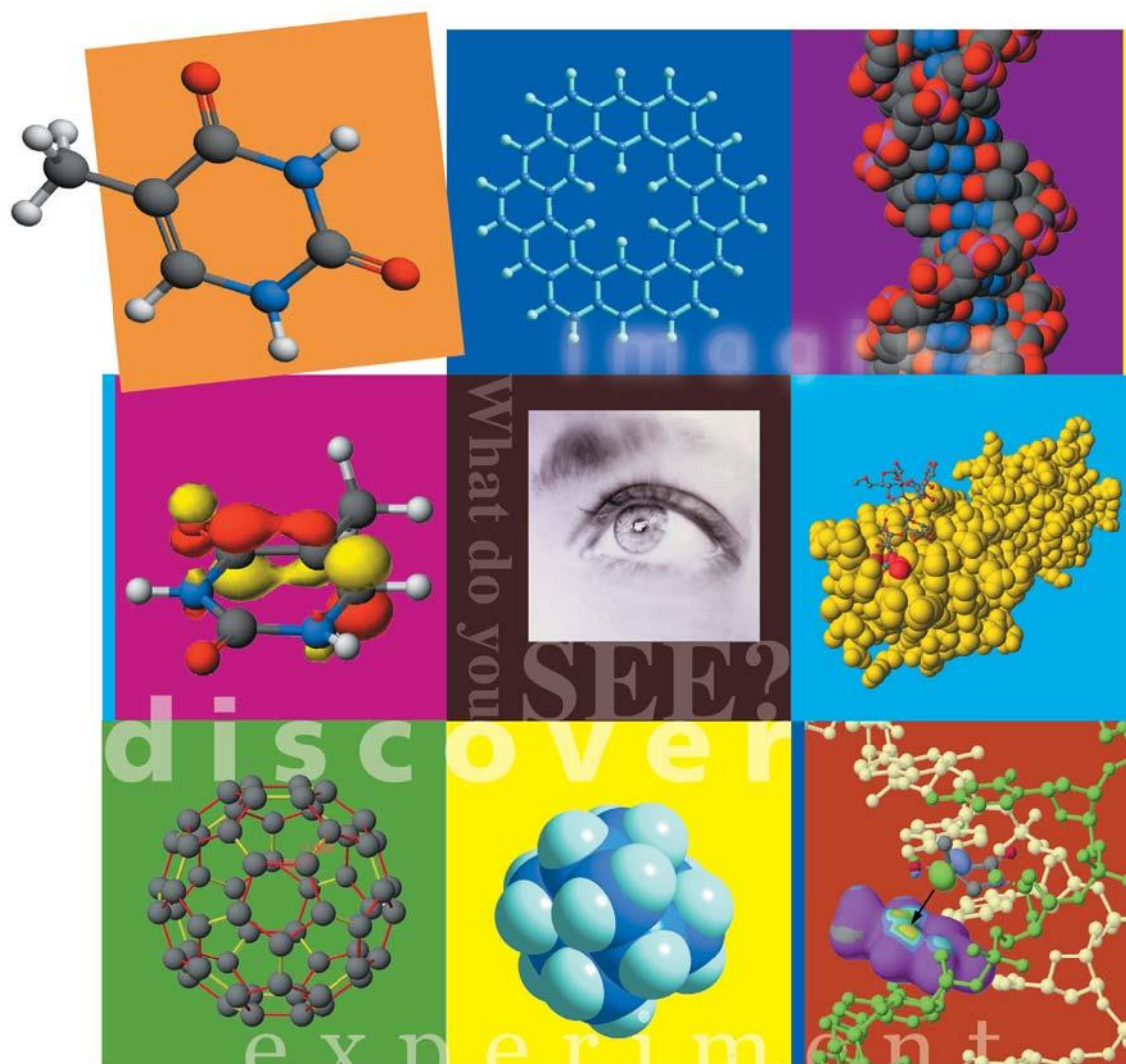# Teaching with SCIGRESS

Exercises on Molecular Modeling
in Chemistry
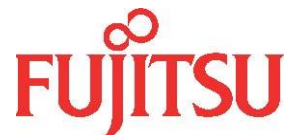
Crispin Wong and James Currie, eds.
Pacific University

Teaching with SCIGRES

Molecular Modeling in Chemistry

Crispin Wong
and
James Currie
Editors

Pacific University
Forest Grove, Oregon

FUJITSU

Copyright statement

Table of Contents

Exercises

Exercises

I. Building Basics

Build:

$$CH_3CH_2CH_2CHCH_3$$

with substituents O and H (drawn as O H above the fourth carbon).

1. Choose **Start | Programs | SCIGRESS Suite | SCIGRESS**. The application opens, containing an empty document window referred to as the workspace.

2. Click on the Drawing pencil tool icon. Your cursor turns into the Drawing pencil cursor shown to the right.

3. Select the arrow button in the Element Type box located on the left of the Style bar. A drop down list of elements appears. Select **C-Carbon**. The default hybridization is sp^3.

4. Look in the Hybridization box. If the hybridization is not sp^3, click on the arrow. From the drop down menu, **select sp3-tetrahedron**. The carbon atom is now sp^3 hybridized.

5. Click once in the workspace to deposit a single carbon (C) atom. Do not double click unless you want multiple atoms.

6. Click on the existing C atom, hold the button down and drag away (click-drag). Release the mouse button. Another C atom appears connected to the first C atom with a with a single bond. Add three more C atoms by doing a click-drag. You should have a string of five C atoms.

7. Return to the Element Type list and select O-oxygen. Make the atom sp^3 hybridized.

8. Click-drag from the C next to the end. A red O atom is connected to the C atom.

9. Click on the Select tool icon. The cursor changes into an arrow shown to the right.

10. Click in the workspace but not on any atoms or bonds. This selects all atoms and bonds in a molecule.

11. Select **Action | Beautify | Comprehensive** (**ctrl + b**). If the skeleton structure was built correctly, the correct number of hydrogen atoms is present and the angles are proper for sp^3 hybridization.

The Beautify menu has the following correction abilities:

Valence - Adds hydrogens to make the correct number of total bonds. For a neutral carbon, that number is four; for a neutral oxygen it is two; a neutral nitrogen three and a positively charged nitrogen is four.

Configuration - Changes the hybridization to match the number and type of bond. For example, an atom with a total of four attachments or electron pairs is sp^3, three total attachments is sp^2 and two total is sp.

Rings -Gives ring structures the proper shape.
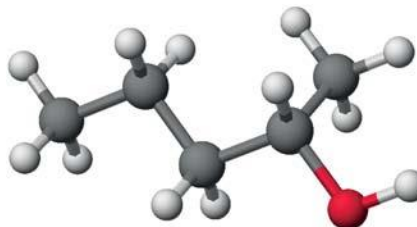
12. Click **File | Save as…** . Save the molecule with the name "2-pentanol" in the appropriate location. The molecule saves as a chemical sample ".csf" file type.

II. Viewing the Structure

Open the molecule "2-pentanol" from Exercise I in the Editor.

Translate, Rotate, and Zoom

1. Change the position of the molecule by translating, rotating, and zooming. Experiment with each of these three modes.

   Rotate: Click on the Rotate icon. In the window there is a circle around the molecule and your cursor is an arrow. Click and drag your cursor to rotate the molecule. Movement inside the circle produces rotation around the X and Y axes while movement outside the circle allows rotation around the Z axis.

   Translate: Click on the Translate icon. The cursor appears as a hand. Click and drag the mouse to move the molecule around on the screen.

   Zoom: Click on the Zoom icon. The cursor becomes a magnifying glass. Click and drag the magnifying glass up or down to zoom in or out.

   The following key combination is very useful for repositioning the atom:
   Press the **Ctrl + f** to center the structure and scale the molecule to fit the window. If only one atom is selected, the structure is centered around the selected atom but no scaling is done.

   Press i to return the molecule to the initial scale and orientation.

2. Enter the Select tool mode.

Rotate

Translate

Zoom

3. Select **View | Styles | Atoms...** The Atom Attributes dialog box opens. The options in the Atom Attributes box allow you to change the way the atoms and bonds are displayed.

4. Click on the **Label** tab to display the list of label options.

5. Click in the square next to Atomic Symbol. The square has a check mark inside, indicating that the option is selected. Click **OK**. The element letter appears on each atom.
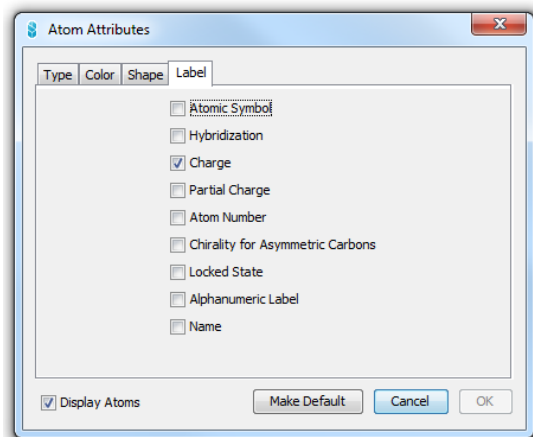
NOTE

Viewing with labels enables easy inspection of charges. However, the structures are more suitable for formal reports without the labels.

6. Select **View | Styles | Atoms...** Click on the **Shape** tab to display the list of options. Locate the "Scale Spheres such that" section. The default view is usually scaled to 0.2 times the van der Waals radius. To view the atoms at their true size, set this number to be 1.0. This mode is known as the space-filling view.

A quick way to reach the space-filling mode is to click on the space-filling icon on the toolbar. The icon is a short cut for the command **View | Atoms & Bonds | Space Filling.**

Experiment with the different view options.
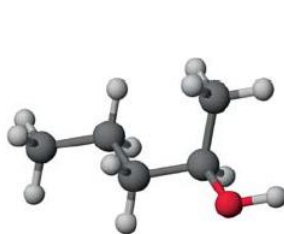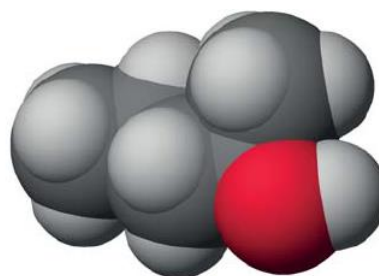
Scale = 0.2   Scale = 1.0

7. Return the molecule to "0.2" atom scaling before continuing with other exercises.

III. Modifying structures

Open the "2-pentanol" file from Exercise I, and change the structure to:

$$\overset{\overset{\displaystyle O}{\|}}{CH_3CH_2CH_2CCH_3}$$
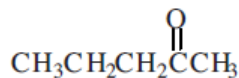
1. Click on the C-O bond using the Select tool to select the bond. Click on the arrow in the Bond type box on the Style bar. A pop up menu appears. Select Double; the C-O bond is now double. The sp^3 hybridization is wrong for atoms with double bonds.

2. Click on the C atom connected to the O atom. Select **sp2 - trigonal plane** in the Hybridization box.

3. Repeat this process for the O atom and set it also for sp^2 hybridization.

4. Select **Action | Beautify | Comprehensive.**

5. Save 2-pentanone structure. You will use it later in the tutorial.
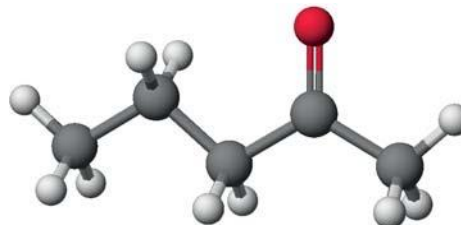
IV. Further modifications

In this section we will use several methods to add or change atoms. These methods are generally interchangeable. Start with "2-pentanol" from Part III and modify it in the following manner:

1. Click on the Select tool icon to enter the Select tool mode.

2. Select the hydrogen (H) atom that you want to replace with a bromine (Br) atom.

3. Click on the arrow in the Element type box on the Style bar. In the pop up menu, select **Periodic Table**. A Periodic Table opens in a dialog box.

4. Select the element **Br**. Click **OK**. The H atom is replaced with a Br atom. Look in the Hybridization box; make sure the hybridization is sp^3.

5. Check the Charge box; make sure the charge is 0. If the charge is not 0, you can change the charge by clicking the up arrow to get a positive charge and the down arrow to get a negative charge.

6. Select the O atom. In the Charge box, change the charge to **+1**.

7. Select **Action | Beautify | Comprehensive**. The missing hydrogen is added and the bond distances and angles are corrected.

📖 NOTE
In this case Beautify would have corrected the hybridization but this exercise was intended to show how to change the hybridization.

8. Select in order, the H atom connected to the O atom, followed by the O atom and then the C atom attached to the O.
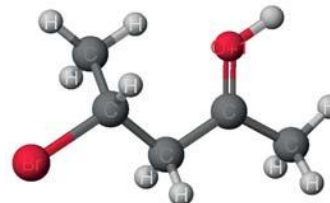
9. Select **Action | Geometry Labels | Bond Angle**. A Bond Angle dialog box opens, displaying the bond angle in degrees. You can label, record or adjust the angle. In this case, you will label the angle. Select **Define Geometry Label**, and then **Unlock Geometry**. Unlock geometry allows the angle to change as you modify the structure. Click **OK**. The label appears on the bond angle.

   In the same manner, you can label the atom distance by selecting two bonded atoms and select **Action | Geometry Labels | Atom Distance**.

10. Select **File | Save as**. The Save as command leaves the last saved version of the file intact, and saves the current molecule as a different file. Name the file "modifications."

V. Using the Fragment Library

Build:
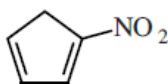
1. Open library window (**Window | Show Explorers | Library Window**) (Library window is open by default).

2. Open the Cyclic folder. A list of SCIGRESS documents is displayed.

3. Drag & drop **cyclopentane** molecule into blank worksheet window.

4. Select **File | Save as**. In the resulting Save As dialog box locate the folder where you want to store the molecule. Name the file "nitrocyclo." Click **Save**.

5. Rotate the molecule so you can see all the C–C bonds.

6. Change Two C–C single bonds to C=C double bonds. Select the bond with the select tool. In the Bond type box, select **Double.** (alternatively drawing pencil tool can be used to connect atoms with additional bonds)

7. Select **Action | Beautify | Comprehensive**.

8. Open the Fragment Library.

9. Select **Functional groups** from the folder list. Find the **nitro** group in the files and highlight it. Drag & drop functional group next to cyclopentadien molecule.

10. Click in the workspace but not on any molecules or bonds to select the nitro group.

11. Select one of the double bonded C atoms adjacent to the CH_2 and shift-click (or Ctrl-click) on the nitrogen (N) atom. Both the C and the N atom are selected.

12. Select **Action | Connect Atoms**. A bond connects the selected atoms.

13. Click in the workspace but not on any atoms or bonds. All atoms are selected.
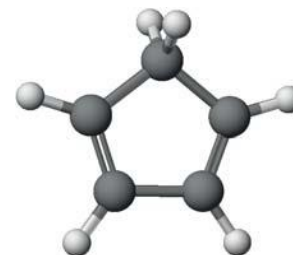
14. Select **Action | Beautify | Comprehensive.**
15. Save the molecule.
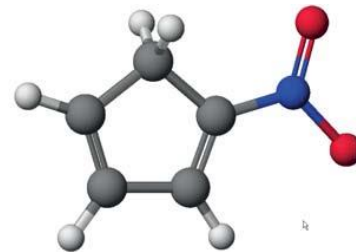
VI. Computing Properties

To predict properties of different models, there are various calculations that you will need to run for your molecules. Experiments are the link between the molecular model and the different SCIGRESS computational applications.

Optimizing molecular geometry is a common calculation. There are two primary options for optimizations: Molecular Mechanics and MOPAC.

Molecular Mechanics

Molecular Mechanics is a computational application that treats molecules like weights connected by springs. Optimizations depend on the bond types and atom hybridizations. Since simple classical physics is used for the optimization, no information about electronic properties or reactivity is determined.

MOPAC

An application that uses a quantum mechanical approach can give the electronic properties of a molecule. MOPAC is such a program and we can use it to find properties such as heats of formation, orbital energies and shapes, and the distribution of electrons.

Optimize a molecule

1. Open the molecule "nitrocyclo" in the Workspace.

2. Select **Experiment | Run**. An Experiment dialog box opens.

3. Choose **Category**: **chemical sample**.

4. Choose **Property**: **optimized geometry**.

5. Choose **Procedure**: **MO-G PM3**, which uses MOPAC and PM3 parameters to optimize the molecule.

6. Click **Start**. A calculation window appears and information about the progress of the optimization scrolls in the window. When the calculations end, MOPAC reports a heat of formation. Other properties are also calculated and stored in the molecule file.

Generate MOs and Other Surfaces

Now that we have electronic information from the MOPAC calculations, we can compose pictures of the molecule surfaces and individual molecular orbitals (MOs).

1. Click in the molecule window to activate it.

2. Select **Analyze | Electronic Structure** | **HOMO & LUMO.**

3. Return to the molecule window. As a second surface example, tabulate the electrophilic susceptibility.

4. Select **Analyze | Electronic Structure | electrophilic susceptibility**.

VII. Visualizing Properties

We will now look at some of the properties that we calculated using MOPAC.

1. Open the molecule "nitrocyclo" in the Workspace.
2. Locate the surface "EFonD" and select it. A check mark next to the surface indicates that it is selected. Your molecule is displayed with a surface colored to show the areas of maximum electrophilic reactivity. Red areas indicate where electrophiles (H^+, NO_2^+, Br^+...) have the highest reactivity.
3. Select the HOMO surface by double clicking it in the Model Explorer, and deselect the electrophilic surface by also double clicking it. The electrophilic surface disappears and the molecule displays highest occupied molecular orbital (HOMO). As a general rule, when you deselect a surface, the surface disappears from the window. All selected surfaces are displayed in the molecule window, allowing you to view more than one surface at a time.
4. Save the molecule.

VIII. Energy Maps

A change in a molecule's potential energy occurs whenever a bond angle or distance is altered. If this energy is plotted against the changing property, we get a potential energy coordinate or map. If the energy change during a reaction is plotted against the physical change taking place, we get a potential energy map that we usually refer to as a reaction coordinate.

Find the energy change as we rotate the dihedral angle.
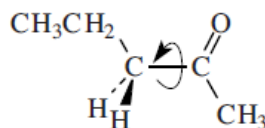
1. Open the file "2-pentanone" from exercise III.

2. Select the O atom. Shift-click on the C atom attached to the O. Shift-click on the next two C atoms with two H atoms bonded to it (see diagram to the right). All four atoms are selected.
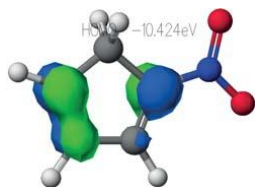
3. Select **Action | Geometry Labels | Dihedral Angle**. Adialog window opens.

4. Click **Define Geometry Label** and **Search** between **-180** and **+180** using **24** steps.

5. Click **OK**. A search label reading 0.00˚ appears on the molecule.
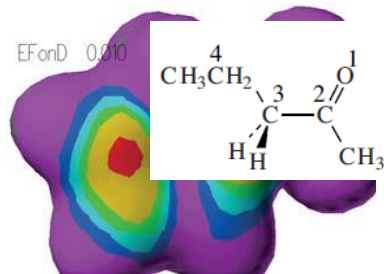
We can now calculate the energy for each of the angles specified by the label.

6. Select **Experiment | Run**. In the Experiment window, select Category: **conformations**, Property: **potential energy map**, Procedure: **standard procedure**.

7. Click **Start**. A window opens, showing the progression of the calculation. When the calculation is done it opens the map window on the left and molecule window on the right.

The map file is saved in the same location as your molecule with the same name as the molecule only with the ending .map.

8. Click on the reaction coordinate line in the left window and observe the corresponding changes in the molecule. Click on the molecule window to activate it. Enlarge and rotate the molecule so you can see the changes.

In some experiments you may need to adjust the axes. To do this, right click on the axis label and choose **Properties**. An Axis Attributes dialog box appears. Click on the Range tab. Adjust the range as needed.

Part 1: General Chemistry

1 Calculating the Geometry of Molecules and Ions

Contributed by Kevin Johnson and James Currie, Pacific University

Overview

In this experiment you will use SCIGRESS to create computer representations of molecules and ions and to optimize their structures.

Recommended Exercises

Complete Pre-lab Assignment: Molecular Modeling before beginning the experiment.

Background

One of the important achievements in chemistry is our ability to predict the bulk properties of a compound based on what we know of the microscopic structure of molecules and ions. Molecular geometry provides much of the information upon which these predictions are made.

From lectures you have learned to use several models to describe molecules. For example Lewis dot structures delineate connectivity and bonding relationships between atoms of a molecule or ion and can predict bond energies and bond lengths. However, Lewis dot structures do not predict shape.

Lewis Dot Structure

The most basic theory that predicts geometry is the valence-shell electron-pair repulsion (VSEPR) model. The VSEPR model is based on the premise that the electron pairs around a central atom of a molecule arrange themselves in space to minimize the repulsive interaction between regions of negative charge. This model generates the five basic electron pair geometries: linear, trigonal planar, tetrahedral, trigonal-bipyramidal, and octahedral. The molecular geometry then follows from the connections to other atoms via bonding pairs and from non-bonding electron pairs yielding eleven possible molecular geometries. To qualitatively predict variations in bond angles associated with double bonds or non-bonding pairs, the theory allows for these electron pairs to fill a larger volume of space, and thus distort the predicted molecular geometry toward smaller bond angles.

VSEPR Tetrahedral Geometry
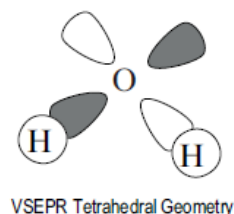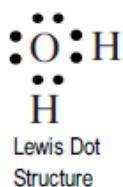
While Lewis dot structures predict bonding and VSEPR theory successfully predicts molecular geometry, neither theory takes into account the relationship between atomic orbitals and electrons in bonds. Valence-bond theory approaches this by describing bonding as the overlap of atomic orbitals from two bonded atoms and explains molecular geometry using hybridized atomic orbitals.

A more comprehensive model of molecular bonding applies the Schrödinger equation to the electrons and nuclei that make up a molecule. This forms the basis of molecular orbital (M.O.) theory. For any but the simplest of molecules, M.O. theory requires application of quantum computational methods discussed below.

4 Oxygen sp^3 Hybrid Orbitals
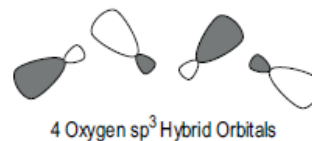
As scientists we use theories and models to describe natural processes that give rise to observable phenomena. Given a choice between competing theories, that which is least complex yet still satisfactorily answers a given problem is preferable. Thus if we are interested in identifying the shape of a molecule and any expected distortions, VSEPR theory will suffice. However, if we are interested in quantitatively predicting values for bond angles and bond lengths, we must resort to a more sophisticated theory and use computational methods based on M.O. theory.
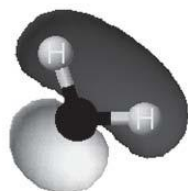
H_2O Molecular Orbital

Computational Methods

All computational methods are based on the postulate that molecules take the form with the lowest (i.e. most stable) energy. Computational chemistry software works by calculating the energy of a given molecular geometry, then adjusting the atom positions toward lower energy configurations. This process, called optimization, is iterated until a minimum energy configuration is achieved.

Computational methods vary in the procedures used for calculating energy and geometry optimization. The two main types of computational methods, classical and quantum, differ in how the energy of a given geometry is calculated.

Classical methods such as Mechanics approximate bonds as "springs" and atoms as "balls" and use classical physics to calculate molecular energy. Quantum methods, such as those in MOPAC, numerically solve the Schrödinger equation using either *ab initio* or semi-empirical calculations. *Ab initio* methods calculate energy from approximated atomic orbitals and are very computationally intense. Semi-empirical calculations are limited to valence electrons and make assumptions about the atomic orbitals of atoms based on experimental – or empirical – evidence, and therefore require less computation. You will be using semi-empirical calculations for this experiment.

Modeling Section

Exercise 1: With the Instructor

Find the SCIGRESS folder on the hard drive; open it and double click on the program SCIGRESS. When the SCIGRESS opens you will have the Projects window in which you can choose to open recent project or to open a new one.

Build Ammonia NH$_3$

1. Having the project opened create a new blank object by choosing "File | New Object | Blank" and then select the Drawing pencil tool.

   ❏ Click on the Drawing pencil tool icon.

2. Select an N atom.

   ❏ Click on the arrow button in the Element type box at the left of the style ribbon. A drop-down list of elements is displayed. Select a N-Nitrogen atom. The hybridization is *sp³*; look in the Hybridization box to verify that the N atom is *sp³*.

3. Click once near the center of the window. A N atom appears where you clicked.

4. Select an H atom in the same manner as the N atom.

5. Place the pencil tool over the center of the N atom. Click and drag away from the atom, then release. This creates an H atom bonded to the N atom.

6. Repeat this two more times to create two more H atoms bonded to the N atom.

In the future you may need to:

Insert Atoms Not on the Atom List

1. Open the periodic table.

   ❏ Click on the arrow button in the Element Type box and choose **Periodic Table**...

2. Click on the symbol of the atom you need.

Delete Atoms

1. Choose the Select tool. The cursor appears as an arrow.

   ❏ Click on the Select tool icon.

2. Click on the atom you want to delete. Make sure that only the atom or atoms you want to delete are selected.

3. Press the delete button on your keyboard or select **Edit | Delete.**
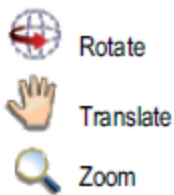
Rotate

Translate

Zoom

Translate, Rotate, and Zoom

1. Change the position of the molecule by translating, rotating, and zooming. Experiment with each of these three modes.

   ❏ Select the icons on the tool palette. Click and drag the cursor in the molecule window to perform the action corresponding to the icon

Fix the Geometry

NOTE

The Beautify commands apply simple VSEPR rules to the molecular geometry. A Comprehensive Beautify command automatically adds hydrogen atoms to complete the valence shell of an atom. In building NH3, if you omit the hydrogen atoms the Beautify command would add them.

1. Notice that the geometry of the NH_3 molecule you created is not correct.

   ✎ According to the VSEPR rules, what shape should it be?

2. Choose the Select tool.

3. Select the entire molecule by clicking inside the window but not on the molecule.

4. Select **Action | Beautify | Comprehensive**

   ✎ Is the molecule now in the shape you expect for NH_3?

5. Use the Rotate mode of the cursor to look at the three dimensional shape of the molecule.

Measure the Bond Lengths and Angles

TIP

When selecting atoms, make sure that the cursor is in the center of the atom.

1. Click on a H atom and hold down the Shift or Ctrl key while you click on a N atom. This selects both atoms.

2. Choose **Action | Geometry Labels | Atom Distance…** A Set Atom Distance dialog box appears; it shows the separation between the atoms in Angstroms.

3. Click **Cancel** after you record the distance in your lab notebook.

4. Select three atoms that form an angle in order: the end, the center, and the other end atom. In that case, you will select H-N-H. Be sure to select them in the correct order.

TIP

Geometry labels can be displayed quickly using **Ctrl + l** shortcut.

5. Choose **Action | Geometry Labels | Bond Angle…** A Set Bond Angle dialog box appears; it shows the angle between the atoms in degrees.

6. Record H-N-H angle measurement in your lab notebook.

   ✎ Is this result consistent with your VSEPR prediction?

   ✎ Does it account for the lone-pair on the nitrogen atom in ammonia?

Save the Molecule

1. Select **File | Save as…**. A dialog box appears.

2. Name the file "NH3", and make sure that it is in the correct folder/location. (Your instructor should specify a designated place). Click **Save**, but keep the molecule window open.

Optimize the Structure

1. Select **Experiment | Run**... The Run Procedure box opens. The program is automatically set to do the optimization calculation.
2. Click **Start.** This brings up a window that displays the progress of the calculation. When the calculation is complete, the dialog box displays "Calculation Done." The structure is now optimized.
3. Close the dialog box and return to the Workspace. Do NOT select the Beautify command after optimizing the structure. It will return the molecule to the previous VSEPR geometry.
4. Measure and record in your lab notebook the bond lengths and bond angles.

Exercise 2: On Your Own

1. Build and optimize a CH_2Cl_2 molecule.
2. Record the predicted bond lengths and angles after beautifying the structure and after optimizing the structure.

Exercise 3: With the Instructor

Build a Nitrate Ion (NO_3^-)

1. Open the Workspace and select **File | New Project**. An empty window appears.
2. Place an N atom with three connected O atoms in the window.
3. Add the formal charges on the appropriate atoms.
   ❏ Select the atom. In the Charge text box type the charge for the atom, or press the up arrow button or down arrow button to display charge values in increasing or decreasing order, respectively.
4. Repeat this process for all atoms with nonzero formal charges.
5. Optional: Turn on the atom labels.
   ❏ Select **View | Styles | Atoms...** Click on the **Label** tab Click the box next to Atomic Symbol and Charge.
6. Change the order of the appropriate bonds. The number of bars between the atoms represents the bond order.
   ❏ Select the arrow button in the Bond Type box and choose the bond type.

7. Select **Action | Beautify | Comprehensive.**
8. Measure the bond lengths and angles.
   ✎ Are these results consistent with what you expect for the nitrate ion?
9. Save the structure with the file name "NO3-."
10. Optimize the structure.
11. Return to the Workspace to measure and record the optimized bond lengths and angles.
   ✎ How is the structure different before optimizing compared to after optimizing?

Exercise 4: On Your Own

1. Build and optimize the structure for CO_3^{2-}.

2. Record the predicted bond lengths and angles after beautifying the structure and after optimizing the structure.

Assigned Molecule

1. Build and optimize the structure that your instructor assigned to you.

2. Record the predicted bond lengths and angles after beautifying the structure and after optimizing the structure.

3. Orient the molecule into a good position. Press **Ctrl + f** to center the molecule in the window. Zoom to fill the window with the molecule.

4. Print a copy of the optimized molecule file and staple it to your worksheet.

View Orbitals

1. Close any open files in the Workspace.

2. Open the file **Carbon.csf** from the location specified by your instructor.

3. In the Model Explorer box, there is a list of atoms and a list of surfaces (molecular orbitals) for the molecule that is open.

4. Expand the surface list and double click to select **HOMO-1 surface**.

5. Record what type of atomic orbital the molecule has.

6. Select **HOMO**, and deselect **HOMO-1.** Record what type of atomic orbital the molecule has.

7. View **LUMO** and **LUMO+1** in the same manner. Record the type of atomic orbital. You should rotate the molecule to make sure you see the correct type of orbital. Only select one surface at a time because it is difficult to view the orbitals simultaneously.

   ✎ What do you notice about the directions of the three p-orbitals?

8. Close the carbon file and open the phosphorus file.

9. View the surface files **HOMO-4** to **HOMO** in the same manner as the carbon files. Record the atomic orbital subshell for each file. Remember to rotate the molecules to view the orbitals fully.

10. Close all the files.

View Interesting Molecules

1. Open the folder **Small Molecules** from the location your instructor specifies.

2. Open the H_2 molecule.

3. Expand surface list in the Model Explorer box and double click on **HOMO.**

4. Select the **LUMO** surface and deselect the **HOMO** surface. You can also view both surfaces together by selecting the both. One H_2 file is the bonding and the other H file is the antibonding. Identify which orbital is the σ and π orbitals.

5. Open the C_2H_4 molecule and view the surfaces **HOMO-2** to **HOMO** in the same way as the H_2 molecule.

6. Identify the σ and π orbitals.

7. Open the C_2H_2 molecule. Examine the orbital files **HOMO-2** to **HOMO**. Record the σ and π orbitals.

8. Open the CO_3^{2-} molecule. View the surface file **HOMO-5**. Recall that you can draw the carbonate ion with three resonance structures, where the double bond is between the central carbon atom and each of the three oxygen atoms. This molecular orbital of the carbonate ion is a π - orbital.

   ✎ What is the term for a π−orbital that extends beyond two atoms?

9. Close the **Small Molecules** folder and open the **Big Molecules** folder.

10. Open and examine the three molecules in the folder.

    ❏ Buckyball is an allotrope of carbon with the molecular formula C_{60}; it was discovered in the 1980s.

    ❏ DNA is the molecule that contains genetic codes for living organisms.

    ❏ Taxol is an anti-cancer drug derived from the Pacific Yew tree.

Report Guidelines

Fill out the report form and turn it into your instructor.

Instructor Notes

We use this experiment half way through the first semester of general chemistry. It is an introduction to SCIGRESS and a good way to visualize molecular geometry. The directions are explicit and require no previous knowledge of SCIGRESS.

The students complete a pre-lab worksheet before coming to class. We encourage students to keep notes in their lab notebook, documenting the experiment as if they are doing a wet lab. A lab report worksheet is distributed at the end of the lab session. During the lab, we follow a format where we do one exercise with the students, and then they do one exercise on their own. You can choose whether or not to follow that format.

The default settings in MOPAC are usually set to do an optimization. However, you may want to double check the settings on the machines you will be using. The difference between PM3 and AM1 calculations should not be noticeable on simple molecules. Our computers are set for a PM3 default optimization, but you can choose which parameter to use.

Before the lab, we created the molecular orbital files for the students to view with MOPAC. You can generate the molecular orbitals and molecules, or you can download the files from the website
http://www.chem.pacificu.edu/cache.

Typical Results

Answers to the Pre-lab assignment:

A) Ammonia NH_3: Shape : Trigonal Pyramidal, H-N-H bond angle 109.5˚ (107˚).

B) Dichloromethane CH_2Cl_2: Shape: Tetrahedral, H-C-H bond angle: 109.5˚, Cl-C-Cl bond angle: 109.5˚, H-C-Cl bond angle: 109.5˚.

C) Nitrate ion NO_3^-: Shape: Trigonal Planar, O-N-O bond angle: 120˚.

D) Carbonate ion CO_3^{2-}: Shape: Trigonal Planar, O-C-O bond angle: 120˚.

Answers to Lab Report:
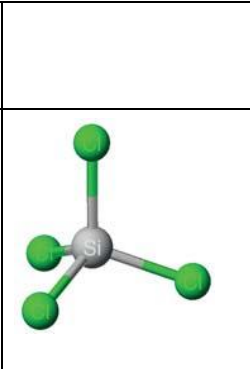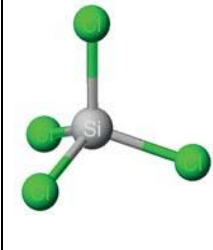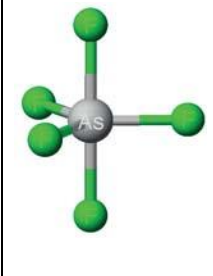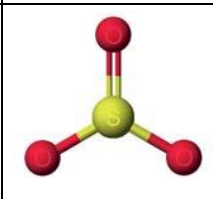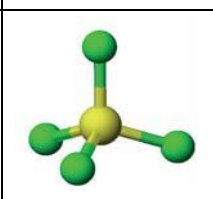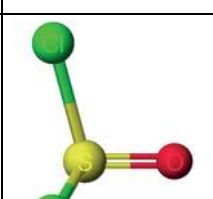
NH_3: N-H bond length: 0.999-1.000 Å, H-N-H bond angle: 107.98˚-108.16˚.
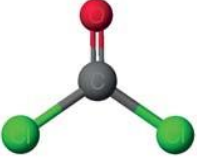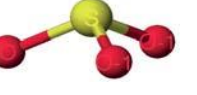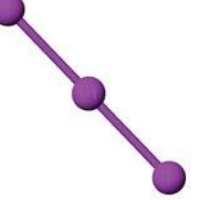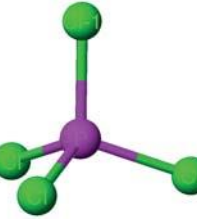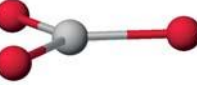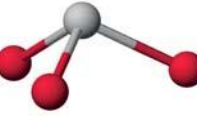
CH_2Cl_2: C-H bond length: 1.102-1.103 Å, C-Cl bond length: 1.756-1.760 Å, H-C-H bond angle: 108.33˚-108.71˚, Cl-C-Cl bond angle: 107.86˚ -108.02˚, H-C-Cl bond angle: 109.85-110.20˚.

NO_3^-: N-O bond length: 1.238-1.239 Å, O-N-O bond angle: 119.93-120.05˚.

CO_3^{2-}: C-O bond length: 1.293 Å, O-C-O bond angle: 119.88-120.13˚

Individual assignment:

Molecule/Ion		Typical results (PM3)	
		Bond length(s)	Bond angle(s)
$SiCl_4$		2.042 Å	109.47˚
AsF_5		1.685 Å 1.661 Å	90.00˚ 120.00˚
PCl_5^{2-}		2.120 Å 2.230 Å	89.5˚ 94.6˚
SO_3		1.384 Å	120.00˚
SF_4^{2+}		1.425 Å	109.47˚
$SOCl_2$		S-Cl 2.08 Å S-O 1.48 Å	97.5˚ 104.7˚

COCl$_2$		C-Cl 1.737 Å C-O 1.198 Å	124.20° 111.60°
ClO$_2^-$		1.606 Å	111.21°
SO$_3^{2-}$		1.640 Å	106.16°
I$_3^-$		2.727 Å	179.23°
PCl$_4^+$		2.025 Å	109.47°
SeBr$_3^-$		2.390 Å	119.96°
SeBr$_3^+$		2.417 Å	100.02°

Orbitals

Atoms

- Carbon: HOMO-1: s, HOMO: p, LUMO: p, LUMO+1: p.
- Phosphorus: HOMO-4: s, HOMO-3: s, HOMO-2: p, HOMO-1: p, HOMO: p.

Molecules

- H : HOMO: σ, LUMO: σ*.
- C$_2$H$_4$: HOMO-2: σ, HOMO-1: σ, HOMO: π.
- C$_2$H$_2$: HOMO-2: σ, HOMO-1: π, HOMO: π.
- CO$_3^{2-}$: HOMO-5: delocalized π.

Pre-lab Assignment- Molecular Modeling

Name: _____

Draw Lewis dot structures and label the nonzero formal charges on atoms for the following four molecules and ions. Predict the shape and bond angles for each. If more than one resonance structure exists draw all possible resonance structures. Record your answers to these questions in your lab notebook. You will need to use them during the lab experiment.

A) Ammonia NH₃

Shape:	H-N-H bond angle:

B) Dichloromethane CH₂Cl₂

Shape:	H-N-H bond angle:
	Cl-C-Cl bond angle:
	H-C-Cl bond angle

C) Nitrate ion NO₃⁻

Shape:	O-N-O bond angle:

D) Carbonate ion CO₃²⁻

Shape:	O-C-O bond angle:

Lab Report- Molecular Modeling

Name: _____

Record all molecular geometry measurements for MOPAC optimized structures.

A. Class assignments:

NH_3	N-H bond length:	H-N-H bond angle:

CH_2Cl_2	C-H bond length:	H-C-H bond angle:
	C-Cl bond length:	Cl-C-Cl bond angle:
		H-C-Cl bond angle:

NO_3-	N-O bond length:	O-N-O bond angle:

CO_3^{2-}	C-O bond length:	O-C-O bond angle:

B. Individual assignment:

Assignment:	Shape:
List bond lengths:	List bond angles:

C. Orbitals

Atoms: Label each file name with the orbital subshell; *s, p,* or *d*.

Carbon

File	Subshell
HOMO-1	
HOMO	
LUMO	
LUMO+1	

Phosphorous

File	Subshell
HOMO-4	
HOMO-3	
HOMO-2	
HOMO-1	
HOMO	
LUMO	
LUMO+1	
LUMO+2	

Molecules: Label each file name with the molecular orbital type: σ, σ*, π, or delocalized π.

H₂

File	Subshell
HOMO	
LUMO	

C₂H₄

File	Subshell
HOMO-2	
HOMO-1	
HOMO	

C₂H₂

File	Subshell
HOMO-2	
HOMO-1	
HOMO	

CO₃²⁻

File	Subshell
HOMO-5	

2   Kinetics of Substitution Reactions

Contributed by James Currie and Kevin Johnson, Pacific University

Overview

In this experiment, you use computer modeling to look at examples of substitution reactions involving halogen atoms on a central carbon atom. You model a series of three substitution reactions where either the entering halogen (E) or the leaving halogen (L) is held constant and the other halogen is varied to be Cl, Br, or I.

Recommended Exercises

Complete Substitution Kinetics Pre-Lab Report before beginning the experiment.

Background

A substitution reaction is simply a replacement reaction in which one atom or group of atoms replaces another atom or group of atoms. A substitution reaction can occur several ways, but in this activity we will be
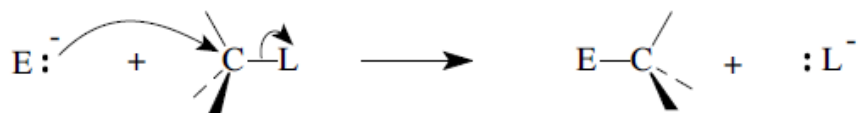
looking at a nucleophilic substitution.

In a nucleophilic substitution reaction, a substance capable of donating a pair of electrons (called the nucleophile) forms a bond to an atom at the expense of a bond to an existing atom or group. We depict a nucleophilic reaction using arrows to show the movement of electrons during bond formation or breakage.

As with any reaction, there is a change in enthalpy associated with a substitution reaction. The two
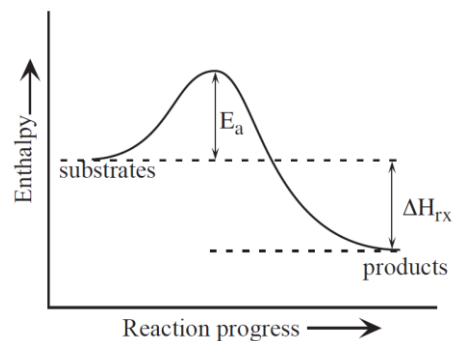
substances must collide with sufficient energy to pass over a barrier known as the activation energy (Eₐ). The structure at the highest energy point on the reaction pathway is the activated complex or transition state. The energy diagram (right) shows a reaction coordinate for a simple single step reaction such as a nucleophilic substitution.
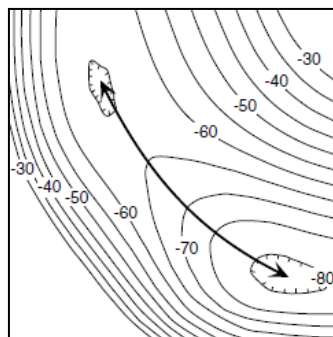
In this activity, we are going to use computer modeling to find the lowest energy pathway that the molecules take as they move from reactants to products. Using an energy map to find the lowest energy pathway is like using a topographic map to find the easiest route across a mountain range. The energy map will be a contour map with enthalpy contours instead of elevation. Converting the route from 3D contour map to a 2D pathway showing enthalpy as a function of reaction progress (time) gives us the reaction pathway usually found in textbooks. We will be able to look at the enthalpy of the starting molecules and products to determine the heat of reaction. We should be able to find the transition state (the highest energy point on the reaction path) and the activation energy.

Activation energies are a principal factor in determining a rate of reaction.

The Arrhenius equation relates the rate constant (k) to the activation energy (E_a), temperature (T), the gas constant R and a constant specific for the reaction (A).

$$k = Ae^{-E_a/(RT)}$$

If you know the activation energy, you can find the rate constant using the Arrhenius equation. For two similar reactions (where A is the same for both), the relative rates are approximately given by:

$$k_1/k_2 = e^{(E_{a_1}-E_{a_2})/(RT)}$$

Studying relative rates allows us to understand why some reactions are faster than others. In this experiment, you can use the above equations to compare the relative rates of the different halogen substitution reactions.

Modeling Section

Your instructor will assign you either a constant entering halogen (E) or a constant leaving halogen (L). You should model three different reactions holding you assigned group constant and varying the other group with Cl, Br and I.

1. Open SCIGRESS.

2. Select the Drawing Pencil tool.

3. Select a carbon atom from the Element type menu.

4. Click once in the workspace to deposit a single carbon atom.

5. Select **Action | Beautify | Comprehensive** with the Select tool. This converts the single carbon atom to a properly built methane molecule (Fig. A).

6. (Optional) Turn on the atom labels.

   ❏ Select **View | Styles | Atoms...** In the Atom Attributes dialog box, click on the **Label** tab. Select **Atomic Symbol**.

7. Add the leaving halogen (L) to the molecule.

   ❏ Select the hydrogen screen right of the carbon atom using the Select tool. Go to the Element menu and click on **Periodic Table** and double click on the leaving halogen atom. The hydrogen atom should change color to represent the halogen (green for Cl, red for Br, or violet for I) (Fig. B).

8. Return to the Select tool and select the entire molecule.

9. Select **Action | Beautify | Comprehensive** (Fig. C).

10. Go to the Periodic Table in the Element menu.

11. Select the second halogen atom (entering halogen E) and change the charge to -1.

    ❏ Select the halogen you need. Click **OK**. Change the charge to -1 by clicking on the down arrow in the Charge box.

12. Click on the carbon and hold down the mouse button as you draw away from the carbon in the opposite direction to the first halogen atom (Fig.D). **DO NOT BEAUTIFY THE MOLECULE AGAIN**.
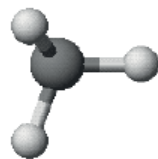
Figure A

Figure B

Figure C

entering halogen   leaving halogen
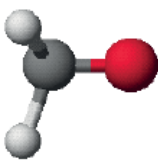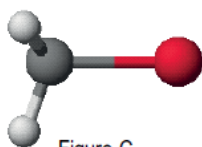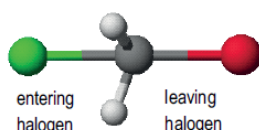
Figure D

Building an Energy Map

You now need to set up a search that varies the two halogen distances and computes the energy at each point. The search label produces a three dimensional equivalent to a contour map (see example at right). In this case the contours represent enthalpy.

The resulting map is called a potential energy surface. The reaction pathway is the path through the valley over the lowest possible energy barrier and on to the products. The search of bond distances is intended to find this lowest energy pathway from reactants to products. For a **C-Cl** bond a reasonable search would be between **1.55** and **3.00Å** in **40 steps**.

31
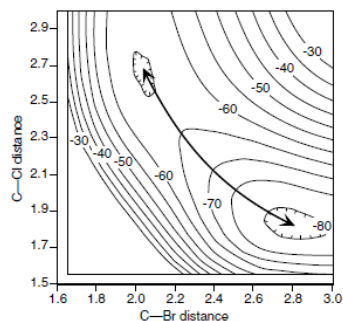
For the larger **Br** or **I** atoms, we should search between **1.60** and **3.00Å** in **40 steps**.

To setup a search label
1. Click on one of the halogen atoms and while holding the Shift key down, click on the carbon atom. When selecting the atoms ALWAYS select the halogen atom first.
2. Select **Action | Geometry Labels | Atom Distance**. In the Atom Distance box that appears, click **Define Geometry Label | Search** and enter the appropriate distances and steps.
3. Repeat this process for the other halogen atom.
4. Save this molecule in the appropriate location using your name as the file name. Since you have built this molecule with too many bonds on the carbon, you get a warning message, but ignore this warning and continue.

To produce the potential energy surface

1. Select **Experiment | Run...** The Run Procedure dialog box opens.
2. Select **conformations** in the category menu. Select **potential energy map** in the property menu, and use the exhaustive search with MO-G PM3 (2 labels) from the procedure menu.
3. Click **Start**. The calculation begins in a separate window. It takes less than a minute to complete the calculations. When the calculations complete, two windows open: the energy map on the left and the molecule on the right. The initial view displays a wide range of energies, so you need to narrow the focus.

To view the energies

1. Rotate and examine the potential energy map and the molecule (click in the appropriate window to make it active).

   ❏ Select the Rotate tool icon. Click and rotate the map.

2. Click on the potential energy surface at various points and note the positions of the atoms that correspond to the point.
3. Click at the minimum point on one side of the transition state. To determine if it is the starting or ending point, look at the molecule and note which carbon-halogen bond is shorter. If C-L is shorter, it is the starting point. If C-E is the shorter bond, it is the ending point.
4. Record the energy and bond distance for each carbon-halogen bond in the starting, transition and ending state.

   For the bond distance, labels on each carbon-halogen bond tells you its length. The shortest carbon-halogen bond length (C-L or C-E) is the "bonded" halogen at that particular point.

   ❏ The energy appears at the bottom of the molecule window:

   ✎ Why is the transition state sometimes called the saddle point?

5. (Optional) To enlarge the molecule for easier reading of the labels, use the zoom tool.

function.   ❏ Click on the Zoom tool icon.

6. Repeat this process with the remaining two reactions keeping the assigned halogen constant.

Report Guidelines

Fill out the lab report worksheet and hand in to your instructor.

Instructor Notes

We do this modeling experiment during second semester general chemistry
after completing the text chapter on kinetics. Students should be reminded
that a mechanism is determined experimentally but modeling is useful in formulating a hypothesis. We do a wet kinetics lab on the bromination of acetone in the next lab session. This exercise assumes that the mechanism is S_N2.

A helpful analogy is to relate the abstract concept of a reaction coordinate to a topographical map and its relationship to observed terrain. We introduce the experiment with slides of actual topographical maps showing mountain passes and juxtaposing the topographic maps with reaction coordinates. We also stop students during their work to discuss what they are seeing on the computer model. The ability to rotate the 3D potential energy surfaces sideways to make them look like a 2D reaction coordinate is also useful.

We assign each student either a constant Cl, Br, I nucleophile (E) or a constant Cl, Br, I leaving group (L). With their assigned group constant, the students model three reactions using Cl, Br and I as the variable group. There are six possible reaction sets. For example, one possible set is a constant E=Cl, where the variables are L=Cl, Br, I.

It is very important that students follow the directions for building their molecules. This type of optimized search is prone to failure if certain atom positions approach a straight line. We prevent this problem by building the molecule in a defined order and by using the 'maintain orientation' option. This option creates dummy atoms as reference points which are included in the numbering. To ensure that the settings are correct, the optimized search experiment can be duplicated and then modified to have specific settings. Before the lab session, you should make this experiment option and name it 'exhaustive search with MO-G PM3 (2 labels)'.

We suggest using 40 points for each search label because it gives a smoother map. The activation energies for the I/Cl and I/Br reactions are

very small and difficult to see well with too few points. A 41x41 search only takes less than a minute on a newer PC. Students might want to expand the potential energy map even further than suggested both by changing map axes and by scaling the window so they can see the energy change more clearly.

Typical Results

The following represents data that is relatively easy to obtain for students (E=entering halogen, L= leaving halogen).

Data for Methyl Halides

E	L	ΔH_f(sm)	ΔH_f(ts)	ΔH_f(prod)	$\Delta H^{\ddagger}$	C — E(Å)	C — L(Å)
Cl	Cl	-74.80	-64.70	-74.80	9.90	2.23	2.20
Br	Br	-72.60	-69.50	-72.60	3.10	2.37	2.37
I	I	-63.79	-58.40	-63.79	5.20	2.47	2.47
Cl	Br	-65.35	-64.62	-81.07	0.73	2.28	2.11
Br	I	-62.26	-61.36	-74.47	0.68	2.30	2.23
Cl	I	-55.01	-54.90	-84.02	0.11	2.20	2.30

Further Discussion

Each instructor will need to consider whether to encourage discussion about some of the more complex issues.

An interesting topic for more advanced discussions is he effect steric hindrance has on molecules. What if the starting halide was chloroethane instead of chloromethane? Would you expect a higher or lower E_a? What about for 2-chloropropane? Pictures of each substance are shown below.
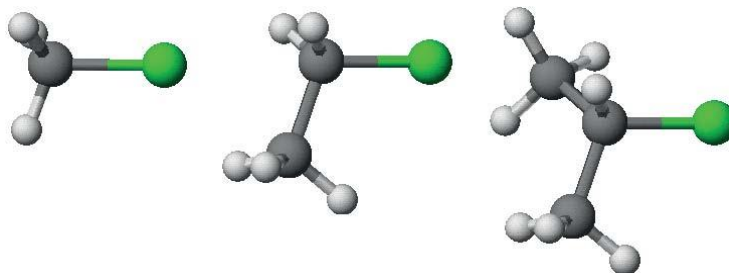
Just as in the real reactions, the reaction models are a bit tricky as the substitution increases. Blank spaces in the table represent modeling problems that do not work well as one large potential energy surface calculation. For these, the regions near the transition state and energy wells need to be done separately to avoid "side reactions."

Simulations modeled in a solvent give more interesting and true to reality surfaces but at the expense of much longer computation times. Students in general chemistry usually do not have the experience to appreciate the subtle difference.

Data for Ethyl halides

E	L	ΔHf(sm)	ΔHf(ts)	ΔHf(prod)	ΔH‡	C— E(Å)	C--L(Å)
Cl	Cl						
Br	Br	-80.83	-78.04	-81.17	3.0	2 .46	2 .39
I	I	-66.23	-63.20	-66.23	3.0	2 .66	2 .60
Cl	Br	-78.06	-72.74	-86.61	5.4	2 .42	2 .26
Br	I						
Cl	I	-64.70	-61.27	-86.73	3.4	2 .26	2 .49

Data for 2-Propyl halides

E	L	ΔHf(sm)	ΔHf(ts)	ΔHf(prod)	ΔH‡	C— E(Å)	C— L(Å)
Cl	Cl						
Br	Br						
I	I						
Cl	Br	-89.07	-80.02	-94.13	9.0	2 .42	2 .33
Br	I						
Cl	I						

Substitution Kinetics Report

Name: _____

Fill in the data below using the potential energy surfaces that you computed using SCIGRESS. From the data, calculate the predicted heat of reaction for each of the processes. For comparison calculate the heats of reaction using the bond enthalpy data in the text.

E=entering halogen L=leaving halogen

E	L	ΔHf reactants kcal/mol	ΔHf transition state	ΔHproducts kcal/mol	Ea forward kcal	Ea reverse kcal	ΔHrx (Scigress) kcal/mol	ΔHrx (text) kcal/mol

1. Considering the energy map, which reactions are endothermic and which are exothermic?

Plot the above energies on the graph below.

2. Fill out the table below. Considering the C—E and C—L distance, and the position of the transition state on the surface relative to the reactants and products, do the transition states look more like the reactants or products?

E	L	C--E distance	C--L distance	Reactants or Products?

3. What are the relative rates constants for the three reaction you calculated? Let the slowest be 1.0. Calculate these using the Arrhenius equation. Assume the prefactor remains constant and T = 298K, and use R=1.987 cal/mol-K.

E	L	Relative Rate
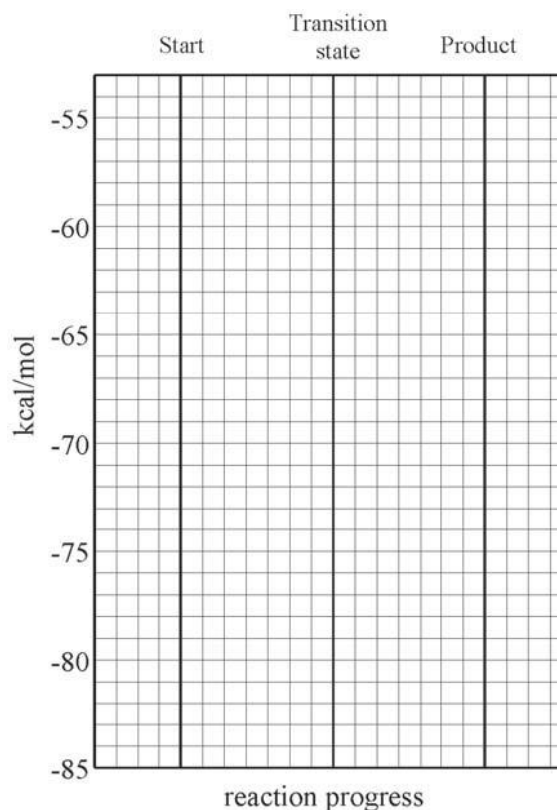
Substitution Kinetics Pre-Lab Report

Name: _____

The chemical reaction below was studied using SCIGRESS.

The following energies were calculated:

starting materials:-78.1 kcal/mol
transition state:-72.7 kcal/mol
products-86.6 kcal/mol

These values were plotted on the graph below:
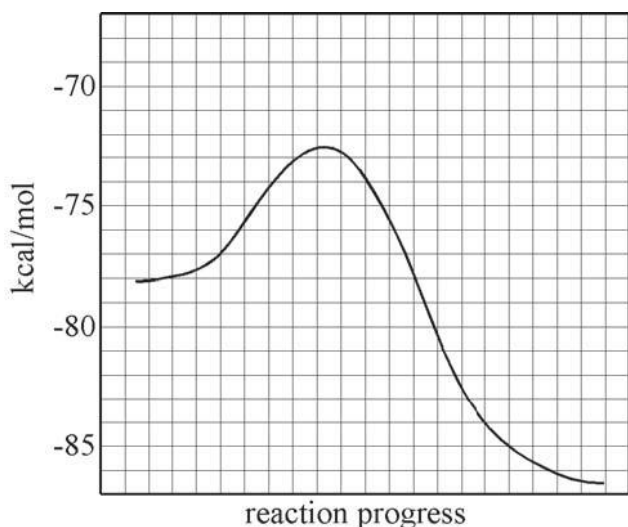
What is the heat of reaction for the forward process?

_____ kcal/mol

Reverse process? _____kcal/mol

What is activation energy (E_a) for the forward process?_____kcal/mol

Reverse process? _____kcal/mol

Now assume the activation energy was 8.0 kcal/mol, and the starting and ending energies were the same as above. Draw the resulting reaction coordinate on the graph below.
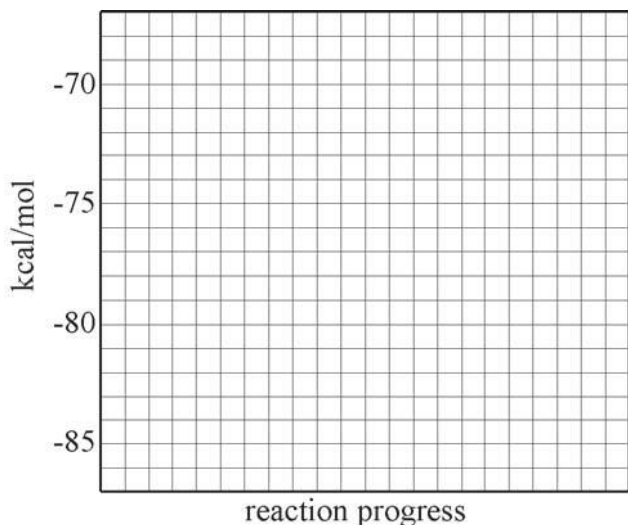
What would be the energy of the transition state?
_____kcal/mol

Would the reaction rate be faster or slower than the previous reaction?_____

3 Sunscreen and Ultraviolet Absorption

Contributed by Crispin Wong and James Currie, Pacific University

Overview

Recently, a cosmetic company hired you to help them develop a sunscreen mixture for a new line of products. To accomplish this task, you will model molecules and the ultraviolet (UV) spectra of sunscreen chemicals with SCIGRESS. Using the spectra you generated, you will design a sunscreen. In the third part of the experiment, you will research other commercial sunscreens and analyze them using a UV spectrophotometer.

Recommended Exercises

Complete Exercises I, II, III, IV, V, VI, and VII before beginning this experiment.

Background

Every day we are exposed to the sun's harmful ultraviolet rays. Awareness about the damage sun exposure causes has increased over the years and a push for sun protection has emerged. One such safety precaution is the use of sunscreen. Most sunscreens utilize chemicals that have chains of alternating single and double bonds resulting in delocalized π bonds. These are called conjugated systems and they are responsible for absorbing harmful ultraviolet light.

Figure 1: The conjugated system in oxybenzone, a common sunscreen chemical, helps absorb ultraviolet light.

Ultraviolet light is subdivided into three different ranges: UVA (320-400 nm), UVB (290-320 nm) and UVC (100 - 290 nm). UVC light is harmful, but most of it is absorbed by atmospheric gases and the ozone layer before reaching the earth's surface. Figure 2 shows the absorptions of atmospheric gases and ozone in the UVC range. Continued depletion of the ozone layer could eventually increase UVC radiation.
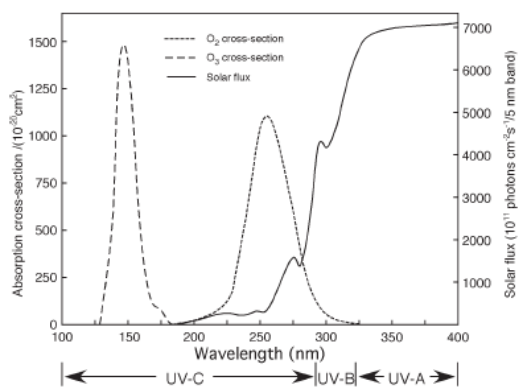
Figure 2: UVC radiation is absorbed mostly by atmospheric gases and ozone. Most of the UVB and UVA radiation reach the earth and our skin, relatively unfiltered.

UVA and UVB wavelengths can penetrate the atmosphere; we try to avoid exposure to these rays. UVA radiation causes immediate redness on skin that disappears after a few hours. It acts through oxygen free radicals and may be linked to skin cancer. UVB is responsible for the more detrimental delayed sunburn and/or tanning. Tanning is a result of UVB light stimulating melanocytes in your

skin which produce the dark colored pigment, melanin. UVB radiation acts directly on biological molecules, causing mutations that are potentially cancerous.

A Sun Protection Factor (SPF) quantifies sunscreen efficacy against UV rays.

$$\text{SPF value} = \text{MED}_{\text{(protected skin)}}/\text{MED}_{\text{(unprotected skin)}}$$

The minimal erythema dose (MED) is the minimum dose of UV energy that produces lasting redness on skin. Protected skin is defined as 2 mg/cm^2 of sunscreen covering the skin. This could equate to as much as 1 fluid oz. over your whole body.

SPF was intended to measure a sunscreen's ability to block UVB light, which is the primary cause of delayed sunburns. The more noticeable effects of UVB mask the effects of UVA rays, and so SPF does not represent UVA blockage well. Researchers are developing a system to measure UVA protection, such as the protection factor A (PFA). PFA is based on the minimal response dose (MRD), which is the smallest UVA dose that produces redness or tanning on skin.

Part I: Model Sunscreen Chemicals

The cosmetic company Make-Me-Pretty is developing a line of products that touts its ability to block UVA and UVB light. They have hired you to lead the research and development for these new products. The product must contain transparent chemicals; physical blockers such as zinc oxide and titanium dioxide will not be considered.

📖 NOTE

Physical blockers like zinc oxide and titanium dioxide are opaque compounds that reflect all wavelengths of light.

To make this task efficient and successful, it is essential to keep your budget to a minimum. One relatively inexpensive tool that aids in developing hypotheses and plans is computer modeling. Modeling allows you to compile and compare information with ease. In this situation, you can model the UV spectra for various sunscreen chemicals. Before you order expensive chemicals to test in the laboratory, you can use the modeled UV spectra to decide which chemicals you want to include in the sunscreen mixture for the cosmetic products.

As time allows, you may also want to research the available sunscreen chemicals. Information such as how chemicals affect the skin or how much of a chemical the FDA allows in a product, may be pertinent to your task.

In this first section, you will model sunscreen chemicals using SCIGRESS. Consult your instructor to find out which chemicals are available for this experiment and which ones you will model on the computer.

1. Build a molecule for each of the assigned sunscreen chemicals in the Workspace.
2. Save the molecules and close the molecules in the Workspace.
3. In Project Explorer right click on your project name and select New object. A New Object window appears. Choose Spreadsheet from the Object Type list. Spreadsheet opens, a table appears on the screen with the first column labeled "chemical sample".
4. Drag and drop a chemical sample to the "chemical sample" column.
   ❏ You can do this one chemical sample at a time or select all the chemical samples you need and drag them into the "chemical sample" column.

5. Repeat step 4 until you have entered all of the molecules in the chemical sample column.
6. Double click in the second column heading cell.
   ❏ An Procedure Browser dialog box opens. In the first column click Property of. In the Category column choose **UV/vis spectra.** In the third column highlight **excitation energy.** In the procedure column highlight **standard procedure.**
7. Click **OK** to close the dialog box. Columns B and C say "UV/vis all" and "excitation energy (cm-1)" respectively.
8. Highlight the cells in the second and third columns and select **Evaluate | Cells**. It may take a while for the calculations to complete. When the calculations finish, wavelengths appear in column two and the corresponding energies in column three.
9. Save the Spreadsheet in a designated location.
10. Open the molecule and UV spectra.
    ❏ Return to the Workspace. Open one of the sunscreen molecules. The molecule should appear in one window, and the molecule's UV spectrum should appear in a UV-visible Transitions window. If the UV spectrum does not appear select | **UV-visible Transitions**.
11. Adjust the scale of the UV/vis graph. The range you want to look at for UVA/UVB absorption is between 250-400 nm.
    ❏ Double click on the x-axis or the y-axis. A View Axis dialog box appears. Change the Wavelength Range (x-axis) to go from250 to 400 nm. Adjust the Intensity Range (y-axis) to best represent the data. Click **OK**.
12. Examine the UV spectrum.
    ✎ At what wavelength does the chemical have a maximum absorbance (λ max)?
    ✎ Is this in the UVA or UVB range?
13. Open another sunscreen molecule and its UV spectrum as you did in step 10. The UV spectrum appears in the same grid as the UV spectrum of the other open molecule. Examine the new UV spectrum, and compare both spectra in the window. You may need to readjust the y-axis.

    If you continue to open the UV spectra of different molecules, you can view their spectra simultaneously. To remove a spectrum from the grid, close the corresponding molecule window.

Part II: Design a Sunscreen

Now that you have modeled the UV spectra of compounds that appear in sunscreens, think about the wavelengths of light that they absorb. There are compounds that are better at blocking UVA, and compounds that block UVB more. A good sunscreen will block both types of ultraviolet light.

Your next assignment is to design a sunscreen. Your employer, Make-Me-Pretty, has asked you to find a combination of three or less chemicals that can accomplish the goal of blocking both types of ultraviolet light. In addition to recommending a combination of chemicals, you also need to recommend a ratio for the chemicals you choose.

Use the UV spectra that you modeled in Part I to look at chemical combinations

that will cover the UVA/UVB range. Notice that some chemicals have stronger absorbance than others. Take this into account when you are finding a recommended ratio.

1.   Find a combination of chemicals that cover the ultraviolet wavelengths using the UV spectra displayed in the Visualizer. Note the maximum absorbance of each chemical and predict the absorbance curve of the mixture.

2.   Verify the chemical combination in the lab using the UV Spectrophotometer. Stock solutions of the sunscreen chemicals (10 mg in 100 mL of absolute ethanol) have been prepared for you. Your instructor will show you how to use the UV Spectrophotometer in your lab.

3.   Use a micropipet or calibrate a dropper to measure out small amounts of the chemicals into a 5 mL volumetric flask. Because of the varying UV absorbance of each chemical, the amount of chemicals you put in should add up to anywhere between 0.5 mL and 2.0 mL. For example, if you want to mix chemical A (strong absorber) and chemical B (weak absorber) in a 1:2 ratio, 0.5 mL of chemical A and 1.0 mL of chemical B (1.5 mL total chemicals) in 5 mL ethanol will probably be sufficient.

4.   Fill the 5 mL volumetric flask to the fill line with absolute ethanol and swirl the mixture.

5.   Run the sample in the UV spectrophotometer. Be sure that you zero the machine before you take the sample. Also, rinse out the cuvet with ethanol before inserting the sample to avoid contamination.

6.   If the absorbance is too high when you take a reading with the UV spectophotometer, decrease the amount of chemicals. If the spectrum is weak, increase the amount of chemicals in the mixture. In either case, remix the solution from scratch. Do not add more ethanol or chemicals to the existing solution because you want the chemical proportions to remain intact.

It may take a few tries to get the concentration right. Consult your instructor if you have problems finding a good concentration for your UV spectrophotometer.

Part III: Compare Commercial Sunscreens

There are many sunscreens on the market today. Your job now is to compare the UV spectra of different sunscreens--potential competition for your company's products. You can examine a number of aspects related to UV absorption, SPF, and different chemicals. Below is a list of possible options to explore. The choices are not limited to this list, so you can be creative and test new ideas.
   ❏ Different SPF sunscreens manufactured by the same company
   ❏ The same SPF sunscreens manufactured by different companies.
   ❏ Different SPF sunscreens manufactured by different companies.
   ❏ Different chemical combinations in same SPF sunscreens.
   ❏ Tanning lotions vs. sunscreen lotions.
   ❏ Aged sunscreens vs. new sunscreens.

Make the extracts for the sunscreens you want to compare. Run the spectra on a UV spectrophotometer. To make comparisons easier you might want to print out the UV absorbances on the same graph.

Make Sunscreen Extracts

1. Weigh out 0.10 g of sunscreen into a 250 mL beaker.
2. Measure 100 mL of water in a graduated cylinder and add it to the beaker.
3. Stir the sunscreen/water mixture until it is a uniform white suspension. Make sure that sunscreen is not sticking to the side and bottom of the beaker. If you can, use a magnetic stir bar to mix the solution up for 5-10 min. or until it is a uniform white.
4. Pipet 1.0 mL of the sunscreen mixture into a 50 mL beaker.
5. Measure out 9.0 mL of absolute ethanol and add it to the 50 mL beaker.
6. Swirl the ethanol/sunscreen mix for about a minute.
7. Run a UV analysis on the sample. If it is too concentrated, dilute the sample with more ethanol. Make sure you keep track of the dilution factor.

It is important to take into account the dilution factor when comparing UV spectra. If they do not have the same dilution constant, you need to correct for concentration to depict the true absorbance relationship between the sunscreens.

Report Guidelines

Consult your instructor about the report requirements.

Instructor Notes

In second semester general chemistry, we dedicate the last three lab sessions to unique group projects. The sunscreen lab is one possible project. Students are able to exercise a greater degree of autonomy in researching an aspect of chemistry.
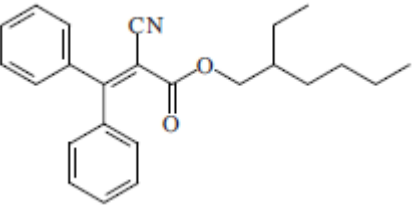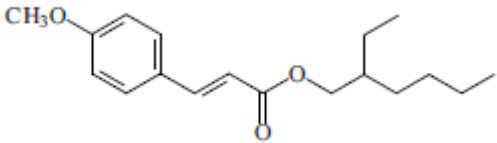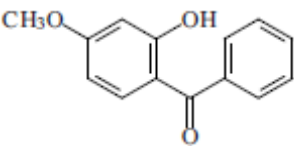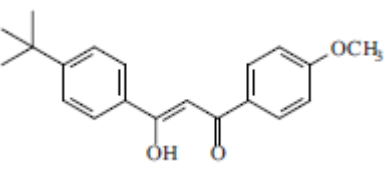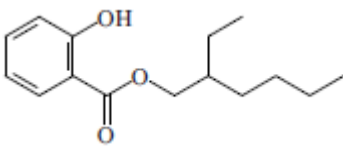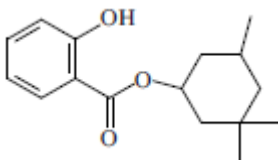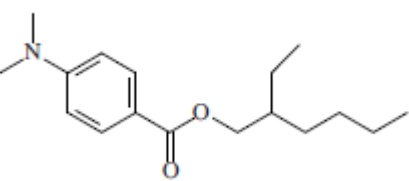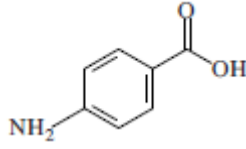
This experiment has purposely been designed to be flexible; you can choose which parts you want to highlight in your lab. It could be a more complex project, where students assemble a report complete with research and experimental results. On the opposite end, the lab could be pared down to a simple computer exercise generating UV spectra. This lab is designed for flexibility but in any case, let the students know which parts they should and should not do.

Since this experiment is an independent group project, our students research sunscreen chemicals and find the corresponding structures. However, this may not be conducive to large labs or a short period of time. Structures of common sunscreen chemicals are included on page p 3-34. You can hand out this sheet to aid students in their modeling. There are a number of ways to go about modeling the structures. Depending on what chemicals you can obtain, the students may model only those chemicals. Each student can build one structure and share with others, or they can build their own set.

We had difficulty modeling octocrilene with the suggested MM/INDO1 parameter. It works best to optimize octocrilene in MOPAC followed by a UV calculation in ZINDO. You may want to create a new experiment where the first calculation is a MOPAC optimization followed by a UV visible transitions calculation using ZINDO at current geometry.

Parts II and III assume students have access to an UV spectrophotometer. In some cases, allowing a large amount of students to use the machine is not feasible. Alternative solutions may range from grouping students together to use the UV spec, to running different spectra, then copying and handing them out to students.

At the end of the project, the students create posters clearly displaying information they collected. They are also required to write a detailed report, covering their results and conclusions.

Structure	Name (s)	CAS number
	Octocrilene, 2-ethylhexyl 2-cyano-3,3-diphenyl-2-propenoate	6197-30-4
	Octyl methoxycinnamate, 2-Ethylhexyl (2E)-3-(4-methoxyphenyl) prop-2enoate	5466-77-3
	Oxybenzone, Benzophenone 3, 2-Hydroxy-4-meth-oxybenzophenone	131-57-7
	Avobenzone, Parsol 1789, 4-tert-Butyl-4'-methoxydibenzoylmethane	70356-09-1
	Octyl salicylate, 2-Ethyl hexyl salicylate	118-60-5
	Homosalate, 3,3,5-trimethylcyclohexyl salicylate	118-56-9
	Padimate O, Octyl Dime-thyl-PABA, Octyl Dimeth-ylaminobenzoate, 2Ethylhexyl 4-dimethylami-nobenzoate	21245-02-3
	PABA, p-aminobenzoic acid (not used anymore)	150-13-0

Part 2: Organic Chemistry

Evaluations of Conformations of Menthone and Isomenthone

Contributed by James Currie, Pacific University

Overview

In the laboratory, you will determine the experimental percent composition of a menthone and isomenthone equilibrium mixture. In the second part of this experiment, you will use SCIGRESS to evaluate the relative stability of the best conformation for each of the two geometric isomers.

Recommended Exercises

Complete Exercises I, II, III, VI, and VI before beginning this experiment.

Background

Menthone and isomenthone are in equilibrium under acidic conditions. Since there are two chiral centers and only one is changing during the equilibrium, menthone and isomenthone are diasteromers of each other. The change in chirality changes the *cis, trans* geometry. A careful look at the structures shows that menthone is *trans* and isomenthone is *cis*.

Although the double bond distorts the chair form of the cyclohexane ring, groups are still in positions that approximate axial and equatorial. For the *cis* isomenthone, one group must be axial and the other equatorial. The isopropyl or the methyl group can assume either position but the ring flips to permit the larger isopropyl group to always be equatorial. As a result, the isopropyl position is equilibrating but the net effect is the methyl shifting from the equatorial position in menthone to the axial position in isomenthone.

(-)-menthone	**(+)-isomenthone**
Optical rotation = -30°	Optical rotation = +92°
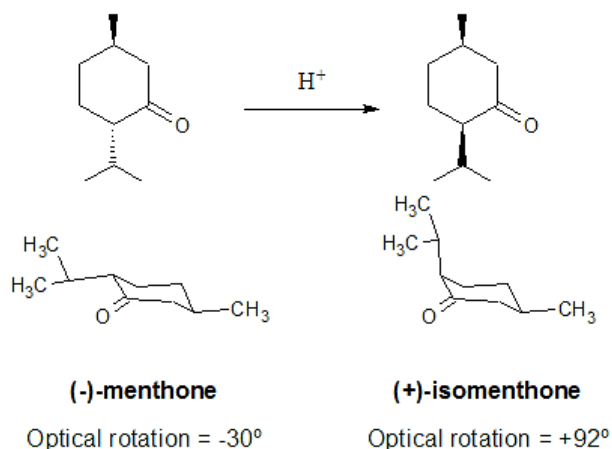
Since the molecules are chiral we can determine the equilibrium composition by using a polarimeter to measure the optical rotation of the mixture. The specific rotation is calculated from the observed rotation using Equation 1. Because we know the specific rotations of the pure isomers we can determine the isomer ratio from the observed rotation of the mixture. You can use Equation 2 to find the mole fraction of menthone in the mixture.

-30°   +92°
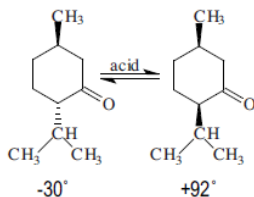
$$[\alpha] = \frac{\alpha}{l \cdot c}$$

where l=length (dm) and c= conc. (g/mL) in 10 mL

Equation 1

$$[\alpha] = [\alpha_A] \cdot X_A + [\alpha_B](1 - X_A)$$

where α_A and α_B = specific rotations of the pure isomers and x_A = mole fraction of A

Equation 2

Laboratory Section

For additional details about the procedure, hazards, and waste disposal, consult the text.

Mix 3 mL of (-)-menthone with 5.0 mL of glacial acetic acid and 5.0 mL of 1M HCL. Reflux for 30 minutes and then let the mixture cool to room temperature. Bring the pH of the solution to 10 using 3M NaOH. Extract the mixture with two portions of diethyl ether. Dry the ether extracts with magnesium sulfate and evaporate the ether in a pre-weighed vial.

To measure the optical rotation, mix all of the product in 10 mL absolute ethanol for a 1 dm tube or 25 mL for a 2 dm tube. Take the optical rotation reading of the product and a blank. Find the specific rotation of the mixture and the mole fraction of menthone.

Modeling Section

1. Build menthone and isomenthone in their best conformations. You may want to start with the cyclohexane fragment. Be sure the hybridization of each atom is correct.
2. (Optional) View the R or S designation on the chiral centers. Menthone is R, S or S, R and isomenthone is R, R or S, S.
   ❏ Select **View | Style | Atoms...**. The Atom Attributes box opens. Click on the **Label** tab and select **Chirality for Asymmetric Carbons.**
3. Save each molecule you build as menthone or isomenthone.
4. Set up a dihedral search label between -180 and +180 using 48 steps for the H-C-C-H angle. Select the isopropyl H first (see diagram to the right). This will find the optimal position for the isopropyl group.
5. Compute an optimized conformation search.
   ❏ Select the experiment settings: **conformations**, **potential energy map** and **exhaustive search with MM2.**
6. View the resulting map. There are two windows displayed: the map in the left window and the molecule in the right.

-168.39 S
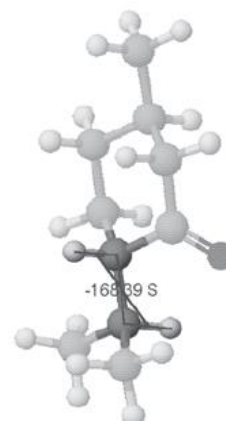
H-C-C-H Dihedral Angle.

📖 **TIP**

If you cannot find the map file, look for file types that fend in ".map," not ".csf." They are labeled as energy.map in the filename.map folder.
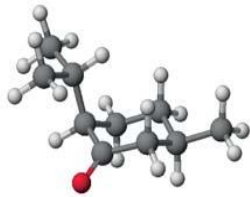
❏ The map opens in the workspace after the computations are complete.

7. Click on the graph in the map window to change dihedral angle and check corresponding energy value. Note the lowest energy position for the isopropyl group.

8. Save the lowest energy conformation from the menthone map as "last name.menthone" (e.g. Smith.menthone), and the lowest energy conformation from the isomenthone map as "last name.isomenthone" (e.g. Smith.isomenthone). (**File | Save Current Structure As…**)

9. Click in the molecule window and rotate the molecule. Compare the amount of eclipsing or direct atom contact as the group rotates.

10. View the molecule with the full van der Waals radius. Notice that in the higher energy conformations the hydrogens appear to be touching.

   ❏ Click on **View | Atoms & Bonds | Space Filling**.

11. Return to the ball and stick mode.

   ❏ Click on **View | Atoms & Bonds | Ball and Cylinder**.

12. Close the map file.

Optimize the Structures

13. Open each molecule, one at a time, in the Workspace.

14. Perform a geometry optimization.
   ❏ Optimize geometry using MM (MM2).

15. Note the energies and the distribution of steric strain after each Mechanics calculation.

16. Copy the table from Mechanics into your report. Typical results from a molecular mechanics calculations on one conformation of isomenthone are shown below. Before copying, hit the "i" key to return them to the proper scale. Shrink the window down to the size of the molecule and orient the molecules in positions that are appropriate for seeing the structure clearly.

Chair isomenthone (axial i-pr)	
Convergence reached, program stopped normally. stretch = 0.751 angle = 1.471 stretch bend = 0.196 dihedral = 7.005 improp torsion = 0.012 van der Waals = 4.459 electrostatics = 0.000 hydrogen bond = 0.000 The energy of the final structure is 13.8935 kcal/mol. Molecular Mechanics optimized structure can be found in chair-isomenthone.a-ipr.	

Report Guidelines

Report the actual yield, theoretical yield and percent yield. Also report the specific rotation of the product. Copy and paste the final structures into your word processor document. Also include a discussion of the following topics:

✎ Locate the major sources of strain in each structure. In this case dihedral and van der Waals strain are likely to be the most important.

✎ From the modeling results, predict which of the two structures is most favored, by how much energy (kcal/mol), and briefly discuss the reasons for the order.

TIP

Hit i at any time to return the molecule to its initial position.

NOTE

A Molecular Mechanics calculation produces relative energies compared to a molecule of the same composition but with ideal bond angles, distances and through space interactions. Thus the results are the net strain in the molecule.

✎ Calculate the actual percentage of menthone and isomenthone based on the experimental specific rotation of the equilibrium mixture.

✎ Compare your experimental results with those obtained through a Mechanics calculation. Which is predicted and which is actually favored?

Instructor's Notes

We use the menthone/isomenthone equilibrium experiment as the first modeling project in organic chemistry. We schedule this project about one third of the way through the first semester. This gives us time to assign a tutorial so we are sure that everyone is adequately prepared in the fundamental modeling tasks. At the same time, if SCIGRESS site license is available, students may wish to install the software on their personal computers. Since this is our first modeling project, we usually take some class time to run through a sample calculation for a similar molecule using a projection system in the lab.

The laboratory portion takes about an hour and a half to isolate the product. Since we have only one polarimeter, students come back on their own time to do the specific rotation measurement. Alternatives such as pooling the product for a group or entire class measurement are reasonable for departments where independent work is not possible. For more details on the laboratory procedure refer to the references at the end of this experiment.

Typical Results

In the experimental portion, students typically obtain 1-2g of equilibrium mixture. When this is dissolved in a 10 mL volumetric flask it gives rotations of 1-2° which translates to ~60-65% menthone.

When you use Mechanics to search 48 dihedral angle points, there are two low energy valleys in the rotational potential energy diagram for (-)-menthone. The use of fewer points in the search might lead students to conclude that the lowest energy is the local minimum at about 180°. If the H-C-C-H dihedral angle of bond leading to the isopropyl group is chosen, the heat of formation local minima are 12.58 kcal at about 68° and 12.75 kcal at about 165°. Selecting the conformer at 68° and optimizing gives an energy of 12.56 kcal. Inverting the chiral center will produce the isomenthone isomer with a heat of formation of 13.88 kcal after optimization. Thus there is a difference of 1.32 kcal in favor of menthone.
We can make an estimate of the expected isomer ratio by making a few assumptions. If we assume that the change in entropy during isomerization is very small then $\Delta G_{rx} \approx \Delta E_{steric}$. Furthermore, we can relate the equilibrium constant (K) to our data (ΔE_{steric} = -1.32 kcal) via the equations:

$$\Delta G_{rx} = -RT \ln K \qquad\qquad \text{Equation 3}$$

$$\left(\Delta H_{rx} \approx -RT \ln \frac{[menthone]}{[isomenthone]}\right) \qquad\qquad \text{Equation 4}$$

Solving for the ratio of isomers gives a prediction of 1.4:1 or approximately 90% menthone in the final mixture. The class results of 60-65% menthone are in reasonable agreement.

If you want to take a bit more time, you can obtain slightly better results using MOPAC (PM3), as seen in the table below.

	Menthone Energy	Isomenthone Energy	ΔE	Predicted Menthone%
Mechanics	12.56	13.88	1.32	90%
MOPAC (PM3)	-78.94	-80.12	1.06	86%

References

Barry, J. *J. Chem. Ed.* **1973**, *50,* 292. Lehman, J.W. *Operational Organic Chemistry*, 3rd ed.; Prentice Hall: Upper Saddle River, NJ, **1999**; 136-142.

The Evelyn Effect

Contributed by James Currie, Pacific University

Overview

The Evelyn effect refers to the observation that the product composition changes during the course of a reaction. The change in composition suggests that a mechanism preference is changing over time; a change in mechanism conflicts with the basic principles of reaction chemistry. You will do a laboratory experiment to see if you can observe the Evelyn effect and then you will use modeling to explain your results.

Recommended Exercises

Complete Exercise I, II, III, IV, V, and VI before beginning this experiment.

Background

The dehydration of 2-methylcyclohexanol or 4-methylcyclohexanol is a common experiment in organic chemistry laboratories. Analysis of the distilled alkene products is useful in validating Zaitzev's rule and in deducing the mechanism as E1 or E2. Because of the presence of products that require rearrangement, most chemists assumed that an E1 mechanism was operative.
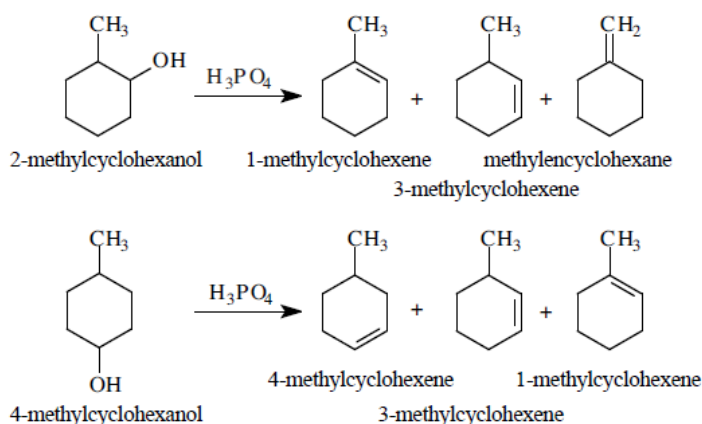
This assumption came under scrutiny in 1994 when Professor David Todd[1] was testing this experiment and interrupted the alkene collection midway to have lunch with the department secretary, Evelyn Jacoby. When he returned, he resumed the distillation but kept the second half products separate from the first. When he

---

[1] Todd, D. *J. Chem. Ed.* **1994**, *71*, 440.

analyzed the two fractions, he was surprised to find different alkene ratios. The reaction products and pathways to the products seemed to have changed over time. Dr. Todd called this phenomenon the "Evelyn effect."

The Evelyn effect goes against the fundamental principle in organic chemistry-- reactions always go by the lowest energy pathway. The favored product should remain so throughout a reaction run at constant conditions. In cases that have multiple paths due to competing mechanisms with close activation energies, we get complex product mixtures. However, even if multiple products are formed from multiple paths of similar energy, the product distribution should not change in constant conditions.

To put this principle in perspective, we can consider an analogy to a more familiar setting. Imagine that there are two mountain routes to a city on the other side of the range. One of these routes may have the lower elevation and therefore be easier to navigate. If these two routes are greatly different in altitude, all the traffic will want to travel via the easiest route. If they are close in altitude, we may expect some flow through the slightly higher route but it should not exceed the lower. The preference for the lower route will not change.

In the chemical world, the mountain pass altitude is analogous to the activation energy, the traffic to the concentration, the speed to the temperature, and the route to the reaction mechanism/pathway. The presence of the Evelyn effect appears to violate this basic principle of chemistry in the same way that a mountain pass changing its altitude would violate our common sense view of topography.

Reaction details

When either 2-methyl- or 4-methylcyclohexanol is heated in the presence of acid, the first step in the mechanism is protonation of the hydroxyl group followed by elimination. If the mechanism is concerted (E2), no rearrangements are possible and a simple product mixtures results. However, if the reaction follows an E1 mechanism, it forms a planar carbocation and rearrangements could lead to many additional products.
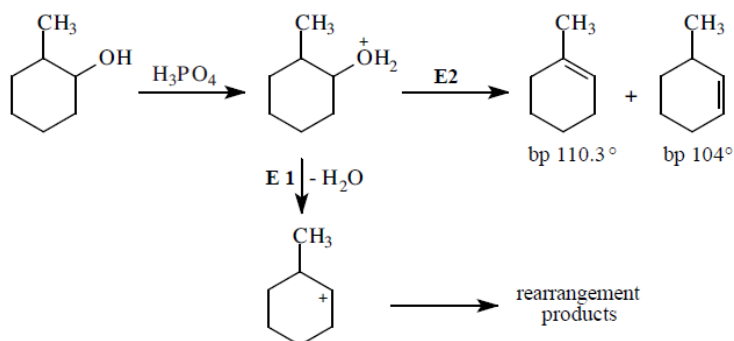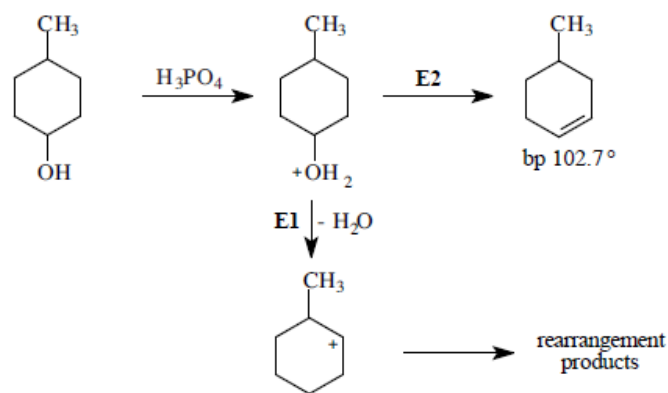
Possible rearrangement products:

| bp 110.3° | bp 104° | bp 102.7° | bp 103° |

| bp 97° | bp 113-17° | bp 106.3° | bp 97.8° | bp 98.3° |
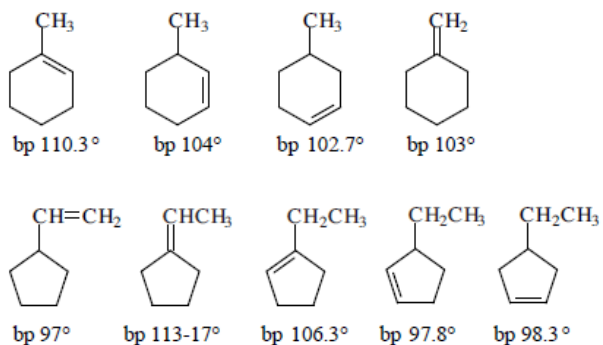
The alkene products are volatile compared to the methylcyclohexanols and we assume that once formed, they distill rapidly from the reaction mixture. This means that the reverse reaction, the addition of water to form isomeric methylcyclohexanols, does not play a role in the ultimate product mix. We can conclude that if rearranged products occur, they come from an intermediate carbocation, by way of an E1 mechanism. Because we assume there is no reverse reaction, we can focus on two issues that will help us understand the Evelyn effect: the presence or absence of rearrangements and the reaction's obedience to Zaitzev's rule.

We could conclude that rearrangement occurs during the dehydration of 2-methylcyclohexanol if we find evidence of methylenecyclohexane or 4-methylcyclohexene. Unfortunately a typical gas chromatograph does not separate these closely boiling compounds from one of the expected isomer, 3-methylcyclohexene. If we do the reaction with 4-methylcyclohexanol, we still cannot separate any rearranged 3-methylcyclohexene from the expected 4-methylcyclohexene, but we can easily see any doubly rearranged 1-methylcyclohexene. Usually the presence of any rearranged product is enough to conclude that a carbocation is involved and the mechanism is E1. As we shall see, such hasty judgments are not wise.

If Zaitsev's rule is followed in the deydration of 2-methylcyclohexanol, we should get a higher yield of the more substituted 1-methylcyclohexene. For an E1 mechanism, stereochemistry of the alcohol does not matter because the intermediate is a planar carbocation. However in an E2, the requirement for *anti*-elimination means that stereochemistry plays a major role. To illustrate this point we will look at the dehydrohalogenation of neomenthyl chloride and its

stereoisomer, menthyl chloride.

CH₃CH₂O⁻
CH₃CH₂OH

neomenthyl chloride   78%   22%

menthyl chloride   100%
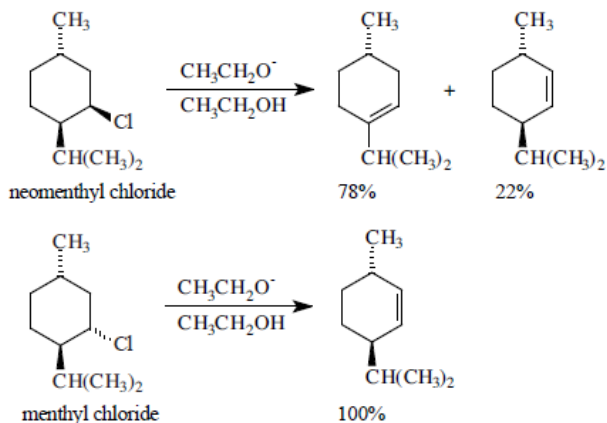
Neomenthyl chloride gives the expected preference for the more substituted alkene. Surprisingly, menthyl chloride gives 100% of the least substituted alkene, apparently violating Zaitzev's rule. To understand this we must look at the stereochemical requirements of the E2 mechanism and how they impact this reaction.

An E2 mechanism requires that the two groups being eliminated be *anti*-periplanar or 180° from each other. In a cyclohexane ring two adjacent atoms can only achieve 180° if they are both axial. When neomenthyl chloride is in its most stable conformation, the two alkyl groups are equatorial and the chlorine atom is axial. Two hydrogens are *anti*-periplanar leading to the two possible products. In menthyl chloride, placing the chlorine atom in the axial position forces the isopropyl group into the axial position. Only one hydrogen is *anti*-periplanar and the elimination of this hydrogen leads to the least substituted alkene. So while the menthyl chloride reaction violates Zaitzev's rule, it does so because stereochemical restrictions permit only one product.
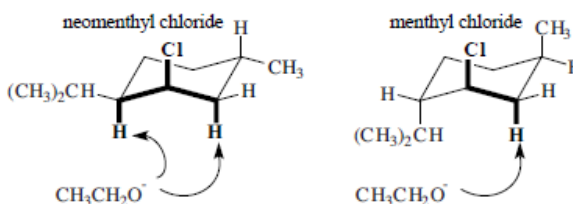
It is important to note that stereochemistry affects the outcome of this reaction only if the mechanism is E2. Dehydrohalogenations in strong base are always E2, as is the case with neomenthyl and menthyl chloride. But we can turn this statement around and say that if we see an effect of stereochemistry, then the mechanism cannot be E1 and must be E2. Could stereochemistry be an explanation of the Evelyn effect? Maybe, but not by itself because the Evelyn effect is concerned not with product preferences but rather emphasis in product preferences.

The real clue comes from the observation by Professor Todd that the commercial 2-methylcyclohexanol and 4-methylcyclohexanol are actually mixtures of *cis* and *trans* isomers. Thus, like menthyl chloride, there could be a difference in the rates of elimination for these two isomers. Indeed two French chemists first reported in

1931[2] that pure *cis*-2-methylcyclohexanol reacts much faster than the *trans* isomer giving greater than 90% of the more substituted 1-methylcyclohexene. Recently, two American chemists[3,4] confirmed the faster reactivity for the *cis*-2-methylcyclohexanol, placing the *cis*/*trans* relative rates at 8.4/1.0, respectively.

A look at the stereochemistry of these two isomers explains this rate preference. In its lowest energy conformation, the *cis* alcohol has the water leaving group in the axial position with *anti*-periplanar hydrogens to either side. Thus stereochemically, either 1-methylcyclohexene or 3-methylcyclohexene are possible with Zaitzev's rule favoring the former. For the *trans* alcohol, the leaving group is equatorial in the lowest energy conformation and an *anti*-periplanar reaction is not possible. However, flipping the ring to the alternate chair form switches all the positions permitting a possible E2 reaction leading to 3-methylcyclohexene.
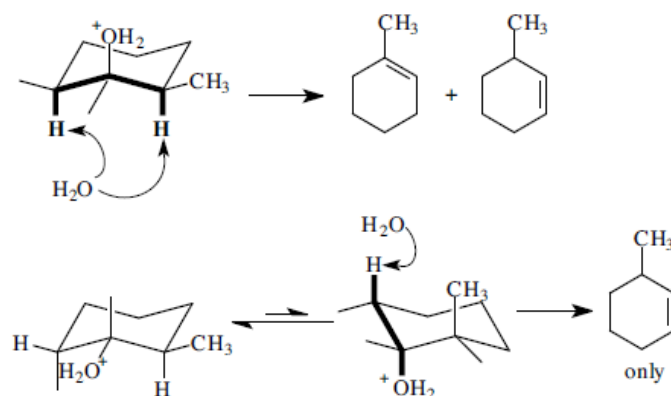
One could argue that the Evelyn effect occurs simply from a change in product preference during an E2 mechanism as the *cis* isomer gets used up and the *trans* begins to dominate. While this statement might be true, it must not be the whole story because, if it was, there would be no 1-methylcyclohexene formed in the later stages. The French chemists found that more than 45% of 1-methylcyclohexene in the products when starting from pure *trans*-methylcyclohexanol.

Now let's turn our attention to 4-methylcyclohexanol. No matter what the mechanism, only 4-methylcyclohexene can be directly formed. Other products can be formed but only after rearrangement during an E1 mechanism. If we see the Evelyn effect during this reaction, it must be from a changing proportion of E1 and E2 mechanisms. In other words, for 4-methylcyclohexanol, the Evelyn effect requires a mixed mechanism. In proposing an explanation for your results, you will need to find an explanation of how the preference for a mechanism can change. The presence of *cis* and *trans* isomers is central to the argument.

Similar to the previous 2-methylcyclohexanol case, *cis*-4-methylcyclohexanol can undergo a direct E2 elimination. The *trans* isomer must flip its cyclohexane ring into a higher energy conformation before the *anti*-periplanar requirement is met. Thus an E2 elimination is harder for the *trans* isomer. You could speculate that a harder E2 will shift the E1/E2 competition toward E1. If this happens what is the expected effect on the product mixture? Keep in mind that once a carbocation is formed, single or multiple rearrangements can follow.

[2] Vavon, G.; Barbier, M. *Bull. Soc. Chem. [4]* **1931**, *49*, 567.
[3] Cawley, J.J.; Lindner, P. E. *J. Chem. Ed*. **1997**, *74*, 102.
[4] Cawley, J. J. *J. Chem. Ed.* **1999**, *76*, 58.
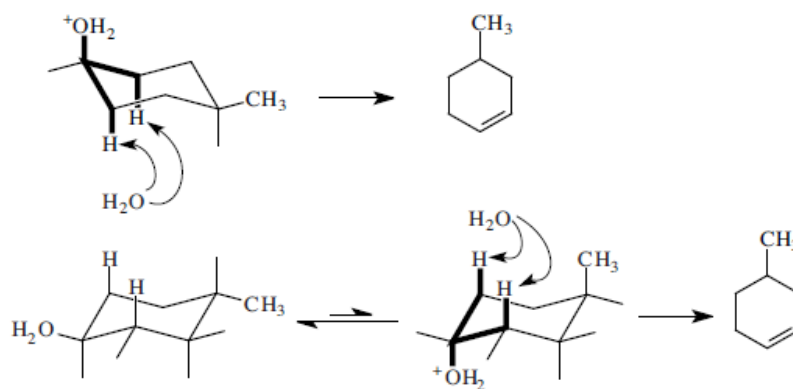
Laboratory Section

For additional details about the procedure, hazards, and waste disposal, consult a text[5].

Your instructor will assign either 2-methylcyclohexanol or 4methylcyclohexanol for the dehydration project. Set up a distillation apparatus using an unpacked vertical reflux condenser to reduce the chances of the higher boiling alcohol ending up in the product. Do not run water through the vertical condenser. Add 150 mmol of the alcohol, 5 mL of 85% phosphoric acid, and acid resistant boiling chips (or a magnetic stir bar) to a 50 mL boiling flask.

Heat the reaction mixture gently giving the products time to form and ascend up the column. Slowly distill the alkene collecting the products into a 10 mL graduated cylinder. Keep the distillation rate down to about one drop per second. Monitor the temperature to be sure it does not exceed 120°, the boiling point of the alkene products. After collecting about 8 mL, pour the distillate into a centrifuge tube and cap. Return the graduated cylinder and resume collecting product. After about 6 mL of additional product has been collected, watch carefully for the end of the reaction. Stop the distillation immediately when product ceases to distill over and white fumes or foaming occur in the reaction flask. Transfer the second distillate to another centrifuge tube and cap.

Wash each distillate with two 5 mL portions of saturated sodium bicarbonate solution. After each wash, draw off the aqueous layer with a Pasteur pipet. Dry each distillate with calcium chloride. Centrifuge if necessary and decant each fraction into a tared vial for analysis. Record the total mass of product and calculate the percent yield.

At a time arranged by your instructor, analyze each fraction by gas chromatography. Products should appear in the order of their boiling points. Alkenes with close boiling points, such as 3-methylcyclohexene and 4-methylcyclohexene may not be separated. The first major peak usually is a mixture of these two products and the last major peak is usually the 1-methylcyclohexene. Your instructor may ask you to confirm the identification of at

---

[5] Lehman, J. W. *Operational Organic Chemistry*, 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2008.

least one alkene by "spiking" with an authentic sample. Be sure this analysis is complete before the next laboratory meeting because you will be expected to share your results.

Modeling Section

Both 2-methylcyclohexanol and 4-methylcyclohexanol are mixtures of *cis* and *trans* isomers. In addition, each of these four isomers can exist in two different chair conformations. To fully explain the Evelyn effect, you will need to model all eight structures that may be involved.

We can make an estimate of the expected isomer ratio by making a few assumptions. If we assume that the change in entropy during a flip to the alternate chair conformation is very small then $\Delta G_{rx}=\Delta E_{steric}$, where ΔE_{steric} is the difference (B-A) in steric energy for the two conformations as calculated by molecular mechanics. Then we can relate the equilibrium constant (K) to our data (ΔE_{steric}) via the equations:

$$\Delta G_{rx} = -RT \ln K \quad \text{R=1.987 cal/mol-K} \qquad \qquad \textit{Equation 1}$$

$$\Delta E_{steric} \approx - RT \ln \frac{[conformation B]}{[conformation A]} \qquad \qquad \textit{Equation 2}$$

For example, if conformation A is 1.00 kcal lower energy, then B/A is 0.185 or 15.6% B. If the preference increases to 2.00 kcal, the fraction of B drops to 3.3% and for 3.00 kcal to 0.63%. Thus A quickly becomes a non-participant in the reaction. You do not actually need to do this calculation but you should keep this example in mind when analyzing the modeling results

Build each structure and optimize their geometry using Molecular Mechanics (MM2). At the end of each calculation record the calculated energy. Check the stereochemistry of the optimized structures looking for the appropriate *anti*-periplanar relationships necessary for an E2 reaction. Complete this modeling exercise before the next laboratory meeting because you will be expected to participate in a discussion of your results.

Copy and paste the results in a word processor. Remember to orient and resize the molecules in a presentable manner.

Report Guidelines

Report the results of your experimental work being sure to include:

✎ a summary of procedure and observations.

✎ yield, theoretical yield, and percent yield.

✎ results of your gas chromatography analysis for each fraction.

✎ a report of the average class results for each starting material. Use those results in your analysis of the mechanism.

Address the following issues in a well-organized discussion leading to logically argued conclusions regarding the mechanisms involved.

✎ Does the data for the 2-methylcyclohexanol dehydration confirm Professor Todd's observation of the Evelyn effect?

✎ Is this effect seen in 4-methylcyclohexanol dehydration?

✎ Is Zaitzev's rule followed?

Sometimes a small product peak is seen between the major methylcyclohexene peaks. Speculate on the identity of this peak.

Consider the modeling results for 2-methylcyclohexanol first. For each of the four isomers compare the energies of the two chair conformations.

✎ Which will predominate? By a lot or a little?

For the dominate conformations, note the stereochemistry of the atoms around the hydroxyl group.

✎ Which structures meet the criteria of having an H and OH *anti*-periplanar for E2?

✎ Which alkene(s), if any, will be formed from an E2 elimination?

✎ Do your results suggest whether *cis* or *trans* should react faster?

✎ If another pathway, such as E1, was possible, would it be most competitive for the *cis* or *trans*?

Write a detailed chronological scenario describing:

✎ what is happening as the reaction proceeds

✎ when the various species are reacting

✎ which mechanisms are predominating

✎ when shifts in mechanism preference might occur.

Repeat the analysis for 4-methylcyclohexanol.

Write out a complete mechanism for the formation of each product that you speculate to be present in the product mixture from BOTH the dehydration of 2-methylcyclohexanol and 4-methylcyclohexanol.

Instructor's Notes
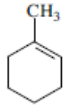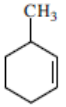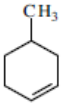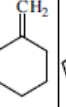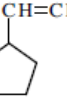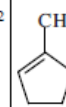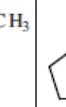
Typical Results

The analysis for this experiment is intimidating and complex but, with time, students can understand it. We spend two weeks on this experiment; during the first week the dehydration are done individually by the students, half doing each isomer. In the interim between the first and second lab, all the gas chromatography and modeling are finished. The second lab is spent discussing the GC and modeling results. We talk in detail about the stereochemical requirements that exist for an E2 mechanism but not for an E1.
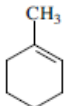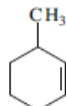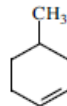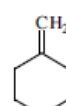
Gas chromatography analysis.

We usually use a fast analysis to so we can accommodate all the students on one dedicated GC during one week. We use an SE30 0.52mm capillary column at 65° for 2 min; heat at 50°/min until 160° and hold for 1 min. Total time about 5 min. Even with much lower temperatures, we have never been able to separate the three isomers boiling in the region of 103°. Typical results from GC analysis are shown in the first row of each fraction. A separate analysis by NMR gave us the percentages we thought they might be useful for your understanding of the experiment.

2-methylcyclohexanol

Fraction	Method	CH_3	CH_3	CH_3	CH_2	$CH=CH_2$	CH_2CH_3	$CHCH_3$
	bp °C	110.3	104	102.7	103	97	106	113-17
1	GC	78%	12%			0.1%	0.6%	0.0%
1	NMR	60%	9%	4%	0.5%	0.1%	0.6%	0.0%
2	GC	60%	30%			1.6%	4.3%	0.6%
2	NMR	60%	16.3%	13.8%	1%	1.6%	4.3%	0.6%

4-methylcyclohexanol

Fraction	Method	CH_3	CH_3	CH_3	CH_2
	bp °C	110.3	104	102.7	103
1	GC	5.4%	94.6%		
2	GC	14.3%	85.7%		

We have been studying this reaction for a long time and believe the evidence is quite persuasive that both E1 and E2 mechanisms are operative. In 2-methylcyclohexanol, the students are not able to separate the 3-methylcyclohexene, 4-methylcyclohexene, and methylenecyclohexane and

therefore have little evidence of how much rearrangement and E1 mechanism occurs. Without more product information, the Evelyn effect could easily be explained simply by considering the unfavorable stereochemical preferences for the *trans* isomer and the resulting faster E2 reaction for the *cis* isomer. The 1-methylcyclopentene is easily observed in the GC during the reactions with 2-methylcyclohexanol but not always correctly identified. If students speculate correctly about its identity, the presence of some E1 is certain.

In order to have a really interesting discussion, it is necessary to include 4-methylcyclohexanol. The 4-methylcyclohexene is the first formed product from either the *cis* or the *trans* isomers. However, the easily observed 1-methylcyclohexene can only come about from an E1 mechanism and the varying amount of this product strongly suggests varying amounts of E1 mechanism. On the other hand, in order for *cis* and *trans* isomers to show different reactivity, and therefore an Evelyn effect, we must have an E2 mechanism. An alternative mechanism proposed by Cawley does not seem to work for the 4-methylcyclohexanol.

The following are the results of a molecular mechanics calculation for the eight structures potentially involved. The first letter (a or e) refers to the axial or equatorial placement of the methyl and the second letter the hydroxyl.

2-Methylcyclohexanol	Steric energy (kcal/mol)	4-Methylcyclohexanol	Steric energy (kcal/mol)
e,e trans-2-MeOH	7.87	e,e trans-4-MeOH	7.52
a,a trans-2-MeOH	9.81	a,a trans-4-MeOH	9.86
a,e cis-2-MeOH	9.76	a,e cis-4-MeOH	9.29
e,a cis-2-MeOH	8.47	e,a cis-4-MeOH	8.09

Stereochemically either chair form of the *cis* isomers have the correct *anti*-periplanar relationships. On the other hand, only the a,a *trans* conformations can undergo an E2 elimination and there is about a 2 kcal preference for the e,e conformation. We can speculate that in the *trans* case, the greater difficulty of producing an E2 mechanism permits the alternative E1 to be more competitive. Thus as the *cis* isomer is consumed and *trans* begins to predominate, we see increasing amounts of E1.

If a good discussion takes place during the second lab, many students will turn in well-reasoned reports with logical conclusions. Nearly all students write passable reports and show considerable understanding of this complex reaction.

Part 3: Physical Chemistry

Experiment 6: Investigating the Resonance
                  Stabilization of Benzene

Experiment 7: Investigation of the
                  Rotational Barrier in
                  N-N-dimethylacetamide

6

Investigating the Resonance Stabilization of Benzene

Contributed by Kevin E. Johnson, Pacific University

Overview

In this experiment, you will use both experimental and computational methods to determine the enthalpic contribution to the resonance stabilization of benzene by comparing the enthalpy of formation ($\Delta H_f°$) of several compounds.

Recommended Exercises

Complete Exercise I, II, III, IV, V, and VI before beginning this experiment.

Background

The classic example of resonance is the simplest aromatic compound, benzene (C_6H_6). A Lewis structure predicts alternating double and single bonds around the six carbon ring. Yet it is known that the bond lengths are all equal at 1.397Å, intermediate between the length of a C-C single bond (1.537Å) and a C=C double bond (1.335Å). This is ascribed to bond resonance, where the bonding predicted by the Lewis structure is delocalized throughout the six carbon cyclic structure. Thus, the resulting bonds are all equivalent.
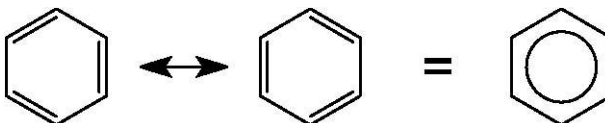
In addition to this structural component, the resonance delocalization of the bonds results in an energetic stabilization, or resonance energy. This is a measurable decrease in the molecule's energy as compared to the hypothetical cyclohexatriene, with fixed alternating single and double bonds. The resonance stabilization of benzene is thus the difference between the enthalpies of formation for benzene and cyclohexatriene.

$$\Delta H_f°_{(resonance)} = \Delta H_f°_{(benzene)} - \Delta H_f°_{(cyclohexatriene)}$$

While it is not possible to experimentally measure the formation energy of cyclohexatriene, its heat of formation ($\Delta H_f°(298)$) can be predicted using both experimental and computational methods.

The first analysis is made by calculating the ΔH_f° for the series of six carbon cyclic compounds, and comparing the trend in ΔH_f° predicted by cyclohexane, cyclohexene, and 1,3-cyclohexadiene to benzene to predict the ΔH_f° of cyclohexatriene.
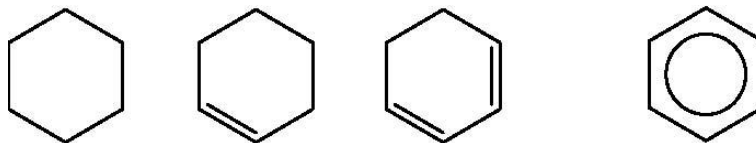
A second method uses the atom and bond equivalence of 1,5,9-*trans,trans,cis*-cyclododecatriene, $C_{12}H_{18}$ (CDDT). CDDT is not aromatic and therefore does not include the resonance stabilization of benzene. Because of this, the ΔH_f° of CDDT equals the sum of the ΔH_f° for cyclohexane and cyclohexatriene.
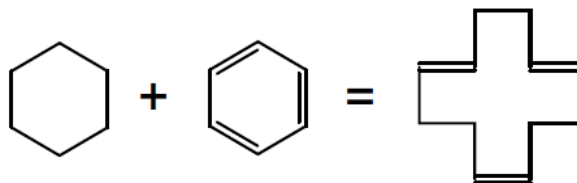
The first and second methods can be accomplished both experimentally and computationally.

A third purely computational method involves calculating energies for molecules with locked geometry. The ΔH_f° for cyclohexatriene can be calculated by locking the alternating double and single bonds. The ΔH_f° can then be used to calculate the resonance stabilization of benzene.

Laboratory Section

Using oxygen-bomb calorimetry measure the heats of combustion for cyclohexane, cyclohexene, 1,3-cyclohexadiene, benzene, and CDDT. With the quantity you obtain from calorimetry, calculate the heats of formation. Use literature values for the heats of formation of the products CO_2 and H_2O.

With these data, you can predict for the ΔH_f° of the hypothetical cyclohexatriene and compare it with the measured ΔH_f° of benzene. The difference of the two numbers will give you the resonance stabilization of benzene.

Literature Values

Use a literature source to lookup the ΔH_{comb}° and ΔH_f° values for the compounds you measure.

Suggested on-line resource: *http://webbook.nist.gov/*

Modeling Section

Comparison of ΔH_f°

Build the series of six carbon cyclic compounds and CDDT. Use Spread-sheet to calculate the enthalpy of formation and to simplify the organization of results. Also consider that the modeling calculations are for isolated and comparison with experimental results should account for the phase change from liquid to vapor.

Modeling hypothetical cyclohexatriene

While only hypothetical, you can model cyclohexatriene and compare the calculated ΔH_f° to that of benzene. Try various approaches to model the hypothetical compound by fixing atomic positions or bond lengths in the cyclic C_6H_6 compound. Be critical of the MOPAC results. Decide which method gives you good results and stick to that method.

One way to generate a starting point for calculations is to place a benzene fragment and use the Beautify command to alter the bond lengths. Either lock atom positions, or bond lengths to maintain the cyclohexatriene geometry during MOPAC calculations. You may choose to lock the C-C bonds or the C=C bonds or both. After each MOPAC calculation, always check the resulting structure to verify that the bond lengths correspond to expectations. Use the ΔH_f° data from only those structures that keep the expected geometry.

Another starting point for calculations that does not start with a biased benzene geometry, is to retrieve a cyclohexane fragment. Delete one hydrogen atom per carbon atom, change alternate bonds to double bonds, and lock the single bond lengths. In order to optimize this restricted structure, the MOPAC calculation may require a change in optimization method (BFGS is more successful than EF for example). Always check the resulting geometry after each MOPAC calculation.

Record the ΔH_f° for the various calculation methods, remember to use only the numbers that produced the correct geometry. Compare these ΔH_f° to that of benzene to estimate the resonance stabilization enthalpy.

Report Guidelines

✎ Report the values from the experimental, literature, and computational investigations of ΔH_f° for the six carbon cyclic compounds and CDDT.

✎ Use and explain the trends observed for the six carbon cyclic compounds to estimate the resonance stabilization energy of benzene. A graphical presentation may be useful.

✎ Analyze the data from CDDT in comparison to cyclohexane and benzene, recognizing that the bonding and atom elements of CDDT amount to the sum of cyclohexane and cyclohexatriene.

✎ Compare direct calculations of the ΔH_f° of cyclohexatriene with optimized benzene.

✎ Summarize the conclusions for each method of calculating the resonance stabilization of benzene. Also, discuss the relative benefits of each method.

Instructor Notes

The first project of our physical chemistry/thermodynamics laboratory course is this computational and experimental investigation of resonance stabilization. The experimental component introduces not only the technique of oxygen bomb calorimetry, but also statistical handling of data and error analysis. Students work in pairs and each group is expected to provide for the class one set of data for all the measured compounds. The class then analyzes the collective results to make a conclusion using the experimental data.

The experimental procedure requires very careful attention to individual measurements of sample mass and temperature rise in order to obtain the precision necessary to accurately predict the stabilization energy of benzene. Literature values of ΔH_f suggest a stabilization enthalpy of –33 kJ/mol while organic textbooks quote –36 kJ/mol. Student results from our labs are in the range of –50 kJ/mol.

The computational component of the experiment is intended to introduce students to the use of computational methods to complement experimental work. I encourage students to explore the potential of the SCIGRESS tools to solve this problem. A straight-forward calculation of heats of formation for the assigned compounds and trend analysis is the first step. I encourage students to investigate the effect of parameter choice when working in pairs. One member uses PM3 and the other AM1 to critically assess which may be the better tool for this study.

Typical Results

Chemical sample	# double bonds	AM1 ΔHf ° (kcal/mol)	PM3 ΔHf ° (kcal/mol)	Literature
C_6H_{12}	0	-38.6	-30.9	-29.49
C_6H_{10}	1	-10.1	-5.0	-1.08
C_6H_8	2	17.6	20.6	25.4
C_6H_6	3	22.0	23.4	19.81
Cyclohexatriene	3	39.6	40.5	
$C_{12}H_{18}$	3	13.1	21.6	21.91
Analysis Method	Stabilization Energy			
Trend Analysis		-23.8	-23.0	-33.4
$C_{12}H_{18}$-C_6H_{12}		-29.7	-29.1	-31.6
Cyclohexatriene		-17.6	-17.1	

Resonance Stabilization Analysis
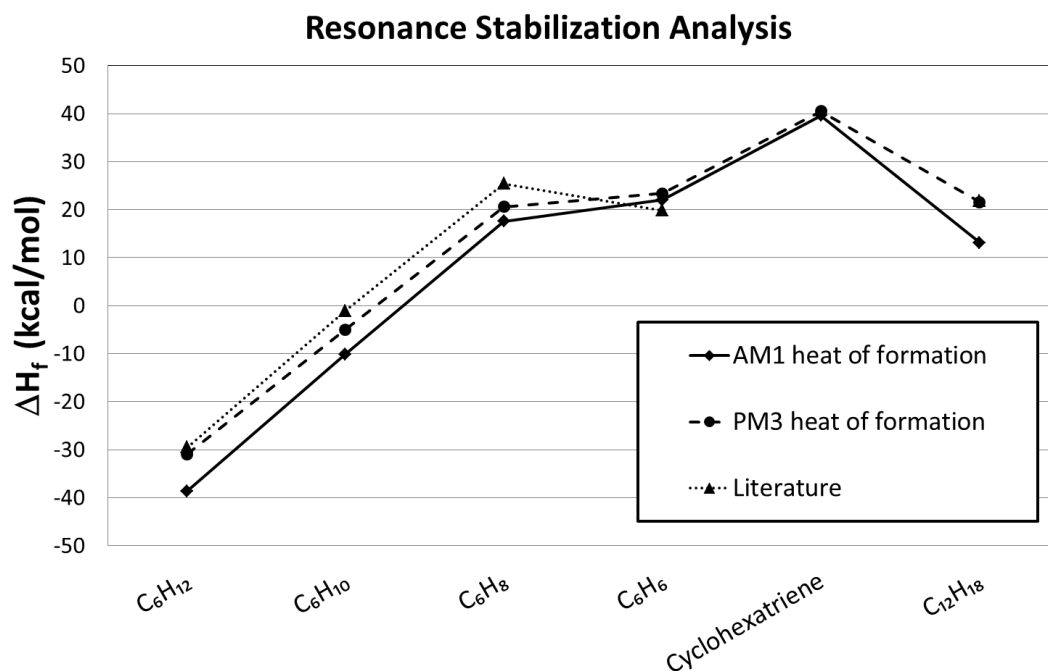
The trend analysis simply takes the linear equation determined by the three unconjugated C_6 cyclic compounds and predicts a ΔH_f° cyclohexatriene with three unconjugated double bonds. The difference between the predicted value and benzene is the stabilization energy. The ΔH_f° of cyclohexatriene can also be predicted by taking the difference between CDDT and cyclohexane. That value compared with benzene provides another route to calculating the stabilization energy.

The purely computational method of calculating a molecule with restricted geometry is best accomplished by locking the single bond length in the unoptimized benzene molecule. The results of this computational method are the furthest from the literature values. These results may not be so surprising considering the experimental methods only predict an energy for the cyclohexatriene based on parameterization of real compounds.

Overall the AM1 parameters are somewhat better than PM3. However, both yield systematic errors in the calculations of these compounds. An important point for students to realize is that these semi empirical computational methods are not likely to produce values that exactly match experimental or literature results. Instead they can be used to make reasonable comparisons between compounds with similar structural and elemental properties.

References

Halpern, A. M. *Experimental Physical Chemistry: A Laboratory Textbook*, 2nd ed.; Prentice Hall: New York, 1997; pp 139-143.

7

Investigation of the Rotational Barrier in N-N-dimethylacetamide

Contributed by Kevin E. Johnson, Pacific University

Overview

In this project, you will investigate the barrier rotation about the carbonyl carbon-nitrogen bond in amides in two ways: experimentally, using temperature dependent nuclear magnetic resonance (NMR), and computationally, using dihedral search labels and MOPAC. In this investigation you will use N,N-dimethylacetamide (NNDMA) as a model compound to measure and compute the energetic barrier to rotation.

Recommended Exercises

Complete Exercises I, II, III, IV, VI, and VIII before beginning this experiment.

Background

Amide bonds exhibit a rotational barrier at room temperature. A simple resonance argument that concludes that bond has partial double bond character explains this rotational barrier. The Lewis structure diagram below illustrates a model in which the non-bonding pair of electrons on the nitrogen is partially donated to the C-N bond and the electrons in the carbonyl π bond are partially shifted to the oxygen atom.
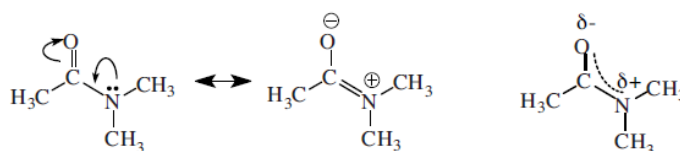
A better model explains the bonding as the formation of a delocalized bond across the O-C-N atoms in the molecule.
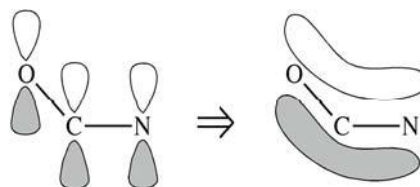
Both models predict partial double bond character for the C-N bond. Therefore rotation about that bond implies breaking the double bond. This is the origin of the

rotation barrier about the C-N bond.

The rotational barrier gives rise to a low energy internal rotation which results in a vibrational mode that can be calculated. The energy of this mode is indicative of the curvature of the potential well the rotational barrier creates.

Additionally, the predicted hybridization of the N atom is intermediate between sp^3 trigonal pyramidal geometry with no double bond character and sp^2 trigonal planar with full double bond character. These geometric restrictions are important in structural confirmations of proteins. The amide – or peptide – bond restricts the possible confirmations of a protein.

Laboratory Section

The room temperature proton NMR spectrum of N,N-dimethylacetamide (NNDMA) exhibits three singlet peaks with equal area. The peak at approximately 2 ppm is from the protons on the acetyl-methyl (bonded directly to the carbonyl carbon). Because they experience different magnetic environments, the two methyl groups bonded to the nitrogen give rise to two peaks at approximately 3 ppm. The peak slightly downfield is *cis* to the oxygen, while the peak slightly upfield is *cis* to the acetyl-methyl.

With enough thermal energy available to break the double bond, rotation about the C-N bond leads to an exchange between the two possible configurations of the two amide-methyl groups. Since the two sides of the equilibrium are chemically equivalent, there is no preference for a particular configuration. However, the exchange rate does depend on temperature, and this can be measured using temperature dependent NMR.
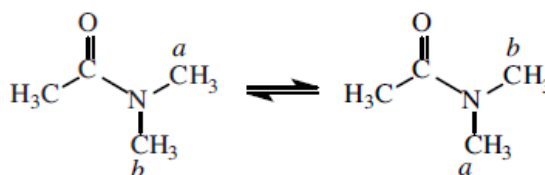
At low temperatures the exchange rate between the two amide-methyl groups is slow, and in the NMR spectrum the protons on each group appear separately. At high temperatures the methyl groups exchange rapidly and on the time scale of the NMR measurement they experience the average magnetic environment. This leads to a single peak for the two methyl groups in the NMR spectrum. At intermediate temperatures the two peaks broaden and coalesce into a single peak, which narrows at higher temperatures.

The temperature dependence of the spectra can be described by the following equation:

$$g(v) = K \frac{\tau(v_A - v_B)^2}{\left[\frac{1}{2}(v_A + v_B) - v\right]^2 + 4\pi^2\tau^2(v_A - v)^2(v_B - v)^2}$$

$g(v)$ = spectral intensity as a function of frequency

K = adjustable intensity factor

τ = mean lifetime in each magnetic environment

v_A and v_B = resonance frequencies of two stationary environments

At long τ the spectrum approaches two peaks and at short τ a single peak. The lifetime is inversely proportional to the rate constant, $k = \tau^{-1}$.

By measuring the spectrum of NNDMA over a range of temperatures (for a 200MHz spectrometer T=40-120°C is sufficient) and fitting the resulting spectra to the intensity equation above, it is possible to determine the temperature dependence of the rate constant k.

An Arrhenius plot of these data will yield the experimental activation energy for the rotational exchange of the two amide-methyl groups.

Modeling Section

1. Build a model of N,N-dimethylacetamide and **select Action | Beautify | Comprehensive.**
2. Check the hybridization of the nitrogen atom. The beautify command should set the nitrogen hybridization to sp^2, reflecting the presence of the amide bond.
   ❏ Select the N atom and look in the Hybridization box.
3. Select one of the C atoms in the amide-methyl groups; next select the N atom, and then the remaining C atoms bonded to the N in any order.
4. Select **Action | Geometry Labels | Improper Torsion** and record the angle. This is the planar value. To determine what value will result for an sp^3 hybridization use the drop down menu to change the hybridization to sp^3 While this is not a good model for the amide bond, it will allow you to compare the MOPAC results with sp^2 vs. sp^3 geometries.
5. Select only the N atom and select **Action | Beautify | Geometry**. Again measure and record the improper torsion as before. This is the sp^3 value.
6. Return the molecule to the proper geometry by selecting **Action | Beautify | Comprehensive.**
7. Save the molecule and proceed to a MOPAC calculation.
8. Set up a geometry optimization and append a force calculation in MOPAC. When complete, note the calculated energy.
   ❏ Run a calculation using **chemical sample, IR transitions**, and the selected parameter set (AM1 or PM3). The experiment manager automatically appends an optimization followed by a force calculation. However, if you want specific options for the optimization you need to create a new experiment procedure. In such a case, consult your instructor.
9. Open the molecule after the MOPAC calculations finish.

   ❏ Click on the molecule window in the Workspace.

📖 NOTE
You can use the improper torsion at the N atom to evaluate the degree to which the bonding at the N atom is planar.

📖 NOTE
When using MOPAC, choose either AM1 or PM3 parameters, but be consistent. Also, calculations are more successful if the optimization options Precise, BFGS and NLLSQ are chosen.

10. Choose the atoms in the following order: first the methyl C *cis* to the O, the N, the carbonyl C and finally the O atom.
11. Measure the dihedral angle.
   ❏ Select **Action | Geometry Labels | Dihedral Angle** and record the angle.
12. Measure the improper torsion as before.
13. View the IR/Vibrational Spectra. The rotational barrier will generate a mode with very low energy.
   ❏ Select **Analyze | IR Transitions.**

14. Select the lowest four modes one at a time by clicking on them. In the molecule window look for atom motions that appear to be rotation about the C-N bond.

   ❏ If you double click on a mode, a View Transition Attributes box opens; it contains the numeric value for that mode. While it is not likely that these can be measured experimentally, the value does indicate something about the barrier.

15. Click in the molecule window to activate it.
16. Select the amide bond dihedral angle as before (methyl C, N, carbonyl C, O). In the Dihedral Angle window set up a search label from -180° to 180° in 72 steps (5° per step).
17. Save the molecule and open a new experiment.
18. Generate an energy map.
   ❏ Select **conformations, potential energy map,** and the same parameters as before. Run the experiment. When the calculation is complete, the molecule and map will open in the in the Workspace. Note the dihedral angle (labeled on the molecule) and the Energy at the two maxima.

   Because two barriers are present and the net result of the rotation over either barrier is the same; the two transitions lead to the same outcome. In order to account for the double barrier and determine an effective energetic barrier, one does not average the barrier energies; rather the rates (or in this case the rate constants) are additive. Use the two calculated activation energies and calculate the exponential portion of the Arrhenius equation at 298K. Sum these "rate constants" and use this value to calculate an effective energetic barrier by once again applying the Arrhenius equation.

Report Guidelines

   ✎ Report the experimental results, including the Arrhenius plot of the measured lifetimes. The report should include a calculated value for the rotational barrier.

   ✎ Report the geometric and energetic outcomes of the optimization calculation, and the calculated effective barrier to rotation.

   ✎ Discuss the origin of this barrier, and how modeling can be used to understand the origin of the experimental observations.

Instructor Notes

This project incorporates components of kinetics and spectroscopy. The internal rotations of amide bonds have energetic barriers such that the rate of internal conversion by rotation is accessible to NMR spectroscopy. The internal rotation about the amide bond can be thought of as a kinetic process with an activation energy. This is where a connection can be made to computational methods: students can easily generate a search parameter on the dihedral angle about the amide bond. Additionally the model of resonance can be investigated by comparing the geometry about the N atom with sp^2 and sp^3 geometries. Determination of the vibrational mode associated with the internal rotation reinforces the idea that a quasi-stationary state is observed at low temperatures with NMR and that thermal energy is required to access higher energy states within the potential well.

The presence of a double barrier can be accounted for by calculating an effective barrier. This is done by recognizing that the rates of the two barrier crossing events are additive.

$$E_{a\ \text{eff}} = -\ln\left(e^{-E_{a1}/RT} + e^{-E_{a2}/RT}\right)RT$$

Typically the lower barrier dominates.

PM3 parameters yield an effective barrier of 3.2kcal/mole, while AM1 parameters yield an effective barrier of 7.8kcal/mol. Experimental results for N,N-dimethyl acetamide yield a rotational barrier of 13.9kcal/mol. The map file generated using PM3 parameters is illustrated below.
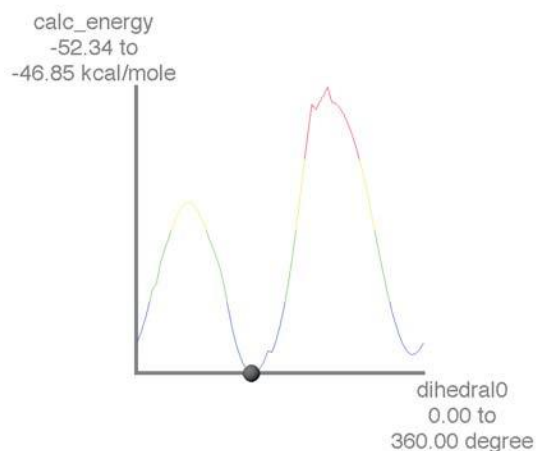
The sampling time frame for NMR is such that the transition between slow exchange and fast exchange occurs at temperatures close to room temperature. This allows the observation of both exchange rate regimes a curve fitting analysis that can extract kinetic data as a function of temperature, which when analyzed using an Arrhenius plot yields the experimental activation energy.

Further Discussion

This project can be extended in many directions. Two possible extensions include: the experimental and computational investigation of other compounds containing amide bonds to determine the effect of substituents on the delocalized bond; and a computational study of the molecular orbitals involved in the formation of the bond. The latter can be initially investigated via an extended Hückel calculation performed on the.map file. After a Hückel calculation the .map file will contain calculated eigen energies for each of the molecular orbitals. The HOMO can be identified by checking the occupancy value of a molecular orbital in internals. Unfortunately these can not be plotted simultaneously, and are therefore difficult to assess.

References

Brown, K.C.; Tyson, R.L.; Weil, J.A. Conformation Interchange in Nuclear Magnetic Resonance Spectroscopy. *J. Chem. Ed.* **1998***, 75* (Dec), 1632.

Gasparro, F. P.; Kolodny, N.H. NMR Determination of the Rotational Barrier in N,N-dimethylacetamide: A physical chemistry experiment. *J. Chem. Ed.* 1977, *54* (Apr), 258.

Stewart, W.E.; Siddall, J.H. III. Nuclear Magnetic Resonance Studies of Amides. *Chem. Rev.* **1970***, 70* (Oct), 517.

Part 4: Inorganic Chemistry

Polarities of Small Molecules

Contributed by William Jordan, Pacific University

Overview

In this experiment you will use SCIGRESS to compute the dipole moments of structurally related molecules. You will compare the different dipoles to generalize about the dependence of molecular polarity on electronegativity, bond angle, bond distance, bond number and distribution of lone pairs.

Recommended Exercises

Complete Exercises I, II, III, and IV before beginning this experiment.

Background

Molecular polarity is an important property in determining solubility, chemical reactivity, melting point and boiling point. Molecular polarity is a consequence of bond polarity and molecular structure. The bond dipole moment is a simple product of charge and bond distance, $m = qd$. Bond polarity increases with increasing electronegativity difference between the atoms joined by the bond and with increasing bond distance.

If bond polarities are opposed as in CO_2, they cancel and the molecule is nonpolar. However, when bond polarities have components in the same direction, they add to give a net molecular polarity. Bonds which are separated by a small angle have particularly large vector additions in a single direction. Since electrons carry charge, localization of lone pairs in particular orbitals may also influence molecular polarity

Modeling Section

1. Build models of NH_3, NF_3, NCl_3, PH_3, PF_3 and PCl_3. It is convenient to build one molecule, modify it and save it as another structure.
2. Label one X-N-X angle and the remaining N-X bond. If you label the first molecule you build, the labels remain when you convert it to another molecule.
3. Use Spreadsheet to determine the dipole moment of each molecule and the partial charge on the central element. Also retrieve the labeled bond angle and bond distance for each molecule. Run the calculations using AM1 parameters.

Report Guidelines

Paste the results into a word processing document and explain the following:
✎ Does dipole moment correlate with electronegativity difference?

✎ Does dipole moment correlate with difference in partial charges?
✎ Are the results explained, in part, by the bond angles and bond lengths?
✎ Is there an apparent influence of electron lone pairs on dipole moments?

Instructor Notes

As can be seen from the results summarized below, there is no single factor which correlates with bond dipole moment. In spite of a greater difference in electronegativity between N-X in NF_3 than in NH_3, NH_3 has a much larger dipole moment. The determining factor would appear to be the lone pair on nitrogen which is opposed to the combined polarity of the N-F bonds. PH_3 has a much larger dipole moment than NH_3 due, presumably, to the larger bond distance of the P-H bond and the narrower bond angle in the case of PH_3. Similar arguments can be invoked to explain the remaining pairwise differences.

Typical Results

	Dipole Moment	Difference in EN	Difference in partial charge	Bond Lenght	Bond Angle
NH_3	1.85	0.84	0.53	1.00	109.1°
NF_3	0.04	-0.94	-0.39	1.36	102.6°
NCl_3	1.00	-0.12	0.00	1.71	110.4°
PH_3	2.29	-0.01	0 .05	1.36	96.5°
PF_3	1.66	-1.79	-1.54	1.54	98.0°
PCl_3	0.01	-0.97	-0.79	1.92	105.4°

76

9 | Structures of Molecules Which Exceed the Octet Rule

Contributed by William Jordan, Pacific University

Overview

In this experiment you will use SCIGRESS to model compounds with two or more possible molecular geometries depending upon placement of multiplebonds or lone pairs. Energies will be computed to evaluate the most stable structures.

Recommended Exercises

Complete Exercises I, II, III, IV before beginning this experiment.

Background

Molecules in which the central atom has eight or fewer valence electrons have unambiguous structures (e.g. CO_2 is linear, H_2O is bent, NH_3 is pyramidal, etc.). However, when the central atom exceeds the octet rule and when electron pairs are used differently (e.g. different bond orders, lone pairs), two or more structures may be possible. For example, two alternative arrangements for XeF_4 can readily be envisioned depending upon the relative placements of the two lone pairs in the octahedral layout:

| Electron pair distribution | Molecular geometry | Electron pair distribution | Molecular geometry |
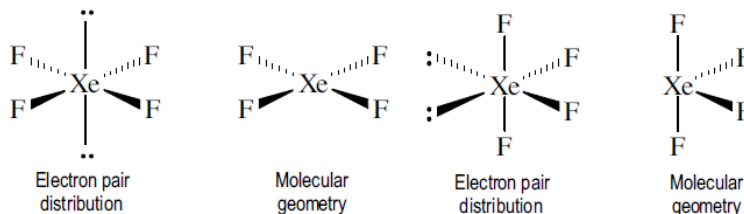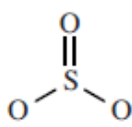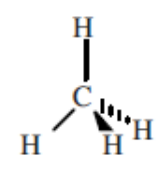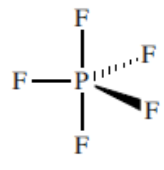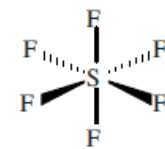
The Valence Shell Electron Pair Repulsion method (VSEPR) successfully predicts that the square planar geometry on the left will be the more stable structure as it minimizes repulsions between the lone pairs. When using the VSEPR method, the following rules should be kept in mind:

📖 NOTE

The VSEPR approach to predicting molecular structure can be reviewed in any general chemistry text.

1. Lone pairs (LP) repel more strongly than bond pairs (BP) so the strength of repulsive interaction increases in the order; BP-BP < LP-BP < LP-LP.
2. Generally, 90° interactions are least favorable, so the most stable arrangements will involve minimizing the number of 90° interactions between LPs and both LPs and BPs.
3. Multiple bonds repel more strongly than single bonds so 90° interactions involving multiple bonds should be avoided.

Number of electron groups	Distribution of electrons	
2	Linear	O=C=O
3	Trigonal Planar	
4	Tetrahedral	
5	Trigonal Bipyramidal	
6	Octahedral	

Only atoms which have valence d-orbitals can exceed the octet rule. Thus, the - third period element phosphorous ($3s^23p^33d^0$) can form six bonds as in PF_6 whereas the second period analog nitrogen ($2s^22p^3$) can form no more than four bonds as in NH_4^+.

When executing molecular orbital computations using SCIGRESS for molecules which use d-orbitals in bonding, it is important that the application have parameters for d-orbitals. Extended Hückel methods may be applied to all elements, while ZINDO has parameters for the transition metals of the fourth and fifth periods as well as most elements in the first three periods. MOPAC, MNDOd and AM1 contain d-orbital parameters for a few elements, while the remaining parameter sets lack d-orbital capabilities.

Modeling Section

ClF$_3$

1. Figure out the three reasonable geometries of ClF$_3$ based on VSEPR principles.

2. Build models of the possible structures in the Workspace and beautify them. If the bond angles and distances aren't close enough to the expected ones, the model may beautify to an alternative geometry.

3. Save the structures. Copy and paste the structures into a word processing document prior to further optimization.

4. Create a new Spreadsheet and enter the three structures into the Chemical Sample column.

📖 NOTE

Use the Action menu to set the bond angles.

5. Program the second column by selecting **Property of: chemical sample, energy total, MM/MO-G AM1 geometry**. Note the method description at the bottom of the window.

6. Evaluate the total energy and record which of the structures is most stable.

7. Copy and paste the structures into a word processing document after optimization is complete.

SOF$_4$

1. Work out the two reasonable geometries of SOF$_4$ based on VSEPR principles. Be sure to minimize formal charges by creating a double bond in the most appropriate place.

2. Build the two models of SOF$_4$. One of the structures may beautify to the other. If so, don't beautify, but optimize each structure using MM geometry (MM2).

3. Save the structures. Copy and paste these structures in a word processing document prior to further optimization.

4. Open a Spreadsheet document and optimize the structures as in the case of ClF$_3$ above.

5. Copy and paste the structures into a word processing document after optimization is complete.

Report Guidelines

✎ For each compound, paste copies of the structures before and following optimization into a word processing document.

✎ Note the most stable structure in each case and explain the results.

Instructor Notes

Typical Results

Two of the three structures for ClF_3 optimize to the same result. The lowest energy form following ZINDO-optimization is the distorted T-structure which minimizes 90° interactions between the two lone pairs (equatorial) and the remaining electron pairs.

Following Optimization via Molecular Mechanics

All fluorines equatorial
(Both lone pairs axial)   **T w o fluorines equatorial**
(One lone pair axial)   **One fluorine equatorial**
(No lone pairs axial)
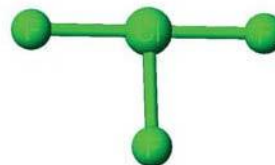
Following ZINDO-optimization
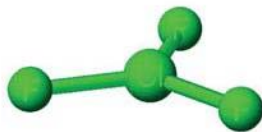
Total energy= -90.487 calculated via ZINDO   **Total energy= -90.505 calculated via ZINDO**
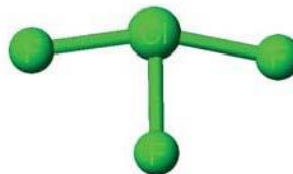
There are two reasonable structures for SOF_4, one with the S-O bond in an axial position and one with the same bond in an equatorial position. The latter is preferable as the number of 90° repulsions involving the multiple bond is minimized. It should be noted that molecular orbital methods (MOPAC and ZINDO) ignore the location of the double bond inserted in the Editor. Thus it may be of interest to note the bond orders computed for the S-F and S-O bonds.
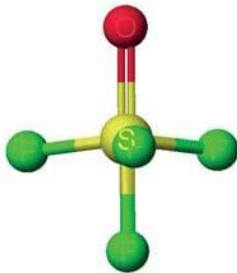
Total energy= -129.026 calculated via ZINDO   **Total energy= -129.073 calculated via ZINDO**
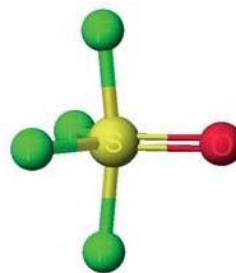
10 Comparison of Gas Phase Acidities of Binary Hydrogen Compounds

Contributed by William Jordan, Pacific University

Overview

In this exercise you will use a SCIGRESS workstation to compare the relative acidities of protonic acids. To compare the acidities you will calculate the heats of deprotonation reactions for hydrogen compounds of neighboring elements in the periodic table. You can then correlate the acidites with periodic properties.

Recommended Exercises

Complete Exercises I, II, III, IV, and VI before beginning this experiment.

Background

Gas phase proton affinity is defined as the enthalpy of dissociation of a proton from a hydrogen compound in the gaseous phase.

$$HB(g) \longrightarrow H^+(g) + B^-(g) \quad \Delta H = \text{gas phase proton affinity}$$
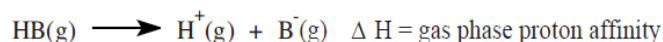
Since a spontaneous reaction is favored by a negative ΔH, the higher the gas phase proton affinity, the less acidic is the hydrogen compound. Gas phase proton affinity is a measure of intrinsic molecular acidity as it is uncomplicated by solvent effects which have a significant influence on acidity in aqueous solutions. Since the aim of this exercise is to compare acidities, the complete ΔH of the reaction will not be calculated. Since H^+ is common to each reaction, the differences between the calculated ΔH values will be the same as the differences in gas phase proton affinities. Enthalpies of formation will be computed using MOPAC routines.

$$\Delta H_{calculated} = \Delta H_f(B^-) - \Delta H_f(HB)$$

Modeling Section

1. Build models of CH_4, NH_3, H_2O, SiH_4, PH_3, H_2S and their conjugate bases.
2. Use the Spreadsheet to compare the gas phase proton affinities by calculating the difference in heat of formation (MOPAC optimization with PM3 parameters) of the acid and its conjugate base. Also calculate the partial charges on the atoms of the acids and their conjugate bases.

Report Guidelines

Paste the results into a word processing document. Discuss the results, including responses to the following questions.

✎ Is there a trend in the gas phase acidities with respect to position in the periodic table?

✎ How can the trend or trends be explained?

✎ Do the acidities agree with the partial charges? Which partial charges?

Instructor Notes

Typical Results

The calculation results, together with the electronegativity of the nonmetal, are presented below. Acidity of binary acids, HB, increases with increasing atomic number of B both within a period ($CH_4 < NH_3 < H_2O$), and within a group ($H_2O < H_2S < H_2Se$). The expected trend is observed, with the exception of hydrogen sulfide, which should be more acidic than phosphine. This illustrates a point which should be understood. Semi-empirical computations do not yield precise agreement with experimental observation.

HB	ΔH	Partial Charge (H)	Bond Distance	EN of B
CH_4	-64.5	0.07	1.11	2.55
NH_3	-41.4	0.13	1.00	3.04
H_2O	-35.9	0.19	0.96	3.44
SiH_4	15.3	-0.16	1.46	1.90
PH_3	14.7	0.01	1.36	2.19
H_2S	14.9	0.05	1.32	2.58

Further Discussion

The exercise can easily be extended to a group of oxyacids, permitting a correlation of acidities with inductive effects of groups bonded to the central nonmetal. Sample data are presented below.

Acid	ΔH	Partial Charge (H)
$H_2SO_3F^+$	187.1	0.414
HSO_3F	48.4	0.304
H_2SO_4	45.7	0.295
H_2SO_3	24.6	0.254
HNO_3	55.3	0.270
HNO_2	29.6	0.223
$HClO_4$	102.0	0.262

11   Crystal Lattices and Close Packing Sites

Contributed by William Jordan, Pacific University

Overview

In this exercise you will enter lattice parameters obtained from crystallographic studies into a SCIGRESS document so as to represent the crystal structure of a compound in different ways. This will help to identify close packing sites which, in turn, aids in understanding why different compounds crystallize in different lattice types.

Recommended Exercises

No recommended exercises.

Background

Many substances crystallize in lattices based upon close packing schemes. In close packing, each atom or ion is considered to be a hard sphere. Each layer of spheres is arranged to pack the spheres as close together as possible. This will be the case if the spheres are in contact and if each row is offset from the neighboring row by half the diameter of a sphere.

Close packing in three dimensions is achieved if the spheres of any one layer lie directly above the depressions in the layer immediately beneath it. Within a layer, there are depressions in the center of a triangular array of three spheres in contact with one another. Two close packing patterns are possible. If the third layer is aligned above the first, then the third and first layers are identified as A layers. The offset layer sandwiched between them is a B layer. The resulting pattern is ABABAB, referred to as hexagonal close packing (hcp).
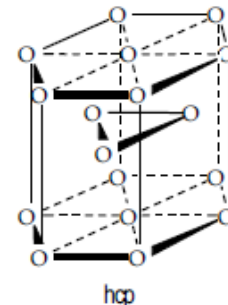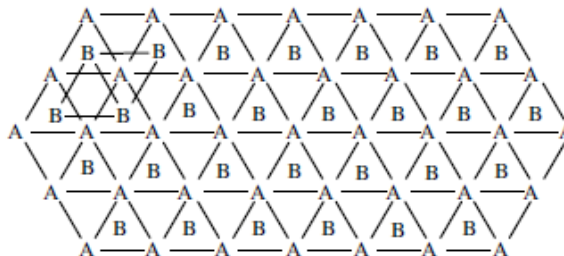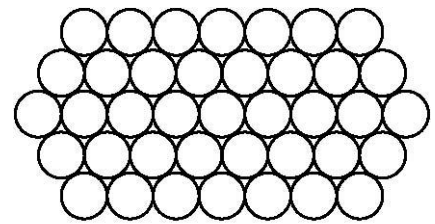
An alternative to hexagonal close packing occurs if the third layer is offset from both the A and B layers, resulting in an ABCABC pattern known as cubic close packing (ccp).

ccp
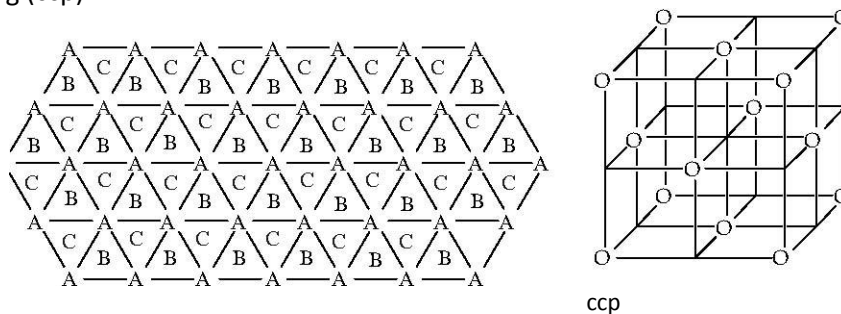
In each close packing scheme, there are two types of spaces or holes between adjacent layers of spheres. A tetrahedral hole exists between a sphere and a triangular array of spheres in the layer directly below it. An octahedral hole exists between a triangular array of spheres in one layer and the triangular array centered directly below it in the adjacent layer.

Tetrahedral Hole in Center   **Octahedral Hole in Center**
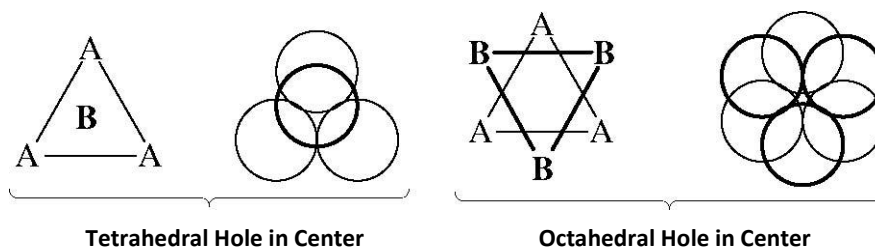
Metals tend to crystallize in close packed structures (ccp and hcp). Ionic compounds tend to crystallize in lattices based on close packed structures with the large ion (usually anion) in a close packed array and the counterion occupying holes (e.g. NaCl is ccp of Cl- with Na+ occupying octahedral holes).

Modeling Section

Aluminum Lattice.
1. Open the Workspace.
2. Select **Action | Crystallize**. A Crystal Shape box opens. Select the **Main** tab.
3. Scroll through the **Space Group** list and select **F23**. This identifies the crystal system (face centered cubic) of metallic aluminum.
4. Enter a=4.04958 in the Angles section of the dialog window. The values for b, c, α, β, and γ remain unchanged.
5. Select **Infinite Lattice** in the Build box.
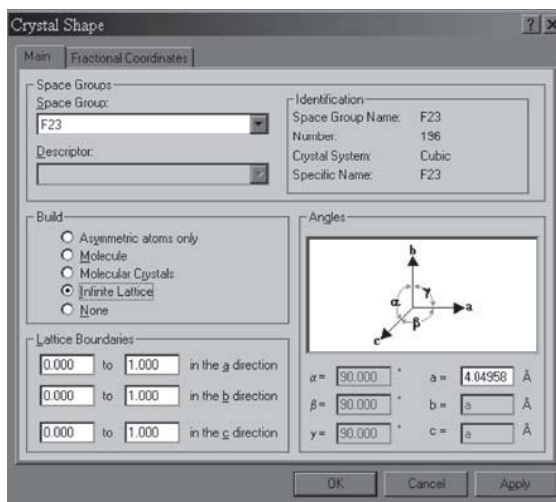
6. Click the Fractional Coordinates tab, and enter the following values from the table below for atoms 1 to 4. Click OK and a crystal structure appears in the Workspace.

atom	symbol	x	y	z
1	Al	0	0	0
2	Al	0	0.5	0.5
3	Al	0.5	0	0.5
4	Al	0.5	0.5	0

7. Select **View | Crystal Boundaries | Show Lattice Boundaries** (this option is turned on by default).

Viewing the lattice

1. Rotate and view the lattice from different angles. Find the arrangement of layers (ABC or AB). View the lattice looking from one corner of the cube to its most distant diagonal.
2. Find a tetrahedral hole. Select the atoms which surround the tetrahedral hole. Alter the atom size by changing the van der Waals scale to better view the tetrahedral hole.
   ❏ Select **View | Styles | Atoms...** In the resulting Atom Attributes dialog box, click on the Shape tab. Type in a new scale value (0-1).
3. Narrow the window so that the boundary around the lattice is small. Copy the lattice and paste into a word processing document.
4. Find an octahedral hole and select the atoms which surround the hole. As in the preceding step, narrow the window, copy and paste into the same word processing document.
5. Orient the lattice to show that the octahedral hole exists between two staggered triangular arrays of atoms. Paste the lattice into a graphics application, and draw lines between the atoms to identify the two triangular arrays. Paste the result into the word processing document.

Sodium Chloride Lattice

Many ionic lattices result from a close packed array of the large ion with the small ion occupying either octahedral or tetrahedral holes. Proceed with the NaCl lattice as in the case of the aluminum lattice.

1. Select P4(2)32 from the space group list.

2. Edit the cell parameters by entering a = 5.64056.

3. Open **Fractional Coordinates** and enter the following values:

atom	symbol	x	y	z		atom	symbol	x	y	z
1	Na	0	0	0		14	Na	0.5	1	0.5
2	Na	0.5	0.5	0		15	Cl	0.5	0.5	0.5
3	Na	0.5	0	0.5		16	Cl	0.5	0	0
4	Na	0	0.5	0.5		17	Cl	0	0.5	0
5	Na	1	1	1		18	Cl	0	0	0.5
6	Na	1	0	0		19	Cl	0.5	1	0
7	Na	0	1	0		20	Cl	1	0.5	0
8	Na	0	0	1		21	Cl	0	1	0.5
9	Na	1	1	0		22	Cl	1	0	0.5
10	Na	0	1	1		23	Cl	0.5	0	1
11	Na	1	0	1		24	Cl	0	0.5	1
12	Na	0.5	0.5	1		25	Cl	1	0.5	1
13	Na	1	0.5	0.5		26	Cl	1	1	0.5
						27	Cl	0.5	1	1

4. Select **Infinite lattice** in the Build box.
   Click **OK** to close the Crystal Shape dialog box. The sodium chloride lattice appears in the window. However, the chloride ions will appear smaller than the sodium ions. You can fix this in the following way:

   ❏ Select **Edit | Preferences | Periodic Table Settings...** Select the Charge tab. Highlight Cl in the box to the right. In the left list the van der Waals radius and charge appears for a Cl ion. Click **Change**, and a small dialog box opens. Change the radius to 1.67Å. Click **OK**. Do the same for Na ion (1.16 Å). Click **OK** to return to the workspace.

5. View the NaCl lattice from different perspectives by rotating the lattice. Note that the Na atoms form a face-centered cubic array with the Cl atoms in the positions of octahedral holes. In actuality, the Cl atoms also form a face-centered cubic array with the Na atoms in the positions of octahedral holes.

View an Extended Lattice

6. Go to the Crystal Shape window.

7. Edit the lattice boundaries so they range from 0 to 2 in the a, b and c directions.

   ❏ Click on **Infinite lattice**, and click **OK** to close the dialog box
   Select **View | Crystal Boundaries** and select **Show Unit Cell** and **Show Lattice Boundaries** (this option is turned on by default).

8. There should now be eight unit cells within the lattice with one of the unit cells outlined.

9. View the lattice from different angles. Choose one or more orientations to copy and paste into your word processing document. Explain how the figure demonstrates that both the Cl and Na ions form ccp arrays.

Report Guidelines

The aim of the report is to demonstrate an understanding of close packing in atomic and ionic lattices. Select orientations of the lattice, or portion of the lattice, which clearly illustrate close packed layers, octahedral holes and tetrahedral holes. If, for purposes of clarity, you wish to eliminate part of a lattice, simply delete selected atoms or ions. Alternatively, you may select only the atoms you wish to emphasize. It may be useful to paste a structure into a graphics document to add arrows, lines and labels for explanatory purposes.

Instructor Notes

Typical Results

The following figures have been pasted from an editor document for the aluminum lattice. In some cases additional graphics were added by first pasting into a chemical graphics document and adding lines and labels before pasting into the word processing document.
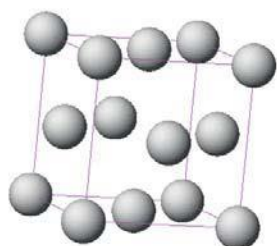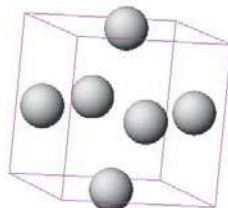
Al unit cell   **Al atoms at faces**   **Octahedral hole**
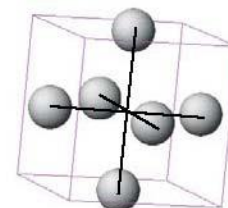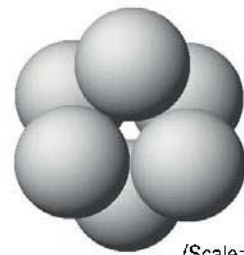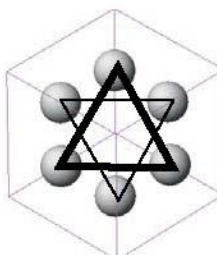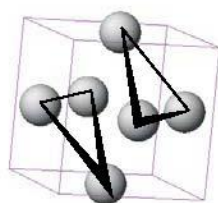
Three different views of octahedral hole (in center).
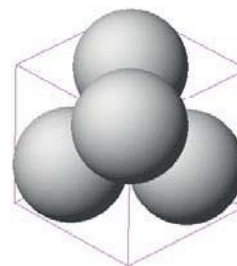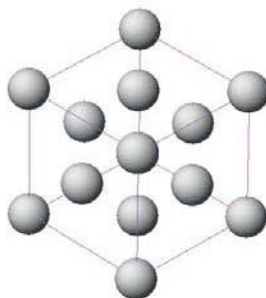
(Scale=1)

Tetrahedral hole in center of cluster (scale=1)
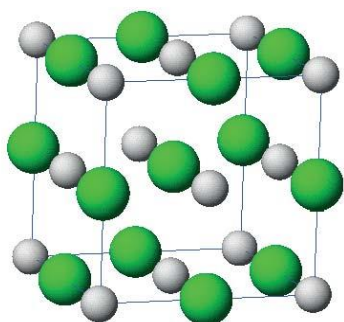
NaCl lattice with Cl⁻ central
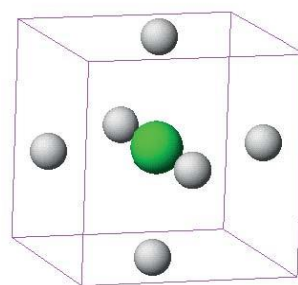
Chloride ion in an octahedral
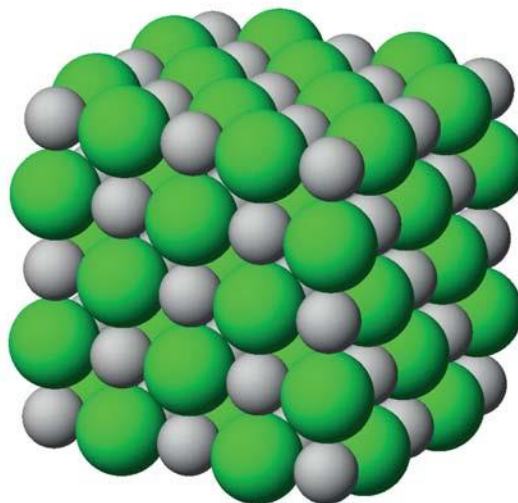
NaCl lattice viewed along alternating close-packed layers of Na⁺ and Cl⁻ ions (8 unit cells, scale=1)

12 Identification and Execution of Symmetry Operations

Contributed by William Jordan, Pacific University

Overview

In this exercise you will learn how to locate elements of symmetry in molecules and how to demonstrate the results of symmetry operations.

Recommended Exercises

Complete Exercises I, II, III, and IV before beginning this experiment.

Background

Molecular symmetry is of great value in understanding molecular bonding and spectroscopy. The symmetry of a molecule is expressed as a collection of symmetry operations or elements of symmetry. A molecule with a particular set of symmetry elements is said to belong to a particular point group. Molecular orbitals and vibrations within molecules also have symmetry properties.

Symmetry operations are manipulations of a molecule which result in indistinguishable representations of the molecule; that is, a molecule looks no different before and following a symmetry operation. The operation may exchange two identical atoms, or it may leave one or more atoms unmoved. The symmetry operations are as follows:

Identity (*E*): The identity is a trivial symmetry operation which involves doing nothing to the molecule. It is mentioned here only for the sake of completeness. All molecules possess an identity and, in some molecules, the identity is the only element of symmetry.

Axis of Rotation (C_n): An n-fold axis of rotation is a rotation about an axis by 360°/n. Often, but not always, rotation axes correspond to bond axes. For example, SF_6 possesses a C_4 axis which passes through opposing S-F bonds.

Rotation by 90° about the axis rotates each fluorine atom perpendicular to the axis into a position previously occupied by another fluorine atom.
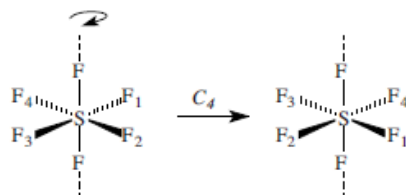
SF$_6$ also possesses 3 C_3 axes of rotation which do not pass through bonds. In the following representation, the C_3 axis is perpendicular to the plane of the paper and passes through the sulfur atom.
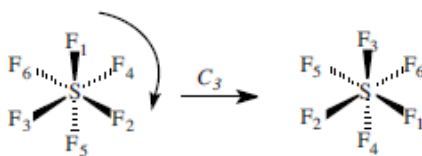
Rotate a molecule

1. Center the molecule in the window by selecting the entire molecule and pressing ctrl + f on the keyboard. It may resize the molecule in the process of centering it. If the molecule is not centered in the window, it will wobble as the rotation commands are executed.

2. Rotate the molecule until the rotation axis is perpendicular to the screen. You should now be looking directly down the rotation axis, which should be centered in the window.

3. Press r on the keyboard to enter rotation mode.

4. Press **Z** to rotate in the clockwise direction and **z** to rotate counter-clockwise. Rotate around the x and y axes by pressing **x**, **X**, **y** or **Y**. Rotation occurs by 5° each time. Thus nine operations result in a 45° rotation, 12 - 60°, 18 - 90°, 24 - 120°, etc. You may rotate continuously by pressing the Space bar. To halt continuous rotation, press the Space bar again.

Plane of Symmetry or Reflection (σ): A mirror plane of symmetry is an operation in which reflection of an atom from one side of the plane to the other results in an indistinguishable representation of the molecule. Atoms which lie within the plane are unmoved. Examples of mirror planes are given below.

92
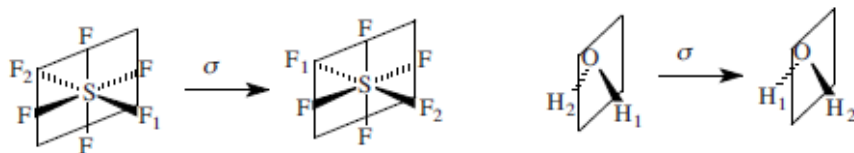
Reflect through a mirror plane

1. Orient the molecule so that the plane is parallel to the right and left sides of the window (perpendicular to the window or screen).
2. Select **Action | Selection | Mirror**.

Improper Rotation or rotation-reflection (S_n): An improper rotation is the result of two operations in sequence; rotation by 360°/n, followed by reflection through a mirror plane perpendicular to the axis of rotation. An example of an improper axis of rotation is shown below.
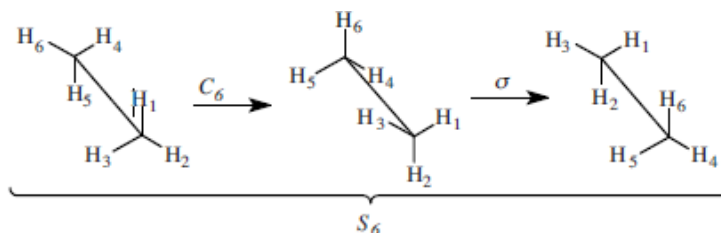
Execute and demonstrate an improper rotation

1. Follow the procedures for rotation and reflection in sequence.
2. To demonstrate the example above, the following steps would be executed:

Action | Selection | Mirror

III   IV

reorient to compare
oto 1st orientation(I)

press y, 6 times

IV   IV
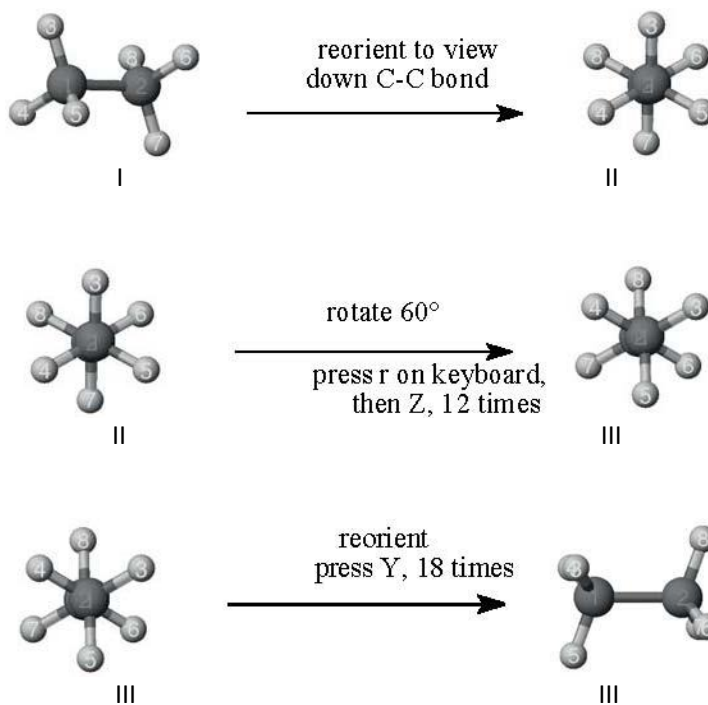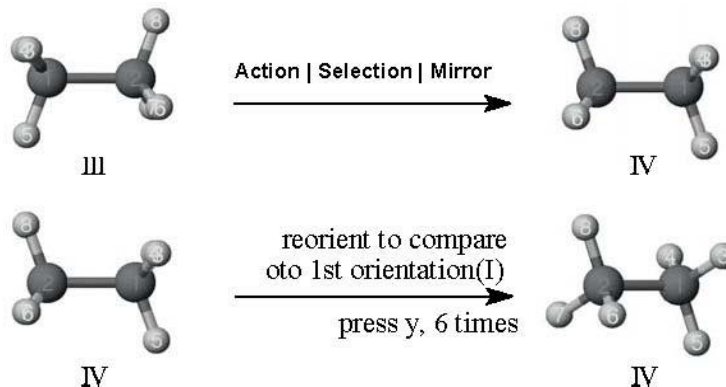
The results of the successive operations of rotation and reflection are shown below.

C_3 σ

I   IV
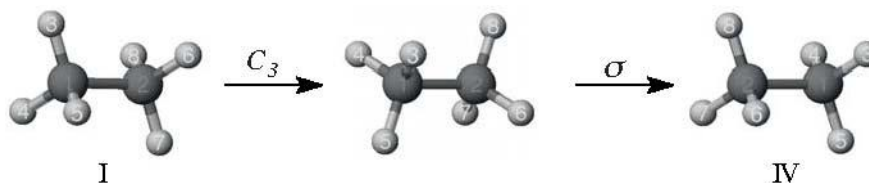
Inversion (*i*): An inversion through the center of a molecule can usually be accomplished by a rotation by 180° followed by reflection through a plane (Mirror operation) perpendicular to the rotation. Note that this is identical to an S_2 as in the example below (rotation by 180° about the C-C axis followed by reflection through a plane perpendicular to the C-C axis).
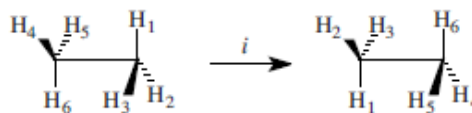
For a planar molecule, an inversion is simply accomplished via rotation by 180° about an axis perpendicular to the plane since subsequent reflection through the plane leaves the molecule unchanged.

Modeling Section

Methane has several elements of symmetry, some of which are challenging to visualize. Build a model of CH_4 and view the atoms by number and the bonds as cylinders. Center the molecule in the window by selecting only the carbon atom and pressing **Control + f**.

For each symmetry element, you should have one document to show the original orientation and one to show the result of the operation. Orient the molecule so as to make apparent the symmetry element (e.g. looking down a rotation axis). Adjust the size of the window so that you can simultaneously view documents showing the original atom placements and those following a symmetry operation. Once you have completed the operation, copy the structures and paste them into a word

94

processing document.

Report Guidelines

Label (e.g. C_3), and give the number of each type of symmetry element in the molecule to accompany the examples generated using SCIGRESS.

Instructor Notes

This exercise could be adapted to numerous other molecules and polyatomic ions.

Typical Results

There are 4 types of symmetry elements in the T_d point group, 4 C_3 rotation axes, 6 σ_d mirror planes, 3 C_2 axes, and 4 S_4 improper axes.
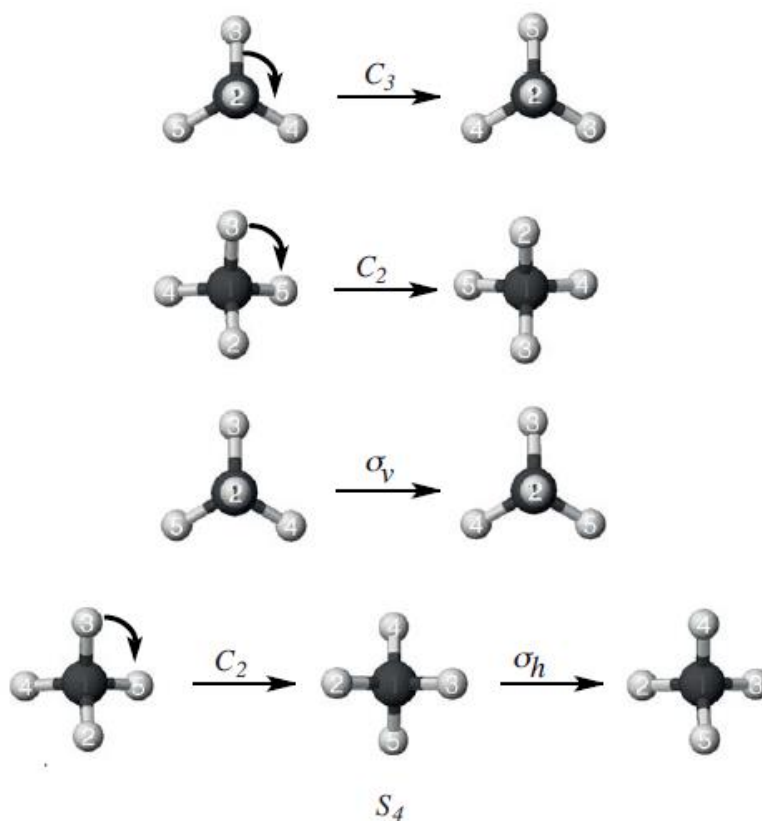
The Relationship Between Infrared Spectra and Molecular Geometry

Contributed by William Jordan, Pacific University

Overview

In this exercise you will use SCIGRESS to compute the infrared spectra of small inorganic molecules. Group theoretical methods will be used to predict the number and symmetries of molecular vibrations possessed by a molecule.

Recommended Exercises

Complete Exercises I, II, III, IV, and VI before beginning this experiment.

Background

One of the applications of symmetry in chemistry is the prediction of vibrations which will be active in infrared (IR) and/or Raman spectroscopy. Such predictions can be confirmed by using MOPAC to generate IR spectra and to examine the vibrations responsible for the IR peaks. As an example, the symmetries of IR active vibrations in a water molecule are worked out below. The same method is applicable to other molecules. The water molecule belongs to the C2v point group for which the character table is as follows:

> **NOTE**
>
> A basic understanding of the chemical applications for symmetry and group theory is necessary. It is advisable to review the relevant material in an inorganic chemistry text or other reference such as F.A. Cotton, *Chemical Applications of Group Theory*, 3rd ed., Wiley, New York, 1990.

C_{2v}	E	C_2	$\sigma(xz)$	$\sigma(yz)$		
A_1	1	1	1	1	z	x^2, x^2, x^2
A_2	1	1	-1	-1	R_z	xy
B_1	1	-1	1	-1	x, R_y	xz
B_2	1	-1	-1	1	y, R_x	yz
Γ	9	-1	3	1		
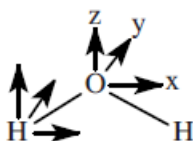
1. Determine each atom's contribution to the character of the transformation matrix under each symmetry operation. If an atom moves, it contributes 0 to the character. If an atom remains stationary, determine the contribution of each coordinate. If a coordinate is unchanged, it contributes 1; if it is reversed in sign, it contributes -1; if it is something in between, it makes a fractional contribution.

   a. The consequences of each operation type are summarized as follows:

E: All atoms remain and all coordinates are unchanged. Each coordinate contributes 1 to the character of the reducible representation. Thus, for water, the character for the E element is 9.

C_2: The hydrogens move and thus contribute 0. Oxygen remains stationary and the z coordinate remains unchanged, contributing 1. However, x and y are both reversed in sign and thus each contributes -1. The character is thus 1-1-1 = -1

$\sigma_v(xz)$ (plane of molecule): All atoms remain unchanged. The z and x coordinates are unchanged, each contributing 1. Each y coordinate is reversed, contributing -1. For each atom, the contribution is 1 and the character is 1+1+1 = 3

$\sigma_v(yz)$: Only O is unchanged. The z and y coordinates contribute 1 each. The x coordinate is reversed and contributes -1. Thus the character is 1+1-1 = 1.

b. The reducible representation is **Γ = 9 -1 3 1**

2. Determine the irreducible representations of the coordinates by making use of a group property.

$$n_i = 1/order \ \Sigma\left|[(\chi_{RR})(\chi_{IR})]\right|$$

where :

n_i = number of irreducible representations (IRs) having a symmetry species i

χ_{RR} = the character of the reducible representation and χ_{IR} = the character of the irreducible representation

For water, the order is 4:

$$n_{A1} = 1/4 \ [(9)(1) + (-1)(1) + (3)(1) + (1)(1)] = 3$$

$$n_{A2} = 1/4 \ [(9)(1) + (-1)(1) + (3)(-1) + (1)(-1)] = 1$$

$$n_{B1} = 1/4 \ [(9)(1) + (-1)(-1) + (3)(1) + (1)(-1)] = 3$$

$$n_{B2} = 1/4 \ [(9)(1) + (-1)(-1) + (3)(-1) + (1)(1)] = 2$$

3. The reducible representation represents the sum of the motions of all of the coordinates: 3 of which result in translation (x, y, z) and 3 of which result in rotation (R_x, R_y, R_z) for a nonlinear molecule. Thus, for water, there are 3 vibrations. If the irreducible representation of a vibration is that of a coordinate (x, y, or z), it will be IR active. Subtract the irreducible representations of translations and rotations.
a. Translation (x = B_1, y = B_2, z = A_1)
b. Rotation (R_x = B_2, R_y = B_1, R_z = A_2)
c. The vibrations which remain:

A_1 3 - 1 = 2 B_1 3 - 2 = 1

A_2 1 - 1 = 0 B_2 2 - 2 = 0

4. Of the 3 vibrations, 2 are A_1 symmetry and 1 is B_1 symmetry. A_1 is the irreducible representation of the z coordinate and B_1 is the irreducible representation of the x coordinate so all three vibrations are IR allowed.

Note that, for a particular point group, **Γ** will depend upon the number of atoms in the molecule.

If a symmetry operation is a multiple (e.g. $2C_3$ in C_{3v}) the contribution to the irreducible representation must be multiplied by that multiple in calculating the number of irreducible

representations.

If a vibration belongs to a degenerate irreducible representation (*E* or *T*), then there will be a single vibrational frequency, corresponding to 2 (*E*) or 3 (*T*) degenerate transitions. Likewise, when accounting for rotations and vibrations, if 2 rotations, say, belong to an *E* irreducible representation, that accounts for 1 *E* irreducible representation.

Modeling Section

Simulation of the Infrared Spectrum of $H_2O(g)$

1. Build a water molecule in the Workspace.
2. Optimize the molecule with MOPAC using PM3 parameters. Append a Force calculation. When you append calculations, MOPAC executes the selected calculations in sequence; in this case, a geometry optimization followed by an IR/vibrational spectrum.
   ❏ Open a new experiment and select **chemical sample, IR transitions**, and **MO-G PM3 FORCE**. The experiment manager automatically does an optimization followed by an IR calculation. Select **Run**.
3. View the vibrational spectrum.
   ❏ Activate the molecule window after the calculations are complete. Select **Analyze | IR Transitions**. An IR spectrum appears in a separate window. Each separate vibration is represented by a peak with a triangle symbol at the minimum.
4. Adjust the x and y axes as needed. For example, the water spectrum should be displayed between 0 and 4000 cm^{-1} and should have no peaks at frequencies below 1000 cm^{-1}. To check frequencies above 4000 cm^{-1}, double click just below the horizontal axis to open the Axis Attributes/ Axis Parameters dialog box.

   ❏ Enter a Wavelength Range of 6000 to 1000 cm^{-1}. Click OK to rescale the spectrum. Rescale the spectrum, as desired, to view parts of the spectrum in greater detail.

5. Resolve overlapping peaks as needed.

   ❏ Click on a curve with the Select tool. A red line appears, defining

6. Click on a peak symbol to view the vibration corresponding to a particular peak. The structure window will rearrange to show the atom motions involved in the vibration. To better view the vibration, click anywhere in the structure window and reorient the structure.

7. Paste the spectrum and vibrations into a word processing document and assign each vibration to an irreducible representation.

N₂F₂

1. Build a molecule of N₂F₂ in the Workspace. Optimize and execute a Force calculation as in the case of the water molecule.
2. View the molecule and determine the point group. Obtain the character table for the point group.
3. Use the methodology given in the Background section to determine the number and symmetries of the vibrations of the molecule. Determine which of the vibrations are IR active.
4. View the computed IR spectrum of N₂F₂. View the vibrations, both IR active and IR inactive, and assign their symmetries (irreducible representations).

Report Guidelines

✎ Paste the spectra and vibrations for water into a word processing document and assign each vibration to an irreducible representation.

✎ Use group theoretical methods to determine the symmetries of the vibrational modes of N_2F_2 and determine which vibrations are IR active.

✎ Paste the spectra and vibrations for N_2F_2 into a word processing document and assign each vibration to an irreducible representation.

Instructor Notes

Typical Results

Water
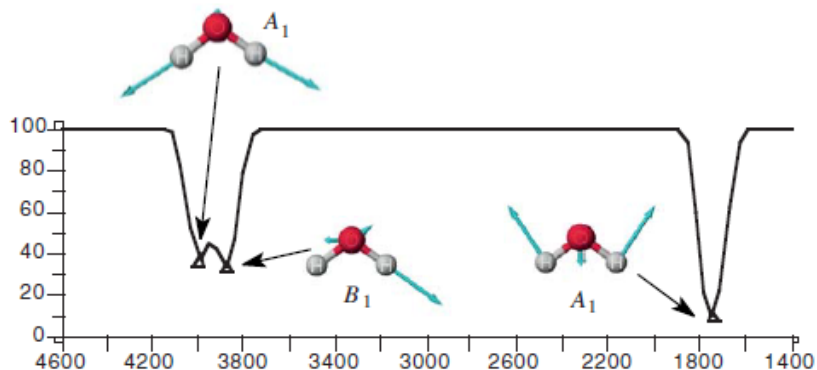
N2F2

Character table for C_{2h} point group

C_{2h}	E	C_2	i	σ_h		
A_g	1	1	1	1	R_z	x^2, y^2, z^2, xy
B_g	1	-1	1	-1	R_x, R_y	xz, yz
A_u	1	1	-1	-1	z	
B_u	1	-1	-1	1	x, y	
Γ	12	0	0	4		

Number of each symmetry species:

$$nA_g = \tfrac{1}{4}\,[(12)(1) + (0)(1) + (0)(1) + (4)(1)] = 4$$

$$nB_g = \tfrac{1}{4}\,[(12)(1) + (0)(-1) + (0)(1) + (4)(-1)] = 2$$

$$nA_u = \tfrac{1}{4}\,[(12)(1) + (0)(1) + (0)(-1) + (4)(-1)] = 2$$

$$nB_u = \tfrac{1}{4}\,[(12)(1) + (0)(-1) + (0)(-1) + (4)(1)] = 4$$

Reducible representations of translations (A_u, $2B_u$) and rotations (A_g, $2B_g$).

Remainder are vibrations: $3A_g$, $1A_u$, $2B_u$

IR active modes: $1A_u$, $2B_u$
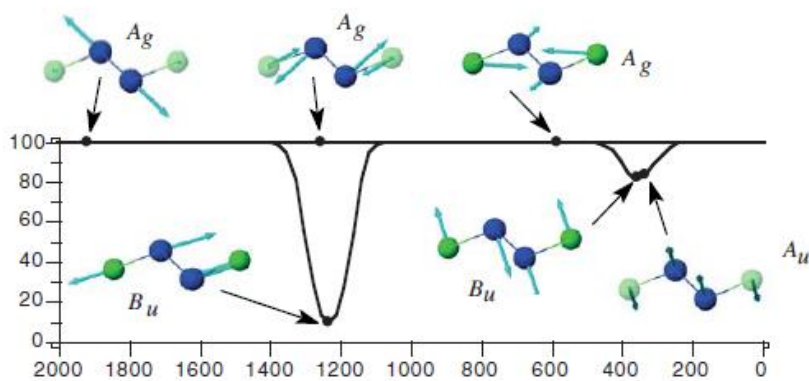
Further Discussion

An additional system worth studying is SF_4. The structure is based on a trigonal bipyrimidal distribution of electron pairs, and since there is one lone pair, there are two structures possible. The two structures do not have the same number of IR active transitions, so that infrared spectroscopy provides an experimental criterion for determining the correct structure. One structure belongs to the C_{2v} point group and prediction of the number of IR active vibrations is straightforward. The other belongs to C_{3v} and here transformation of the atoms under the C_3 operation is more difficult.

SF_4

The structure of SF_4 is based upon a trigonal bipyramidal distribution of electrons, one being a lone pair. Thus there are two possible structures one with the lone pair in an axial position and one with the lone pair in an equatorial position.

Transmittance (%)

Wavenumber (cm-1)

SF4_C3V
SF4_C2V

14   Molecular Orbitals of Simple Molecules
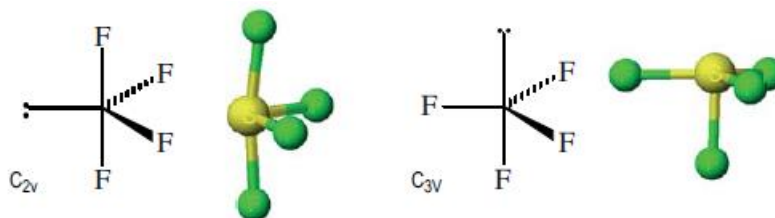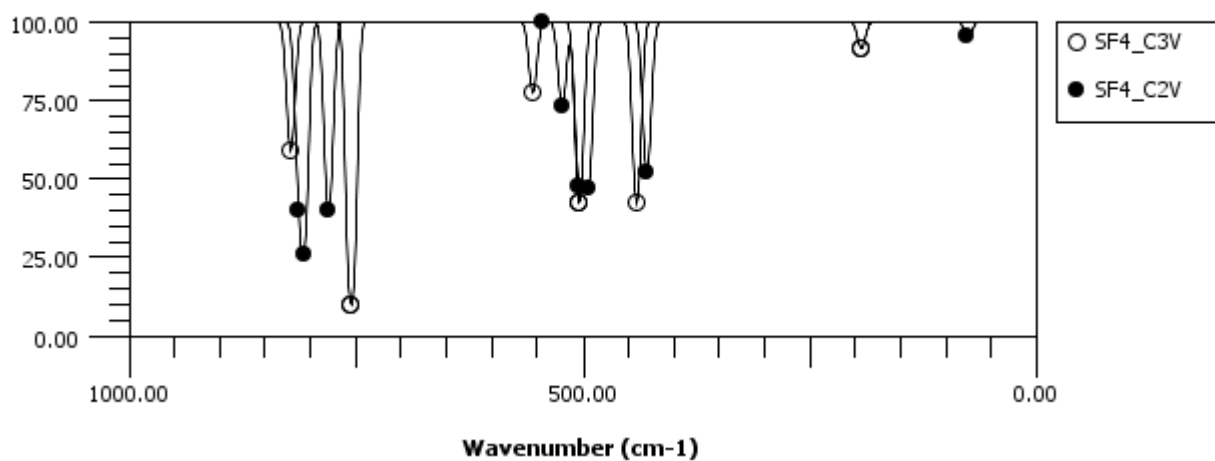
Contributed by William Jordan, Pacific University

Overview

You will use SCIGRESS to compute and view molecular orbitals for small inorganic molecules. By viewing a molecular orbital it is possible to deduce the atomic orbital combination from which it is constructed. The objective is to understand how atomic orbitals combine to form molecular orbitals.

Recommended Exercises

Complete Exercises I, II, III, IV, VI, and VII before beginning this experiment.

Background

Molecular orbital theory is based on the presumption that electrons in molecules reside in molecular orbitals (MOs). MOs are typically spread over two or more atoms. Quantum mechanically, MOs are constructed by combining atomic orbitals (AOs). In the linear combination of atomic orbitals (LCAO) approach, AOs are simply added together. If there is net positive overlap between orbitals, a bonding MO results. If there is net subtractive overlap, an antibonding MO results. If there is no net overlap, the resultant MO is nonbonding. Thus AOs combine only if they have compatible symmetries.

Bonding AO combination   Antibonding AO combination   Nonbonding AO combination

It is also the case that AOs combine more strongly as their energies approach equality. Thus a bonding MO formed from AOs of dissimilar energies will have a greater contribution from the lower energy AO and the corresponding antibonding MO will have a greater contribution from the higher energy AO, as shown below.

103
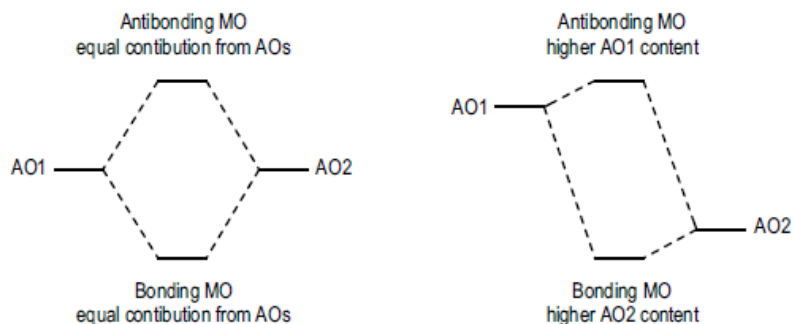
Some AO combinations, together with the resultant MOs, calculated with SCIGRESS software, are shown below.

Bonding AO Combination   Bonding MO   Antibonding AO Combination   Antibonding MO
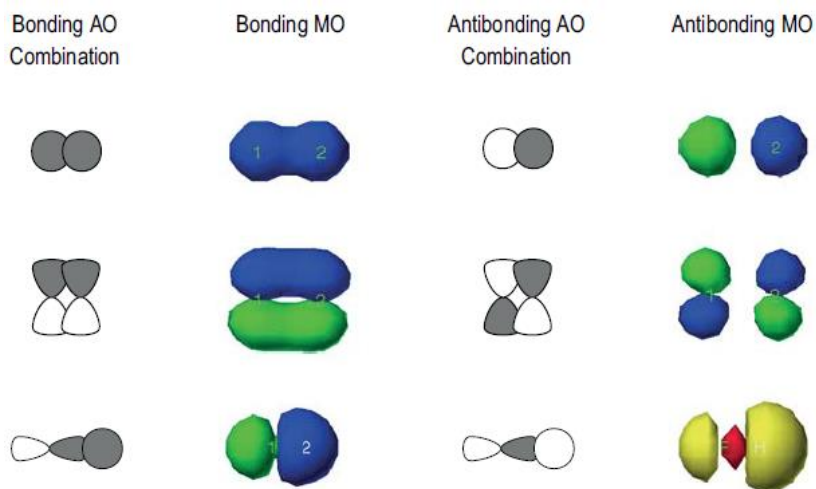
 The number of MOs in a molecule must be equal to the number of AOs in the constituent atoms.

Modeling Section

1. Build the molecules CO, H2O and O_3 (ozone, O=O-O) in the Workspace.
2. Optimize the structures using MOPAC with AM1 parameters.
3. Compute the molecular orbitals for the molecules.
   ❏ Select Analyze | Electronic Structure | HOMO-5 to LUMO+4.
4. View the molecular orbitals after the calculations finish.
   ❏ To view the molecular orbitals doubleclick on the molecular orbital in the Model Explorer.
5. Create a list of the MO energies for each molecule using Spreadsheet. Enter the molecule(s) of interest into the Chemical Sample column. Select Property of: molecular orbital, energy and standard procedure. Select Evaluate Cell.
6. Orient the molecule so that each of the MOs can be viewed clearly and paste a copy of each of the MOs into a report document. Prior to copying, press f to center the molecule in the window. Also, reduce the window to just enclose the molecule so that the pasted image will not occupy more space than necessary.

Report Guidelines

✎ Arrange the MOs vertically in the report from highest (top) to lowest in energy as in an energy level diagram.

✎ Identify the orbital with an appropriate symbol (e.g. $\sigma 2s$) and give its energy.

✎ Identify the MO as bonding, antibonding or nonbonding and give the number of electrons occupying the orbital or set of degenerate orbitals. If you have chemical graphics software, show the atomic orbital combination which gives rise to the MO as in the example below.

✎ Account for the number of molecular orbitals

✎ Explain the bond order.

Instructor Notes

This exercise may easily be expanded to include molecules with other geometries, e.g. NH_3, CH_4. MOs involving d-orbitals will be examined in later exercises. The exercise may also be shortened by omitting ozone. The basic ideas of molecular orbital formation are covered by considering CO and H_2O. However, ozone illustrates delocalization of -bonding over several atoms.

The exercise may be refined by using the results to construct MO energy level diagrams. This may be done using graphics software to scale the relative spacing of energy levels in accordance with the computed MO energies. Examples are given below in the Typical Results.

Typical Results

Carbon monoxide

$\sigma^* 2p$

$\pi^* 2p$

$\sigma 2p$

$\pi 2p$

$\sigma^* 2s$

$\sigma 2s$

Water

antibonding

antibonding

nonbonding

bonding
(slightly)

bonding

bonding
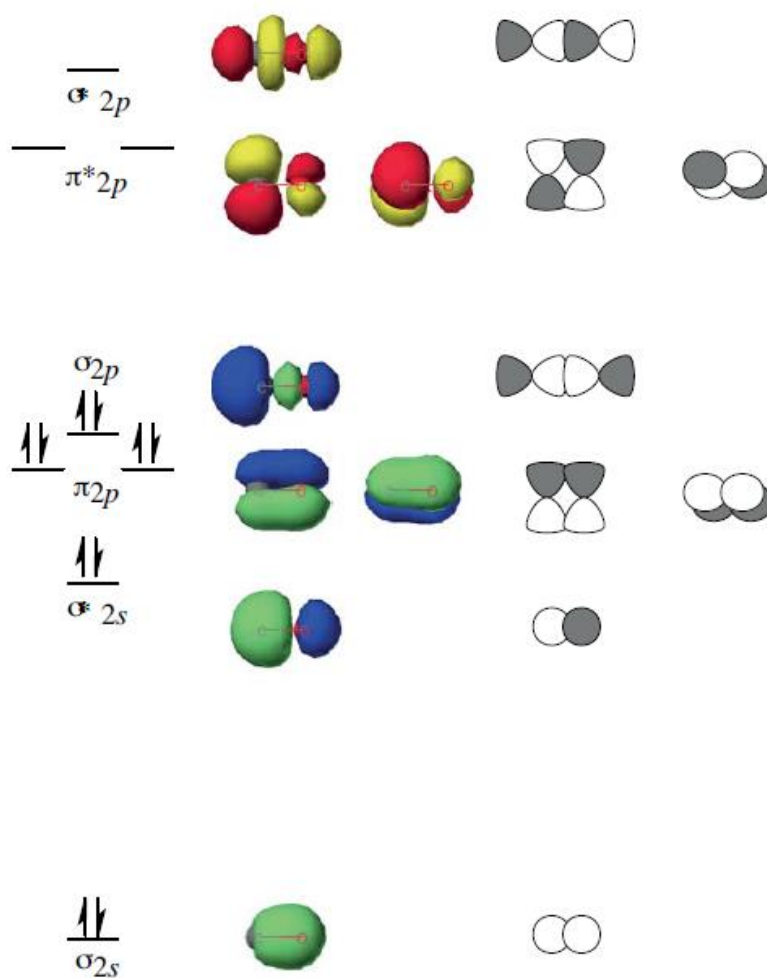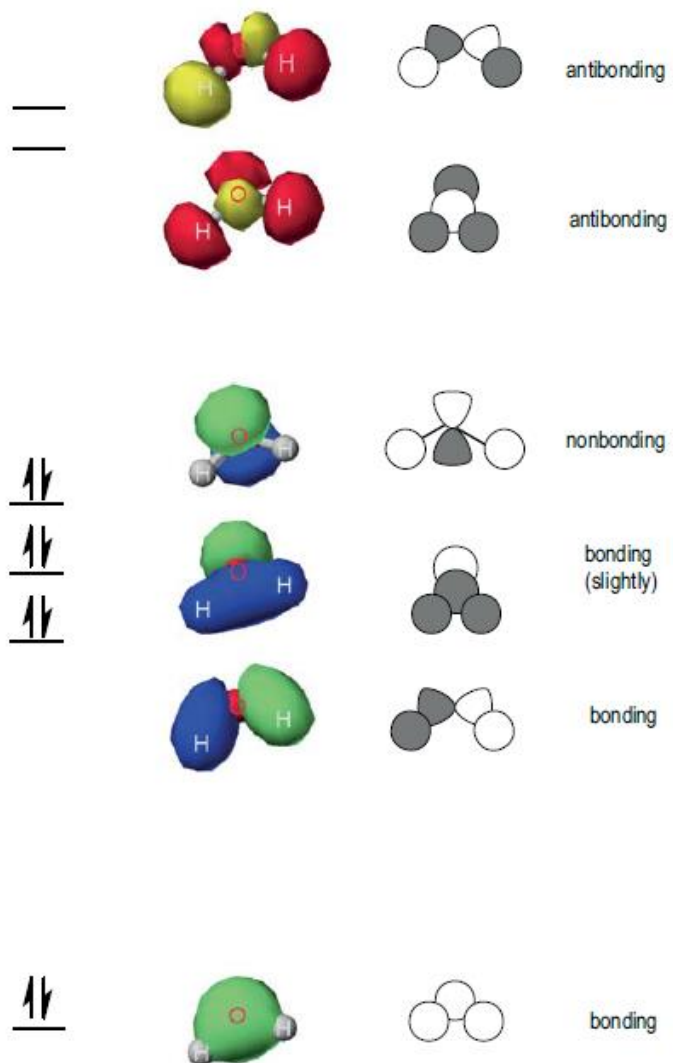
15 Molecular Orbitals of Octahedral Transition Metal Complexes

Contributed by William Jordan, Pacific University

Overview

In this exercise you will use SCIGRESS to compute the molecular orbitals of several iron and cobalt octahedral complexes. From visual inspection of the molecular orbitals, you can determine the involvement of metal atomic orbitals in molecular orbital formation. You will also examine the effect of the metal ion and oxidation number on bonding by comparing MO diagrams for analogous Fe(II) and Co(III) complexes.

Recommended Exercises
Complete Exercises 1, II, III, IV, V, VI, VII before beginning this experiment.

Background

The principles that govern bonding in compounds of the main group elements are also applicable to bonding in transition metal complexes. However, transition metal ions are characterized by partially filled valence *d*-orbitals with no electrons in the valence s and p orbitals. Thus the highest energy electrons are typically in molecular orbitals derived from metal *d*-atomic orbitals.

While molecular orbitals are properties of the molecule, a particular molecular orbital may be localized more on the metal atom or on one or more ligands. Some MOs are localized entirely on the ligands and have essentially no metal contribution. These are derived from ligand atomic orbitals which do not overlap with metal atomic orbitals. The molecular orbitals which result from σ-overlap between metal and ligand orbitals are represented below together with their symmetry labels. Note that metal-ligand π-bonding, possible for some ligands, is not included.

a_{1g} $t1u$

eg $d_{xz}, d_{yz}, d_{xy}, \; t2g$ Set
Nonbonding
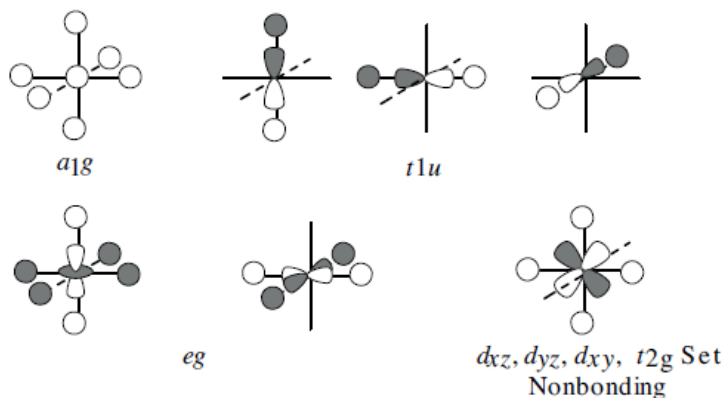
Below is a valence molecular orbital energy level diagram for an octahedral transition metal complex, MX_6. Only σ-bonding is included; ligand orbitals not involved in binding to the metal are omitted. The diagram would look different for a different geometry (e.g. tetrahedral or square planar) or for mixed ligands in an octahedral geometry (e.g. MA_4B_2). If -bonding between metal and ligands were included, the t_{2g} orbitals would no longer be nonbonding and would be used to form a set each of 3 bonding and 3 antibonding orbitals.

Regardless of the metal ion or ligands, the six ligands will provide a total of six pairs of electrons to the σ-bonding system. In the diagram, these ligand pairs are each represented by the letter "L", placed arbitrarily in the six lowest energy molecular orbitals. This leaves the non-bonding t_{2g} and antibonding $e_g{}^*$ orbitals available for occupation by 0 -10 d electrons from the metal ion
[e.g. 6 for Fe(II), 5 for Fe(III)]. Thus if there are more than 6 d electrons in a transition metal ion, the d electrons will have different energies in the complex ion (some in t_{2g} and some in $e_{2g}{}^*$). The energy gap between the t_{2g} and $e_g{}^*$ orbitals is a fundamental property of electronic structure of octahedral complexes and is labeled as Δ_0 or $10D_q$. Energetically, placement of an electron in a t_{2g} molecular orbital is preferable to occupation of an $e_g{}^*$ orbital.
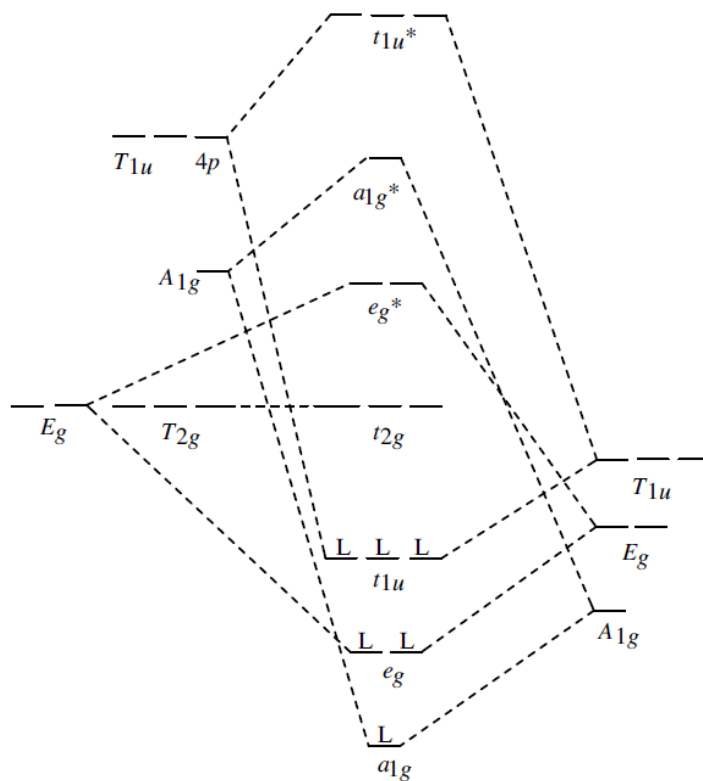
Modeling Section

1. Build models of $[CoF_6]^{3-}$ and $[FeF_6]^{4-}$ and optimize with Molecular Mechanics (MM2).
2. Calculate the energy using ZINDO.
   ❏ Select **Experiment | Run Single Engine | ZINDO**, Job: SCF Energy.
3. Generate the molecular orbitals. Assume low spin (no unpaired electrons).
   ❏ Select **Analyze | Electronic Structure | HOMO-5 to LUMO+4**.

4. View the molecular orbitals after the calculations finish.

✎ The number of MOs should be large; is it the number expected?
❏ Double click on the molecular orbital in the Model Explorer window to activate it.

Report Guidelines

✎ Explain the number of molecular orbitals obtained.

✎ Note MOs which involve metal *d*-orbitals and identify the *d*-orbitals involved. Paste depictions of these MOs into a word processing document together with your comments.

✎ Which MOs contain the metal *s* and *p* orbitals. Are they bonding and/or antibonding, nonbonding? Why?

✎ Describe the highest occupies molecular orbital (HOMO(s)) and the lowest occupied molecular orbital (LUMO(s)).

✎ Compare the results for $[CoF_6]^{3-}$ and $[FeF_6]^{4-}$ and note any significant differences. How can the differences be explained?

Instructor Notes

Typical Results

There should be a total of 33 molecular orbitals for MF_6 since there are nine valence atomic orbitals in the metal ion and four for each of six ligands. Following is a partial molecular orbital energy level diagram for $[CoF_6]^{3-}$. Only the MOs which are formed from metal orbitals are included.

16.97 eV   —— —— ——
              $t_{1u}*$

13.65 eV   ——
              $t_{1g}*$

9.28 eV   —— ——
              e_g*

-1.23 eV   —— —— ——
              $t_{2g}*$

-1.85 eV   —— ——
              e_g

-2.56 eV   —— —— ——
              t_{2g}
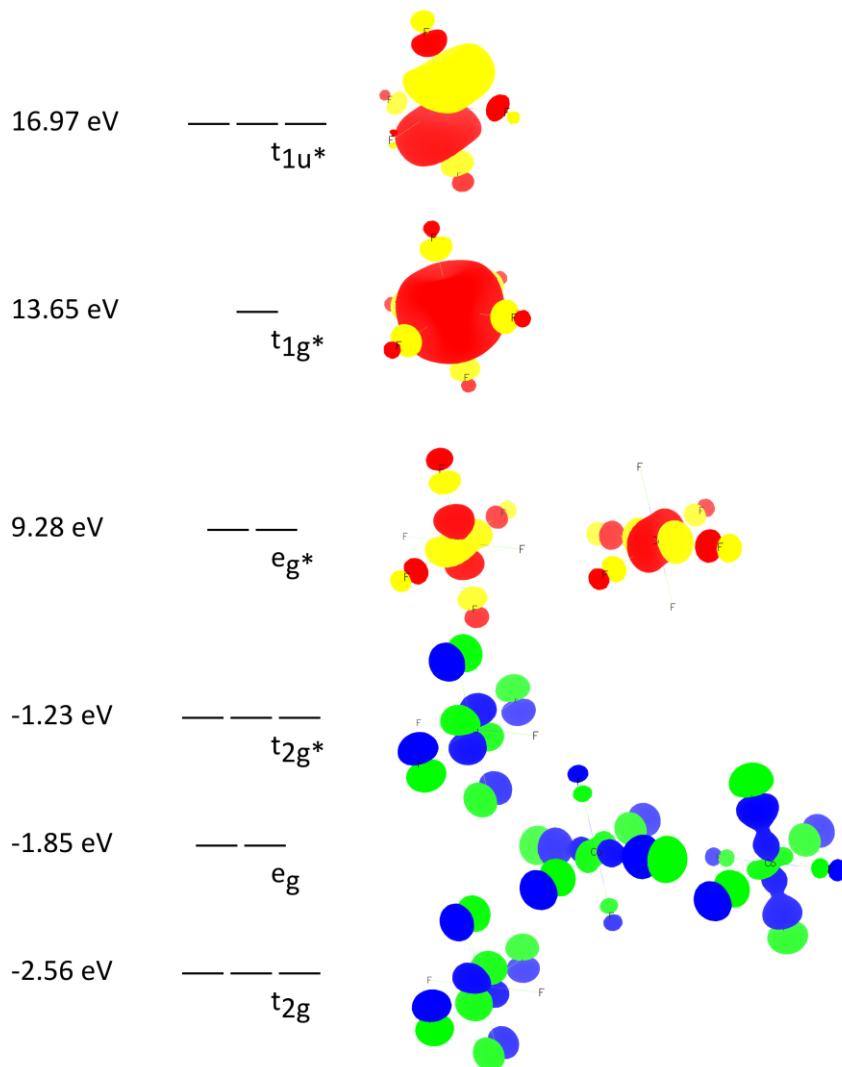
The metal d t_{2g} orbitals enter into the formation of -bonding and antibonding MOs with ligand group orbitals. The result is that the HOMOs are -antibonding MOs instead of t_{2g} nonbonding MOs. This lowers the energy difference between the HOMOs and LUMOs, which is known as Δ_0, or $10D_q$. In the above diagram, the metal s and p orbitals are chiefly represented in unoccupied antibonding MOs.

The energies of the HOMO and LUMO are lower for $[CoF_6]^{3-}$ than $[FeF_6]^{4-}$ see diagram below) in part because the higher charge density on Co(III) than Fe(II) results in lower energy of corresponding atomic orbitals on the complex and because the more highly charged metal ion interacts more strongly with ligand donor atoms.
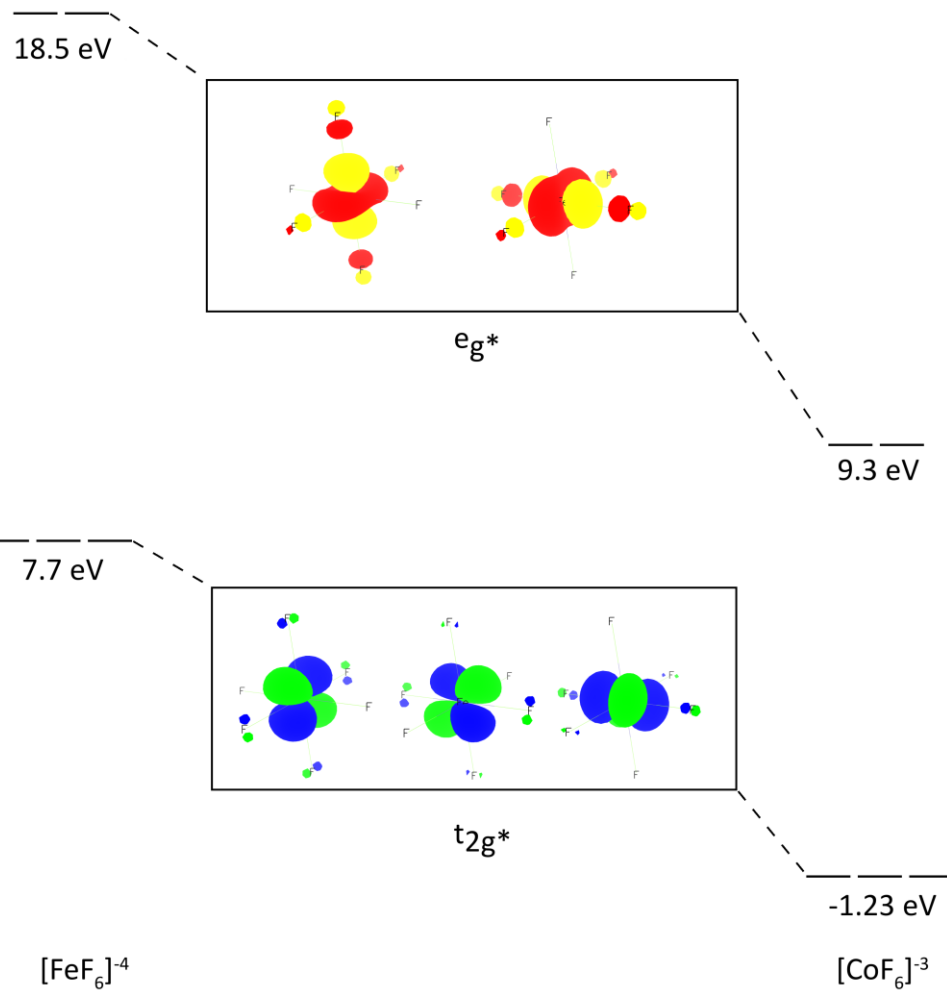
18.5 eV

9.3 eV

e_g*

7.7 eV

-1.23 eV

t_2g*

$[FeF_6]^{-4}$

$[CoF_6]^{-3}$

16   Relative Ligand Field Strengths

Contributed by William Jordan, Pacific University

Overview

In this exercise you will use SCIGRESS to compare the effect of the ligand on the energies of MOs derived from the metal d-orbitals in octahedral transition metal complexes.

Recommended Exercises

Complete Exercises I, II, III, IV, V, VI, VII and Experiment 15
before beginning this experiment.

Background

Most organic compounds and small inorganic compounds are colorless because the energies necessary to excite the lowest lying electronic transitions are in the ultraviolet. In contrast, most transition metal compounds absorb light in the visible region and thus are colored. The relatively low energy needed to excite valence electron of metal complexes is due to the fact that the lowest energy transitions occur between closely spaced electronic states derived from metal *d*-orbitals.
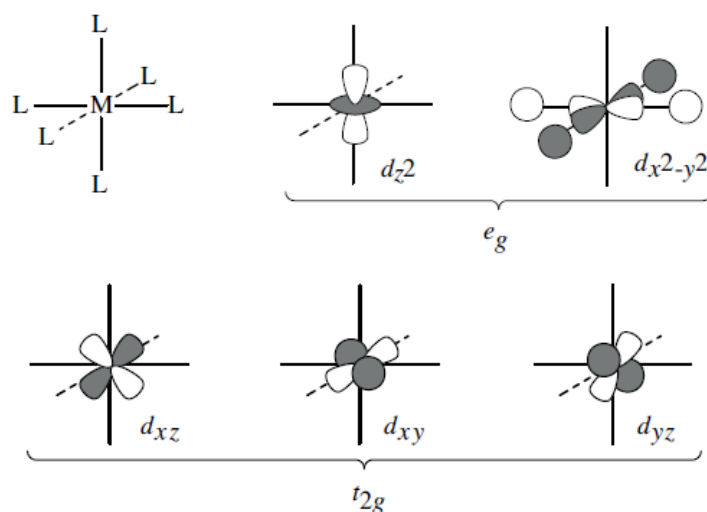
Octahedral coordination is by far the most common geometry by which transition metals and ligands combine. It has long been recognized that, when six identical ligands bond octahedrally to a transition metal ion (O_h point group), the metal *d*-orbitals are split into two groups, a triply degenerate t_{2g} set, and an e_g set at higher energy. The energy difference between the two sets is the octahedral ligand field

splitting energy (Δ_0 or $10D_q$).
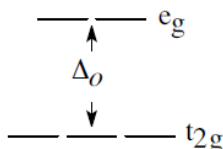
The magnitude of Δ_0 depends upon the metal ion and the ligand. For example, $M(NH_3)_6$ will have different values of Δ_0 where M = Fe(II), Fe(III) and Co(III). The order in which Δ_0 changes as a function of the ligand, is called the spectrochemical series because it is determined experimentally by spectroscopic measurements.

In order for an electronic transition of energy Δ_0 to occur, there must be at least one electron in a t_{2g} orbital and at least one vacancy in an e_g orbital. Thus the t_{2g} and e_g orbitals involved are among the highest occupied molecular orbitals (HOMOs) or lowest unoccupied molecular orbitals (LUMOs).

The simplest case of metal ligand interaction is σ-bonding where each ligand donates an electron pair to a covalent bond with the metal ion. In this case, the t_{2g} orbitals are non-bonding. Why are they non-bonding? When the ligands are able to engage in -bonding to the metal ion, the molecular orbitals derived from metal d-orbital involvement are altered and the value of Δ_0 is affected. States arising from metal *d*-orbitals are split in geometries other than octahedral as well.

Modeling Section

1. Build models of the following octahedral complexes: $[FeF_6]^{4-}$, $[Fe(OH)_6]^{4-}$, $[Fe(H_2O)_6]^{2+}$, $[Fe(NH_3)_6]^{2+}$. Assume that the complexes are low spin (use coordinate bond type to connect ligands with metal ion).
2. Open new Spreadsheet and enter each of the iron complexes into the Chemical Sample column.
3. Program the rest of the columns in the following way:
   - ❏ Column B. **Property of: chemical sample, energy total,** and **MM2**. This method evaluates total energy via a ZINDO calculation following optimization using Molecular Mechanics.
   - ❏ Column C. **Property of: chemical sample**, **HOMO energy,** and **extract from sample**.
   - ❏ Column D: **Property of: chemical sample**, **LUMO energy**, and **extract from sample**.
   - ❏ Column E: Select **Analysis** followed by **analysis equation** and **algebraic**. Set up an algebraic equation to calculate the difference in energy between the LUMO and HOMO for each complex. Enter the equation as Δ=D-C or D-C, where D and C are the column letters of the LUMO and HOMO energies respectively.
4. (Optional) Tabulate the most relevant MOs (from HOMO -2 to LUMO +2) for one or more of the complexes to view the MOs involved.
   - ❏ Select **Analyze | Electronic Structure | HOMO-5 to LUMO+4**

Report Guidelines

Paste your Project Leader results into a word processing document.

- ✎ Compare the energy differences (LUMO - HOMO) to assess relative field strengths of the different ligands. Try to explain why the ligand field strength increases in the order observed. Numerous textbooks on inorganic chemistry and coordination chemistry may be referenced to better understand this phenomenon.

- ✎ Are the t_{2g} and e_g orbitals of the LUMOs and HOMOs bonding, antibonding or nonbonding?

- ✎ Is -bonding involved in any of the complexes? If so, paste copies of the tabulated MOs into the report.

- ✎ How does the -bonding affect Δ_0 in these cases?

- ✎ The method used by Spreadsheet to execute the ZINDO calculation assumes singlet multiplicity. Is that true in the case of Fe(II) complexes?

Instructor Notes

The same exercise can be done for other metal ions. The ligand field strength is most easily assessed for metal ions in which the orbitals of the HOMO t_{2g} set are equally occupied; e.g. Cr^{3+} or V^{2+} (d^3) and Fe^{2+} or Co^{3+} (low spin d^6). Low spin d^6 complexes are the simplest to evaluate because the ground state is a singlet. Thus Project Leader can be used without altering the standard procedure.

If the metal ion has 1 or more unpaired electrons (e.g. 3 for Cr^{3+}), it may be best to first optimize each model using Molecular Mechanics. This can be done in Project Leader by selecting energy steric as the property and using the standard procedure. The determination of the energy using ZINDO should be done separately for each model. Be sure to enter the correct spin multiplicity (quartet for Cr^{3+}) when you run ZINDO. ZINDO uses a self-consistent field (SCF) method for determining energy. There are 3 SCF types used by ZINDO. If the molecule is a singlet (no unpaired electrons), you may use the default type (RHF). If there are one or more unpaired electrons, you should specify ROHF.

Typical Results
FeX$_6$

Ligand, X	HOMO energy (eV)	LUMO energy (eV)	Difference (eV)
OH	8,975	19,657	10,682
F	7,663	18,467	10,804
H2O	-19,114	-8,258	10,856
NH3	-17,371	-6,460	10,911

The above results are consistent with the experimentally determined

spectrochemical series for these ligands; $NH_3 > H_2O > F^- > OH^-$. The closer in energy are the metal and ligand orbitals which combine to form MOs, the larger will be Δ_0. When only σ-overlap is possible, the t_{2g} orbitals are nonbonding and Δ_0 measures the energy needed to promote an electron from the t_{2g} to the antibonding e_g^* orbitals.

Generally, good σ-donor ligands produce higher ligand field splittings. Thus NH_3, the stronger Lewis base, gives a larger Δ_0 than H_2O. When -donation occurs from a filled ligand orbital to a metal ion d-orbital, the t_{2g} orbitals for a bonding and antibonding set and Δ_0 represents the energy difference between the antibonding t_{2g}^* and e_g^* orbitals. In this case, Δ_0 is diminished. Hydroxide and fluoride ions are examples of -donor ligands where hydroxide is the stronger electron donor.

When empty -bonding orbitals are available on the ligand, the metal ion becomes the electron donor and so-called "back-bonding" occurs. In this case Δ_0 represents the energy difference between the bonding t_{2g} orbitals and the antibonding e_g^* set; as a result Δ_0 increases. This situation does not occur with the ligands selected for this modeling exercise.
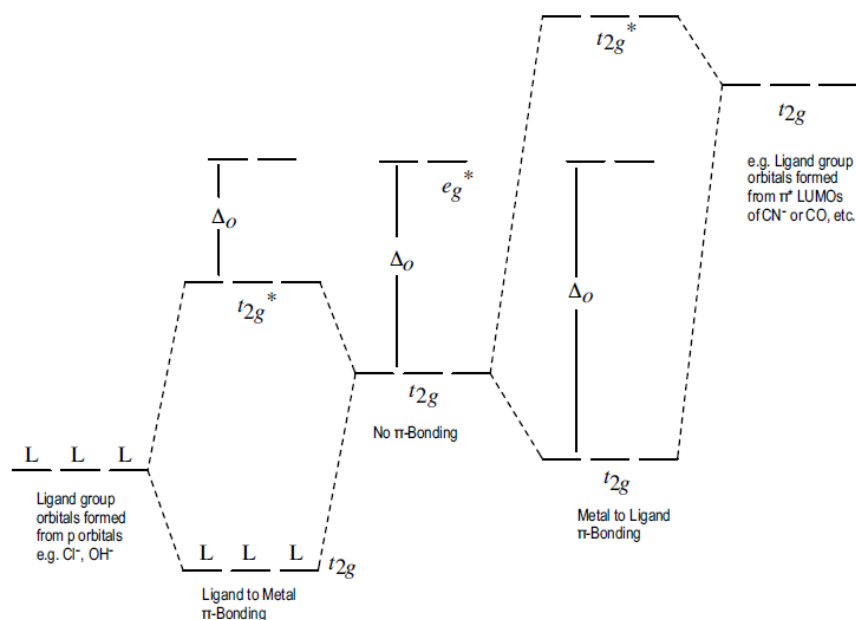
The results for the Cr(III) series are given below. While a similar general trend is observed, the order of hydroxide and fluoride is reversed. While chloride yields a relatively low LUMO - HOMO difference, in agreement with experiment, the LUMO actually corresponds to an MO derived from the metal s-orbital. The metal e_g^* orbitals are calculated at somewhat higher energy.

Octahedral Ligand Field Splitting Energies of Cr(III) Complexes, [CrX$_6$]

Ligand X	HOMO energy (eV)	LUMO energy (eV)	Difference (eV)
Cl	0.61	9.02	8.41
OH	2.88	13.63	10.75
F	1.06	11.90	10.84
H$_2$O	-24.77	-13.54	11.23
NH$_3$	-23.03	-11.42	11.61

Stereoselectivity of Chiral Chelating Ligands in Metal Complexes

Contributed by William Jordan, Pacific University

Overview

In this exercise you use the SCIGRESS workstation to investigate the stereoselectivity of a chiral chelating ligand.

Recommended Exercises

Complete Exercises I, II, III, IV, V, VI, VII, and VIII before beginning this experiment.

Background

Coordination compounds derive chirality from the arrangement of ligands in the coordination sphere. One of the ways in which this occurs is via puckering of a chelate ring into a conformation nonsuperimposable upon its mirror image. Consider 1,2-diaminoethane (ethylenediamine or en) which binds metal ions through each amine nitrogen (N) thus forming a five-membered chelate ring. For example, the complex ion $[FeCl_6]^{3-}$ will react with a molecule of ethylenediamine to form $[Fe(en)Cl_4]^{-}$.
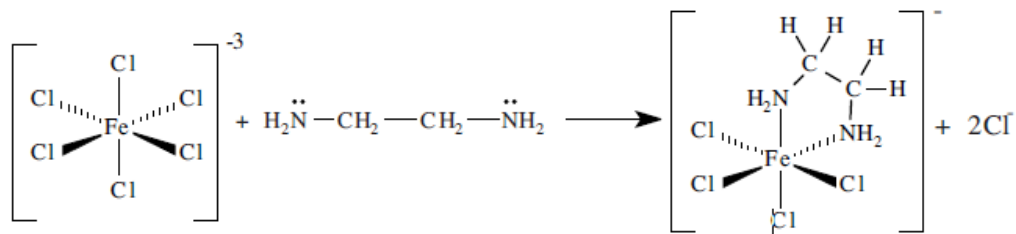
The heterocyclic ring containing two carbon atoms, two amine nitrogens and an iron(III) ion is not planar. The N-Fe-N angle is close to the octahedral value (90°) and the nitrogen and carbon atoms prefer tetrahedral coordination. If the chelate ring is reoriented so that the mean plane of the chelate ring is in the horizontal position as shown below, it can be seen that the chelate ring adopts a chiral conformation (is not superimposable upon its mirror image). By convention, the mirror image conformations are labeled λ and δ.
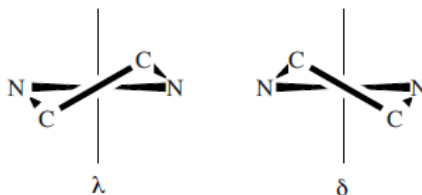
Of course, reaction of ethylenediamine with an iron(III) complex would result in λ and δ isomers in equal proportions (a racemic mixture). While in principle, it is possible to separate the enantiomers, the two conformations are generally interconvertable at room temperature. Here there is sufficient thermal energy to invert the chelate ring, making actual separation impossible under ambient conditions.

If a hydrogen atom on one of the two carbons is replaced by a methyl group for 1,2-diaminopropane (pn for propylenediamine), that carbon becomes chiral. In this case the λ and δ conformations are not equivalent in energy and they would not be synthesized in equal proportions. Thus, the ligand would be stereoselective for one of the two forms.

Modeling Section

1. Build models of $[Co(NH_3)_4R\text{-}pn]^{3+}$ (R-pn = R-1,2-diaminopropane) in two different conformations by doing one of the following:
   Begin from scratch with a Co^{3+} ion and connect ligands to it with the appropriate charge and bond type (coordinate).
   -OR-
   Select a similar compound from the Fragment Library. It is convenient to begin with _(en)3Co_3+ from the Ligand category in the Fragment Library Delete the bonds between N and C for two of the three chelate rings. Delete the unattached fragments also (-CH$_2$ -CH$_2$-).

2. Beautify to obtain the desired compound and reorient so that the ethylenediamine ring is horizontal and oriented toward the front of the molecule, shown in the diagram to the right.

3. Copy and paste the model you have prepared of $[Co(NH)~(en)]^{3+}$ to a new Workspace document. Alternatively, you may duplicate the document using the Save as command. At this point the two models will be identical.

4. Convert one to its mirror image by selecting **Action | Selection | Mirror**
   The two models should now appear as mirror images of one another.
   ✎ Are they superimposable or are they enantiomers?

5. Find the equatorial hydrogen on a carbon atom of the first molecule. Replace the hydrogen with a carbon atom. On the corresponding carbon of the second model replace the axial hydrogen with a carbon.

6. Beautify each model to convert the added carbon to a methyl group. The models ought to look similar to those depicted on the next page, where the hydrogens have been hidden for the sake of simplicity.
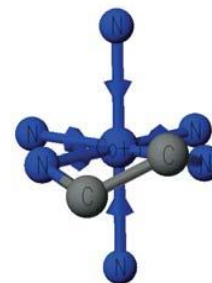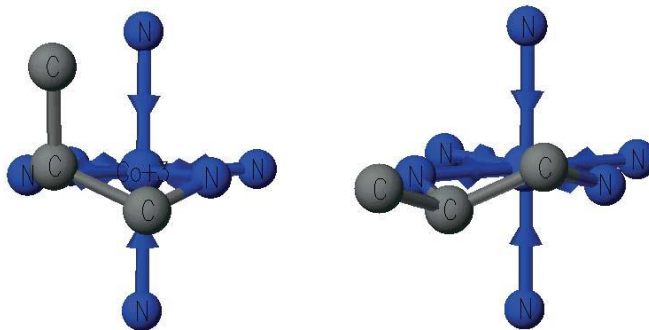
TIP
Remember that
SCIGRESS can label the
chiral centers as R or S.

7. Verify that the configuration about the asymmetric carbon is R and not S in each case. If it is S, you have replaced the wrong H by a CH_3 group.

8. Optimize each structure using Mechanics (MM2) and note the energy in each case.
   ✎ Which structure is more stable? Why?

Interconvert the two conformations of the R-pn chelate ring by conducting an exhaustive search

1. Open the conformation file that has the methyl group in the axial position in the Workspace. Save the document under a new name.

2. Select connecting atoms for the search label in the following order: methyl group carbon, carbon, axial hydrogen. The structure should now appear as shown to the right.

3. Select **Action | Geometry Labels | Dihedral Angle**. The Dihedral Angle window will appear. Click on **Define Geometry Label** and then **Search**.

4. Set up the label to search between **165.000** and **60.000** using **100 steps**. The larger the number of steps, the smoother the potential energy curve is but the calculation consumes more time.

5. Click **Apply**; the step increment will automatically be adjusted to agree with the search limits and the number of steps. Click **Done** to close the dialog box.
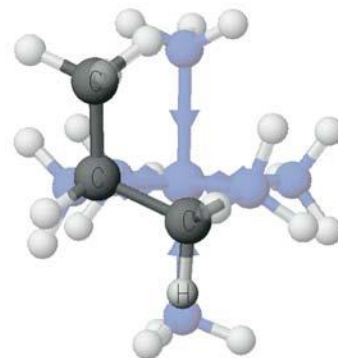
Generate an energy map

1. Select **Experiment | Run**. In the Experiment window select the following options: conformations, potential energy map and exhaustive search with MM2.

2. Click **Start**, and the calculations will begin. When the calculations are done, a plot of potential energy vs. dihedral angle and the accompanying structure will open. The plot and the structure appear in two separate windows which may be activated separately by clicking in one or the other.

3. Click on different points on the plot to view the corresponding structure and energy value. You may also animate the interconversion of conformations by selecting **View | Animate Along Axis**.

Report Guidelines

✎ Paste structures of different conformations of the molecule into a word processing document. Appropriate structures would be those with the methyl group of R-1,2-diaminopropane in axial and equatorial positions. A third structure of interest is the maximum energy conformation in the interconversion of axial and equatorial forms. Also paste the plot of potential energy vs dihedral angle to show the point on the plot corresponding to each conformation.

✎ Identify the lowest energy conformation and explain why this should be most stable.

✎ Why does energy rise and pass through a maximum as the axial and equatorial conformations are interconverted? (Check the bond angles in the chelate ring.)

Instructor Notes

Typical Results

The equatorial conformation is most stable as it minimizes steric interactions between the methyl group and the amine groups at the remaining coordination sites of the complex ion. In the least stable conformation (maximum in the plot of potential energy vs dihedral angle), the chelate ring is almost flat and the bond angles about the ring atoms deviate most from ideal angles.
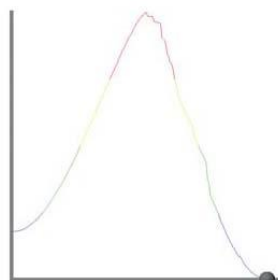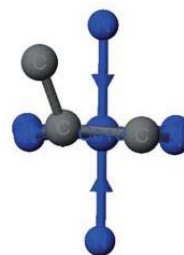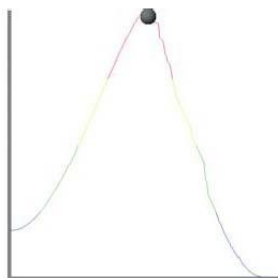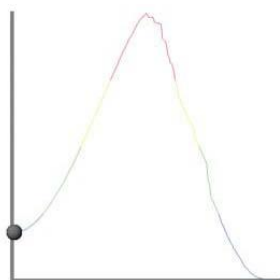
Part 5 Working with Biomolecules and docking

Importing and Cleaning Protein Crystal Structures

This tutorial describes how to import Protein Data Bank (PDB) files, view protein structures and sequences, and build proteins and peptides from their sequence. You'll learn how to:

• import and view PDB molecules

• inspect and correct hetero groups

• add hydrogen atoms

• balance charges

Background

In this tutorial, we will investigate the 1.75Å resolution crystal structure (1G2O) of Mycobacterium tuberculosis purine nucleoside phosphorylase (TB-PNP) complexed with the transition-state analogue Immucillin-H (ImmH) and phosphate. TB kills millions each year and it has been estimated that one-third of the world's population is infected with latent TB.

Analysis of the genome sequence of Mycobacterium tuberculosis (TB) predicted that it expresses purine nucleoside phosphorylase (PNP) which catalyzes the phosphorolysis of purine nucleotides to purine bases and deoxynucleosides to (deoxy)ribosyl 1-phosphate. PNP recycles purines, a crucial function for organisms that do not synthesize purines. It is thought that inhibition of this enzyme will cause physiological changes in TB that will cause the bacterium to enter a latent state thereby preventing the development of active TB in infected individuals.

One problem with this approach is that humans also express PNP. In humans it is known that inhibitors of human PNP have potential clinical use as immunosuppressants. Since there is a risk that a TB-PNP inhibitor would also inhibit human PNP, leading to unacceptable side effects during treatment, there is a need to develop TB-PNP inhibitors that do not bind to human PNP.

Although TB-PNP is evolutionarily related to human PNP, a recent crystal structure shows a significant difference in the hydrogen bonding to ImmH between the mammalian and TB PNPs. It is therefore possible that molecularmodeling can be useful in designing selective inhibitors for TB PNP. The need to develop selective inhibitors is a common problem in drug discovery and development.

The goal of these exercises is to teach you how to use Scigress to design your own selective inhibitors by using peer-reviewed studies from the current scientific literature applied to the important problem of obtaining new drugs for the treatment of TB.

Importing proteins

The Protein Data Bank (PDB) maintained by the Research Collaboratory for Structural Bioinformatics (RCSB) contains tens of thousands of x-ray crystal structures for proteins and other biomolecules.

The structures of proteins and other biomolecules can be downloaded from http://www.rcsb.org/pdb/ as PDB files. PDB files have an extension of.pdb or.ent. You can view the 3D structure of TB-PNP complexed with ImmH using files downloaded from the PDB that

you import in the Workspace application.

In this exercise, you will use the PDB file 1G2O.pdb provided in the Scigress directory Biomolecules and Docking\2-Importing and Cleaning Proteins.

1 Open or create a project. Let's assume it will be named "Biomolecules"

2 In the Workspace application, choose **File | Import**….

3 Navigate to the location of the Scigress application files, for example:

C:\Program Files\Fujitsu\Scigress Suite\Biomolecules and Docking\2-Importing and Cleaning Proteins.

4 Choose PDB (*.pdb), (*.ent) from the Files of Type drop down list.

5 Click 1G2O.pdb.

6 Choose Open to import the PDB file into a Scigress a new 3D window.

Saving a PDB molecule

From now on you will work with the PDB molecule in a Scigress chemical sample file (.csf) rather than the PDB file so that you can retain all of the views and information you create in modeling.

To save an imported PDB molecule as a chemical sample file:

1 Choose **File | Save As**.

2 Save the file as a Scigress chemical sample with the name 1G2O.csf. Please ignore the warning about valence exceptions by pressing OK.

When you save a PDB file in the Scigress .csf format, the chemical sample file includes all the information from the PDB file. For example, residues are stored as Groups and you can use the Select Group tool as a shortcut to select all atoms in the residue. When you run experiments using the imported PDB structure, data such as rendering style and computed atom properties are also added to the .csf file.

In Scigress, the term 'chemical sample' refers to the proteins, ligands, water molecules and other chemical entities that are contained in the model you develop.

To view information about the chemical sample as a spreadsheet:

1 Choose **Analyze | Model Properties Table**. The Chemical Properties spreadsheet appears.

Properties for a PDB molecule are presented in several worksheets available when you click the tabs below the table:

Atoms contains atoms and their properties

Bonds contains bonds and their properties

Crystals contains the space group and other crystallographic information

Groups contains residues and their properties

Notes contains REMARK, NOTE, SOURCE, TITLE, COMPND, HEADER, and other PDB records

Electrons tracks nonbonded electron pairs so that the valence can be checked (Electrons are added when you beautify a structure and do not appear when the PDB file is first opened.) The number and type of worksheets will vary depending on the data present in your chemical sample file.

2 To view comments about the chemical sample, click the **Notes** tab and expand the **Comment** column.

3 Read through the comments to confirm that this is the structure for ImmH complexed with TB-PNP at 1.75Å resolution.

4 Scroll to ID 249 and observe that the first six residues (Met-Ala-Asp-Pro-Arg-Pro) in each of the three chains of TB-PNP were not resolved and are missing from the structure.

5 Scroll to ID 415. The HET records and the HETNAM records identify groups that are not standard amino acids such as ligands, metals, solvent, or ions. There is one IMH and one PO4 for each of the three chains in TBPNP. The names and molecular formula of each HET group is given in the HETNAM and FORMUL records. You will use this information later when cleaning up the HET groups.

Simplifying the structure

Each of the chains in TB-PNP is the same and each has ImmH and PO4 complexed in its active site. In this tutorial you will analyze the binding of ImmH to the active site in TB-PNP. It is helpful to work with the simplest model first, a single chain and its complexed groups. In this tutorial, you will reduce the structure to the monomer by removing Chains B and C.

To reduce the structure to a monomer

1 Click the **Groups** tab in the Model Properties spreadsheet. The Groups worksheet appears.

2 Scroll the window until you see the start of the row containing **Chain B**.

3 Select the first row of **Chain B** by clicking on the left-most cell of the row in the worksheet (ID 263). Row ID 263 highlights.

4 Scroll until you see the last row containing **Chain C** (row ID 786).

5 **Shift-click** the row. All residues in Chains B and C are highlighted indicating that they are selected.

6 Close the Model Properties spreadsheet.

The 3D structure workspace appears.

7 Choose **Edit | Delete**. Chains B and C disappear exposing the HET groups and water molecules

associated with Chains B and C.

8 Using the Select Molecule tool, click any atom in Chain A. Chain A is selected (highlighted) and the rest is grayed.

9 Choose **Edit | Select | Select Neighbors** and choose **Residues, waters, HETs** from the **Select Nearest** pulldown menu. The **Selection Radius** should be 5 Å.

10 Click **OK**. Chain A and all residues, waters, and HET groups within 5 Å of any atom in Chain A are selected.

11 Choose **Edit | Invert Selection**. Chain A and all residues, waters, and HET groups within 5 Å of any atomin chain A are deselected and everything else is selected.

12 Choose **Edit | Delete**. The excess water and HET groups that were associated with Chains B and C disappear and the remaining Chain A is selected and highlighted.

13 With the Select Molecule tool, click the ImmH group at the top of Chain A
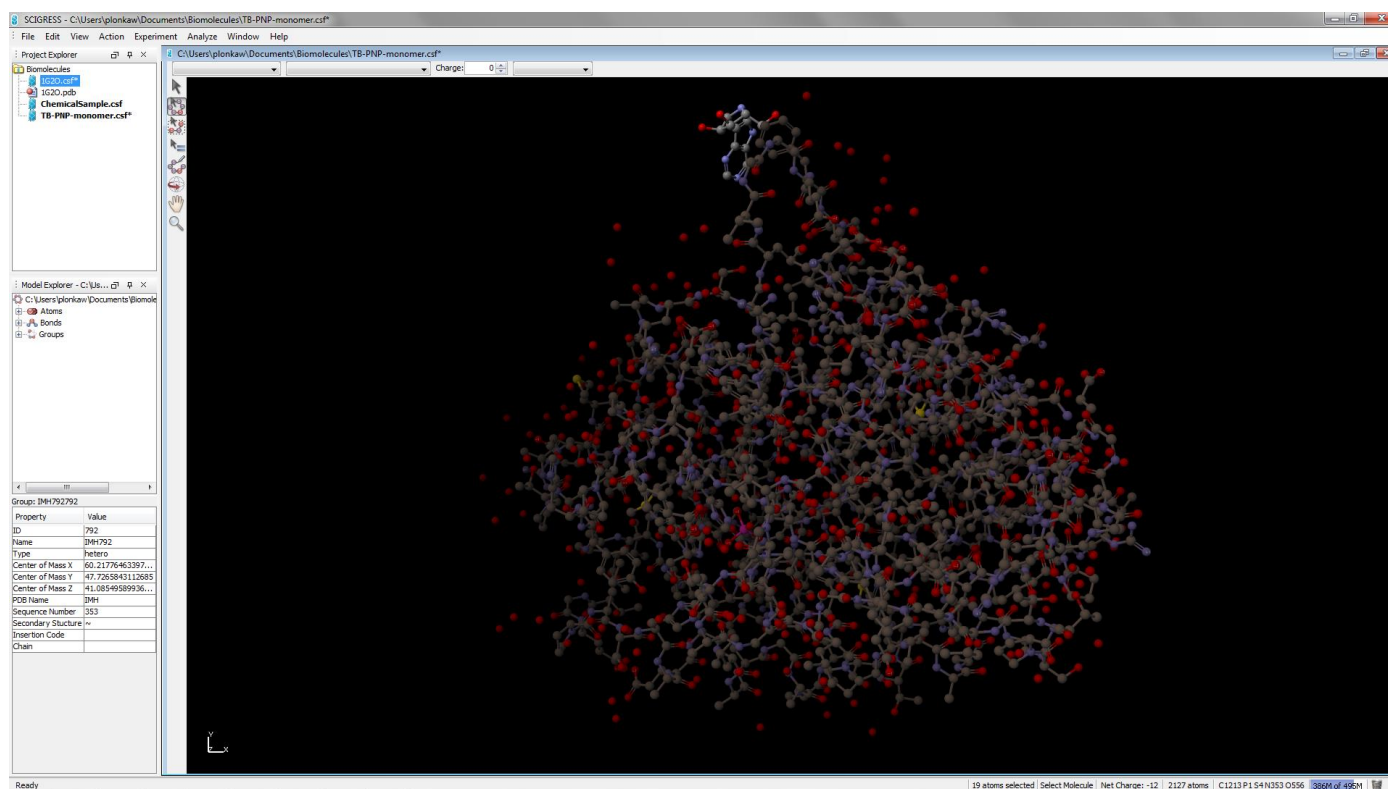
and choose **Edit | Delete**.

The status bar at the bottom of the window shows that 2108 atoms remain, and your sample has the following structure:

14 Choose **File | Save As** and save this file as a chemical sample file with the name TB-PNP-monomer.csf.

Next you will prepare the structure for molecular modeling and analysis.

Cleaning protein structure

When a PDB file is imported in Scigress, some new information is automatically calculated and added to the new.csf file that contains the PDB molecule. However, information required for molecular modeling must still be added, including hydrogen atoms, atom hybridization, and correct bond types for HET groups and non standard residues. If residues are missing or incomplete, it may be necessary to correct their structures. Before you begin, lock atoms to prevent moving them accidentally from their crystallographic positions.

Lock atoms at their crystallographic positions

1 Choose **Edit | Select All**. All atoms and bonds are highlighted.

2 Choose **Action | Lock | Lock**. Selected atoms are locked at their current position in space.

3 Choose View | Colory by | Molecule

Check and correct the bonding in HET groups

1 Choose **View | Fit in Window**. Notice that the protein chain, immucillin, and phosphate groups are displayed in different colors.

2 Choose the Select Molecule tool and click the Immucillin-H (ImmH) ligand. ImmH is selected and appears highlighted; the rest of the structure is grayed.

3 Choose **View | Hide Unselected**. All unselected atoms and bonds disappear. They are hidden from view, but not deleted.

4 Choose **View | Fit in Window**. ImmH zooms to fill the window. (You may need to rotate ImmH to see all the atoms and bonds.)

5 Choose **View | Color by Element**. Carbon atoms are gray, oxygen atoms are red, nitrogen atoms are blue. (Hydrogen atoms are white, but they are hidden in this example.)

Your structure should match the structure shown here.
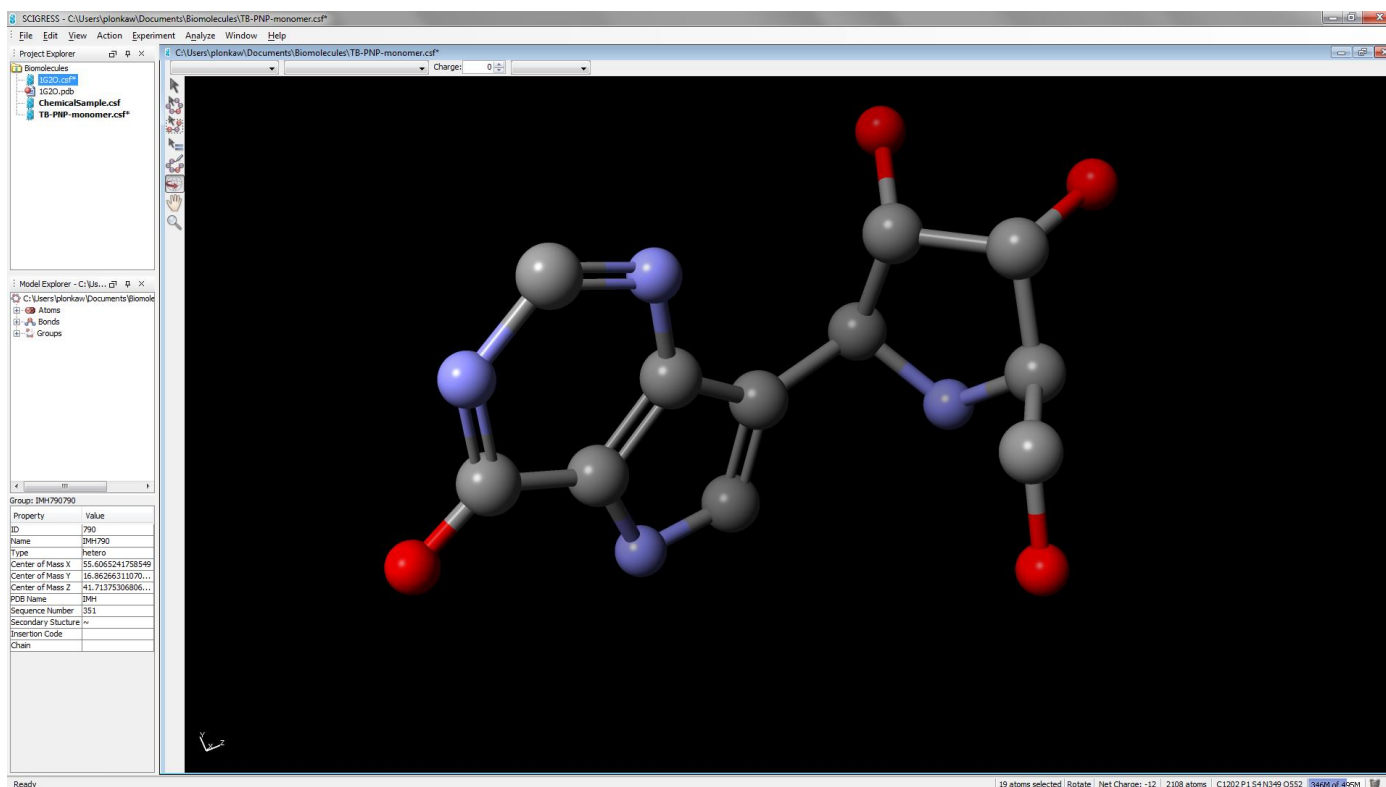
6 To change a bond order, click a bond with the Select tool, then choose a new bond type from the pull-down menu in the style bar. Selected bonds change to the new bond type.

7 Examine ImmH and look for charges that do not agree with the figure.

8 Click the nitrogen atom in the iminoribitol ring to select it.

9 To change the charge to 1, click the up arrow next to the charge box in the style bar. A **+1** appears on the selected nitrogen atom and the total charge on the whole chemical sample decreases to -11.

Add hydrogen atoms and define atom hybridization for ImmH

1 Change to the Select Molecule tool and click to select the ImmH ligand. ImmH is highlighted.

2 Choose **Action** | **Beautify** | **Valence**. Hydrogen atoms, electrons, and the atom hybridization are added. ImmH should now look like this
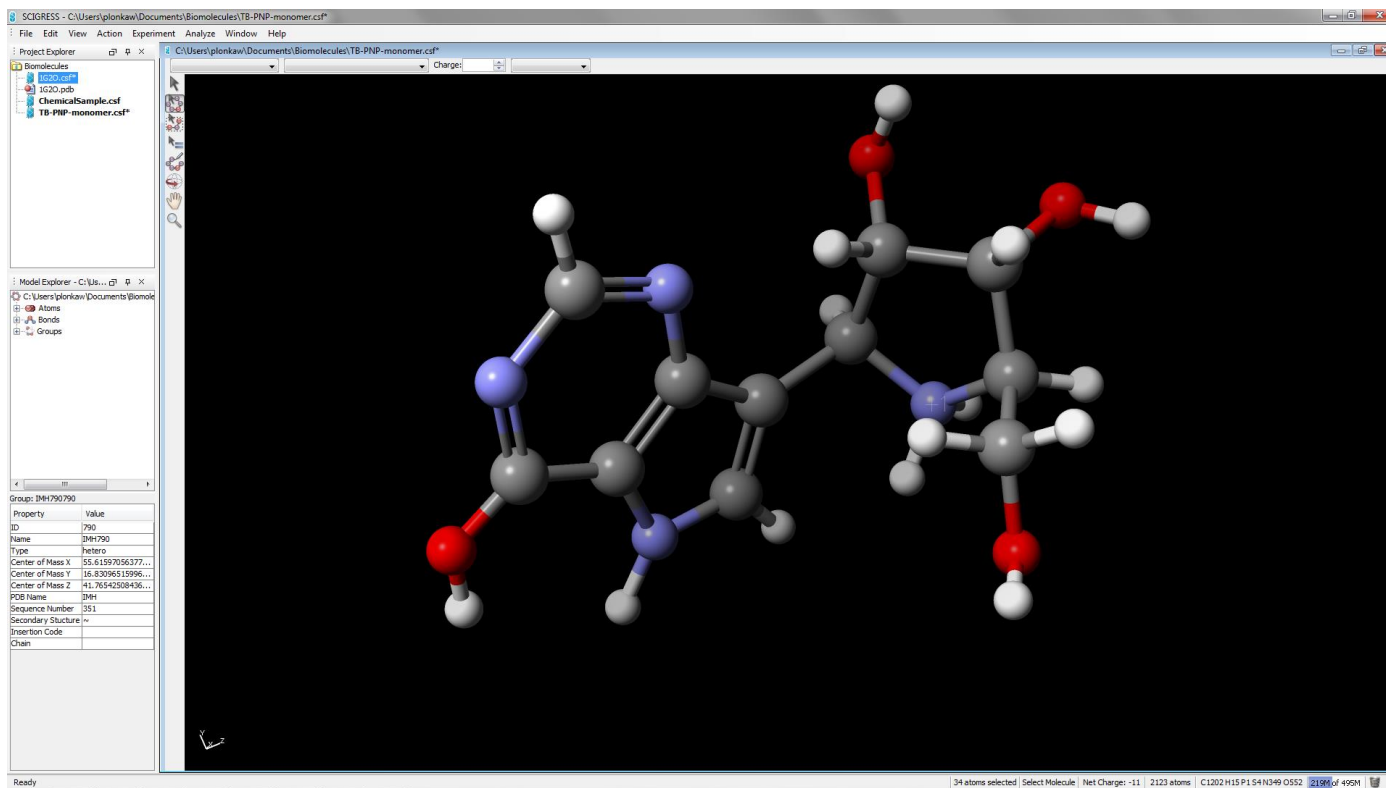
Add hydrogen atoms and define atom hybridization for PO4

1 Choose View | Show All. The full structure is displayed.

2 To clean the PO4 ligand, select it and then choose Beautify | Valence, if necessary.

For this exercise, don't be concerned about which oxygen atoms have the negative charge. The cleaned PO4 should look like this:
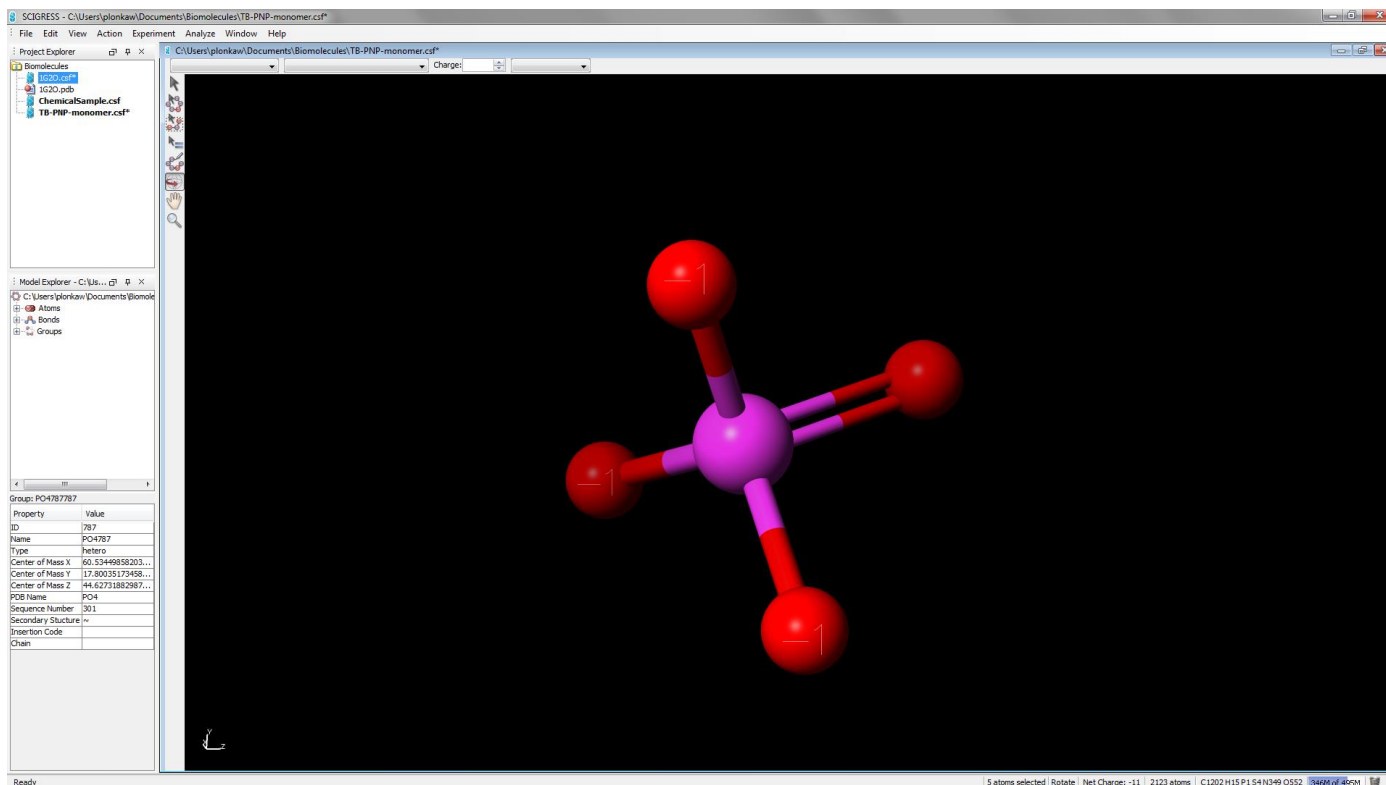
To balance the charge in the protein

Depending upon its environment in a protein crystal, the histidine residue can be neutral or protonated.When Scigress reads a PDB file, histidine is assumed to be neutral unless a charge has been specified in the PDB file. In TB-PNP, histidine should be protonated. The following steps guide you to protonate all histidine residues.

1 Choose **View | Show All**.

All the atoms and bonds in the chemical sample appear.

2 Choose **Edit | Select All**.

All the atoms and bonds are selected.

3 Choose **View | Fit in Window**.

The view zooms out to show all atoms.

4 Choose **Edit | Find**.

The Find dialog opens, enabling you to find and select specific structures.

5 Choose **Atoms** from the **Search for** pull-down.

6 Choose **Name** from the **Where** pull-down.

7 Type ND1 in the text box next to the **Is** item.

8 Click **Find**.

The 10 nitrogen atoms in the ten histidine residues are selected and the rest of the structure is not selected.

9 Choose the Select tool. The style bar (ribbon) at the top of the 3D structure window shows **N - Nitrogen**.

10 Change the **Charge** in the style bar to 1. The charge on each of the highlighted ND1 atoms changes to +1 and the total charge on TB-PNP drops to -1.

To add hydrogen atoms and atom hybridization to the protein

1 Choose **Edit | Select All**. Everything becomes selected.

2 Choose Action | **Beautify | Valence**. In just 5 seconds, hydrogen atoms, electrons, and hybridization are added to the protein and the water molecules.

3 Choose Action | **Beautify | H-Bonds**.

Hydrogen atoms in -OH and water molecules are rotated to maximize hydrogen bonding. Only hydrogen atoms are changed with these Beautify commands; the remainder of the structure remains the same because it is locked.

Save your work as TB-PNP-cleaned.csf

1 Choose **File | Save As** and save your work as a chemical sample file (*csf) with the name TB-PNP-cleaned.csf.

2 Choose **File | Close** to end your work with this chemical sample. Usually, the next step is to refine the position of the hydrogen atoms using molecular mechanics calculations, but we will not do so at this time. You will learn how to optimize structures in a subsequent exercise.

Viewing and Analyzing Proteins, Ligands, and their Complexes

Analyzing the properties of proteins containing thousands of atoms is best accomplished by representing molecular properties with graphics that reduce complexity. Because drug design focuses only on that relatively small portion of the protein structure where the small drug molecule binds, it is important to easily visualize only the active binding sites and their ligands. In addition, finding and visualizing small molecule-protein interactions, such as hydrogen bonding, must be simple, fast, and easy.

For example it is important that you are able to:

• define and manipulate groups of atoms

• display hydrogen bonds and atoms that bump together

• manipulate sequences

• identify active sites using crevice maps and sequence alignment

• color by properties, like hydrophobicity and hydrophilicity

• label atoms that bump together

• extract information from sequence alignments

In this exercise, you will be introduced to the unique Sequence View in Scigress, a powerful analysis tool for navigating and visualizing structures containing tens of thousands of atoms.

You will use the Scigress Workspace to:

• view the protein's accessible surface

• view and analyze the sequence of a protein

• locate the active site

• show and measure hydrogen bonds between the ligand and the protein

• display surface of the binding pocket and look at docking interactions

Viewing the accessible surface

An accessible surface is the surface of a protein that could be touched by a 1.4 Å sphere; a sphere that is approximately the size of a single water molecule.

To create and view the accessible surface

1 Choose **File** | **Import File**, navigate to the folder named 3-Viewing and Analyzing Proteins and Ligands, and open TB-PNP-All- Refined.csf.

TB-PNP-All-Refined.csf opens and displays the cleaned 3D structure.

2 Using the Select Molecule tool, click an atom in the protein. The entire protein become selected.

3 Choose **View** | **Hide Unselected**.

The selected protein remains in view, while other atoms are hidden (but not deleted). To economize the computation and display of an accessible surface, it is useful to include only the portion of the structure that is of interest. Visible atoms, whether or not they are selected, contribute to the accessible surface. The selected atoms define a virtual space, within which the surface is drawn.

4 Choose **Analyze** | Surfaces| **Accessible Surface**.

After a few seconds, a new accessible surface appears along with the surface label, **asf 0.01.**

5 To select and de-select a surface, click and shift-click the center of the label. When the surface is selected (the default condition), it is opaque. When you de-select it by shift-clicking the label, the surface is translucent so you can see both the surface and the molecular model within it.

Mauve (purple-red) areas and blue areas are hydrophillic while the cream area is hydrophobic. Mauve indicates hydrogen acceptors (*e.g.* C=O) and blue indicates areas of hydrogen donors (*e.g.* -NH or more generally -XH).

As expected since proteins tend to fold so that hydrophobic residues are on the interior and hydrophillic groups are on the exterior, hydrophillic regions dominate the outside of the protein. Large patches of hydrophobic areas (cream colored) on the surface of a protein suggest that it is involved in interactions with other proteins. Notice that the cream colored portions of the TB-PNP are located where chains B and C of the trimer touched the left and top of this chain.

This is a crystal of TB-PNP with a bound inhibitor - ImmH - in the active site. You should note that there is no tunnel leading to the active site. TB-PNP has folded around the ligand, closing the tunnel and making it difficult to identify an active site from the solvent accessible surface.

The labels (+1 and -1) showing through the surface are the atom charges.

Before continuing, hide the accessible surface and redisplay the hidden atoms.

To hide the accessible surface

1 Locate the Model Explorer. Click on the "+" sign next to the word surfaces.

2 Double click to remove the check beside asf0.010.
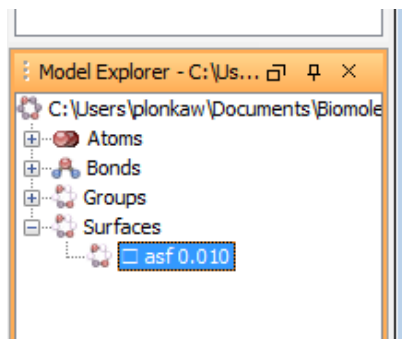
To show the hidden atoms and bonds

1 Choose View | Show All. All atoms and bonds, including those that you had previously hidden, are displayed in preparation for the next step in the exercise.

Viewing the protein sequence

Scigress displays sequence data for a protein structure in the Sequence View window. When you open one or more protein files in the Scigress Workspace application and analyze their sequence, the sequence data are added to the Sequence View window. Only one Sequence View window can be open at one time.

Use the Sequence View window to:

• view the sequence

• color the sequence by property

• mutate residues

• edit the sequence

• build peptides and proteins from residues

• analyze secondary structure

• adjust the conformation or secondary structure of a protein

• align sequences

• superimpose sequences

To view the sequence of TB-PNP

1 From the 3D structure workspace, choose **Action | Protein | Sequence** to display the Sequence View window. SequenceView opens in a new window and displays sequences of residues using 1-letter codes in rows that begin with the chemical sample name.

135

You can also view residues represented as 3-letter codes.

2 Choose **View | 3-Letter-Code**. Residues appear using 3-letter codes and they are colored by chemical type. The sequence is directional, starting from the N terminus of the first residue to the C terminus of the last one. HET groups are always displayed, but water molecules (HOH or DOD) are never displayed in the Sequence View. Sequence View provides its own set of menus and a tool palette

3 Choose Window | Tile. The Sequence and 3D structure windows are displayed side-by-side. Click a window to make it active. The active window determines which set of menus and tool palette are active.

4 Scroll to the bottom of the Sequence View window. An asterisk '*' called the "chain terminal" marks the end of each protein chain. Note that the chain terminal is not part of the structure and you cannot select the chain terminal.

Analyzing the protein sequence

Initially, residues in the Sequence View are colored by chemical type. You can change the coloring method to display other properties.

To color a sequence by property or residue type

1 Click the Sequence View window to activate it and choose **View | Residue Color...** The Residue Colors dialog appears.

2 Click the **Sequence Protocol** radio button under **Scheme**. The **Window** text box specifies the number of adjacent residues that are grouped together for analysis. For example, a window of 7 will result in the sequence being analyzed in blocks of 7 residues to determine properties such as the antigenicity or hydropathy.

3 Choose **Karplus and Schulz Flexibility** from the **Method** pull-down. Note the colors and color codes in the lower part of the dialog.

4 Click **OK**.

The dialog closes and the residue colors change to indicate flexibility based on the local 7-residue sequence. A red residue is predicted to be in a highly flexible portion of the protein. A blue residue is in a relatively inflexible portion of the protein.

You can also color the protein in the 3D structure window to correspond by property.

To color the 3D structure by sequence property or residue type

1 Click on the title bar of the 3D structure window to activate it and choose **View | Color by | Residue**. The Residue Colors dialog appears..

2 Click the **Sequence Protocol** radio button.

3 Choose **Karplus and Schultz Flexibility** from the **Method** pull down menu and click **OK**.

The dialog closes and the atom colors in the 3D structure window change to indicate flexibility based on the local 7-residue sequence. A red residue is predicted to be in a highly flexible portion of the protein. A blue residue is in a relatively inflexible portion of the protein.

Simplifying the protein structure: ribbons

To display the 3D structure as a ribbon
1 If necessary, click the 3D structure workspace title bar to activate the window and choose **View Backbone | Backbone Ribbon | Solid Ribbon**. You see ribbons and tubes colored by the residue flexibility that trace the protein backbone.

You may also color the ribbon by property in the 3D structure window.

To color the ribbon by sequence property or residue type

1 If necessary, click the 3D structure window to activate it and choose **View | Color by | Residue**. The Residue Colors dialog appears.

2 Click the **Residue Type** radio button. The list of choices in the **Method** pull-down menu changes.

3 Choose method **Secondary Structure** and click **OK**.

4 The dialog closes and the ribbon colors change to indicate secondary structure. Helices are yellow, sheets are red, and other regions of the protein are gray.

To view atoms, not ribbons

1 If necessary, click the 3D structure window to activate it and choose **View | Backbone | Backbone Ribbon | None**. The ribbon disappears and the protein atoms are colored by secondary structure.

2 Choose the Select tool and click the background. All visible objects are selected.

3 Choose **View | Color by Element**. All atoms are colored by element type.

Locating the active site

To locate residues in the active site
1 Click the Sequence View window to activate it.

2 If necessary, choose **View** | **3-Letter Code**.

3 Choose the Select tool from the tool palette and click **IMH** (351 at the end of the sequence). The bound ImmH ligand in the 3D structure window is highlighted and the rest of the sample becomes gray. Sequence View provides fast access to 3D structures by residue and group.

4 Click the title bar of the 3D structure window to activate it.

5 Choose Edit | Select | Select Neighbors. The Select Neighbors dialog opens

6 Choose Residues, waters, HETs from the Select Nearest pull-down menu and change the Selection Radius to 3 Å.

7 Click OK.

The dialog closes and atoms within 3 Å of any atom in ImmH are selected and highlighted in the 3D structure window. In addition, residues in Sequence View that contain a selected atom are highlighted and all others are dimmed.

8 Scroll through the sequence in the Sequence View and observe which residues are highlighted. Note that residues near the ligand are often far from each other in the sequence. The three dimensional structure is necessary to understand active sites. Also note that no residues have very high flexibility. The residues associated with the active site have low to

moderate flexibility.

9 In the Sequence View window, press Ctrl while you click IMH to deselect it. In both the Sequence View and 3D structure windows, IMH is deselected and its display is grayed, but other selected residues remain selected.
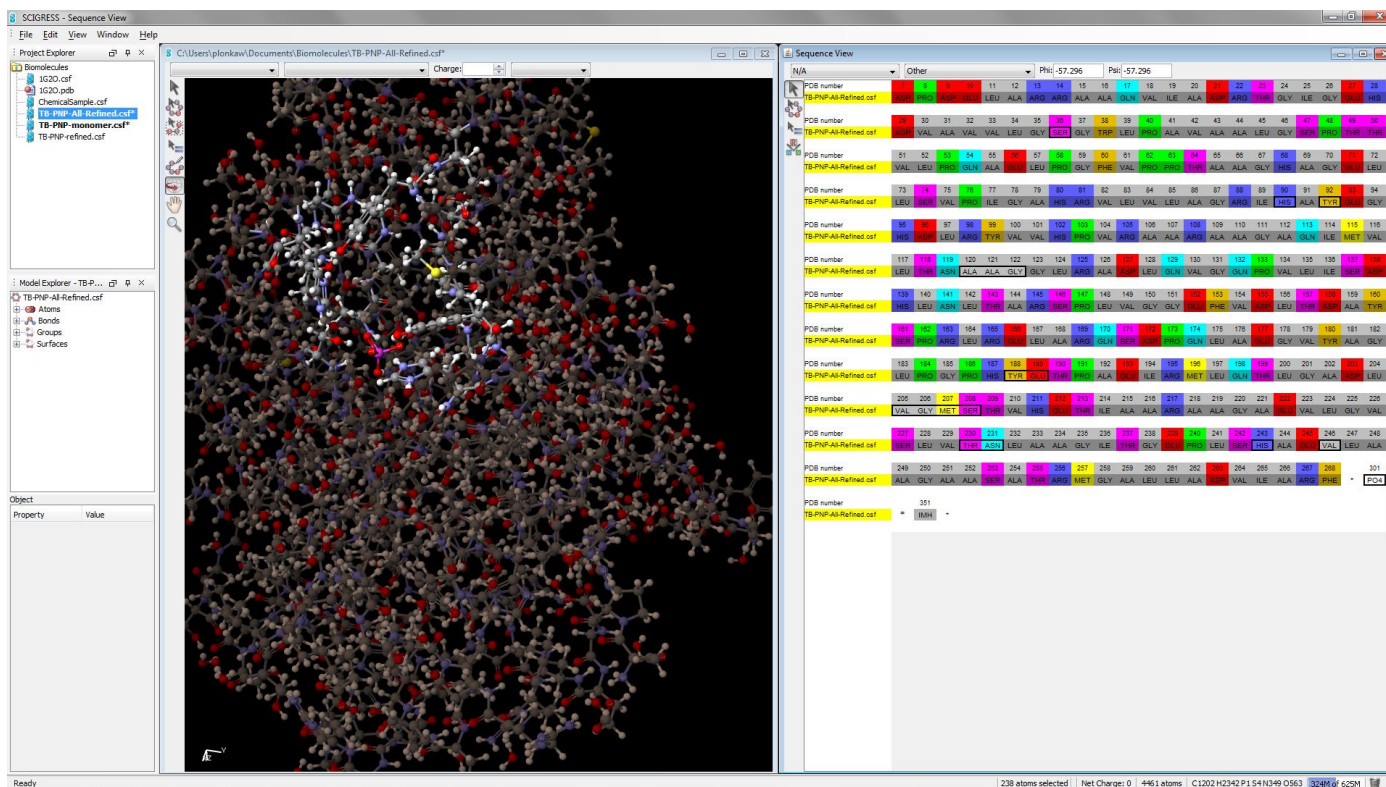
At this point all the residues in the active site are selected. The 16 residues are: Ser36, His90, Tyr92, Ala120, Ala121, Gly122, Tyr188, Glu189, Val205, Gly206, Met207, Ser208, Thr230, Asn231, His243, Val246. PO43- and two water molecules are also selected. Differences in protein-ligand active sites may be important in determining binding selectivity. In Chapter 6, you will examine the structure of human PNP and determine whether the active site residues in TB-PNP are conserved in human PNP.

To group and name the active site

1 With the residues in the active site selected, activate the 3D structure window by clicking its title bar.

2 Choose Action | Group Atoms. The Group Atoms dialog appears.

3 Choose active site from the Group Type drop down menu. Group type is used to specify whether a group is an amino acid, ligand, nucleic acid, hetero group, active site, or none of these.

4 Type the name ActiveSite into the Group Name text box.

5 Click Group>>.

The ActiveSite group is created and added to the Defined Groups list on the right.
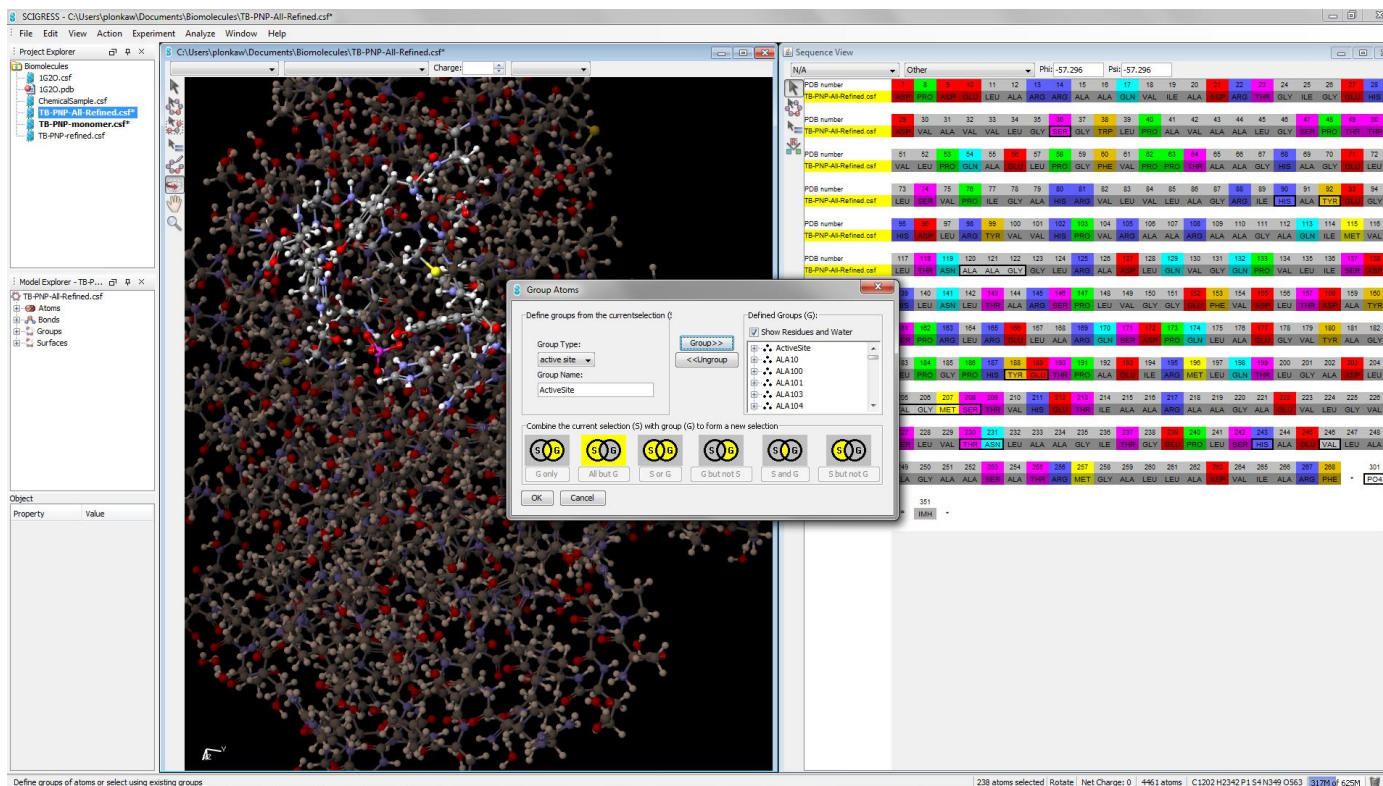
6 Click OK. The Group Atoms dialog closes. In the future, you will be able to select the active site quickly using Edit | Group Atoms or by using the Select Group tool.

Locating the hydrogen bonds

ImmH has specific hydrogen bond interactions with TB-PNP. Understanding these interactions is important when designing new inhibitors. If other interactions are equal, inhibitors with more specific hydrogen-bond interactions will bind better than those with fewer interactions.

To display hydrogen bonds between ImmH and TB-PNP

1 In the Sequence View window, choose the Select tool and click **IMH**. ImmH highlights in both the 3D structure workspace and Sequence View windows.

2 Click the title bar in the 3D structure window to activate it and choose **Analyze** | **Label H-bonds**. Blue distance labels appear, indicating the distance between hydrogen bond donors and acceptors. Each blue distance label measures the length of a hydrogen bond interaction.

3 Click **View** | **Fit in Window**. ImmH fills the window.
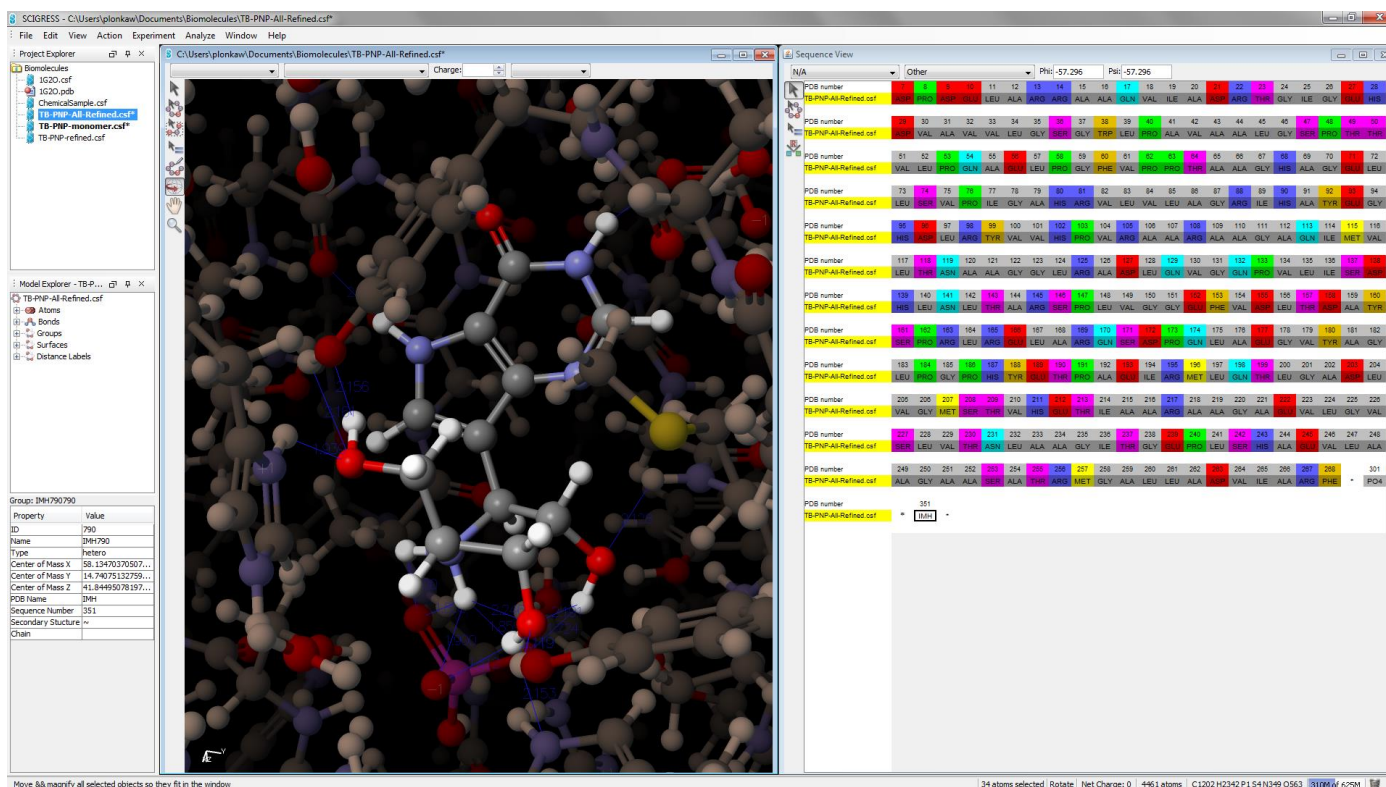
To identify the hydrogen bonding residues in TB-PNP

1 In the 3D structure window, choose the Select tool and click a protein atom at the end of an H-bond distance label to select it.

2 To select additional atoms, shift-click each atom at the end of each of the remaining protein H-bonds to ImmH. Ignore the H-bonds to PO4 3-.

The selected atoms are highlighted and their residues highlight in the Sequence View window. Other residues dim.

3 Scroll through the Sequence View and note the residue type and name. You should identify bonding to at least 5 different residues: Tyr188, Glu189, Met207, Asn231 (2 bonds), and His243. You may identify more. The hydrogen bond toMet207 is relatively long and expected to be weaker than the other hydrogen bonds. Changes in ligand-protein hydrogen bonding can be important in determining binding selectivity. In Chapter 6, you examine the structure of human PNP and determine whether the residues that are important to hydrogen bonding are conserved.

Displaying the active site pocket

The TB-PNP protein active site is a pocket in the protein that contains the bound ligand. The surface of the protein near a bound ligand maps out the pocket. To generate this surface in Scigress, the ligand is selected first and then the accessible surface of the protein enclosing the ligand is drawn. This is the surface of the protein adjacent to the ligand and the surface that forms a pocket around the ligand.

To view the surface of the binding site adjacent to the ligand

1 Click to activate the Sequence View window, choose the Select tool, and then click residue **IMH**. ImmH highlights in the Sequence View window and in the 3D structure window.

2 Activate the 3D structure workspace by clicking its title bar, then choose **View | Show All**. This is to ensure that all atoms are included in the surface determination. Hidden atoms do not contribute to accessible surfaces.

3 Next choose **Analyze** | Surfaces | **Adjacent Surface - Pocket**. After several seconds, a wireframe surface is drawn around ImmH.

4 Choose **View | Hide Unselected**.

ImmH is displayed and all other atoms and bonds are hidden. The adjacent surface1 and ImmH remain visible.

5 Choose **View | Fit in Window**.

The 3D structure window includes this view:
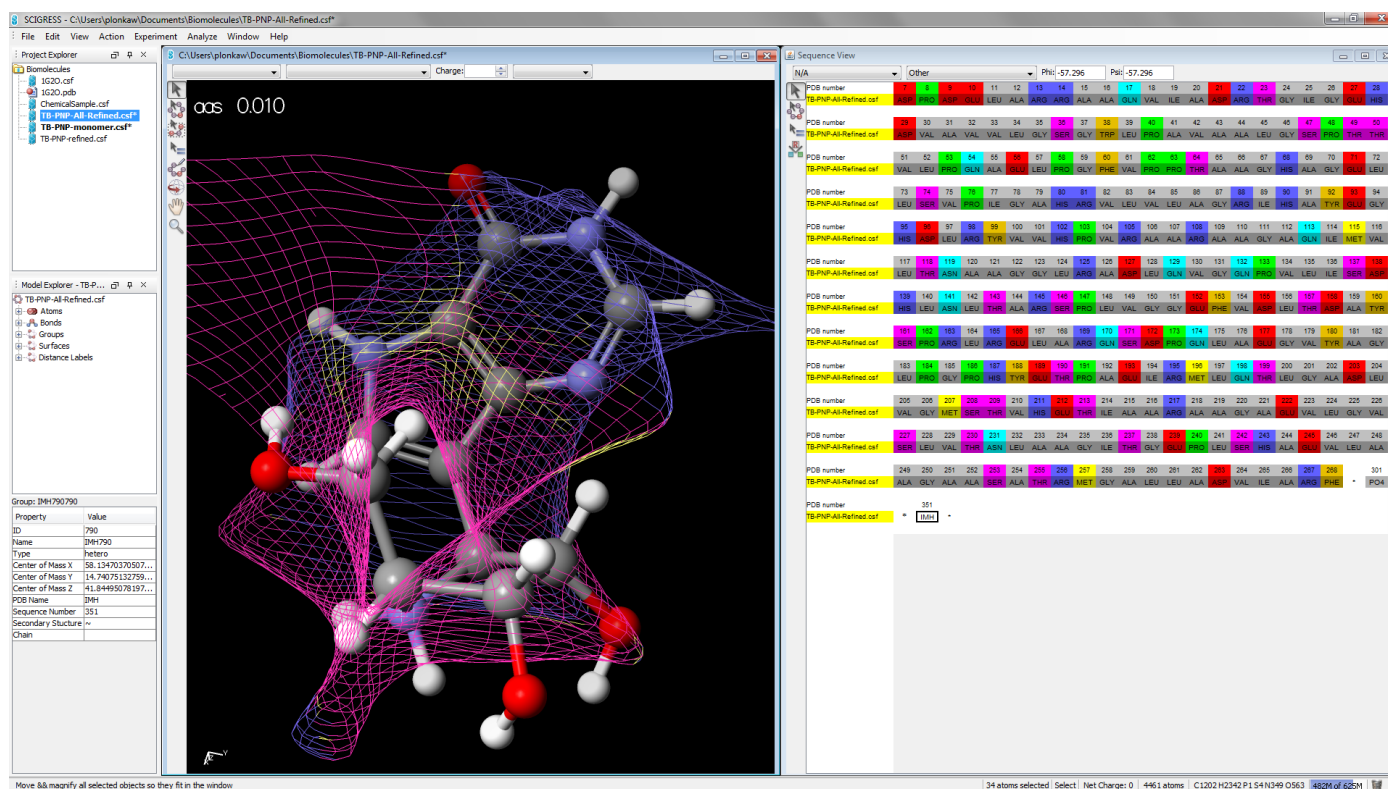
Magenta marks a surface where the protein needs H-acceptors (*e.g.* -C=O), blue marks a surface where the protein needs H-donors, and cream marks the hydrophobic surface. The pocket surface is colored so that it is easy to design ligands. Ligands that bind well should have H-bond acceptors (*e.g.* -C=O) groups touching the red surface and N-H groups with the N-H bond poking through the surface. Note ImmH's hydrogen bond donors hook through the surface and ImmH's hydrogen bond acceptors touch the surface. ImmH's hydrophobic atoms are far from the surface.

Notice extra space near the blue -NH-C=N- region in the upper right-hand corner. This empty

space suggests that ImmH might be modified to yield compounds of higher affinity by the addition of functional groups to the carbon. The new functional group should donate hydrogens to acceptor groups on the protein which do not currently form hydrogen bonds. You will check for this region in human PNP. If it is not present, then you may be able to exploit this region for selectivity. This completes the viewing and analyzing proteins and ligands introductory exercise. As you work with Scigress you will discover many additional views and analysis methods.

Discovering Active Sites of Homologous Proteins by Sequence Alignment

You have developed a model of binding for a lead compound in TB-PNP. Is the lead compound also going to bind to the human PNP leading to a loss of selectivity and potential side effects?

To answer this question, you need to determine how your lead compound might interact with human PNP. Ideally this could be accomplished by obtaining a crystal structure of the lead compound (PNP030) bound to human PNP. Unfortunately, this is often not possible for a variety of technical reasons. In addition, if the ligand is selective it will not bind to human PNP in any case! Molecular modeling represents a relatively rapid approach to determining the likelihood that your lead compound will exhibit binding selectivity. You use modeling to determine how the ligand binds to human PNP, so you can understand whether the ligand will exhibit high selectivity. Modeling provides the basis for understanding the mechanism for selectivity. In the case of PNP, there is a crystal structure for the human PNP enzyme, but there is no ligand bound to it. You must therefore find the active site in human PNP by comparison with the active site with TB-PNP and model ligands in the human PNP active site in order to understand whether the ligand is selective.

In this exercise, you will use Scigress to:

• create a crevice surface of human PNP and quickly scan it for possible binding regions

• align the sequence of human PNP with the sequence of the homologous TB-PNP for which the active site is known

• identify the active site residues in human PNP from the alignment

• create a named atom group for the active site in human PNP

• compare human PNP's active site with the crevice surface

• dock a ligand bound in the active site of TB-PNP into the active site of human PNP

Background

The 2.75-Å resolution crystal structure (1ULA) of *homo sapiens erythrocytes* purine nucleoside phosphorylase (human-PNP)1 is available. In a previous exercise, you identified the active site in TB-PNP by locating the residues adjacent to the bound ligand ImmH. This technique cannot be used with the crystal structure for human-PNP (1ULA) because 1ULA lacks a bound ligand.

Creating a crevice surface is a first quick step for locating regions of a protein that might be good binding sites.2 The crevice surface colors the protein surface by the depth from an enclosing smooth surface. The result is that deep crevices within the protein where ligands will bind are colored blue and other areas are cream colored.

Human-PNP and TB-PNP3 almost certainly evolved from a common ancestor. Both enzymes are trimeric, have a similar 3-dimension fold, and catalyze the same reaction. Once Nature discovers how to catalyze a reaction, the active site residues and their relative geometries are

conserved in evolutionarily related enzymes that catalyze the same transformation. Therefore, active site residues and geometries are usually conserved in evolutionarily related enzymes. Thus, it is reasonable to suppose that the catalytic residues present in the human-PNP active site are the same as those of TB-PNP.

Multiple sequence alignment of human-PNP with other PNP enzymes related to it by evolution can therefore be used to identify the human-PNP active site.

In this exercise, you will identify the residues in human-PNP that align with the active site residues of TB-PNP and use superposition to confirm that the 3D structure of these residues in human-PNP are the same as that in TB-PNP. The alignment and superposition allow you to identify the active site in human- PNP.

Finally, you will compare the location of your active site with the crevice map to check for consistency in the methods. In the following exercise, use files provided for you in the Biomolecules and Docking directory 5-Discovering an Active Site from a Homolog.

Using the crevice surface to scan for potential binding sites

To create and view the crevice surface

1 From the Scigress Workspace application, choose **File** | **Import File**, navigate to Biolomolecules and Docking directory 5-Discovering an Active Site from a Homolog, and open 1ULA.csf.

The 3D structure window for human-PNP (1ULA) opens.

2 Choose **Analyze** | Surfaces | **Crevice Surface**. After a few seconds, the crevice surface appears. Rotate the surface to examine it.
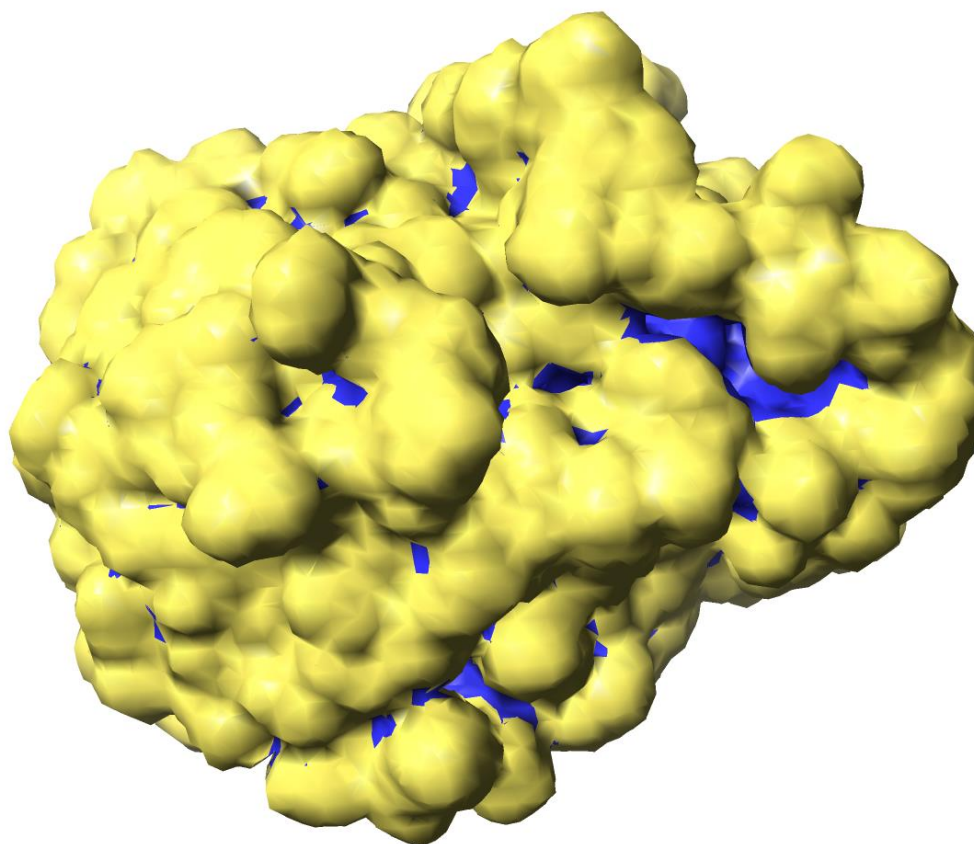
The crevice surface is an accessible surface colored by the distance from a smoother enclosing accessible surface. Blue is used to identify regions far from an enclosing smoother outer surface.

Notice the largest crevice or "lake" on the surface, shown above. This is a region of the surface that might be a good ligand binding region because of its large complex shape and its depth.

The large irregular shape makes it more likely that a site of this shape is unique to the 1ULA protein and does not occur in any other protein. Consequently, a ligand bound to this entire complex region would be selective.

The depth of the crevice suggests that there may be many binding interactions on the bottom and sides of the crevice leading to strong binding. At the end of this exercise, you will compare the active site located by homology to the region suggested by the crevice surface.

To close the crevice surface

1 When you have finished examining the crevice surface, use Model Explorer double click to remove the check beside **acs 0.0100**. The crevice surface disappears revealing the protein chemical sample, but remains part of the chemical sample file that you can display at any time.

Aligning sequences

First, we will prepare human-PNP and TB-PNP for alignment by analyzing them in Sequence View.

To prepare human-PNP and TB-PNP for alignment

1 Choose **File | Import File** and open TB-PNP+PNP030.csf. TB-PNP+PNP030.csf opens and displays the cleaned 3D structure.

2 Choose **Action | Protein | Sequence**. The Sequence View window opens and the sequence of TB-PNP is displayed.

3 Choose **Window | 1ULA.csf**. The 1ULA.csf 3D structure window becomes the active window.

4 Choose **Action | Protein | Sequence**. The Sequence View window comes to the front and the sequence of human-PNP (1ULA) is displayed below that of TB-PNP. (One Sequence View window displays the sequences for all open chemical samples.) Next, align the sequences either automatically or manually, as follows.

To automatically align human-PNP and TB-PNP (Option 1)
1 In the Sequence View window, choose **Edit | Align**. The Align Sequence dialog appears.

2 Select 1ULA (human-PNP) as **Sequence1** and TB-PNP+PNP030 as **Sequence2** and click **OK**. Gaps appear in human-PNP and in TB-PNP that align the sequences according to the maximum scoring alignment using the BLOSUM501 substitution matrix in the Needleham-Wunsch alignment algorithm2.

To select the active site residues in TB-PNP

1 Choose **Window | TB-PNP+PNP030.csf**. The 3D structure window for TB-PNP+PNP030 comes to the front.

2 Choose **Action | Group Atoms**. The Group Atoms dialog appears.

3 Choose **ActiveSite** from the **Defined Groups** list. ActiveSite appears in the **Group Name** text box.

4 Click **G Only** and then click **OK**. The Group Atoms dialog closes and active site residues in TBPNP+ PNP030 are selected and highlighted. The remainder of the chemical sample is grayed.

5 Click in the title bar of the Sequence View window. The Sequence View window comes to the front.

6 If necessary, click the Sequence View Select tool. The active site residues for TB-PNP+PNP030 are highlighted and outlined with black boxes.

Next you will select the corresponding residues in human-PNP.

To select matching residues in human-PNP

1 In the Sequence View, choose **Edit | Match Selection**. The Match Sequence Selection dialog appears.

2 Choose TB-PNP+PNP030 as the **Master Sequence** and 1ULA as the **New Selection**.

3 Click OK. Residues in 1ULA are selected and a black border appears around the newly selected

residues. Note that the selected residues in 1ULA and TBPNP+ PNP030 are identical with two exceptions.

4 To examine the selected residues in the 1ULA 3D structure workspace, choose Window | 1ULA.csf. The 1ULA.csf 3D structure workspace comes to the front. Notice that the selected residues form an open pocket ready to receive a ligand. This is different from the pocket in TB-PNP+PNP030 which has already closed around the PNP030 ligand. Apparently, as a ligand binds, the PNP active site pocket closes to wrap the extended residues around the ligand.

To create an active site group in human-PNP

1 In the 3D structure window for human-PNP, 1ULA.csf, choose **Action** | **Group Atoms**. The Group Atoms dialog opens.

2 Choose **active site** from the **Group Type** pull-down menu.

3 Type ActiveSite-human into the **Group Name** text box, choose **Group>>** and then click **OK**. A named group is created for the active site in 1ULA.csf.

4 Choose **File** | **Save**. 1ULA.csf is saved including the gaps you entered in the sequence and the named active site.

You identified the active site residues by alignment of two proteins. Would you have reached the same conclusions if multiple sequences had been used? In the next section, we examine the results of a multiple sequence alignment to answer this question.

Validating the alignment using multiple sequences

Y. pestis PNP -------------MATPHIN-AEMGDFADVVLMPGDPLRAKFIAETFLQD 36
E. coli PNP -------------MATPHIN-AEMGDFADVVLMPGDPLRAKYIAETFLED 36
H. influenzae PNP --------------MTPHIN-APEGAFADVVLMPGDPLRAKYIAETFLQD 35
V. cholerae PNP -------------MATPHIN-AQMGDFADVVLMPGDPLRAKYIAENFLDN 36
S. aureus PNP ------------MKSTPHIKPMNDVEIAETVLLPGDPLRAKFIAETYLDD 38
H. sapiens PNP -MENGYTYEDYKNTAEWLLSHTK--HRPQVAIICG*S*G-LGGLTDKLTQAQ 46**
B. taurus PNP -MQNGYTYEDYQDTAKWLLSHTE--QRPQVAVICG*S*G-LGGLVNKLTQAQ 46
M. musculus PNP -MENEFTYEDYETTAKWLLQHTE--YRPQVAVICG*S*G-LGGLTAHLKEAQ 46
B. subtilis PNP| -MK-----DRIERAAAFIKQNLP--ESPKIGLILG*S*G-LGILADEIENPV 41
M. tuberculosis PNP MADPRPDPDELARRAAQVIADRTGIGEHDVAVVLG*S*GWLPAVAALGSPTT 50++
P. aerophilum PNP --------MVKLTNPPKSPKELGFDEFPSIGIIGG*S*G--LYDPGIFENAV 40
Y. pestis PNP VREVNNVRGMLG----------FTGTYKGRKISVMG--HGMGIPS----- 69
E. coli PNP AREVNNVRGMLG----------FTGTYKGRKISVMG--HGMGIPS----- 69
H. influenzae PNP VVEVTNVRNMLG----------FTGTYKGRKISIMG--HGMGIPS----- 68
V. cholerae PNP AVQVCDVRNMFG----------YTGTYKGRKISVMG--HGMGIPS----- 69
S. aureus PNP VEQFNTVRNMFG----------FTGTYKGKKVSVMG--SGMGMPS----- 71
H. sapiens PNP IFDYGEIPNFPRSTVPGHAGRLVFGFLNGRACVMMQ--GRF*HMY*EGYPLW 94**
B. taurus PNP TFDYSEIPNFPESTVPGHAGRLVFGILNGRACVMMQ--GRF*HMY*EGYPFW 94
M. musculus PNP IFDYNEIPNFPQSTVQGHAGRLVFGLLNGRCCVMMQ--GRF*HMY*EGYSLS 94
B. subtilis PNP| KLKYEDIPEFPVSTVEGHAGQLVLGTLEGVSVIAMQ--GRF*HFY*EGYSME 89
M. tuberculosis PNP VLPQAELPGFVPPTAAGHAGELLSVPIGAHRVLVLA--GRI*HAY*EGHDLR 98++
P. aerophilum PNP EVQIHTPYGLPSDN-------VIVGRVAGRVVAFLPRHGRG*HKY*PPHKIP 83
Y. pestis PNP CSIYAKELITDFGVKKIIRVGSCGAVRTDVKLRDVVIGMGACTDSKVNRM 119

148

E. coli PNP CSIYTKELITDFGVKKIIRVGSCGAVLPHVKLRDVVIGMGACTDSKVNRI 119
H. influenzae PNP CSIYAKELITEYGVKKIIRVGSCGTVRMDVKVRDVIIGLGACTDSKVNRI 118
V. cholerae PNP CSIYVTELIKDYGVKKIIRVGSCGAVNEGIKVRDVVIGMGACTDSKVNRI 119
S. aureus PNP IGIYSYELIHTFGCKKLIRVGSCGAMQENIDLYDVIIAQGASTDSNYVQQ 121
H. sapiens PNP KVTFPVRVFHLLGVDTLVVTN*AAG*GLNPKFEVGDIMLIRDHINLPGFSGQ 144**
B. taurus PNP KVTFPVRVFRLLGVETLVVTN*AAG*GLNPNFEVGDIMLIRDHINLPGFSGE 144
M. musculus PNP KVTFPVRVFHLLGVETLVVTN*AAG*GLNPNFEVGDIMLIRDHINLPGFCGQ 144
B. subtilis PNP| KVTFPVRVMKALGVEALIVTN*AAG*GVNTEFRAGDLMIITDHIN---FMGT 136
M. tuberculosis PNP YVVHPVRAARAAGAQIMVLTN*AAG*GLRADLQVGQPVLISDHLN---LTAR 145++
P. aerophilum PNP YRAN-IYSLYMLGVRSIVAVS*AVG*SLRPDYAPGDFVVPDQFVDMTKGREY 132
Y. pestis PNP RFKDH-----------DYAAIADFEMTRNAVDAAKAKG--VNVRVGNLFS 156
E. coli PNP RFKDH-----------DFAAIADFDMVRNAVDAAKALG--IDARVGNLFS 156
H. influenzae PNP RFKDN-----------DFAAIADFDMAQAAVQAAKAKG--KVVRVGNLFS 155
V. cholerae PNP RFKDH-----------DFAAIADYKMVKAAEEAAKARG--IDVKVGNLFS 156
S. aureus PNP YQLPG-----------HFAPIASYQLLEKAVETARDKG--VRHHVGNVLS 158
H. sapiens PNP NPLRGPNDERFGDRFPAMSDAYDRTMRQRALSTWKQMGEQRELQEGTYVM 194**
B. taurus PNP NPLRGPNEERFGVRFPAMSDAYDRDMRQKAHSTWKQMGEQRELQEGTYVM 194
M. musculus PNP NPLRGPNDERFGVRFPAMSDAYDRDMRQKAFTAWKQMGEQRKLQEGTYVM 194
B. subtilis PNP| NPLIGPNEADFGARFPDMSSAYDKDLSSLAEKIAKDLN--IPIQKGVYTA 184
M. tuberculosis PNP SPLVG-------GEFVDLTDAYSPRLRELARQSD------PQLAEGVYAG 182++
P. aerophilum PNP TFYDGPR----TCHIQIGLEPFTQEIRQILIETAKKYN--RTHDGGCYVC 176
Y. pestis PNP ADLFYTPDPQMFDVM-EKYGILGVEMEAAGICGVAAEFGAKALTICTVSD 205
E. coli PNP ADLFYSPDGEMFDVM-EKYGILGVEMEAAGIYGVAAEFGAKALTICTVSD 205
H. influenzae PNP ADLFYTPDVEMFDVM-EKYGILGVEMEAAGIYGVAAEYGAKALTICTVSD 204
V. cholerae PNP AELFYTPDPSMFDVM-DKYGIVGVEMEAAGIYGVAAEYGAKALAICTVSD 205
S. aureus PNP SDIFYNADTTASERW-MRMGILGVEMESAALYMNAIYAGVEALGVFTVSD 207
H. sapiens PNP VAGPS*FE*TVAECRVL-QKLGADA*VGMS*TVPEVIVARHCGLRVFGFSLI*TN* 243**
B. taurus PNP LGGPN*FE*TVAECRLL-RNLGADA*VGMS*TVPEVIVARHCGLRVFGFSLI*TN* 243
M. musculus PNP LAGPN*FE*TVAESRLL-KMLGADA*VGMS*TVPEVIVARHCGLRVFGFSLI*TN* 243
B. subtilis PNP| VTGPS*YE*TPAEVRFL-RTMGSDA*VGMS*TVPEVIVANHAGMRVLGISCI*SN* 233
M. tuberculosis PNP LPGPH*YE*TPAEIRML-QTLGADL*VGMS*TVHETIAARAAGAEVLGVSLV*TN* 231++
P. aerophilum PNP IEGPR*FS*TKAESRIWREVFGCDI*IGMT*LVPEINLARELGMCYGLIALV*TD* 226
Y. pestis PNP HIRTGEQTT---AAERQTTFNDMIEIALESVLLGDNA------------- 239
E. coli PNP HIRTHEQTT---AAERQTTFNDMIKIALESVLLGDKE------------- 239
H. influenzae PNP HIRTHEQTT---AEERQLTFNDMIEIALDSVLIGDAL------------- 238
V. cholerae PNP HIKTGEQTT---SEERQNTFNEMIEIALDSVLIGDQAGY----------- 241
S. aureus PNP HLIHETSTT---PEERERAFTDMIEIALSLV------------------- 235
H. sapiens PNP KVIMDYESLEK*AN*H*EE*V*LAAGKQAAQKLEQFVSILMASIPLPDKAS---- 289**
B. taurus PNP KVIMDYESQGK*AN*H*EE*V*LEAGKQAAQKLEQFVSLLMASIPVSGHTG---- 289
M. musculus PNP KVVMDYENLEK*AN*H*ME*V*LDAGKAAAQTLERFVSILMESIPLPDRGS---- 289
B. subtilis PNP| AAAGILDQP--*LS*H*DE*V*MEVTEKVKAGFLKLVKAIVAQYE---------- 271
M. tuberculosis PNP LAAGITGEP--*LS*H*AE*V*LAAGAASATRMGALLADVIARF----------- 268++
P. aerophilum PNP YDIWVPHQP--*VT*A*EA*V*EKMMTEKLGIIKKVIAEAVPKLPAELPKCSETL 274

The table above shows the Clustal-W multiple sequence alignment of purine nucleoside phosphorylase (PNP) homologs (proteins related by evolution). The "**" line is the sequence of human-PNP and the "++" line is the sequence of TB-PNP. A multiple sequence alignment is made to discover portions of evolutionarily related proteins (homologs) that are unchanged by

149

evolution. These unchanged or conserved residues are likely to be important to the protein's function. In particular, the active sites of enzyme homologs tend to be highly conserved. The multiple sequence alignment shows that the active site of TB-PNP is conserved in human, mouse, bovine, and soil bacteria residues. The TB-PNP active site residues do not appear to be conserved in E. coli, Y. pestis, H. influenzae, V. cholerae or S. aureus PNP. This suggests that these enzymes have a different active site and can be classified into two sub-families by active site. Within the TB-PNP sub-family, it appears that our analysis is correct and we have found the conserved active site residues. Only two of sixteen residues in the active site of TB-PNP are different in human PNP. A strategy in the design of selective inhibitors would focus on the interactions of these two residues with ligands.

Comparing the active site to the crevice surface

Validate the crevice surface by checking that the active site is located near the crevice that you identified as a potential binding site at the beginning of this exercise.

To view the crevice surface

1 From the 1ULA 3D structure window, choose **Analyze | Show Surfaces**. The Show Surfaces dialog opens.

2 Choose the surface **crevice1.acs** and click **OK**. After a few seconds, the crevice surface appears. The crevice surface is opaque, but if you deselect the surface, it becomes translucent.

3 To deselect the crevice surface, shift-click the surface label Notice that the largest lake on the crevice surface is near the active residues.

4 Choose **Analyze | Surfaces**, click to remove the check beside **crevice1.acs** and click **OK**. The crevice surface disappears.
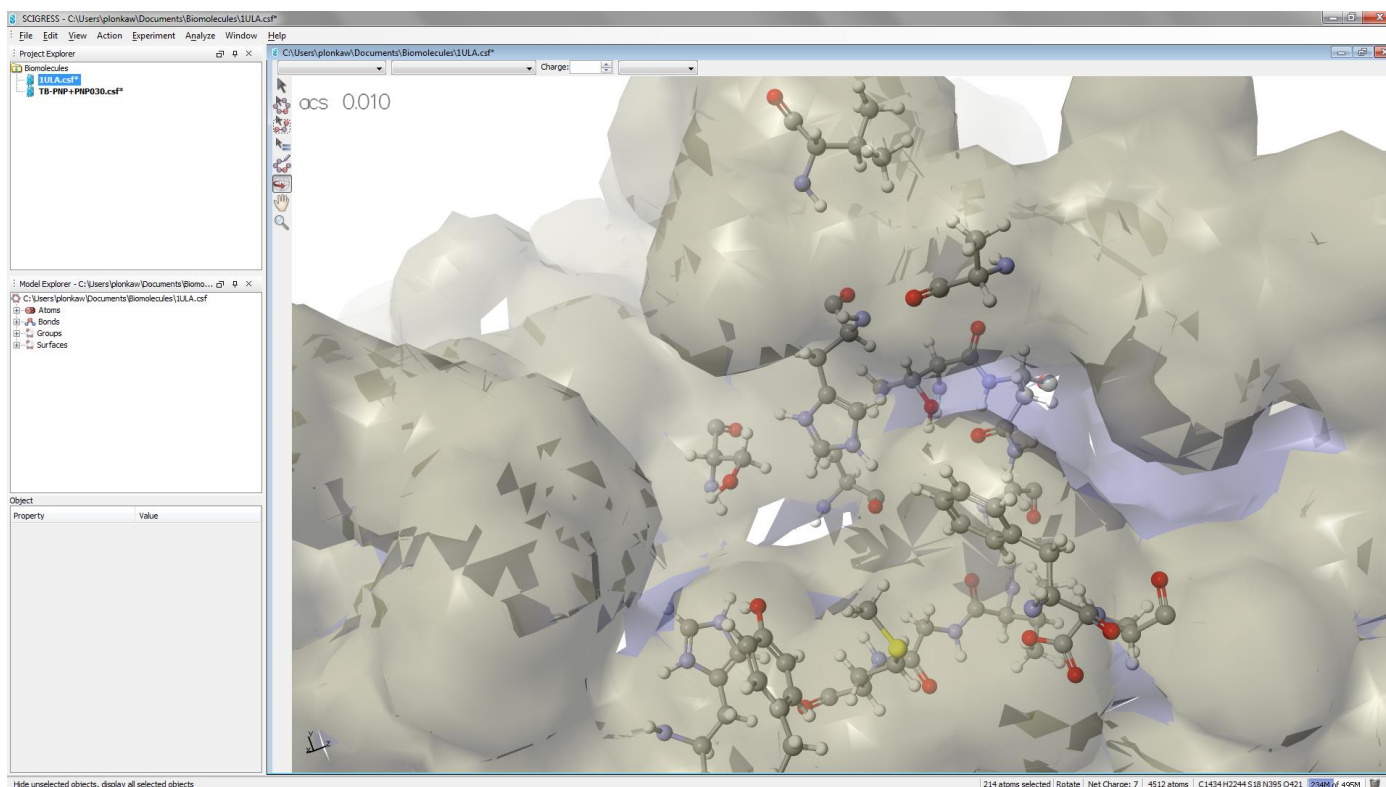
Docking a bound ligand from one protein into a homologous protein

Dock PNP030 into human-PNP by superimposing human-PNP onto TBPNP+ PNP030 and then deleting TB-PNP.

To superimpose human-PNP and TB-PNP

1 Click the Sequence View window The Sequence View window comes to the front.

2 From the Sequence View window, verify that both TB-PNP+PNP030 and 1ULA sequences are displayed, that their sequences are aligned and that only the active site residues are highlighted and outlined with black boxes. If not, use the Edit | Match Selection to reselect the matching residues in 1ULA.

3 From the Sequence View window, choose Edit | Superimpose Sequences. The Superimpose Sequences dialog opens.

4 Choose 1ULA as the Probe Sequence and TB-PNP+PNP030 as the Target Sequence.
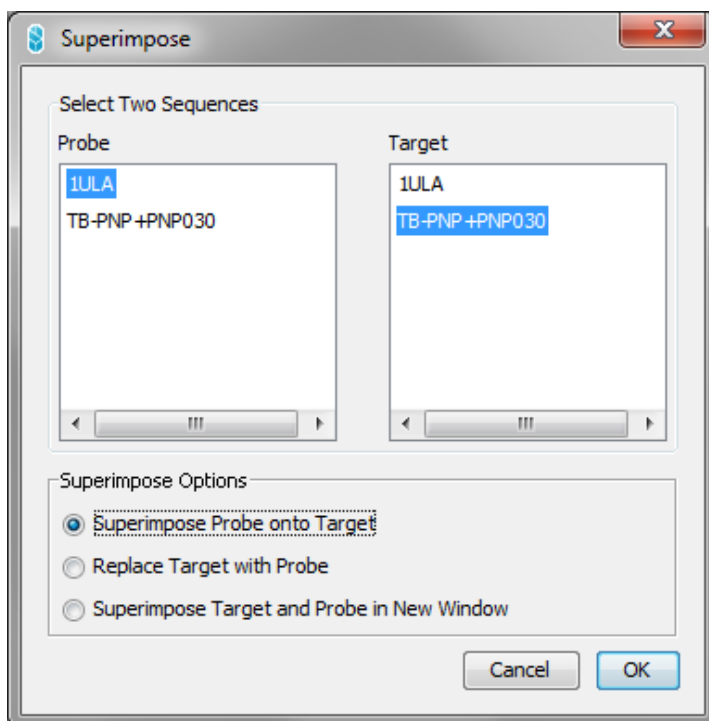
5 Choose Superimpose Probe onto Target and click OK. After approximately 10 seconds, the Superimpose Sequences dialog closes and superposition is complete.

6 Choose Window | TB-PNP+PNP030.csf. The TB-PNP+PNP030 3D structure comes to the front with both TB-PNP and 1ULA superimposed.

7 In the 3D structure workspace, choose the Select Molecule tool and click TB-PNP (TB-PNP is the unselected and dimmed protein). TB-PNP highlights. 1ULA, PNP030, PO4 3-, and the water molecules are grayed.

8 Choose Edit | Delete. TB-PNP disappears and the remaining portions of the structure become selected.

9 Choose View | Color by | Molecule
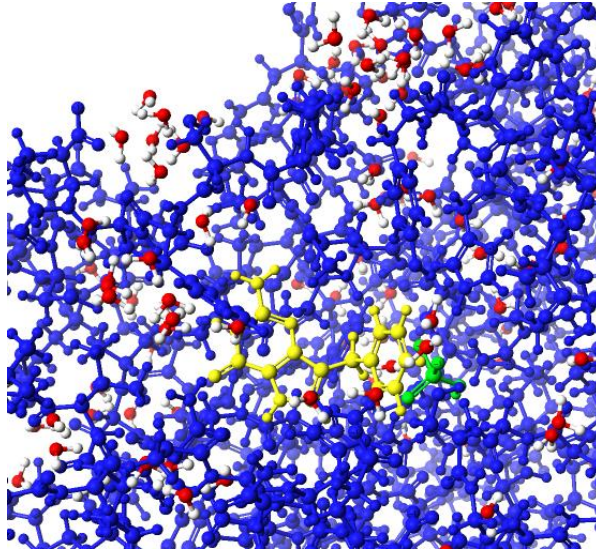
So what?

You have seen how the new ligand, PNP030, could be docked in human-PNP using a superposition method.

Automated Docking and Scoring

This tutorial describes how to automatically dock and score a ligand in an active site from the Workspace application using a potential of mean force (PMF) with a genetic docking algorithm. First, you will validate the docking method by redocking immucillin into its crystal structure in TB-PNP. Then you will dock PNP030 into TB-PNP and compare its score to immucillin.

In this section you will:

• create a copy of immucillin (ImmH) in the TB-PNP chemical sample file

• automatically dock the immucillin copy into the TB-PNP active site

• calculate the RMS distance difference between the docked structure and the original x-ray crystal structure

• dock PNP030 into the active site and compare its score with that of immucillin

Background

Scigress automates the docking of ligands into active sites by using a genetic algorithm with a fast, simplified potential of mean force (PMF).1 The potential of mean force is a knowledge-based approach that extracts pairwise atomic potentials from structure information of known protein-ligand complexes contained in the Protein Data Bank. PMF has been demonstrated to show a significant correlation between experimental binding affinities and its computed score for diverse protein-ligand complexes.2 Similar to empirical force-fields use in Mechanics and Dynamics, PMF uses atom types. PMF has 34 ligand atom types and 16 protein atom types resulting in 544 unique pair potentials. Scigress represents each potential as a numerical table at 0.2 Å intervals in the PMF.prm data file. Docking a ligand into an active site is sensible when both contain only PMF atom types. For ease of use, when Scigress encounters an atom that is not a standard PMF atom type, it assigns it the atom type "PMFnone" and uses the parameters for aliphatic carbon (PMF atom type CF).

The numerical pair potentials contain information for all ligand-protein atom pair distances, including the short distances that are typical of covalent bonding. Docking models used in Scigress assume that the protein and ligand dock non-covalently. It is assumed that bonds are not formed between the ligand and protein because bond formation would cause changes in the atom types and possible substantial changes in the shape of the ligand and active site.

Therefore, non-bonded terms are added to the potential to keep atoms at typical non-bonded distances. The standard PMF implementation uses Amber3 van der Waals potentials for this purpose. The standard Scigress implementation uses specific 6-12 Lennard-Jones potentials for each pairwise interaction.4

Scigress' specific pairwise potentials permit hydrogen bonding and electrostatic interactions which are sometimes prohibited by the Van derWaals potentials.

PMF offers these advantages:

• unlike mechanics based methods, there is no need to add hydrogen atoms to the protein or to worry about the charges in the protein

• it is based on experimental data and can be refined as more data become available

• unlike molecular mechanics, there is no partitioning of the energy into many large and mostly cancelling terms, e.g. Van der Waals and electrostatics

• it is fast and simple

• potentials implicitly include solvent and entropic effects because they are derived from crystal data1

• no knowledge of the binding affinities is required to derive the potentials Genetic algorithms (GA) are a standardmathematical technique for optimizing functions of many variables. They are inspired by evolution. These algorithms encode a possible docking solution in a simple sequence of numbers called a chromosome. Here, the chromosome for a flexible ligand contains translation, rotation, and dihedral angle numbers that describe the position, orientation, and conformation of the ligand in the active site.

The genetic algorithm begins by creating a random population of chromosomes (i.e. random positions, orientations, and conformations of the ligand) and then applying recombination operators to these chromosomes so as to preserve those which are the best or most fit. It preserves the best solutions by evaluating the PMF score of each chromosome and allocating reproductive opportunities in such a way that those chromosomes which have a lower score are given more chances to "reproduce" than those chromosomes which are poorer solutions. The process continues for a specified number of generations (usually 1,000 - 40,000). The use of a Lamarckian genetic algorithm for ligand-protein docking was pioneered by Art Olson's group at Scripps in the program AutoDock.2 Scigress uses the GA for docking because it offers the advantage of relatively unbiased searching and it does not require derivatives of the potential. In Scigress, the docking is embedded in the FastDock compute engine. Similar to AutoDock, FastDock uses a Lamarckian GA in which a minimization is performed so that individuals adapt to the environment in which they are placed.

The goal of this exercise is to introduce how to use Scigress to design your own selective inhibitors by docking flexible ligands into a rigid protein active site. Of course, proteins are not rigid and often flex to accommodate ligands. You can overcome this limitation by allowing the protein active site to flex during docking, but this increases the time and complexity required for the calculations dramatically. It also complicates the analysis and interpretation of the structure-activity relationships. With Scigress you can apply either strategy but in this exercise we will focus only on the fast one.

Validating the PMF method

Before using any new scientific method, you should first validate that the method works in situations where you know the result. Here we will dock immucillin-H into TB-PNP and compare the docked result to the crystal structure of the bound ligand-protein complex.

To dock a copy of immucillin into TB-PNP

1 Navigate to the Biomolecules and docking folder 6-Automated Docking, choosing **File | Import** and select TB-PNP-All-Refined.csf. TB-PNP-All-Refined.csf opens and displays the cleaned 3D structure.

2 Choose **Action | Group Atoms**, select ImmH from the list of **Defined Groups**, and click **G Only**. ImmH is selected and highlighted.

3 Click **OK**. The Group Atoms dialog closes.

4 Choose **Edit | Copy**. A copy of ImmH is placed on the clipboard.

5 Choose **Edit | Paste**. A copy of ImmH is added to the 3D structure workspace. It is selected and highlighted.

6 Choose **View | Color by | Element**. The copy of ImmH is colored by element so that it is easy to see which is the copy and which is the original ImmH.

Atoms and bonds in the ImmH copy are colored according to element type

7 Choose Analyze | Dock into Active Site. The Dock into ActiveSite dialog appears.

The list of Ligands on the left contains the names of the ligand type atom groups in each open chemical sample file. When you copied and pasted ImmH, Scigress automatically gave the copy the new group name ImmH1. The list on the right contains the names of the active site groups in each open chemical sample file.

8 Select the ligand ImmH1 and the target ActiveSite, and change the dialog to match the rest of the choices shown above. These settings will dock the flexible ligand ImmH1 into the rigid active site group ActiveSite in the protein TB-PNP-All-Refined. Docking takes place in the Active Site's window using the fast grid-based method. When docking a ligand into a target, all other groups of group type **ligand** in that target are ignored. For example, in this case, ImmH will be ignored when you dock ImmH1. This allows you to have many ligands overlaid in an active site.
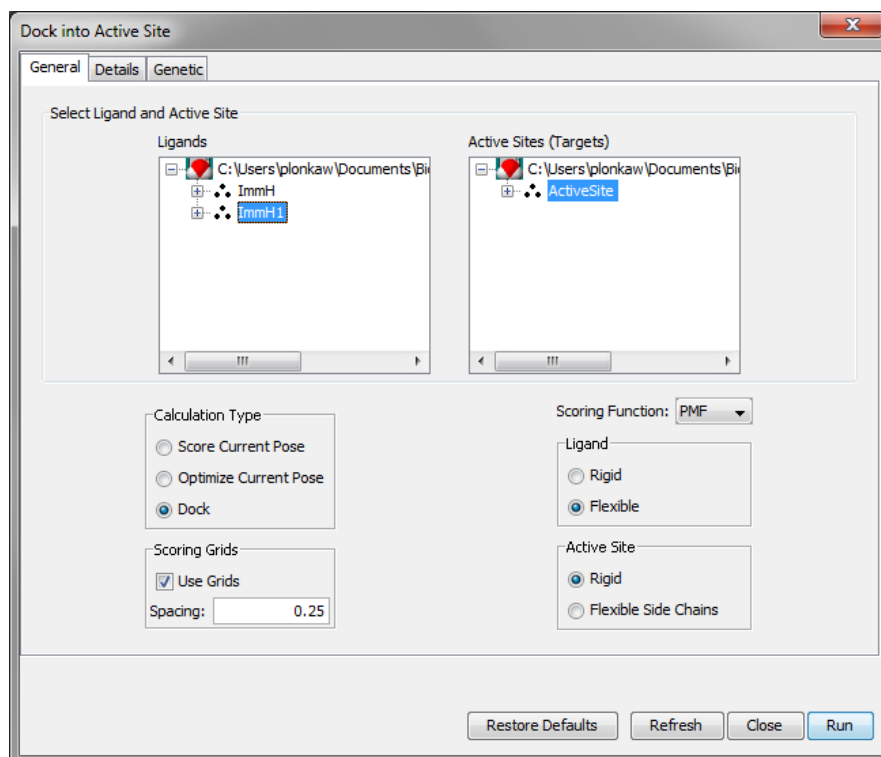
9 Click the Details and later Gentic tab and change options to match those shown:
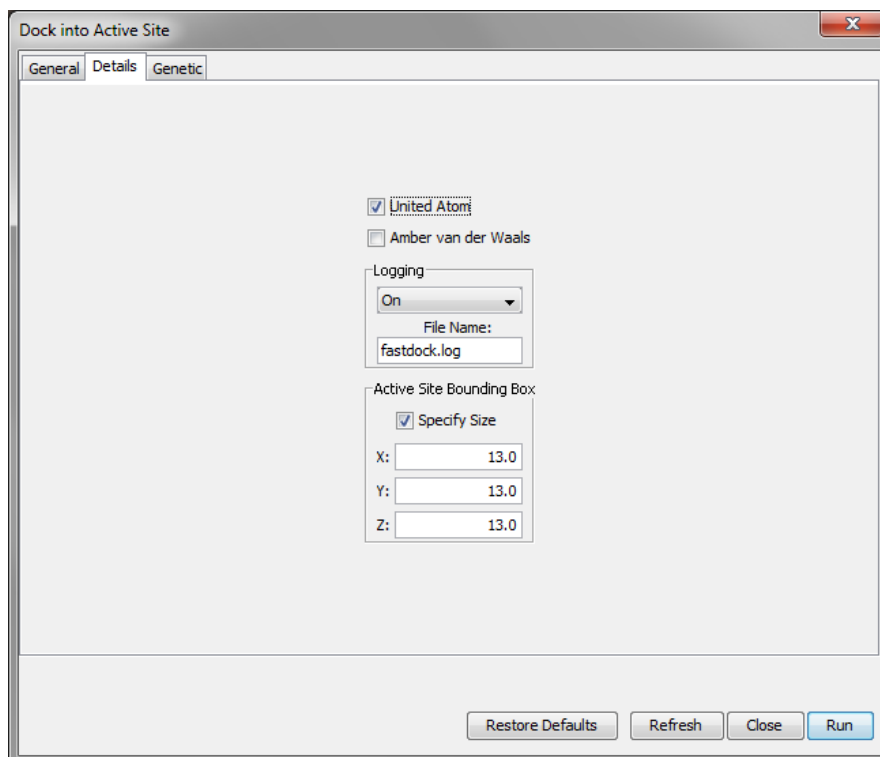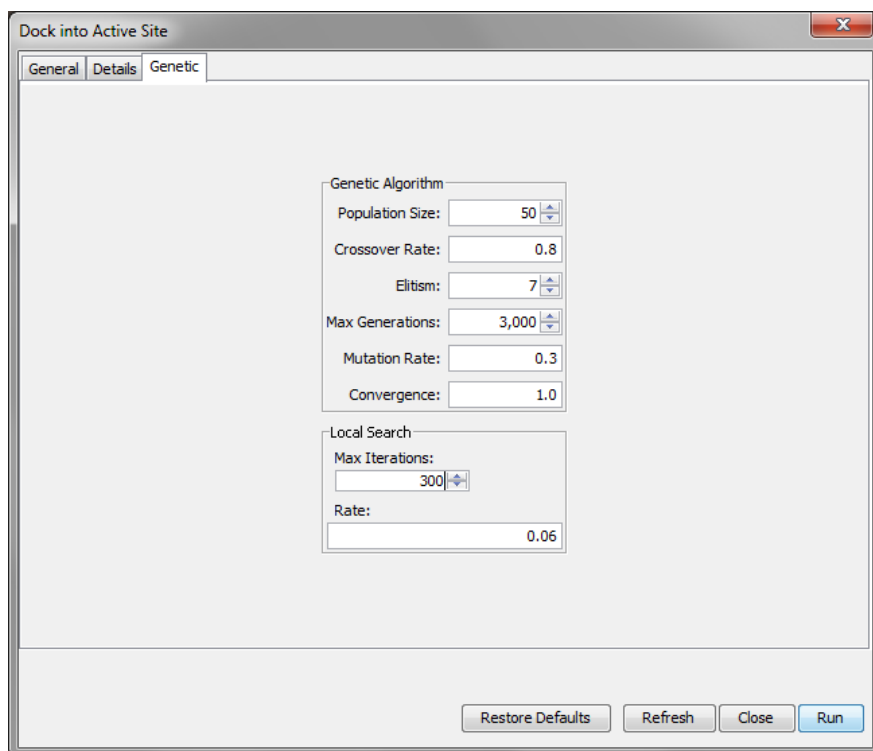
These settings are for docking a ligand into a 13x13x13 Å box located at the center of the target active site using a united atom (explicit hydrogen atoms are not considered) potential of mean force with a docking algorithm that has a population of 50 chromosomes and runs for 3,000 generations. The outcome of the calculation is a map file with different docking geometries (2 in current example)

10 Choose Analyze | New Conformation Analysis window. Select the lowest energy entry (there may be only one) and go to File | Save selected structures as. Click OK. Close the conformation search and map windows.

11 Double click the previously created conformation1.csf window. Choose **Action | Group Atoms**, select ActiveSite from the list of **Defined Groups**, choose **G only**.

12 Choose ImmH2 from the list of **Defined Groups**, click **S or G** and click **OK.** The Group Atoms dialog closes and the active site, and the docked ImmH2 are highlighted in the 3D structure window.

13 Choose View | Hide Unselected.

14 Choose View | Fit in Window. The transparent portion of the protein disappears and the active site and ligands zoom to fill the center of the 3D structure window.

15 Choose View | Color by | Molecule. The docked ligand and active site are clearly visible.
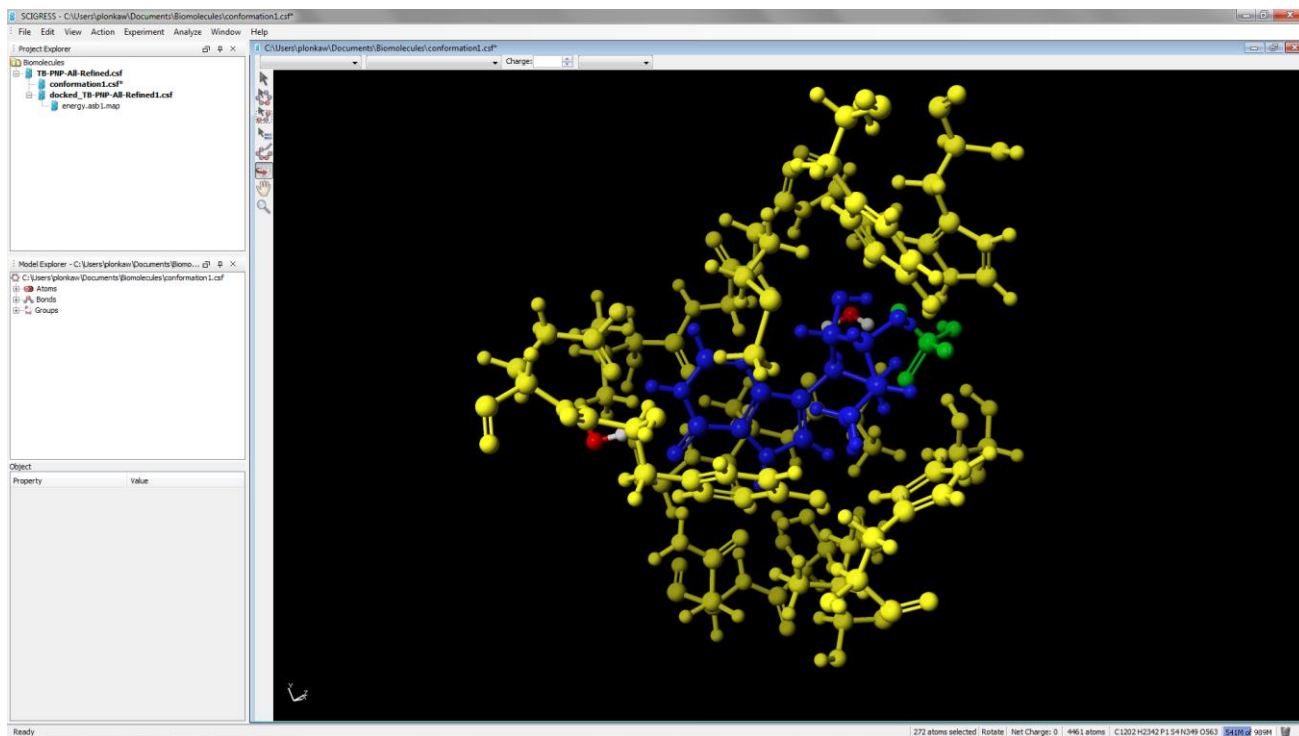
Discovering Inhibitors by Automatically Scoring and Docking Libraries using PMF and quantum docking.

You can use Scigress to screen virtual compound libraries so that you can select those compounds that are most likely to have high binding affinities. This tutorial describes how to build a model of known PNP inhibitor within the active site of bovine PNP (1A9T) and then rank order a library of compounds by docking the compounds in bovine PNP and scoring their optimal pose.

A useful docking method:

• correctly reproduces known crystal structures

• scores the ligands in order of activity

• runs quickly on libraries of compounds

In this section you will:

• import 2D structures of a library of substituted 9-deazaguanines, automatically converting them to 3D structures

• automatically dock each member of the library into bovine PNP and score the optimal pose using LocalSCF quantum semiempirical docking procedure. (V. M. Anisimov, V. L. Bugaenko; J. Comp. Chem., 2009, 30 (5) 784-798; "QM/QM Docking Method Based on the Variational Finite Localized Molecular Orbital Approximation")

• examine the docked structures

Background

One class of compounds that has been extensively characterized as PNP inhibitors is the substituted 9- deazaguanines. Data include measured binding constants, IC50 and EC50 values, as well as high resolution x-ray crystal structures of substrates, products and inhibitors bound to bovine and human PNP.1 In this exercise you will automatically dock a library of 19 compounds into the active site from the bovine PNP crystal structure 1A9T.pdb. Of the several crystal structures available, this crystal structure was chosen because it contained the product substrate that requires the largest pocket. As you will

see, the active site model here differs from the single chain active site model you generated for TB-PNP in that it uses the full trimeric chain structure and contains a capping phenyl alanine residue from an adjacent chain. The availability of these experimental data provides a firm foundation for evaluating computational strategies in deriving structured-guided structure activity relationships that are a necessary part of any drug discovery program. This area has been revolutionized by the development of fast, automated methods for obtaining docked ligands inside proteins. The goal of this chapter is to teach you how to use Scigress to design your own selective inhibitors by docking a library of flexible ligands into a rigid protein active site. One of the consequences of using a rigid protein is that you must be careful to be sure that

the library you screen contains similar sized structures.

Of course, you can overcome this limitation by allowing the protein active site to flex during docking, but this increases the time and complexity required for the calculations dramatically. It also complicates the analysis and interpretation of the structure activity relationships.

With Scigress, you can apply either strategy, but here you will focus on docking a flexible ligand into a rigid active site. In previous sections, you have worked one molecule at a time. Here you will use automated methods to work with a library of compounds.

First, navigate to Biomolecules and Docking / 7-Discovering Inhibitors by Scoring Docked Structures using File | Import File and import the 1AT9.csf file containing the protein with active site. The protein has balanced overall charge (needed for quantum calculations) and active site defined. Those two steps are required to perform mixed PMF/quantum docking.

Setting up the library

You will often find 2D structures for libraries in chemical databases. Chemical databases allow you to export the structures and related data in single SD format file. Read these files directly into Scigress Docking Spreadsheet and it automatically converts the structures to 3D, creates chemical sample files, and places the associated activity data into a table.

To import a library of compounds from SD file:

1 Choose File | Import File and navigate to Biomolecules and Docking / 7-Discovering Inhibitors by Scoring Docked Structures and import the PNP Ligand Library.sdf file. Leave the default import options.

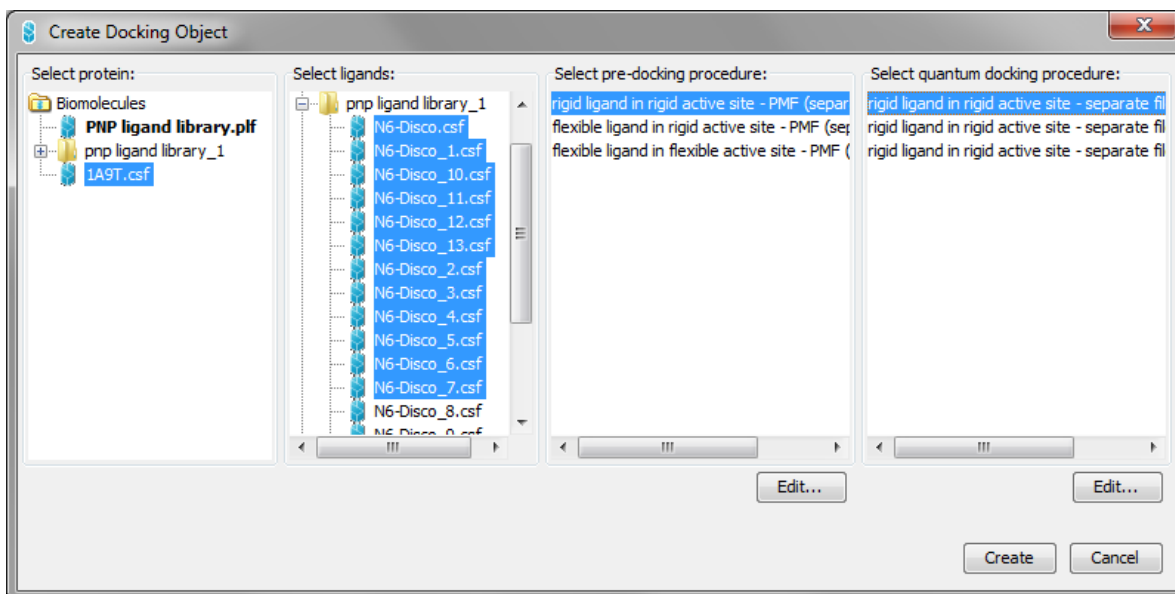2 Right-click in Project explorer to access New Object menu option
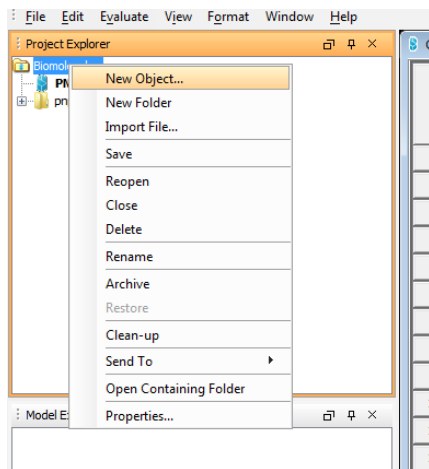
3 Choose the "Docking" from list of object types, and Quantum Docking from the Type list:
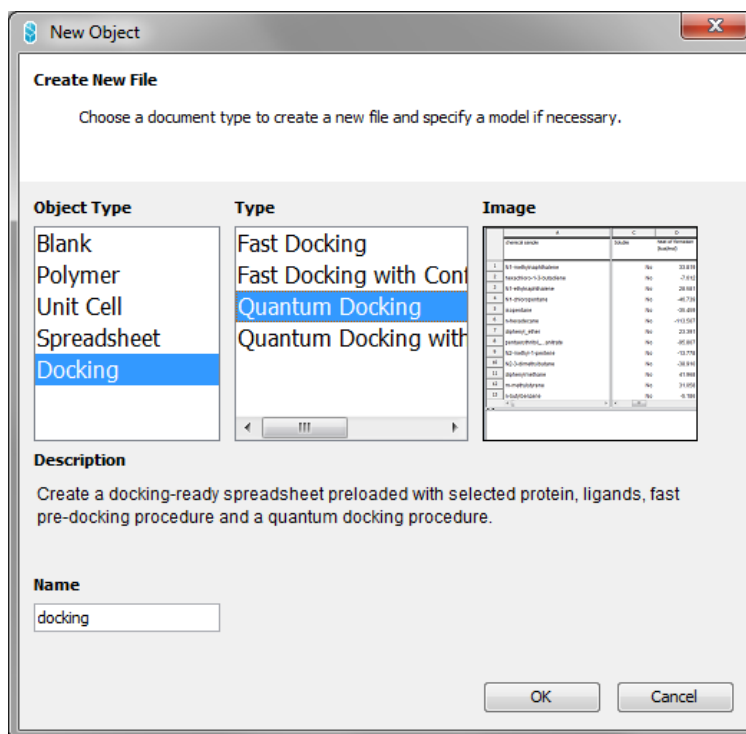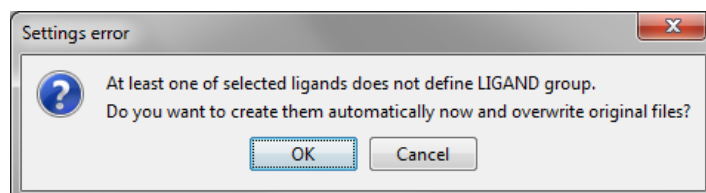
4 Press OK

5 Select 1A9T as the protein, and one or more ligands from the pnp_ligand_library folder. The more you select, the longer the calculations will tak. Select "rigid ligand in rigid active site" for both pre-docking and docking procedures, as shown below:

4 Press Create

5 Press OK in the subsequent warning dialog

Settings error

At least one of selected ligands does not define LIGAND group.
Do you want to create them automatically now and overwrite original files?

OK   Cancel

6 A chemical spreadsheet ready for docking calculations is created

C:\Users\plonkaw\Documents\Biomolecules\docking.plf

	A	B	C	D	E
	chemical sample	dock a ligand into an active site [score] (from A)	quantum scoring (ligand binding energy) [kcal/mol] (from B)	quantum scoring (ligand-receptor complex energy) [kcal/mol] (from C)	
1	N6-Disco				
2	N6-Disco_1				
3	N6-Disco_10				
4	N6-Disco_11				
5	N6-Disco_12				
6	N6-Disco_13				
7	N6-Disco_2				
8	N6-Disco_3				
9	N6-Disco_4				

Explanation Worksheet 1

5 Select the whole worksheet and press Crtl-E to start automated docking of chosen ligands to the active site of the 1A9T protein. The calculation should take a couple of minutes on a modern PC.

The Column "C" contains the ligand-receptor binding energy (docking score). The lower the value, the higher the affinity of ligand to the receptor.

So what?

In this exercise you have learned how to set up automated docking of multiple ligands to a protein. The methods selected "Rigid ligand in rigid active site" were the fastest, but least accurate. Please explore SCIGRESS docking capabilities, by repeating the exercise employing the Flexibility and Conformer generation methods.

Part 6 Basic QSAR analysis in SCIGRESS

This tutorial describes how to perform simple QSAR analysis using automated procedures in SCIGRESS. You'll learn how to:

• import SDF file and define experimental value

• choose QSAR method/descriptor set

• analyze the results

Background

In this tutorial, we will show how to perform simple QSAR analysis – prediction of melting points of fatty acids. Data for this tutorial is based on the publication:

Guendouzi A, Mekelleche SM. Prediction of the melting points of fatty acids from computed molecular descriptors: a quantitative structure-property relationship study. *Chem Phys Lipids*. 2012; 165(1):1-6.

Importing molecules

Importing molecules is a first step of the analysis

1 Create a project "QSAR"

2 In the Workspace application, choose **File | Import**….

3 Navigate to the location of the Scigress application files, for example:

`C:\Program Files\Fujitsu\Scigress Suite\Biomolecules and Docking\8-QSAR Fatty Acids Melting Points`

4 Choose MDL files (*.mdl, *.mol, *.sdf, *.sd) from the Files of Type drop down list.

5 Click fatty_acids.sdf.

6 Check **Create spreadsheet** option.

7 Click **Open**

8 Select **MDL mol name**, and leave **Load additional data** selected.

9 Click **OK**

The molecules in the SDF files were cleaned and pre-optimized, but if necessary spreadsheet procedures can be used to prepare molecules for QSAR.

SCIGRESS will consider any column with numerical values as a **data to predict** but for clarity it is recommended to assign a desired column an appropriate name using **Edit Header** option:
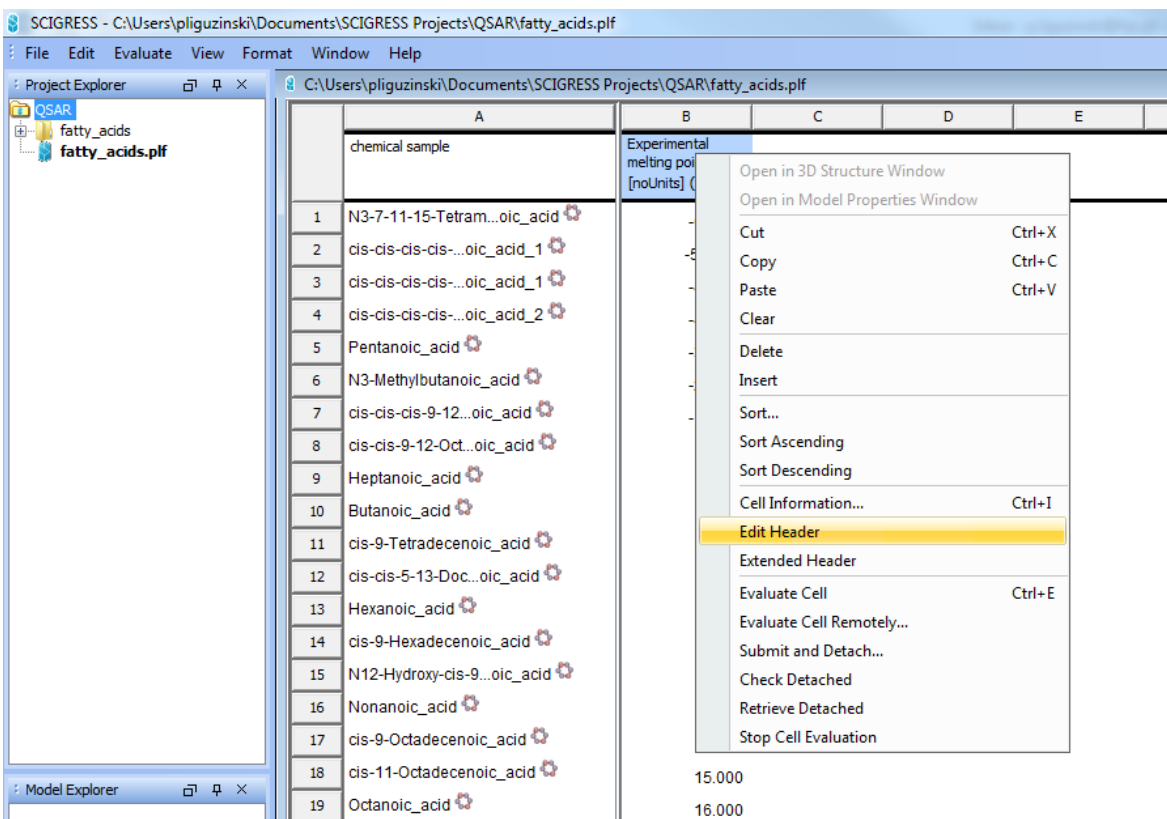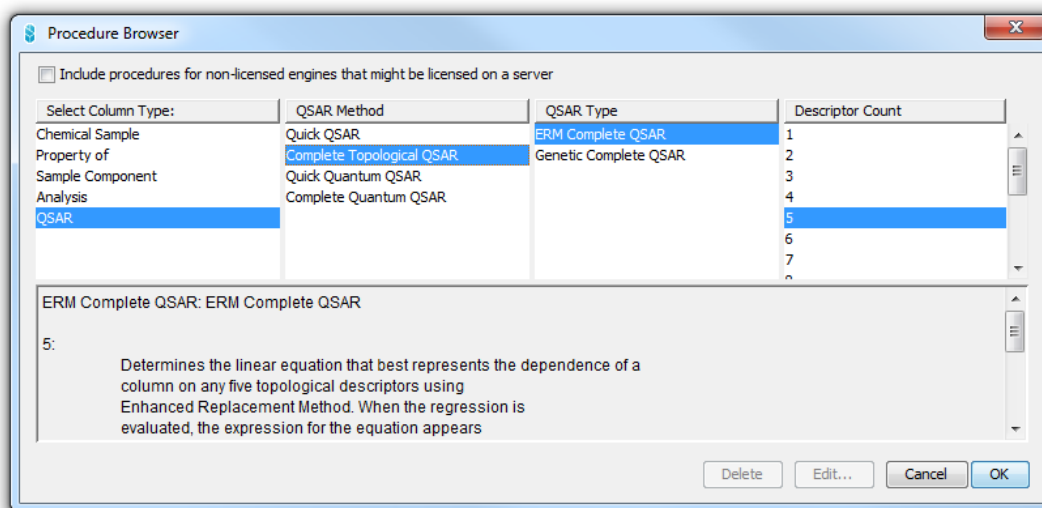
Rename the imported column to "**Experimental melting point**"

Creating QSAR models is a standard spreadsheet procedure:

1 Double click on the first empty column.

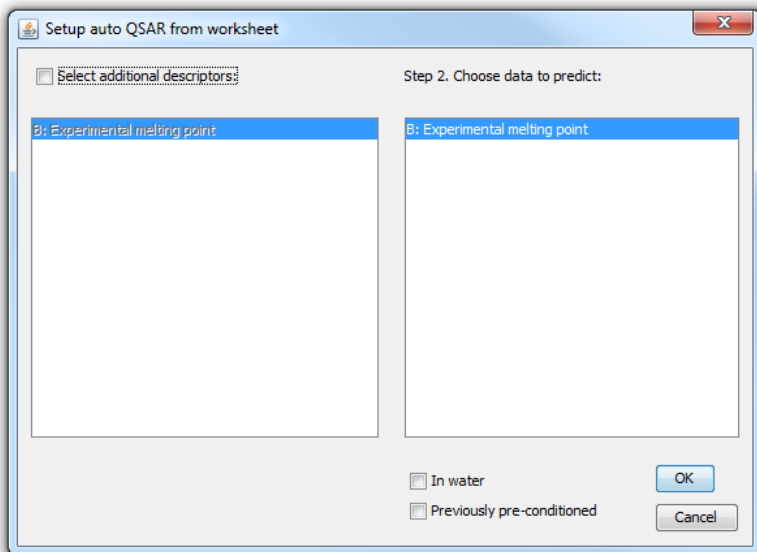**2** Choose **Select Column Type: | QSAR**

3 Select **QSAR Method | Complete Topological QSAR**

4 Leave other parameters on default values (**ERM Complete QSAR | 5**)

5 Press **OK**

6 The following menu will appear. It is used for defining **data to predict** and **descriptors from external sources**:

7 Press **OK**

8 Select "QSAR" column and select **Evaluate Cell** (right-click).

9 A warning message will appear suggesting that data is not uniformly distributed. It can be ignored for tutorial purposes.

10 Press **Yes** (Do you want to continue?)

SCIGRESS QSAR model consists of one spreadsheet column and three spreadsheet tabs (**Report**, **Plot**, **Predict**). After successful model creation SCIGRESS will switch to **Predict** tab. During model creation data outliers are identified and appropriate warning message will be displayed.
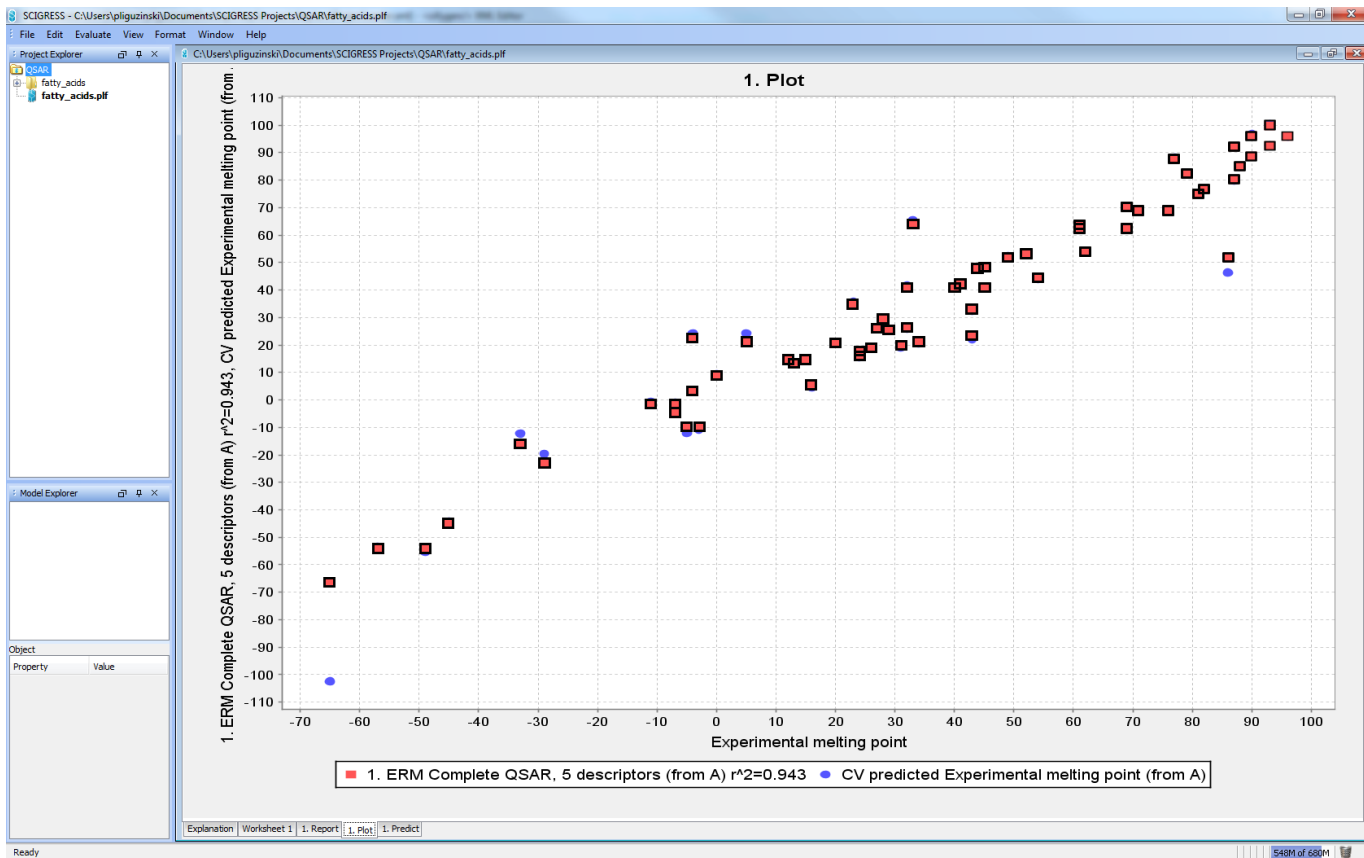
1 Click on **Report** Tab

2 Inspect following data in the **QSAR report**:

- Best QSAR equation

- Applicable prediction range and chemical space

- Analysis of QSAR equation

- List of descriptors evaluated

3 Click on **Plot** Tab

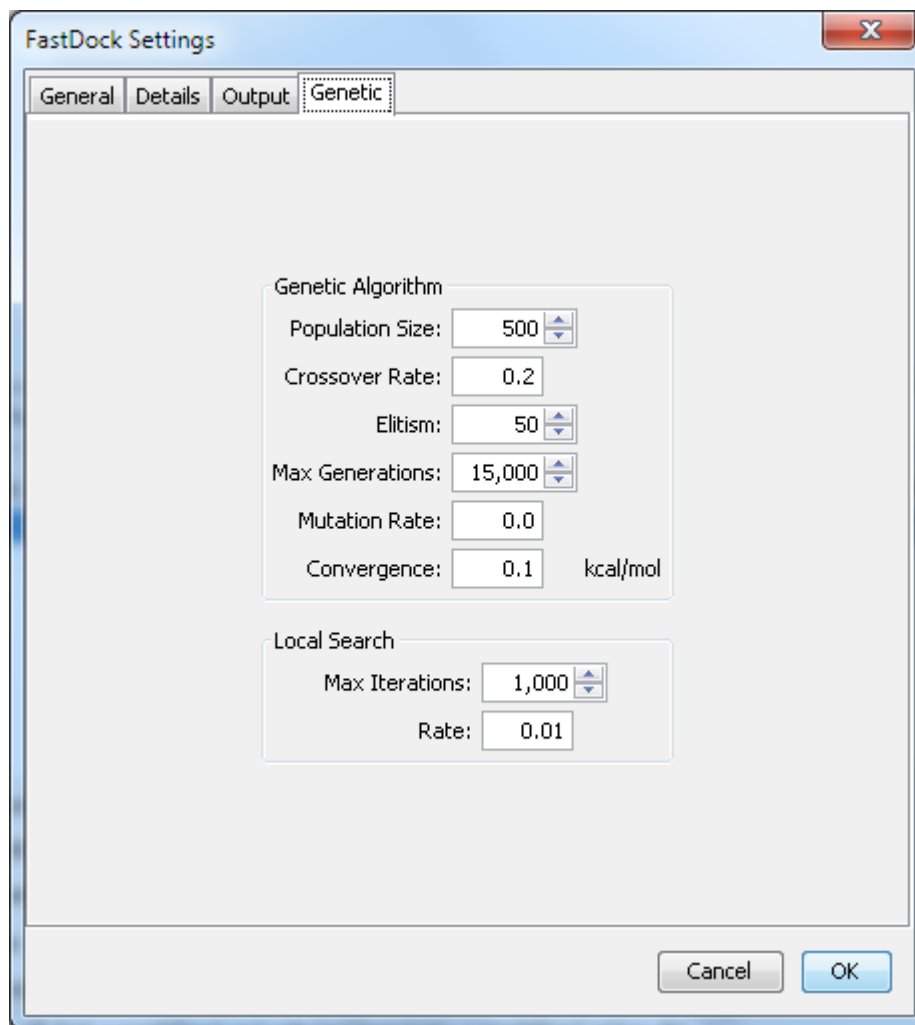**4** Select **ERM Complete QSAR** data by clicking at graph legend:

4 Choose **View | Regression** to display regression equation.

Predict tab allows reusing model for predicting new molecules' properties.

Appendix I Understanding FastDock genetic algorithm settings

The FastDock genetic algorithm settings dialog looks as following:

Population size

The number of genomes or chromosomes in the population.

Crossover rate

After the fitness score for a population is calculated, the most fit are selected and allowed to reproduce. The crossover rate is the fraction of the genes that are crossed during reproduction.

Elitism

The number of best solutions which are retained unchanged in the next generation.

Maximum number of generations

The maximum number of generations for this docking.

Tip: You will usually chose between 1,000 and 40,000 generations. The more flexible the ligand or protein, the more generations that are required to achieve the docking with the lowest possible score. For typical drug –sized ligands, 10,000 generations will give a docked structure within a few kcal/mol of the best possible. Use 40,000 generations or more for publication quality results.

Mutation rate

After selection and crossover of the population, genes are mutated at this rate.

Docking convergence

Continue evolving the population until the score converges to less than this number or the maximum number of generations is reached.

Local search rate

The rate for the local search usually set to about 0.04.

Recommended values of the parameters are:

Population size: 500

Crossover rate: 0.2

Elitism: 50

Max generations: 15 000

Mutation rate: 0.0

Convergence: 0.1 kcal/mol

Max iterations: 1000

Rate: 0.01