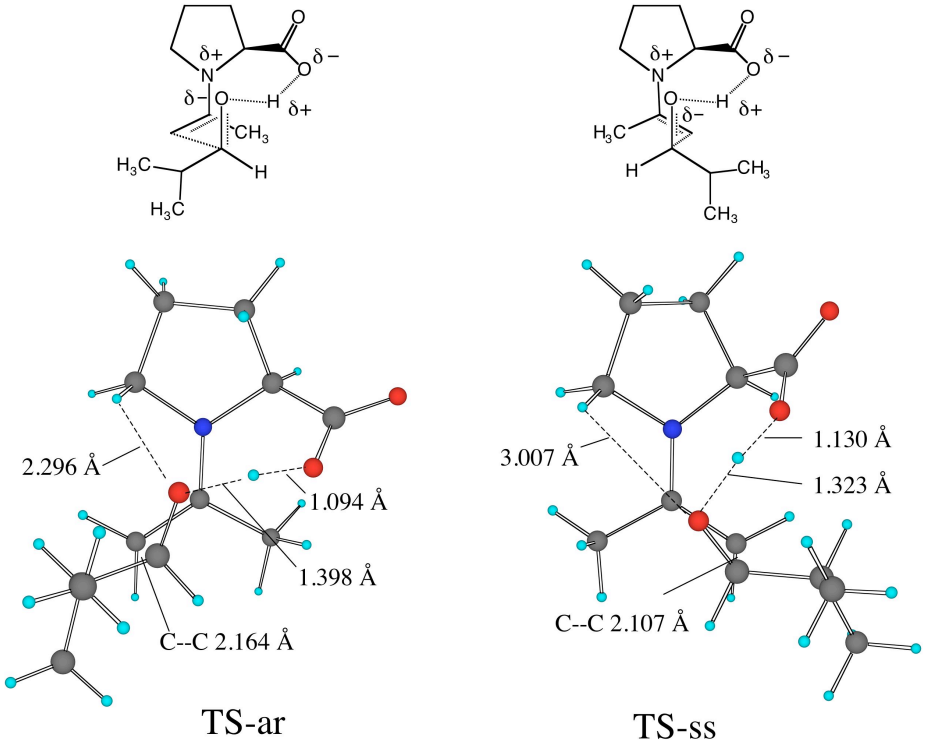


Figure 1

In the original paper by Arnò et al., the B3LYP/6–31G* barriers in the gas phase for the C-C bond-formation process, via **TS-sr**, **TS-ss**, **TS-ar** and **TS-as**, relative to MC are 14.4, 13.9, 9.6 and 11.0 kcal/mol, respectively. These barriers indicate that the reactive channels associated with the anti arrangement of the enamine are more favorable than those for the syn arrangement. In addition, for the anti channels the attack of the active methylene group of the enamine on the re face of the carbonyl group of isobutyraldehyde via **TS-ar** is 1.4 kcal/mol lower in energy than the attack on the si face via **TS-as**. This energetic result is in reasonable agreement with experiments where the (R)-cetol (**4**) is isolated in 96% of enantiomeric excess, ee, (Scheme 1). Along the syn reactive channels **TS-ss** is 0.5 kcal/mol less energetic than **TS-sr**. Therefore, if the C-C bond-formation takes place along the syn arrangement of the enamine, 5-s, the major stereoisomer will be the (S)-cetol in disagreement with experiment. Inclusion of

solvent effects, DMSO, reduces the barrier heights by the larger solvation of the TSs than the reactants but it does not modify substantially the stereoselectivity found in the gas phase.

We will try to reproduce these results concentrating our attention on two of these TSs : **TS-ar** (in the next practical tutorial this will be called **TS-1**) and **TS-ss** (in the next practical tutorial this will be called **TS-2**).

4.b Software for the exercise

In this lab, we will use **Avogadro** and **Gaussian** software packages together with some Windows programs (PuTTY, WinSCP) necessary to run calculations on Unix machines using PC as terminal:

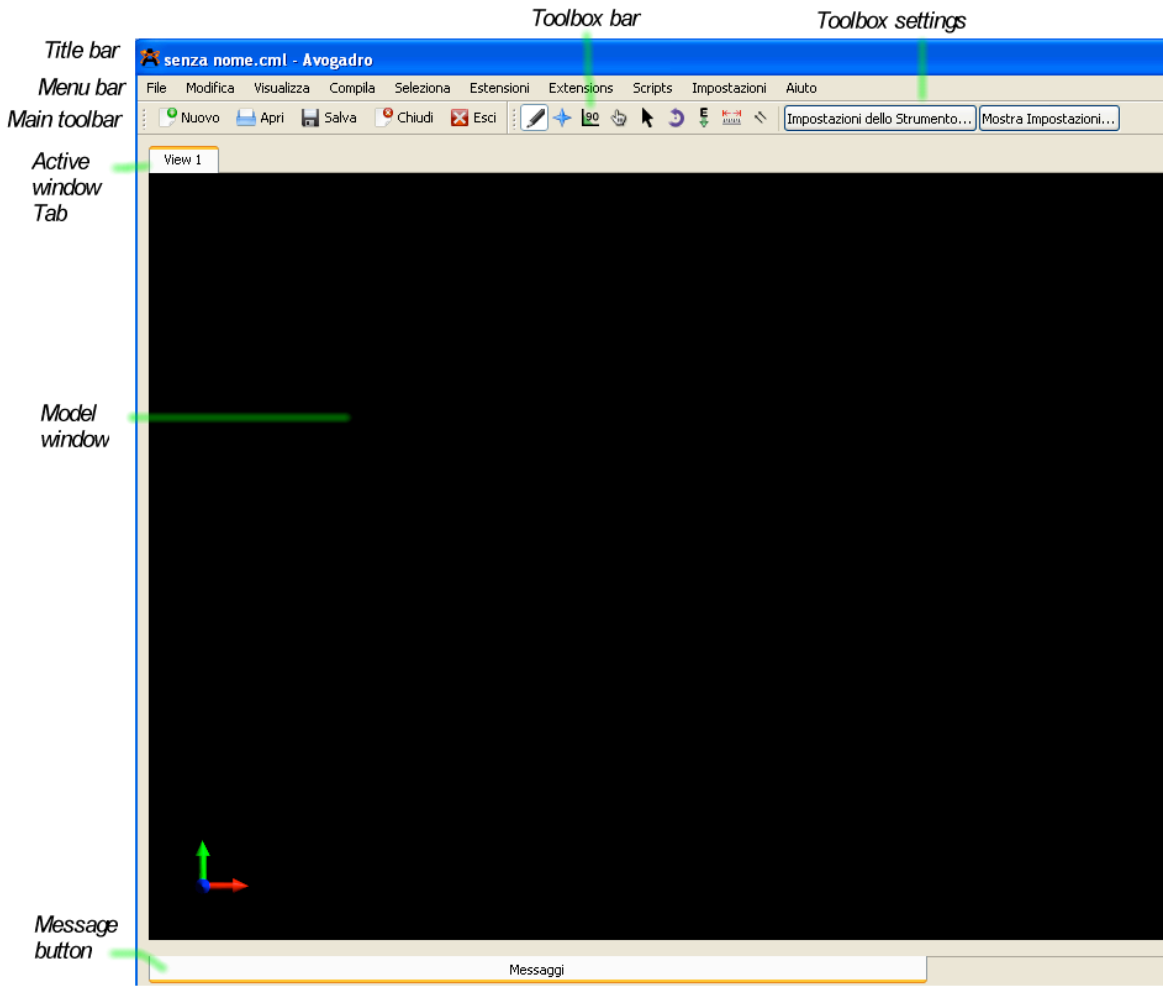
4.b.1 Avogadro

We will use Avogadro to draw our molecule, create input files for Gaussian and analyze Gaussian output

Avogadro is an advanced molecular editor designed for cross-platform use in computational chemistry, molecular modeling, bioinformatics, materials science, and related areas.

Avogadro (a quick tutorial)

The Graphical User Interface (GUI) is the part of Avogadro that you interact with to perform tasks. The GUI consists of a model window, menus, toolbars and dialog boxes. It can also include various panels to change settings. At the bottom of the GUI is a Message button, which displays information about actions performed by the software.



The **Model Window** is the work space where you do your modeling.

All Chem3D commands and functions can be accessed from the menus or toolbars. The toolbars contain icons that offer shortcuts to many commonly used functions. You can activate the toolbars you want from the Window -> Toolbars menu.

The File Menu

In addition to the usual File commands, you use the File menu to import data from different file types and to export images in various file formats.

- Import – a list of various options appear, to select the file type to import among Molecule File, Trajectory, Fetch from PDB, Fetch from chemical structure, Fetch from URL,

- Export – a list of options to export the current scene in Graphics (BMP, JPG, GIF, PNG), Vector Graphics (PDF, PostScript or SVG) or POV-Ray (high quality rendering).

The Edit Menu

In addition to the usual Edit functions (like the important Undo/Redo function) you can use the Edit menu to copy the model, to clear the model window (Clear), and to select or deselect the whole model (Select All and Select None).

The View Menu

The View menu is where you decide various settings about the visualization of the Model window.

- New View and Duplicate View - change view or to replicate the current view. This will open a new tab where you can work with the same model in a different visualization.
- Detach View – look at your model from far away
- Close View – close the current view tab
- Full Screen Mode – put the model on full screen
- Set Background Color – decide the colour of the background (default is black)
- Display Axes and Display Unit Cell Axes – if selected 3D Cartesian axes and unit cell axes (if present) are showed
- Debug Information – if selected it shows some information on the Model window
- Use Quick Render – if selected it enables Avogadro to use a faster graphical rendering, therefore improving performances and saving time while building the model
- Properties – it shows a tab with a list of properties: by clicking on each entry a window opens, showing the selected chemical properties about the model you are working with (bond length and angles, torsions, chemical formula…)

The Build Menu

The Build menu commands assist you in the building of your model by giving a series of useful options

- Change H in methyl – Replace the selected H atoms with CH_3
- Add Hydrogens – Add H atoms to selected atom
- Add Hydrogens for pH – Add or remove acidic hydrogens according to the desired pH for ionisable groups in peptides
- Remove Hydrogens – Remove selected H atoms
- Insert – allow you to insert new models from
    - Fragment: a navigation tab is opened where you can select a molecule, the folder share/avogadro/fragments includes many examples
    - SMILES: type a SMILES string and click OK to insert a molecule written in that notation

    - Peptide: it opens the peptide builder where you can build peptides from the 20 standard amino acids by clicking on the corresponding button or typing the three-letter code. You can also select the kind of structure, stereochemistry and ionisation state of the C- and N-termini
- Cartesian Editor – shows the current atom Cartesian coordinates and allows you to manually edit them
- Super Cell Builder and Unit Cell Parameters – useful to work on periodic systems

The Select Menu

The Select menu includes additional selection options to the Select All and Select None entries in the View Menu (which are repeated here)

- Invert Selection – select the elements complementary to the current selection in your model
- Select SMARTS – select atoms using the SMILES notation
- Select by Element – a periodic table is opened and by clicking any elements all atoms matching it will be selected
- Select by Residue – to be used for peptides and proteins, it allows you to select all residues with a specific residue name (e.g. VAL, TYR…)
- Select Solvent – it selects water molecules labelled with HOH
- Add Named Selection – it can store the current user-defined selection for later recall in the Project Tree (see below)

The Extensions Menu

This menu enables you to optimise the geometry of your molecule and to export an input file for some computational chemistry software like Gaussian, MOPAC and so on. The geometry optimisation use a molecular mechanics force field among 4 possible choices. You can select it using the Molecular Mechanics --> Setup Force Field. Moreover you can build surfaces (Create Surfaces) according to several kind of data (electrostatic potential, molecular orbitals and so on…), visualize vibrational spectra (Spectra) and animate vibrations if data from a frequency calculations are loaded (Vibrations).

The Settings Menu

It is very important for selecting which toolbar has to be visible from the Toolbars entry. The entry Configure Avogadro allows you to access many Avogadro settings and the Project Tree

The Toolbox

The Toolbox contains tools to draw and manipulate your model

 Draw Tool: when it is selected, you can simply left-click in Model Window to draw atoms and bonds. To shows the tool's settings click on the adjacent button Toolbox settings. A window will appear on the left side of the GUI and there you can select the atomic element, the order of bond and optionally saturate or not with hydrogens. Analogous windows are opened for the other tools. Default is carbon, single-bond, H saturation, so clicking once in the Model Window you will get CH_4.

 Navigate Tool: when it is selected, it enables you to rotate, zoom and move your model by holding the left, the central or the right mouse button respectively and dragging.

 Bond Centric Manipulation Tool: it allows you to change the view by keeping a bond plane fixed, to retrieve information about bond geometry and to change bond angles and length.

 Manipulate Tool: with this tool you can pick up atoms and move them y simple click and drag.

 Selection Tool: you can select single atoms or drag over many atoms for multiple selections. A blue glow appear around the selected elements.

 AutoRotate Tool: when selected, by click and drag you can decide an axis of rotation around which the model will rotate automatically, and the length of your line is proportional to the rotation speed you want.

 AutoOptimization Tool: allows the geometry optimisation with Molecular Mechanics to be done in a straightforward way, by selecting method, force field and option in the settings window (on the left side of the GUI).

 Measure Tool: the tool to measure distances, angles and dihedral angles by clicking on 2, 3 or 4 atoms.

 Align Tool: by clicking on 2 atoms, choosing an axis from the settings window on the left and clicking on the button Align, you can put the 1st in the origin and the 2nd aligned along the axis you chose.

Change the display type

By default Avogadro represents the models with the ball-and-stick representation, but you can choose the representation you like most by going to Settings --> Toolbars --> Display types. A tab with many display options will be opened on the left. By selecting one of them you can change how atoms and bonds are showed. Notice that representations can be overlapped unless you deselect the corresponding options.

Drawing surfaces with Avogadro

Provided that you have loaded the required data from a calculation (e.g. from Gaussian output file) you can compute the following surfaces:

- Van der Waals (no need for additional data to your model)
- Electrostatic Potential (data from calculation needed)
- Molecular Orbitals (data from calculation needed)
- Electronic Density (data from calculation needed)

You can change the colour of surfaces according to one of the preceding properties, for example you can draw a Van der Waals surface coloured by electrostatic potential.

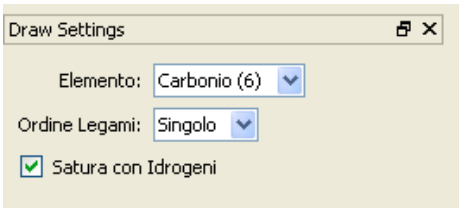
Building Molecules

Molecules can be built using the Draw Tool and its settings (left side of Avogadro GUI window). The option "Saturate with H" connects the appropriate number of hydrogen atoms to the atom you draw by clicking in the Model window. Single click put an isolated atom, eventually with its hydrogen atoms. If you want to create a bond:

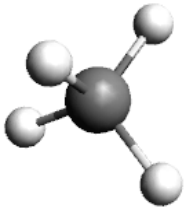
1. Select the element in the settings (optionally disable H saturation)
2. Single left-click in the model window: an atom appears
3. Eventually change the atom type and the bond order (single, double, or triple), then click and hold the mouse button on the atom you drew before
4. Drag until the bond is long enough and release the mouse button.

Now you can go on easily by adding H to saturate (Build menu) and eventually replacing them with other atoms: just select the element in the Draw Tool Settings panel and single click on the H you want to replace: you should see the change of the colour and the dimension of the sphere (given that you are working in the ball-and-stick representation).

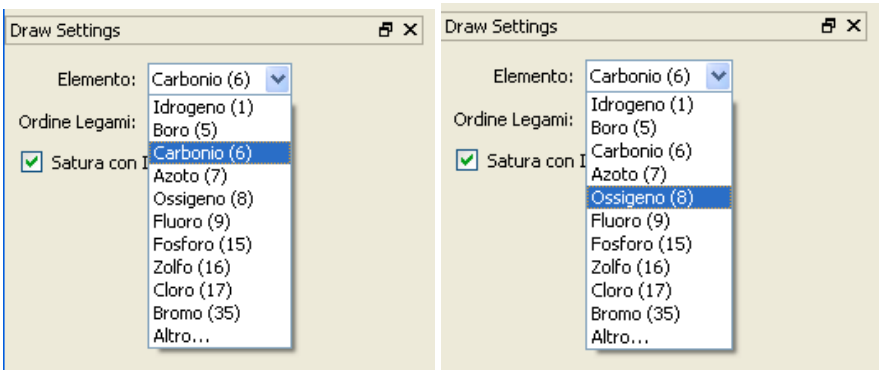
Open Avogadro (or clear any previous molecule), select the Draw Tool. In the settings panel, select carbon, with H saturation on.



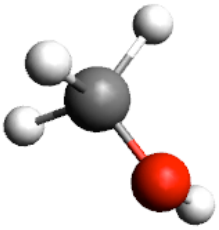
After single click in Model window you will have the methane:



Now change the atom type to Oxygen

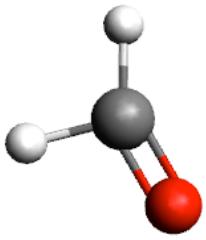


Single click on one H atom of the methane, you should see:



You have drawn the methanol! Now suppose you wanted to draw formaldehyde: in the settings panel, change bond order from single to double while keeping Oxygen as atom

type. Then single click on the C-O bond previously drawn. If everything goes well, you get formaldehyde

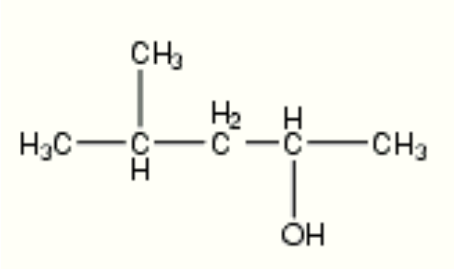


Use the same way to elongate the carbon chain and draw more complicated structures

To change one element into another, disable the H saturation feature, select the new element, then click on the atom you want to switch, and it's done.

Alternative ways to create models: SMILES

Try to build the model of 4-methyl-2-pentanol following the procedure stated above. After that you can try to draw the same structure using the SMILES notation, a standard label way to write molecules.



1. Open a new empty file (click on the New button).
2. From the **Build menu**, choose **Insert --> SMILES**.A textbox appears.
3. In the textbox type *CC(C)CC(O)C* and click *OK*
4. The required structure should appear in the Model window. If you don't see anything there are two reasons:
    a. The view is out of center, go to **View Menu** --> **Center** to put your molecule at the center of the Model window
    b. You typed a wrong string.

The SMILES notation uses capital letters for atom symbols and hydrogen atoms can be implicit. The above string mirrors the common naming of organic molecules: the longest 5-carbon chain

with substituents specified in parentheses next to the atom they attach to. Single bonds are implicit, while double bonds are specified with = and triple with # symbols.

What would be the SMILES string for acetone?

Avogadro can also read output file from Gaussian and you are going to use it to analyze the output from the practical tutorial.

General Avogadro website http://avogadro.openmolecules.net/wiki/Main_Page

4.b.2 Gaussian03

The Gaussian program will be used to perform the TS search calculations

Gaussian is a computational chemistry software capable of predicting many properties of molecules and reactions, including the following:

- Molecular energies and structures
- Energies and structures of transition states
- Bond and reaction energies
- Molecular orbitals
- Multipole moments
- Atomic charges and electrostatic potential
- Vibrational frequencies

• NMR and EPR properties

• Reaction pathways

Computation can be carried out on systems in the gas phase or in solutions, and in their ground state or in an excited state.

More information about Gaussian can be found at *www.gaussian.com* website and in Foresman. J. B., and Frisch. Æ. (1993). Exploring Chemistry with Electronic Structure Methods, Second Edition. U. S. A.: Gaussian, Inc.

General Gaussian site: http://www.gaussian.com/

Gaussian takes a text file with a .com (or .gjc in windows environment) extension as an input. In this input file, the molecular configuration of the molecule is described as well as the specific calculations that you want to perform (geometrical optimization, frequency determination, single point energy, etc) and which methods you want to use (HF, DFT, etc.). After running Gaussian, a text output file is generated with the extension .log (or .out in windows environment) and the same name as the input file. The output file can be browsed to find selected data.

Rather than using text files as output, life is made easy by a graphical interface, such as Avogadro (see section 4.b.1).

Gaussian Input

Gaussian input is always in the form of a unformatted text or ASCII file, usually called 'jobname'.com. The input file includes several different sections:

* **'Link 0' commands**. This section starts with the % sign and is the way Gaussian designates general commands concerning the way in which the program should run, for example, you can specify the amount of memory to be used (%mem), the directory where to write scratch files used by the program (%chk, %rwf, %d2e, %int). For most jobs, you will not need *any* Link 0 command but for the purpose of our lab. we will need to specify to Gaussian to save to your current directory the temporary 'Checkfile', and to call it filename.chk. If you do not do this, the checkfile is kept in a temporary, 'work', directory, and deleted when the job ends.

* The **Route Section**. The first line of this Section needs to start with the pound sign (#). Here you specify which is the type of calculation you want to do, the method and basis set and other options (see the following "Gaussian Commands for the Route Sections"). This is the most important part of the input and needs to be followed by one blank line.

Alternative forms:

#N, Normal print level; this is the default.

#P, Additional output is generated. This includes messages at the beginning and end of each link giving assorted machine-dependent information (including execution timing data), as well as convergence information in the SCF.

#T, Terse output: output is reduced to essential information and results.

* The **Title Section**. This is free for you to choose. You can add a brief description of the calculation.

* The **Molecular Charge and Spin Section**. Here begins the description of the molecule. You need to specify, as integers, the *overall molecular charge* and the *spin multiplicity* (for example, if is a neutral species (charge=0), and it is a spin-Singlet, with S=0 (multiplicity=2*S + 1= 1).

* The **Geometry Specification Section**. Here you specify the molecular system to be studied in Cartesian Coordinates or in Z-matrix format. This section needs to be followed by a blank line.

In general, Gaussian input is subject to the following syntax rules:

• Input is free-format and case-insensitive.

• Spaces, tabs, commas, or forward slashes can be used in any combination to separate items within a line. Multiple spaces are treated as a single delimiter.

• Options to keywords may be specified in any of the following forms:

keyword = option

keyword(option)

keyword=(option1, option2, ...)

keyword(option1, option2, ...)

• Multiple options are enclosed in parentheses and separated by any valid delimiter (commas are conventional and are shown above). The equals sign before the opening parenthesis may be omitted, or spaces may optionally be included before and/or after it.

• Note that some options also take values; in this case, the option name is followed by an equals sign: for example, Opt(MaxCycle=99).

• All keywords and options may be shortened to their shortest unique abbreviation within the entire Gaussian system. Thus, the Conventional option to the SCF keyword may be abbreviated to Conven, but not to Conv (due to the presence of the Convergence option). This holds true whether or not both Conventional and Convergence happen to be valid options for any given keyword.

• Comments begin with an exclamation point (!), which may appear anywhere on a line. Separate comment lines may appear anywhere within the input file.

Gaussian Commands for the Route Sections

It will generally need to specify the following things:

* The theoretical *method*

* The *basis set*

* The *type of job*

Theoretical Method and Basis Set

The methods that you can select are several. Some example include Hartree-Fock (written more simply as "HF"), MP2, the density functional theory approach (for example, "B3LYP"), or the semi-empirical "PM3" method. You will always need to specify a basis set, *except for semiempirical computations*, where no basis is needed. Some example of basis ser are STO-3G, 3-21G, 6-31G, 6-31+G, 6-31+G**, aug-cc-pVDZ.

Job Type and Options

If you do not specify anything else than the method and basis set, Gaussian will perform the default job-type, a *single-point energy computation* at the geometry specified. This can also be done by adding the keyword "SP". Here is a list of most used jobtype **keywords** you will need, and some indications on options you will need:

* **SP**: Single-Point Job

* **Opt:** Geometry Optimisation

* **Freq**: Frequency Calculation

* **IRC**: Reaction path following

Here are listed some keywords that you can use to predict molecular properties

• **Pop**: Atomic charges, Dipole and Multipole moments, Molecular orbitals

• **Cubegen, Prop**: Electrostatic potential

• **Pop=Chelp, ChelpG or MK**: Electrostatic-potential derived charges

• **Prop**: Hyperfine coupling constants (anisotropic)

• **NMR and Freq=(VibRot, Anharmonic)**: Hyperfine spectra tensors (incl. g tensors)

• **Freq**: IR and Raman spectra, Thermochemical analysis

• **NMR**: NMR shielding and chemical shift

• **CIS, Zindo, TD**: UV/Visible spectra:

******PLEASE, REFER TO THE MANUAL (WWW.GAUSSIAN.COM) FOR DETAILS AND MORE KEYWORDS******

Geometry Specification

There are two main ways of specifying the molecular geometry: Cartesian coordinates and Z-matrix coordinates.

Cartesian Geometry Specification

Geometries in Cartesian coordinates are entered simply as follows:

Atom label1 x1 y1 z1

Atom label2 x2 y2 z2

Atom label3 x3 y3 z3

For example, here is an example for H_2O:

O 0.000000 0.113568 0.000000

H 0.753016 -0.454271 0.000000

H -0.753016 -0.454272 0.000000

Note that the unit of distance is Angstrom.

Z-Matrix Geometry Specification

A Z-matrix describes geometry in terms of distances, angles and dihedral angles. For example, here is a valid Z-matrix for H_2O:

O

H 1 0.96

H 1 0.96 2 105.0

or

O

H 1 R2

H 1 R3 2 A3

R2 0.96
R3 0.96

A3 105.0

Note that the decimal point is included in the angle specification (which is in degrees). This is **very important**. Without the decimal point, Gaussian interprets that number as an integer, where it is expecting a real number.

4.b.4 PuTTY, WinSCP

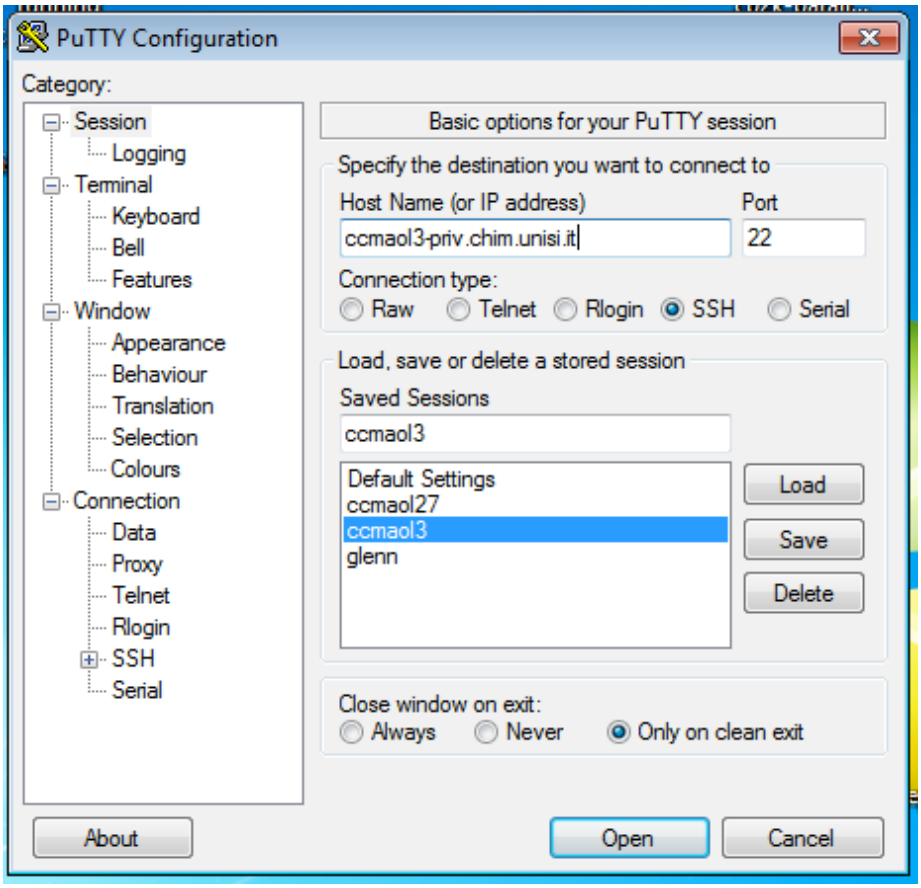
Most computational chemistry software has been written to run under **Unix**. You will use Gaussian on a Unix-based environment here.
The brief tutorial below will teach you about logging on, running programs, managing files and editing files under Unix.

Part of the work will you will be doing will be done on a Unix computer called "**ccmaol3-priv**" which belongs to the group of Computational Chemistry and Photochemistry of Siena.

Loggin on Unix

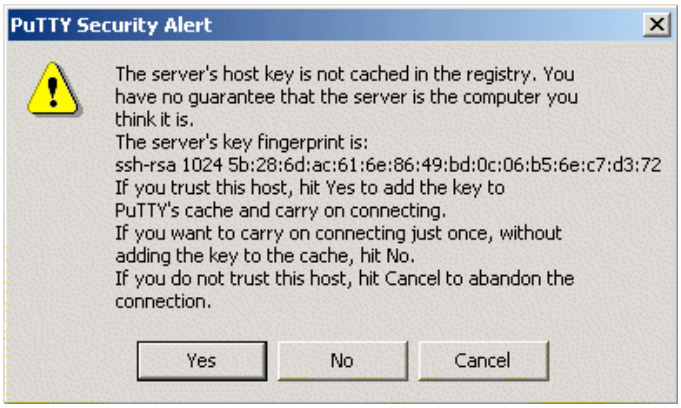
To log on to ccmaol3-priv, you need to follow this procedure:

Start the Windows program called **"PuTTy"** that will allow us to login to a remote session on the Unix systems. This will produce a session configuration window like the one shown below.



As "Host Name" type "**ccmaol3-priv.chim.unisi.it**" and select the SSH "Protocol". Note that you can save your settings to make it easier to connect to machines in the future by typing a session name (for example **studenti@ccmaol3-priv**) in the "Saved Sessions" text field and selecting Save.

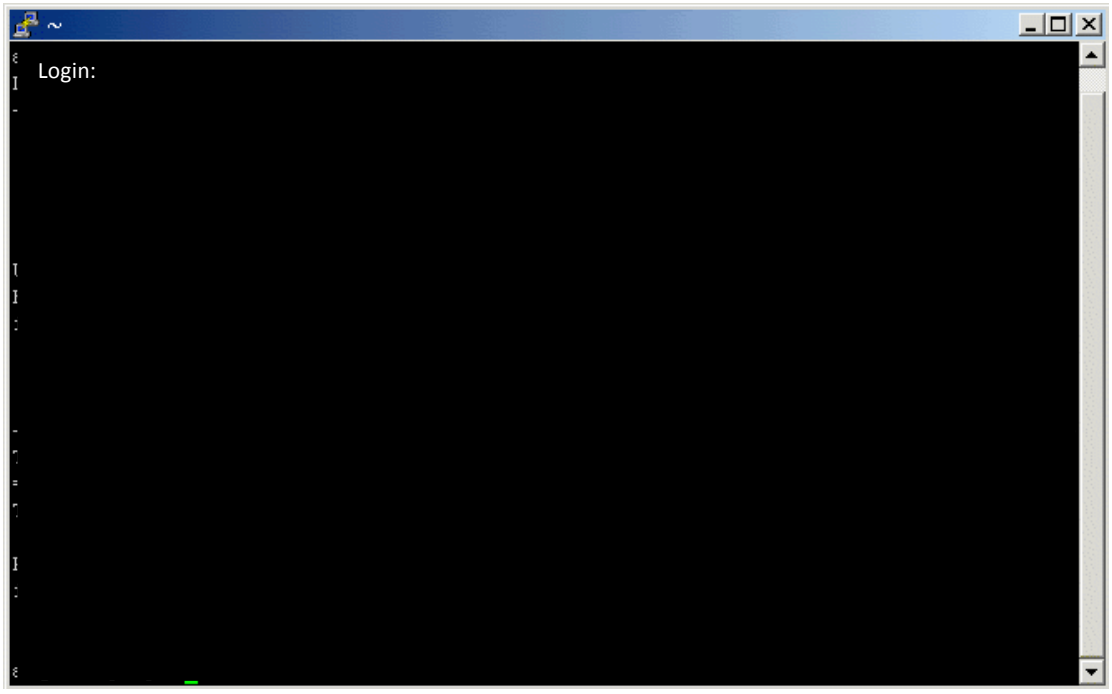
After selecting your host, protocol, click on **Open**. This will connect you to the server and ask if you want to accept the RSA key of the server.



You will want to select **yes**. After doing this, whenever you connect to the same server with PuTTY, it will use that key to identify the server and alert you if it is not the same.

After accepting the RSA key from the server, PuTTY will ask you to input your username. Type "studenti" and press the "enter" key. Then you will be asked for a password - this will be given to you by the demonstrator.

Type that in and press the "enter" key.



If you have done all this correctly, you are now connected to ccmaol3-priv, and have an active Unix window. In this window, you can carry out any text-based commands: list files, start calculations, edit files with vi (see below for all of these).

Running Programs in Unix

Unix programs are usually started by issuing a command at the prompt. Some programs automatically open a new window, whereas some give output to the same window. In either case, you will often need to add an "&" character after the command to make the program run "in the background". This means that while the program runs, you are still able to continue issuing other commands. For example, to run Gaussian 03 on the input file "filename.com" you should type:

* "g03 filename.com &"

To run the "vi" text editor program on the text file filename.com:

* "vi filename.com"

Managing Files in Unix

You will need some basic knowledge to create and manipulate files in Unix. Listed here are some simple commands you may need.

* "ls" gives a list of the files in the current directory.

* "cp filename.com filename_new.com" copies the existing file "filename.com" to "filename_new.com".

* "mv filename.com filename_old.com" moves " filename.com" to " filename_old.com".

* "mkdir newdirectory ". Makes a new directory "newdirectory"

* "cd newdirectory" changes directory. "newdirectory" is the directory you want to change to. Here are some examples of the use of "cd":

* "cd". Changes to your home directory.

* "cd sub_dir". Changes to the sub-directory "sub_dir" which belongs to your current directory when you issued the command.

* "cd ..". Changes to the parent directory of the current directory.

* "rm filename.chk" removes (deletes) the file " filename.chk". **Once this is done, the file is *gone* and cannot be recovered - be careful!!**

* Command history. Within Unix, you can usually retrieve a recent command by pressing the "Up" arrow key at the prompt until you reach that command. The command can also be edited before re-entering it.

Editing Files in Unix

To create the input for the Gaussian jobs you are going to run, you will need to *edit* text files. You will also need to do this in order to read some output files, etc. There are lots of programs you can use to edit files in Unix: for example, Vi, Pico, nedit, Emacs.

In this lab, you may try to use "Vi" that is a screen oriented text editor. This means that it takes up almost the entire screen, displaying part of the file on each screen line, except for the last line of the screen. The last line of the screen is used for you to give commands to vi, and for vi to give information to you. The other fact that you need to understand is that vi is a "modeful" editor, i.e. you are either entering text or you are executing commands, and you have to be in the right mode to do one or the other. You will be in command mode when you first start editing a file. There are commands that switch you into input mode. There is only one key that takes you out of input mode, and that is the <escape> key.

This section will tell you the minimum amount that you need to do simple editing tasks using vi:

To start editing a file, enter the command "vi file_name" and return

The cursor arrow keys should work to move around the file.

The commands to enter new text are:

a Append new text, after the cursor.

i Insert new text, before the cursor.

o Open a new line below the line the cursor is on,

and start entering text.

O Open a new line above the line the cursor is on,

and start entering text.

<escape>

Once you've entered input mode using the one of the

a, i, O or o commands, use <escape> to quit entering

text and return to command mode.

The commands to copy text are:

yy Copy the line the cursor is on.

p Append the copied line after the line the cursor is

on.

The commands to delete text are:

dd Delete the line the cursor is on.

x Delete the character the cursor is on.

The commands to write the file are:

:w

Write the file back to the file with the name that you originally used as an argument on the vi command line.

:w file_name

Write the file back to the file with the name ``file_name''.

The commands to quit editing and exit the editor are:

:q

Quit editing and leave vi (if you have modified the file, but not saved your changes, vi will refuse to quit).

:q!

Quit, discarding any modifications that you may have made.

If you need or want more information type "man vi".

Cutting and Pasting to the Unix Window

* To paste text from the Unix Window to another place in the same window, simply select it by left-clicking and dragging. There is no need to type "Ctrl-C" or anything similar - the text is already in the Unix clipboard. Simply go to the place where you wish to insert the copied text, and click once with the *central* mouse button. If your mouse only has two buttons, click on both buttons *at the same time*.

* To copy text from the Unix Window to a Windows application, first select the text with your mouse. Then click on the icon at the top left of the Unix window, and choose "Edit" then "Copy X Selection" then "To Clipboard" from the menu that appears. You can now paste in the usual way into your Windows application, using "Ctrl-V".

* To copy text from a Windows application into a Unix window, first copy (or cut) the text into the clipboard within your Windows application. Then go to the Unix window, and click on the top-left of the window, and choose "Edit" then "Paste to X Selection" then "From Clipboard" in the menu. Then go to the place where you wish to insert the copied text, and click once with the *central* mouse button, as above.

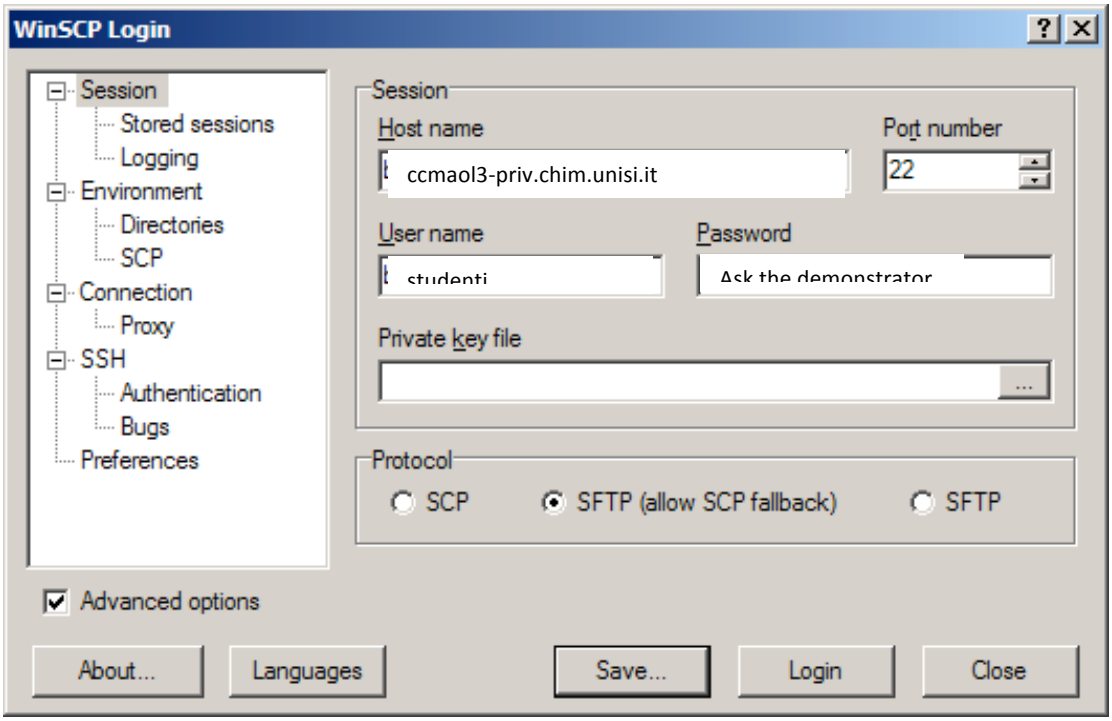
Transfering Files from Windows to Unix

Because editing files in Unix using the "vi" text editor is not simple, we shall use any text editor under the windows environment and transfer our files to the Unix machine using a program colled WinSCP. We can also transfer back the output of our calculations to analyze it under windows using Avogadro.

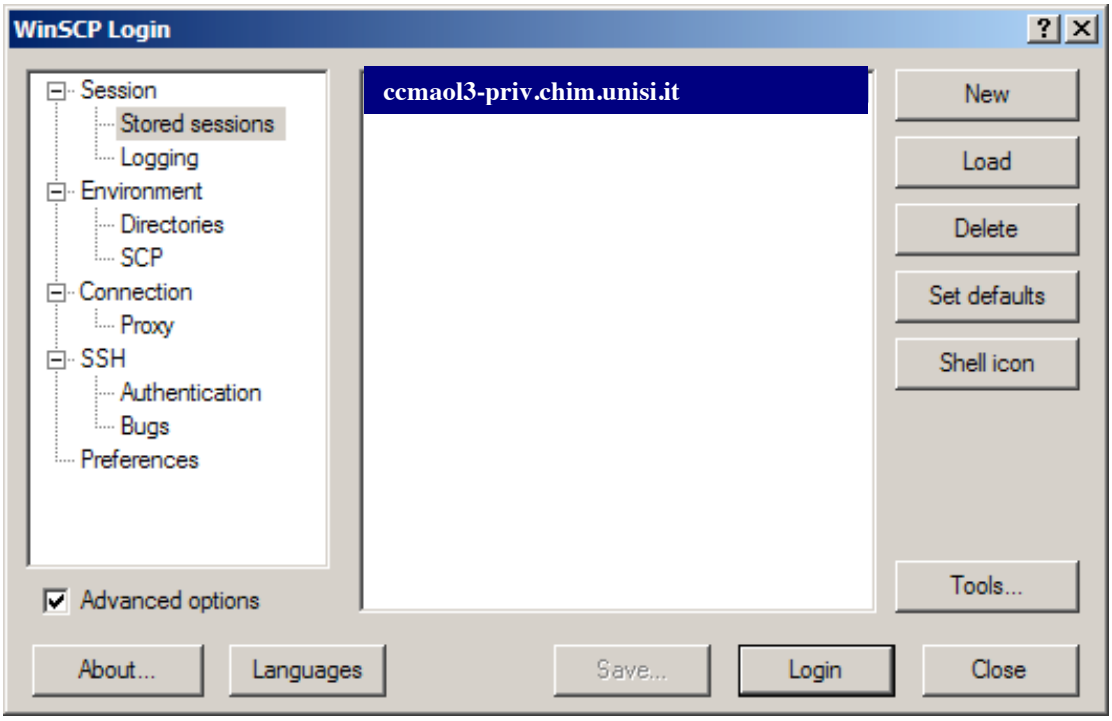
WinSCP is a secure file copying program for Windows

You should see the screen below when WinSCP first starts up. Click Session in the left panel of the screen to configure WinSCP to connect to another system, Fill in the text fields on the

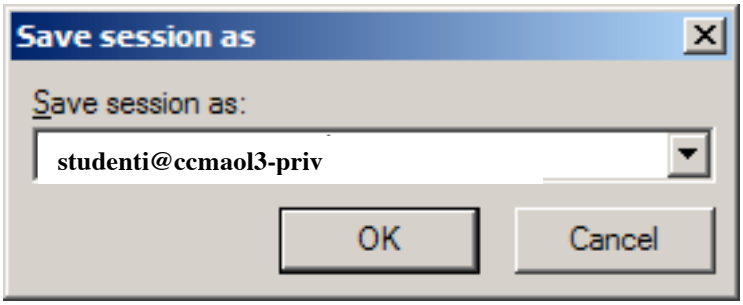
Session screen as shown. You do not need to change the Port number, and you do not need a Private key file.



You can save settings if you won't have to do this every time. So, uncheck the Advanced options checkbox in the bottom left corner of the window, and then select Stored sessions in the left pane, as shown below.



Click Save on the right side of this window. In the "Save session as", enter studenti@ccmaol3-priv, as shown below and click OK.



To connect to ccmaol3-priv, simply select Stored sessions in the left pane of the WinSCP window, then select your connection to in the right side and, finally, hit the Login button at the bottom. You might be prompted for your password again; if so, enter it.

Once connected to ccmaol3-priv, WinSCP will present you a window where on the right side you can find the files present in your home directory on the Unix machine (this is often referred to as the remote file listing). On the left side you have your windows folders. Select the folder where you want to transfer (to or from) files. Transferring files with WinSCP is very simple. You

only need to drag the file from your computer (left side) into the Unix machine (right side), or viceversa. When you're finished just close WinSCP.

4.c Practical Tutorial

This practical tutorial consists of five main steps. The objective of this lab is to understand chemical reactions and transition state theory giving you simple, illustrative calculations. The first four steps are necessary to compute the two TSs structures (**TS-1** (**TS-ar**) and **TS-2** (**TS-ss**)) while the last step contains the comparison of the results obtained for the two cases in order to understand the origin of the stereoselectivity and reproduce the observed product ratio.

Step 1: Create an initial geometry for transition states

Locating a transition state is not a trivial task and is more demanding than performing a geometry optimization, although both procedures involve the finding of a stationary point (zero gradient) on the potential energy surface. What makes TS searches more problematic is the requirement that the desired stationary point have a single "downhill" direction, meaning that motion along one, and only one, direction in 3N-6 dimensional space will lower the total energy. There are several approaches to locate a TS, but usually the quickest way is to have a good guess of the TS structure using chemical intuition and then start a TS search from that point. For example, if we would have to locate the TS for the breaking of a particular C-C bond, an initial geometry with a stretched C-C bond is a reasonably good starting point for a TS search.

In the present case, we will draw a structure resembling a Zimmerman-Traxler type transition state and showing, for **TS-1**, the anti arrangement of the enamine and the attack of the active methylene group of the enamine on the re face of the carbonyl group of isobutyraldehyde, while for **TS-2**, the syn arrangement of the enamine and the attack of the active methylene group of the enamine on the si face of the carbonyl group of isobutyraldehyde. The two models will be characterized by: 1) an almost formed C-C bond between the carbonyl carbon of the isobutyraldehyde and the terminal carbon of the alkene segment of the enamine and 2) a stretched O-H bond of the carboxylic acid of the proline. In fact, the proton is being transferred to the carbonyl oxygen of isobutyraldehyde.

A good starting structure for **TS-1** and **TS-2** are the following:

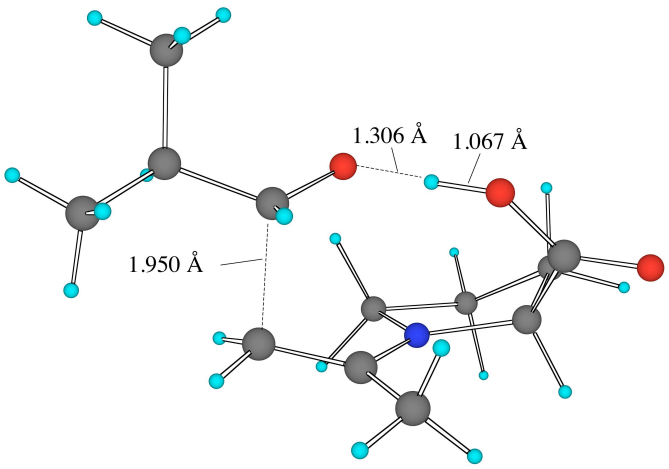


Figure 2: initial guess structure for TS-1

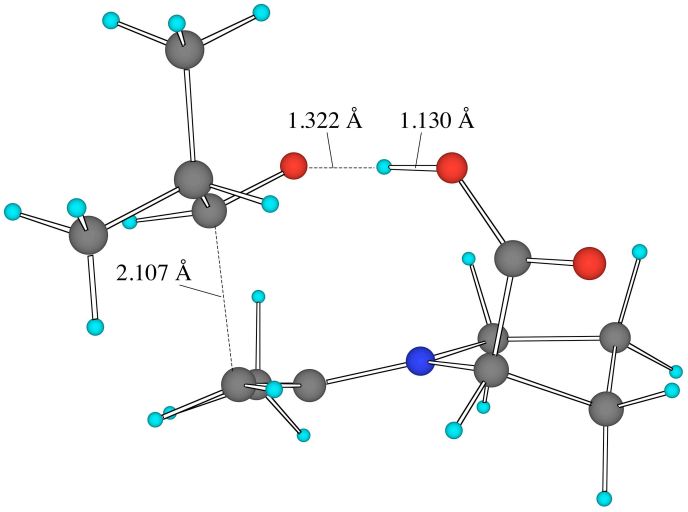


Figure 3: initial guess structure for TS-2

Step 1.a Draw Molecules

To draw the two molecules of interest and generate the text input file for Gaussian we will use Avogadro (refer to the quick tutorial in Section 4.b.1). Once you have created your molecule you can directly save a Gaussian input as described below:

Step 1.b Creating a Gaussian Input File

A Gaussian Input file contains the coordinates and geometry of the model and the Gaussian keywords taken from the settings of the dialog box.

To create a Gaussian Input file:

1.b.1 From the Extensions menu, click on Gaussian…

The Input Gaussian dialog box appears.

Important: we will not choose any properties or methods here! We will edit our input file separately using any text editor (such as word, etc…).

1.b.2. Click Generate.

An input file saves in Gaussian's native .COM format.

Step 2: Prepare the approximate Hessian

Before starting a TS search is necessary to compute the initial Hessian matrix. The initial Hessian must have one negative eigenvalue (i.e., one imaginary frequency) and a suitable eigenvector associated with this eigenvalue. the one with a corresponding mode closely corresponding to the desired motion of the chemical reaction (C-C bond formation and proton transfer, in the present example). We are going to create such a Hessian by doing quick frequencies calculation. Hessian calculations can be quite expensive and because is not required to have an exact Hessian to start a TS search, we can compute the initial Hessian at a lower

level of theory. For example, if one is trying to locate a TS at the RHF/6-31G* level, an initial Hessian computed with a smaller basis set, say RHF/STO-3G, will often work fine.

Step 2.a Editing the .com Gaussian Input File

Open the text file previously generated with Avogadro using a text editor (such as word, for example) and modify the settings such as method, basis set, calculation type etc. as follows:

```
**********************************************************************

%chk=filename.chk
#p rhf sto-3g gfinput pop=full
   freq

Estimating the Hessian for TS-1

0 1
6        2.684531   -0.638608   -2.084797
6       -1.098009   -1.385873    0.589539
7        0.967215   -1.394902   -0.608859
6        1.024100   -2.166397    1.748089
6        0.264526   -1.719231    0.499328
6        2.444422   -1.201880   -0.659824
6        0.326211   -1.169266   -1.937221
6        3.015649   -0.233590    0.460361
8        4.190426   -0.372333    0.776134
8        2.198787    0.674563    1.020861
6        1.512852   -1.185796   -2.921928
1        2.649077    0.448302   -2.063613
1        3.653946   -0.943729   -2.468097
1       -1.672909   -1.392928   -0.322227
1       -1.620807   -1.802781    1.435684
1        1.832995   -2.841200    1.485497
1        0.343005   -2.670240    2.424825
1        1.445498   -1.299045    2.253442
1        2.945196   -2.168604   -0.551768
1       -0.396856   -1.961502   -2.137845
1       -0.186159   -0.205896   -1.944244
1        1.250399    0.784091    0.544828
1        1.717679   -2.204276   -3.243055
1        1.310891   -0.579915   -3.800595
6       -0.876623    0.532041    0.865933
6       -2.330248    1.063715    0.615970
8        0.045768    0.963687    0.072020
```

N.B. Here you will have the cartesian coordinates generated by Avogadro for your geometry

```
6       -2.280522    2.607176    0.732498
6       -3.379677    0.486120    1.590186
1       -0.658780    0.468028    1.948405
1       -2.615245    0.814609   -0.405636
1       -3.249274    3.035294    0.492734
1       -1.541017    3.010034    0.047715
1       -2.014126    2.908402    1.741438
1       -4.326143    1.002687    1.462846
1       -3.059898    0.617590    2.619709
1       -3.550267   -0.570119    1.415772
```

Explanation:

%chk=filename.chk sets the name of the checkpoint file. ***WARNING: do not forget to add this line. We need to store the checkpoint file for the next calculation***

rhf requests a Hartree-Fock calculation.

sto-3g basis set is a minimal basis set. It uses three gaussian primitives per basis function ("3G"). "STO" stands for "Slater type orbitals", and the STO-3G basis set approximates Slater orbitals with gaussian functions.

gfinput is needed to read correctly the output with Molden (regarding the gaussian printing of basis set information the keyword gfinput has to be turned on).

pop keyword controls printing of molecular orbitals and several types of population analysis and atomic charge assignments. The default is to print just the total atomic charges and orbital energies. Options: **pop=None**, no orbitals are printed, and no population analysis is done; **pop=Minimal**, total atomic charges and orbital energies are printed; **pop=Regular**, the five highest occupied and five lowest virtual orbitals are printed, along with the density matrices and a full (orbital by orbital and atom by atom) Mulliken population analysis. Since the size of the output depends on the square of the size of the molecule, it can become quite substantial for larger molecules; **pop=full**, same as the Regular population analysis, except that all orbitals are printed.

freq this keyword asks for frequency analysis. This will calculate the matrix of second derivatives of the energy with respect to the position of the nuclei, and diagonalise it to generate vibrational frequencies.

Remember that there should be **three** blank lines, the before and after the title section and one after the geometry specification

Step 2.b Saving, Transferring on Unix and Running the calculation

On the Unix machine you will be using for the calculation the input files. They need to have the extension ".com". Thus the first step after the editing of your input is to save it as "filename.com".

*** WARNING: Be sure that your file is a text file. This warning is particularly important if you have used Word as text editor because it save files as word document by default. We need to save it as "text only" to be readable in the Unix environment***

Next step will be the transfer of your input file in your home directory on ccmaol3-priv. Before this you will create a special directory with your surname inside the student account where you will loggin on. To do this:

1) Open PuTTy and follow instructions given in section 4.b.4.
2) Loggin on ccmaol3-priv as studenti (ask for the password to the demonstrator)
3) Type mkdir "your surname"
4) Do not close Putty, you will use it again for this step!

Now transfer the input file to ccmaol3-priv. To do this:

1) Open WinSCP and follow instructions given in section 4.b.4
2) Loggin on ccmaol3-priv as studenti (ask for the password to the demonstrator)
3) Once connected to ccmaol3-priv, the WinSCP window has on the right side your home directory (/home/studenti) on the Unix machine. Please select the sub-directory with your surname. Find your input file inside folders in your PC on the left side and drag it in the sub-sirectory on the left. When you have finished just close WinSCP.

Once the input file is on the Unix machine, it needs a transformation to be Unix-compatible, since it was created under Windows. It is enough to type

dos2unix filename.com

and it will be converted. If you don't do it, your Gaussian job will crash.

IMPORTANT: Our ccmaol3-priv workstation is actually a cluster of several machines and is composed of a master (ccmaol3-priv itself) and pure compute nodes. To make our calculations running we will access directly the node and run the calculations there. To do this:

Type in the putty window "ssh nodename". For the nodename ask the demonstrator.

You have to go inside the folder you created before by typing

cd "your surname"

Now you are ready to **run** your calculation typing at the prompt:

"g03 filename.com &"

The calculation will run for about 5 minutes.

Step 2.c Interpreting the output

Gaussian writes its output to a file called the *logfile*, which has the same name as the input file, but the '.log' extension. It writes to this file progressively as the job advances. You can look at the .log file typing "vi filename.log. Use the "CTRL+F" and "CTRL+B" command to go forward and backword in your file.

The output is quite verbose, and the present discussion will focus on the most important part of the information you can find in the logfile.

The first lines of the output are taken up by system details and copyright information, then you will see the following 'header' containing information about the program and its version:

```
 ****************************************
  Gaussian 03:  IA64L-G03RevC.02 12-Jun-2004
                6-Sep-2006
 ****************************************
```

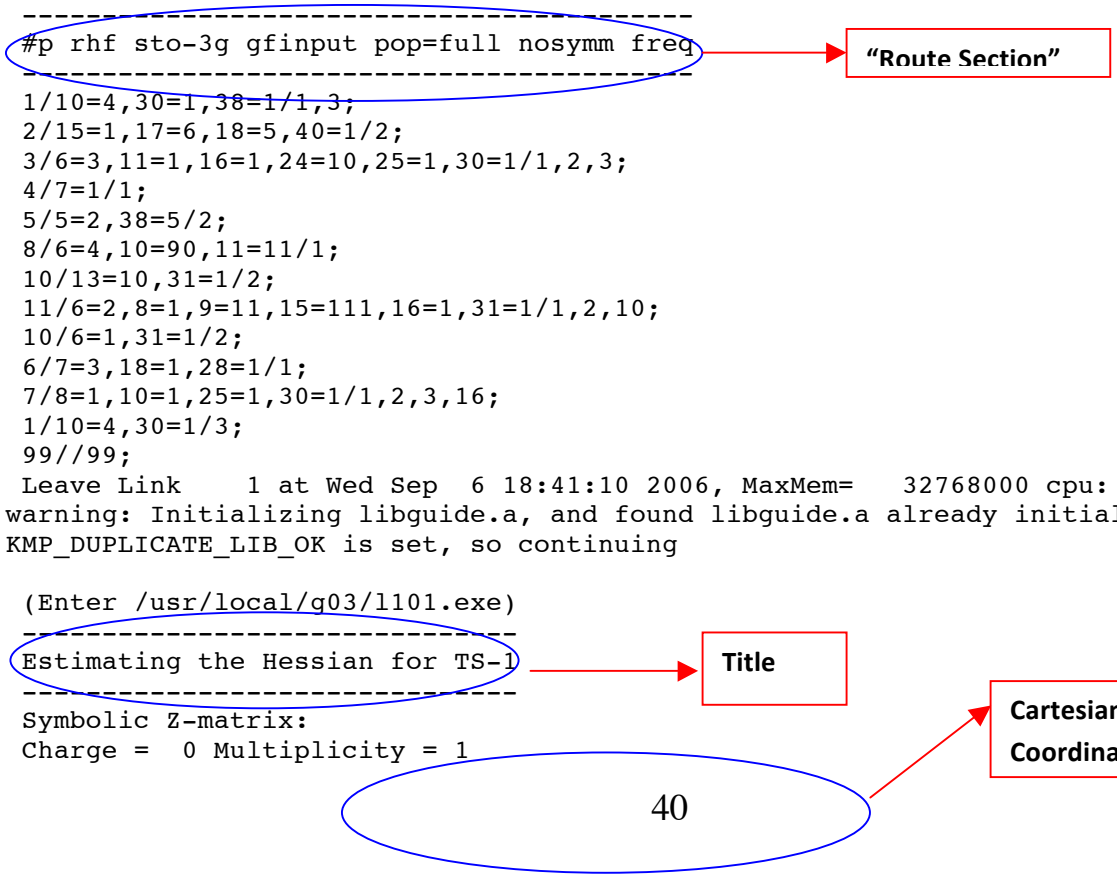
Shortly after that, your Command Section and Job Title will be echoed.

Then, you will see lots of details about your geometry. In the following block, for example, you will see the Cartesian Coordinate representation of your molecule.

```
 ----------------------------------------------------------------------
```

```
 %chk=TS-1-hess
  %mem=250Mb
  ----------------------------------------
  #p rhf sto-3g gfinput pop=full nosymm freq
  ----------------------------------------
  1/10=4,30=1,38=1/1,3;
  2/15=1,17=6,18=5,40=1/2;
  3/6=3,11=1,16=1,24=10,25=1,30=1/1,2,3;
  4/7=1/1;
  5/5=2,38=5/2;
  8/6=4,10=90,11=11/1;
  10/13=10,31=1/2;
  11/6=2,8=1,9=11,15=111,16=1,31=1/1,2,10;
  10/6=1,31=1/2;
  6/7=3,18=1,28=1/1;
  7/8=1,10=1,25=1,30=1/1,2,3,16;
  1/10=4,30=1/3;
  99//99;
  Leave Link    1 at Wed Sep  6 18:41:10 2006, MaxMem=   32768000 cpu:       0.2
warning: Initializing libguide.a, and found libguide.a already initialized, but
KMP_DUPLICATE_LIB_OK is set, so continuing

  (Enter /usr/local/g03/l101.exe)
  ------------------------------
  Estimating the Hessian for TS-1
  ------------------------------
  Symbolic Z-matrix:
  Charge =  0 Multiplicity = 1
```



```
 6                     2.68453  -0.63861  -2.0848
 6                    -1.09801  -1.38587   0.58954
 7                     0.96722  -1.3949   -0.60886
 6                     1.0241   -2.1664    1.74809
 6                     0.26453  -1.71923   0.49933
 6                     2.44442  -1.20188  -0.65982
 6                     0.32621  -1.16927  -1.93722
 6                     3.01565  -0.23359   0.46036
 8                     4.19043  -0.37233   0.77613
 8                     2.19879   0.67456   1.02086
 6                     1.51285  -1.1858   -2.92193
 1                     2.64908   0.4483   -2.06361
 1                     3.65395  -0.94373  -2.4681
 1                    -1.67291  -1.39293  -0.32223
 1                    -1.62081  -1.80278   1.43568
 1                     1.833    -2.8412    1.4855
 1                     0.34301  -2.67024   2.42483
 1                     1.4455   -1.29905   2.25344
 1                     2.9452   -2.1686   -0.55177
 1                    -0.39686  -1.9615   -2.13784
 1                    -0.18616  -0.2059   -1.94424
 1                     1.2504    0.78409   0.54483
 1                     1.71768  -2.20428  -3.24306
 1                     1.31089  -0.57992  -3.8006
 6                    -0.87662   0.53204   0.86593
 6                    -2.33025   1.06371   0.61597
 8                     0.04577   0.96369   0.07202
 6                    -2.28052   2.60718   0.7325
 6                    -3.37968   0.48612   1.59019
 1                    -0.65878   0.46803   1.94841
 1                    -2.61525   0.81461  -0.40564
 1                    -3.24927   3.03529   0.49273
 1                    -1.54102   3.01003   0.04772
 1                    -2.01413   2.9084    1.74144
 1                    -4.32614   1.00269   1.46285
 1                    -3.0599    0.61759   2.61971
 1                    -3.55027  -0.57012   1.41577
****************************************************************************
```

Then the calculation will start.:

```
Standard basis: STO-3G (5D, 7F)
 AO basis set in the form of general basis input:
  1 0
 S   3 1.00       0.000000000000
      0.7161683735D+02  0.1543289673D+00
      0.1304509632D+02  0.5353281423D+00
      0.3530512160D+01  0.4446345422D+00
 SP   3 1.00       0.000000000000
      0.2941249355D+01 -0.9996722919D-01  0.1559162750D+00
      0.6834830964D+00  0.3995128261D+00  0.6076837186D+00
      0.2222899159D+00  0.7001154689D+00  0.3919573931D+00
 ****

...

SCF Done:  E(RHF) =  -736.288020069     A.U. after   16 cycles
```

Basis set specification for each atom

```
             Convg  =    0.6531D-08             -V/T =  2.0085
             S**2   =    0.0000
 KE= 7.300821293264D+02 PE=-4.206736100217D+03 EE= 1.499555568751D+03
 Leave Link  502 at Wed Sep  6 18:41:19 2006, MaxMem=   32768000 cpu:       5.1
warning: Initializing libguide.a, and found libguide.a already initialized, but
KMP_DUPLICATE_LIB_OK is set, so continuing
 ********************************************************************

           Population analysis using the SCF density.

 ********************************************************************
Molecular Orbital Coefficients
                     1        2        3        4        5
                     O        O        O        O        O
EIGENVALUES --   -20.18564 -20.16774 -20.14409 -15.42362 -11.15896
   1 1   C  1S          -0.00001  -0.00000  -0.00000  -0.00000  -0.00002
   2        2S           0.00003  -0.00002  -0.00001   0.00027  -0.00006


...

(Enter /usr/local/g03/l701.exe)
 Compute integral second derivatives.
 ... and contract with generalized density number  0.
 Leave Link  701 at Wed Sep  6 18:46:15 2006, MaxMem=   32768000 cpu:       0.5
warning: Initializing libguide.a, and found libguide.a already initialized, but
KMP_DUPLICATE_LIB_OK is set, so continuing

 (Enter /usr/local/g03/l702.exe)
 L702 exits ... SP integral derivatives will be done elsewhere.
 Leave Link  702 at Wed Sep  6 18:46:15 2006, MaxMem=   32768000 cpu:       0.0
warning: Initializing libguide.a, and found libguide.a already initialized, but
KMP_DUPLICATE_LIB_OK is set, so continuing

 (Enter /usr/local/g03/l703.exe)
 Compute integral second derivatives, UseDBF=F.
 Integral derivatives from FoFDir, PRISM(SPDF).
 Symmetry not used in FoFDir.

...

 Full mass-weighted force constant matrix:
Low frequencies  -748.4009  -0.7464  -0.4355  -0.1681  -0.0002  0.0003
Low frequencies     0.0005   43.7044   50.8752
 ******    1 imaginary frequencies (negative Signs) ******


...

 Item               Value     Threshold  Converged?
 Maximum Force            0.000017     0.000450     YES
 RMS     Force            0.000003     0.000300     YES
 Maximum Displacement     0.001726     0.001800     YES
 RMS     Displacement     0.000452     0.001200     YES
 Predicted change in Energy=-9.795718D-09
 Optimization completed.
    -- Stationary point found.
 GradGradGradGradGradGradGradGradGradGradGradGradGradGradGradGrad

 Leave Link  103 at Wed Sep  6 18:47:51 2006, MaxMem=   32768000 cpu:       0.1
```

SCF results

Print of the MO coefficients

Beginning of the frequency calculation

To check if the Hessian has only one negative eigenvalues look for "Low frequencies"

Convergence criteria

```
warning: Initializing libguide.a, and found libguide.a already initialized, but
KMP_DUPLICATE_LIB_OK is set, so continuing

IN NATURE THERE ARE NEITHER REWARDS OR
PUNISHMENTS -- THERE ARE CONSEQUENCES.

                              -- ROBERT GREEN INGERSOLL
 Job cpu time:  0 days  0 hours  4 minutes 23.6 seconds.
 File lengths (MBytes):  RWF=     68 Int=      0 D2E=
1
 Normal termination of Gaussian 03 at Wed Sep  6 18:47:51
```

All the Gaussian job terminate with a phrase and if the job has finished correctly with the phrase "Normal termination"

The output contains also a section with the "**Thermochemistry**" pr it at steps 4 and 5.

IMPORTANT: If our convergence criteria are not met, this is normal and not an indication that something went wrong. One normally does an Hessian calculation at a stationary point as part of a vibrational analysis to obtain harmonic frequencies and normal modes. The above message appears when the input geometry is *not* a stationary point (i.e., the gradient is not sufficiencly close to zero), in which case a vibrational analysis is not meaningful. In our case, we are not really interested in the frequencies and normal modes; we simply want to use the resulting Hessian in our TS search. Since we computed the Hessian at a guessed geometry, it is not at all surprising that the geometry is not a stationary point.

If your calculation contains a large negative frequencies in the section "Low Frequencies" it means that the Hessian is good for starting our next TS search (in the ouput, in fact, it is also evidenced that
****** 1 imaginary frequencies (negative Signs) ******).

Before actually starting the TS search, it is advisable to plot the mode corresponding to the imaginary frequency in order to verify that this mode reasonably describes our reaction coordinate.

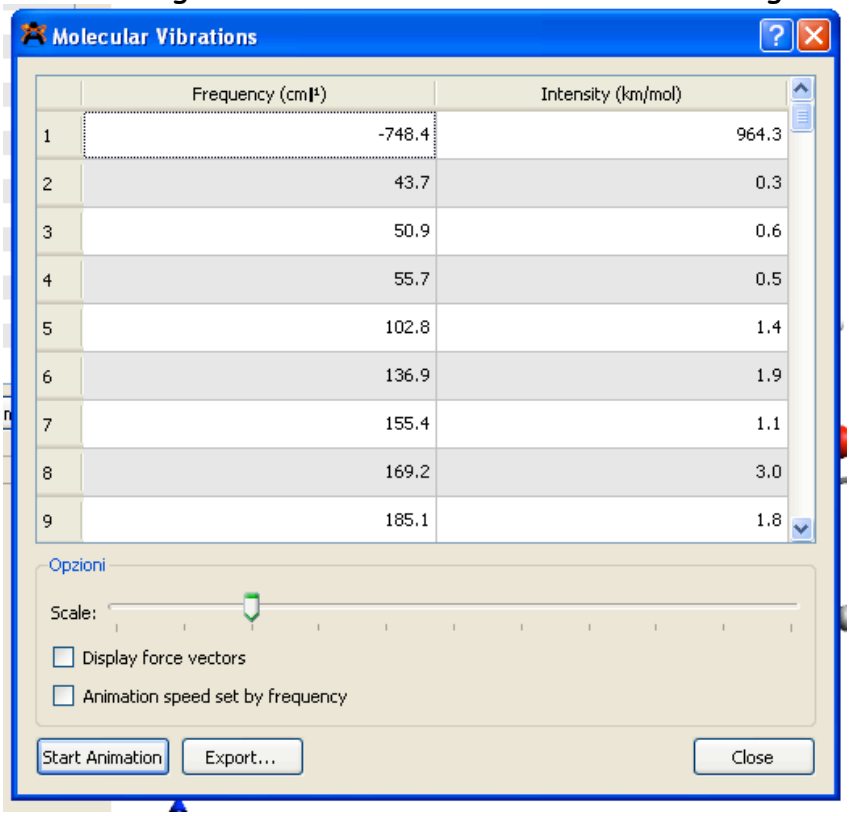
Step 2.d Displaying the normal mode

To display the normal mode corresponding to the frequency we will use Avogadro under the windows enviroment. First of all you have to transfer on your PC the "filename.log" file using WinSCP.

Now you can start Avogadro on your PC. The Avogadro GUI will appear (refer to Section 4.b.1)

Proceed as follow:

1) Click "Open" in the toolbar
2) Be sure the option "All files" is activated, otherwise select it
3) Choose the file. Search for the output file in your directory.
4) Under the "Extensions" menu, click "Vibrations".
5) A dialog box like the following should appear:



6) Click on the negative frequency and click on Start Animation button. The molecule will start to move. If you don't see it move the dialog box away from Model window. Optionally you can select Display force vectors to view the force on each atom.
7) Check if the movement is the required one for the reaction under study.
8) Click on Stop Animation in the same dialog box, then click Close.
9) Close Avogadro.

If the movement correctly corresponds to the reaction coordinate under study go to step 3. Otherwise change the guess structure and re-compute the frequencies.

Step 3: Search for the transition state

In this step we actually run the TS search. We have to modify the previous input file and run the job.

Step 3.a Copying and Editing the previous Gaussian Input File

Make a copy of your previous "filename.com" in "filename_new.com" Open the text file and modify the settings such as method, basis set, calculation type etc. as follows:

```
*****************************************************************

%chk=TS-1-opt
%mem=250Mb
#p B3LYP 3-21G* gfinput pop=full
   opt(TS,readfc,noeigentest)

Search for TS-1

0 1
6        2.684531   -0.638608   -2.084797
6       -1.098009   -1.385873    0.589539
...

1       -3.550267   -0.570119    1.415772
*****************************************************************
```

N.B. Here you will have the same cartesian coordinates of your previous input! In fact, you have not optimized anything at this point!

Explanation:

%chk=filename_new.chk. ***WARNING: if you want to save your old checkpoint file you need to change the name here. This requires also that you copy (on the Unix machine, i.e. using puTTy) the old "filename.chk" to "filename_new.chk". If you forget it the job will stop because it needs to read information from it!***

gfinput, pop=full: see Step 2.a

B3LYP is one of the energy functionals of the density functional methods.

DFT ACCURACY**: quantum mechanical methods have matured into a hierarchical set of methods suitable for very accurate results for small molecules (a few heavy atoms) to rather approximate methods that can be applied to molecules with thousands of atoms. The absolute accuracy of hybrid density functional theory methods such as B3LYP is only moderate (mean average error ~ 3 kcal/mol; maximum error ~ 20 kcal/mol). However, the method can be used to compute activation energies on reactions involving around 25 heavy atoms with accuracies of a few kcal/mol and stereoselectivities with errors of less than 0.5 kcal/mol; the method and related newly developed functionals are the methods of choice for the exploration of problems in organic stereoselectivity

3-21G* is a split valence basis set. The standard basis set used for this kind of studies is the standard and well-tested 6-31G* basis set. Here we will use the smaller 3-21G* basis set thus sacrificing a bit of accuracy for speed and simplicity.

Opt keyword requests that a geometry optimization be performed. The geometry will be adjusted until a stationary point on the potential surface is found. Options: **Opt=TS** requests optimization to a transition state rather than a local minimum; **Opt=noEigentest** keyword avoids testing for eigenvalues. Quite often, in fact, the initial guess for a transition state will give more than one negative eigenvalues, and crash; **Opt=readfc** keyword feeds the Hessian computed at step 2 (if it is okay!) into the TS search calculation.

Step 3.b Saving, Transferring on Unix and Running the calculation

See Step 2.b

The calculation will run for about 50 minutes. Use this time to setup the needed calculations for the other TS

Step 3.c Interpreting the output

For a geometry optimisation, the energy gradient will be computed, and the geometry updated. Thus, a new information will be printed in the output compared to the previous computation. There will be, in fact, a convergence test, the results of which appear in the output as follows:

```
**************************************************************

       Item               Value     Threshold  Converged?
 Maximum Force            0.012664     0.000450     NO
 RMS     Force            0.001718     0.000300     NO
 Maximum Displacement     0.108228     0.001800     NO
 RMS     Displacement     0.025875     0.001200     NO
**************************************************************
```

All the four criteria need to be converged for the job to end. If they are not, there will be a new energy computation at the new geometry, and so on. As the optimisation proceeds, the Values should all decrease, and should start to drop below the Thresholds. Some of the 'NO's will be replaced by 'YES's. After the geometry has converged, there will be a message such as:

```
**************************************************************

       Item               Value     Threshold  Converged?
 Maximum Force            0.000030     0.000450     YES
 RMS     Force            0.000004     0.000300     YES
 Maximum Displacement     0.001046     0.001800     YES
 RMS     Displacement     0.000238     0.001200     YES
 Predicted change in Energy=-4.439922D-09
 Optimization completed.
```

```
    -- Stationary point found.
                        ----------------------------
                        !   Optimized Parameters   !
                        ! (Angstroms and Degrees)  !
--------------------------                         -------------------
! Name  Definition            Value          Derivative Info!
-------------------------------------------------------------------------
 ! R1    R(1,6)                  1.5573         -DE/DX =    0.0    !
 ! R2    R(1,11)                 1.5444         -DE/DX =    0.0    !
 ! R3    R(1,12)                 1.0953         -DE/DX =    0.0    !
 ! R4    R(1,13)                 1.0906         -DE/DX =    0.0    !
 ! R5    R(2,5)                  1.391          -DE/DX =    0.0    !
 ! R6    R(2,14)                 1.0826         -DE/DX =    0.0    !
 ! R7    R(2,15)                 1.0825         -DE/DX =    0.0    !
 ! R8    R(3,5)                  1.3452         -DE/DX =    0.0    !
*************************************************************************
```

The final geometry will be presented in several ways, and will be followed by the wavefunction information.

If the TS search locates a stationary point, we need now to verify that it is indeed a saddle point with a subsequent Hessian calculation (Step 4).

You can also plot Molecular Orbitals with Avogadro, but first you need to convert the chk file into a formatted chk file readable by the software. In the folder on the Unix machine where you have the chk file, type

formchk filename.chk

If you list the files (command ls) you will see a filename.fchk. Transfer it to your Windows PC with WinSCP and open it with Avogadro. Go to Extensions menu --> Create surfaces. In the dialog box you can select Molecular Orbitals for Surface type and also for Colour, and on the right side the list of MOs appears, with HOMO and LUMO labelled. Choose the HOMO for example, click on Calculate. After few seconds a coloured surface has appeared around your model. You can always switch it off using the Display Type panel (section 4.b.1), by deselecting Surfaces option.

Step 4: Calculating frequencies at the transition state

Step 4.a Copying and Editing the previous Gaussian Input File

Make a copy of your previous "filename_new.com" in "filename_new2.com" Open the text file and modify the settings as follows:

```
**********************************************************************

%chk=TS-1-freq
%mem=350Mb
#p B3LYP 3-21G* gfinput pop=full
   nosymm freq geom=check

Calculating Frequencies at TS-1

0 1
**********************************************
```

N.B. Here you have a blank line instead of your geometry because it is read from the checkpoint file!

Explanation:

%chk=filename_new2.chk. ***WARNING: consider the same warning for the checkpoint file as in section 3.a***

At this point you should know the meaning of all the keyword except one:

geom=check simply requests that the geometry is taken from the checkpoint file. If the previous calculation is a geometry optimization, the last optimized geometry will be used for this calculation.

Step 4.b Saving, Transferring on Unix and Running the calculation

See Step 2.b

The calculation will run for about 60 minutes

Step 4.c Interpreting the output

For a frequency job, as discussed earlier (see Step 2.c), there will be a section in the output summarising the results of the frequency run. This will look like this:

```
**********************************************************************

Full mass-weighted force constant matrix:
Low frequencies --- -218.9253  -14.4074   -9.3989   -6.3363   -0.0006   -0.0004
Low frequencies ---    0.0007   42.5760   48.2305
 ******    1 imaginary frequencies (negative Signs) ******
 Diagonal vibrational polarizability:
       49.9223586     120.3780146      25.2210809
 Harmonic frequencies (cm**-1), IR intensities (KM/Mole), Raman scattering
 activities (A**4/AMU), depolarization ratios for plane and unpolarized
 incident light, reduced masses (AMU), force constants (mDyne/A),
 and normal coordinates:
                     1                 2                 3
                     A                 A                 A
 Frequencies --  -218.9220          42.5108           48.1681
 Red. masses --     6.3538          3.4809            4.0634
 Frc consts  --     0.1794          0.0037            0.0056
 IR Inten    --   139.0935          0.6129            0.8042
**********************************************************************
```

This gives you the symmetry assignment of each vibrational mode, its frequency (in wavenumbers), its IR and Raman intensity, the motions of the atoms to which it corresponds, etc. The output also gives you the total Zero-point energy, in atomic units and related thermodynamical properties clustered in a section under the heading "Thermochemistry". There you'll find the Zero-point energy, entropy, partition function, moments of inertia and other useful thermodynamic data.

```
**********************************************************************

 Zero-point correction=                0.322556 (Hartree/Particle)
 Thermal correction to Energy=                    0.339092
 Thermal correction to Enthalpy=                  0.340036
 Thermal correction to Gibbs Free Energy=         0.278980
 Sum of electronic and zero-point Energies=           -745.908379
 Sum of electronic and thermal Energies=              -745.891843
 Sum of electronic and thermal Enthalpies=            -745.890899
```

```
 Sum of electronic and thermal Free Energies=         -745.951955

                       E (Thermal)             CV                S
                      KCal/Mol        Cal/Mol-Kelvin    Cal/Mol-Kelvin
 Total                   212.783             63.522           128.504
 Electronic                0.000              0.000             0.000
 Translational             0.889              2.981            42.164
 Rotational                0.889              2.981            32.539
 Vibrational             211.006             57.560            53.801
 Vibration  1               0.595              1.980             5.139
 Vibration  2               0.595              1.978             4.891
 Vibration  3               0.596              1.974             4.496
 Vibration  4               0.603              1.951             3.496
 Vibration  5               0.613              1.918             2.874

...

 Vibration 25               0.942              1.065             0.477
 Vibration 26               0.945              1.057             0.471
 Vibration 27               0.966              1.016             0.437
*******************************************************************
```

Step 4.d Displaying the normal mode

To display normal modes see Step 2.d

Intrinsic reaction coordinate (IRC) paths should be traced in order to check that each transition structures actually connect the two associated minima of the proposed mechanism

Step 5: Comparing results for TS-1 and TS-2 to confirm the expected stereoselectivity.

The geometry of the two transition states **TS-1** (**TS-ar**) and **TS-2** (**TS-ss**) should resemble the ones already plotted in Figure 1 of the introduction.

→ The total energy for TS-1 is -746.230934939 Hartrees

(Sum of electronic and thermal Enthalpies= -745.890899)

→ The total energy for TS-2 is -746.226851325 Hartrees

(Sum of electronic and thermal Enthalpies= -745.887248)

Thus, the transition state involving the re attack on the anti enamine (TS-ar) is about 2.5 kcal mol^{-1} lower in energy than the transition state for the si attack on the syn enamine (to convert Hartrees to kcal mol^{-1} multiply for 627.51). This energetic difference agrees well with the experimental results where the (R)-cetol (obtained if the C-C bond formation takes place along the anti arrangement of the enamine) is isolated in about 96% ee.

Houk et al. determined that for three reactions studied, the calculated gas-phase enthalpies of activation have the smallest error (standard deviation ± 0.4 kcal/mol) when compared with the experimental enantioselectivities (**Bahmanyar, Houk et al., JACS, 2003, 125, 2475-2479**).

The predicted product ratios from ΔH_{298} values (search for "Sum of electronic and thermal Energies" in the ouput), in our case confirm the stereoselectivity of the reaction:

Structure	Product Ratio	
	Calculated	Experimental
TS-ar	ΔH_{298}=0.0	~96-99%
TS-ss	ΔH_{298}=2.3	~4-1%

4.d Summary of the exercise

Case study:

"Direct Aldol Reaction between Acetone and Isobutyraldehyde"

Steps of the exercise:

1. Create an initial geometry for transition states with Chem3D

2. Prepare the approximate Hessian

3. Search for transition states

4. Calculating frequencies at transition states

5. Comparing results for **TS-**1 and **TS-2** to confirm the expected stereoselectivity.

4.e Parachute Section

If you cannot converge the two TSs, you may wish to see and analyze the inputs and outputs that are on your PC (ask the demonstrator for the directory!).

For TS-1 (TS-ar):

Step of the calculation	Input	Output
2	TS-1-hess.com	TS-1-hess.log
3	TS-1-opt.com	TS-1-opt.log
4	TS-1-freq.com	TS-1-freq.log

For TS-2 (TS-ss):

Step of the calculation	Input	Output
2	TS-2-hess.com	TS-2-hess.log
3	TS-2-opt.com	TS-2-opt.log
4	TS-2-freq.com	TS-2-freq.log