

Colby College Molecular Mechanics Exercises
MOE (Molecular Operating Environment) Exercises
September 2011

Thomas W. Shattuck
Department of Chemistry, Colby College
Waterville, Maine 04901



**A Teaching License for the classroom use of MOE has been kindly granted by
the Chemical Computing Group**

1010 Sherbrooke St. West
Suite 910
Montreal, Quebec
H3A 2R7
Canada

www.chemcomp.com

Please, feel free to use this tutorial in any way you wish ,
provided that you acknowledge the source
and you notify us of your usage.

Please notify us by e-mail at twshattu@colby.edu
or at the above address.

This material is supplied as is, with no guarantee of correctness.

If you find any errors, please send us a note.

Table of Contents

MOE: Molecular Operating Environment:

- 1: Building and Minimizing
- 2: Conformational Preference of Methylcyclohexane
- 3: Geometry (or How Does Molecular Mechanics Measure Up?)
- 4: Building More Complex Structures: 1-Methyl-trans-Decalin
- 5: Conformational Preference for Butane
- 6: MM3
- 7: Comparing Structures
- 8: Plotting Structures
- 9: Conformational Preference of Small Peptides
- 10: Dynamics in Small Peptides
11. Solvation and β -Cyclodextrin
12. Docking: β -Cyclodextrin and β -Naphthol
- 13: Henry's Law Constants and Gibb's Free Energy of Solvation
14. Distance Geometry
15. Protein Structure and Gramicidin-S

An accompanying Introduction to Molecular Mechanics is available at:

<http://www.colby.edu/chemistry/CompChem/MMtutor.pdf>

MOE: Molecular Operating Environment

Introduction

MOE is a molecular modeling program, which is specifically designed to handle large biological molecules. MOE is designed to use several different force fields and semi-empirical and *ab initio* quantum mechanics calculations.

General Notes:

The following exercises are designed to be done in order. Detailed instructions given in earlier exercises will not be repeated in later exercises. If you have questions, turn to this tutorial or use the MOE online tutorials and manuals. The MOE manuals have many interesting examples that extend well beyond the skills taught here. MOE is actually very easy to learn. Follow these instructions carefully until you get the feel of the program. Then try new things. Don't hesitate to explore MOE on your own.

Overview: HIV-Protease and Indinavir

Our first goal is to give you a quick tour of MOE and some of the capabilities for biomolecular visualization. In this exercise we are going to start with the global view and then focus on smaller and smaller portions of the structure. As we delete more atoms, however, we will look in greater detail at the molecular structure and interactions. The structure for this HIV-protease complex is taken from the Protein Data Bank (1HSG.PDB).¹ These atom positions were determined using X-ray crystallography. This structure and similar structures were used to design the drug Indinavir, which is now commonly used for the treatment and prevention of AIDS infections. The molecule that is docked in the binding site of a protein is called the ligand, which is Indinavir in this case.

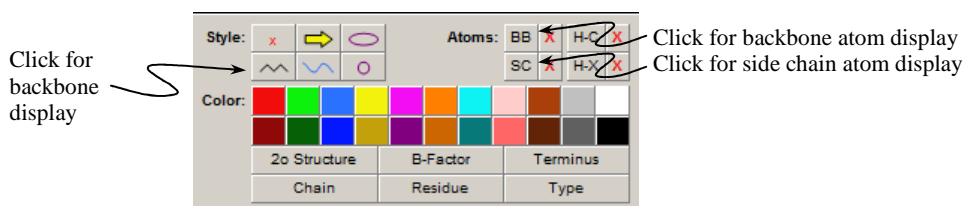
1. Pull down the **File** menu and choose **Open**.


- On Windows systems type in "c:\Documents and Settings\All Users\Documents\moefiles" and press Enter to switch into the proper subdirectory. Click the CWD button to select this directory as your default.
- On OSX systems type in "Documents/moefiles" and press Enter to switch into the proper subdirectory. Click the CWD button to select this directory as your default.

Choose "1hsgcartoon.moe."

2. The HIV-protease should be shown in cartoon form, with the ligand rendered in Van der Waals spheres. The yellow areas show regions of the protein that are in the "beta pleated sheet" form. This form for the protein backbone produces flat extended regions. The red regions are in a more compact helical form. This view shows the drug neatly tucked under two folds of beta pleated sheet. To see these regions in more useful molecular form, continue on.








3. Click on the **Ribbon** button in the bar that runs across the bottom of the screen:



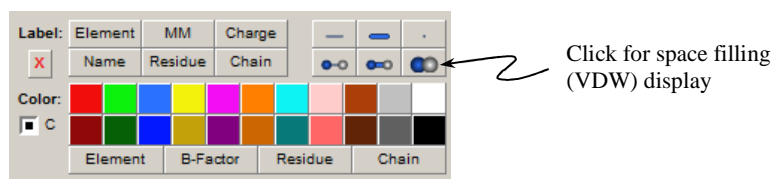
Click on the  button to display the protein backbone. (Try out the other buttons to see the effect, but return to the backbone display with 2 σ structure coloring.) The atom positions in the protein chain backbone will now be shown. The color-coding will be the same as in the original cartoon. Rotate the structure to note the helical arrangement of the red alpha helical regions. The action of the mouse buttons and track ball is given below.







Rotations, Translations, and Zooming To change the orientation, size, and position of the molecule, you can use either of two methods, (1) using the mouse or (2) using the track ball (the soccer ball). To use the mouse, position the cursor in the main window and hold down the center mouse button. Dragging with the center mouse button reorients the molecule. If you wish to rotate the molecule only around the axis perpendicular to the screen, hold down the middle mouse button and drag the cursor in the periphery of the molecule window.

Middle: Reorient molecule—xyz rotation
 Middle and drag in periphery of viewing area: rotate around z only
 Shift-Middle: xyz translation
 Ctrl-Middle: zoom in and out or use the roller
 Shift-Alt-Middle: Translate selected atoms
 Alt-Middle: Rotate selected atoms
 Alt-Left: change dihedral angle between two selected atoms

Alternatively, you can use the track ball. Dragging with the left mouse button on the track ball rotates the molecule. To translate the molecule, click on the  button in the lower-left hand side of the track ball box, the button will switch to a . Dragging with the left mouse button over the track ball then translates the molecule. To switch back to rotation, click on the  button. In other words the  and  buttons toggle between the two states. To zoom the molecule drag with the left mouse button over the thumbwheel, , that is located just below the track ball. Clicking on  at the right of the screen will allow you to start fresh with a centered molecule.

You can also change the atom rendering. Click on the **Atoms** button in the bar that runs across the bottom of the screen:



Click on the  button for ball-stick rendering. Try all six of the rendering options. Click on the  overlapping spheres button to return the display to space filling rendering. Finally, click on the thin  line-rendering button. Click the  button in the Style group to remove the backbone trace. Then click on the  and  buttons to show all the non-hydrogen atoms. Now all the non-hydrogen atoms will be shown. What we want to do now is to reduce the complexity by focusing on just one particular part of Indinavir and its interaction with the protein.

4. Make sure the rendering style for all atoms is the line style. To ensure that all the atoms are labeled by element, click on the **Atoms** button and then select the button in the Color selection group in the bottom row of the Atoms dialog box. Pull down the **Selection** menu and choose **Ligand**. Pink squares should now highlight just the ligand atoms. Click on the **Atoms** button and choose the light green color. In general note that:

An action is applied to selected atoms only. If no atoms are selected, the action is applied to all atoms.

Pull down the **Selection** menu, slide right on **Extend**, and choose **Near (4.5A)**. Pull down the **Selection** menu, slide right on **Extend**, and choose **Residue**. Now all the amino acids near the ligand should be highlighted. Pull down the Selection menu and choose **Invert**. Click on the button at the right-side of the screen. Click on OK in the Delete dialog box that follows. Now only the amino acids close to the ligand should be shown.

5. Click anywhere in the black background to make sure no atoms are selected (no pink boxes). To see the fit of the ligand in the binding pocket of the enzyme, click on the **Atoms** button and choose space filling display, . Notice that the ligand and the enzyme are in close contact, with little or no empty space remaining. This is because the medicinal chemists who designed Indinavir built the drug to efficiently pack into the available space, and also the flaps of the enzyme are flexible and collapse around the ligand.

6. The ability of the enzyme to bind Indinavir is the combination of many small specific interactions. We will focus now on just one of these, a hydrogen bond.

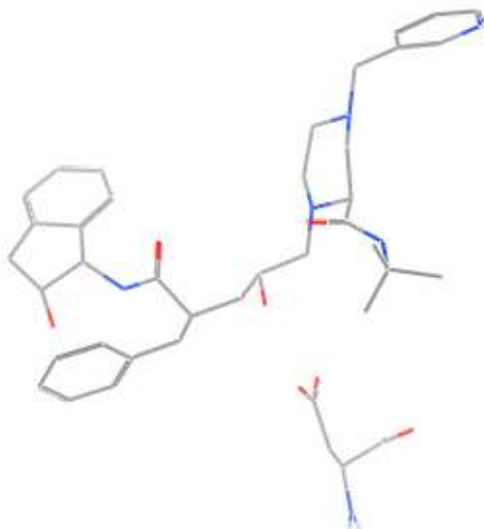
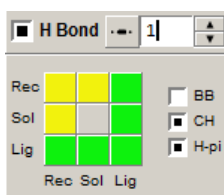


Figure 0.1. Indinavir and aspartate 26 from HIV protease.

Click on the **Atoms** button and choose thin line rendering. Pull down the Window menu and choose **Sequence Editor...** First click on the first "ASP" in the sequence (amino acid 3). Click right on this "ASP", click on **Select**, and then click on **Atoms**. A few atoms in the enzyme should now be selected. Now click on the "MK1" (the last entry). Once again, click right on this "MK1", click on **Select**, and then click on **Atoms**. The ligand should also now be selected along with the single ASP. Close the **Sequence Editor** window. Pull down the **Selection** menu and choose

Invert. Click on the button on the right- side of the screen. Now only Indinivir and one ASP should be on the screen. Click on the Atoms button and choose in the Color group. The screen should appear as shown in Figure 0.1.

7. Notice that there are no hydrogens in either molecule. This lack is because X-ray crystallography is not sensitive to hydrogen atoms. We must use molecular mechanics to fill these in. Structures without hydrogens are very unrealistic, because most of the important interactions between molecules are mediated through the hydrogens. However, before we add the hydrogens we need to fix the position of all the current atoms, since the position of these non-hydrogen atoms are fairly well known from the X-ray experiment. Use the mouse to drag a selection box around all the atoms on the screen (all atoms should have pink boxes). Then select the Constrain button at the right-side of the screen and select **Fix**. To add the hydrogens, pull down the **Edit** menu and choose **Hydrogens** and slide right to choose **Add Hydrogens**. Click anywhere in the black background to make sure no atoms are selected (no pink boxes). Now click on the **Minimize** button to apply molecular mechanics minimization for all the hydrogens. Click on the **Atoms** button and choose ball and stick rendering. The OH on the central part of the ligand should swing around to interact with the COO⁻ on the ASP, Figure 0.2. To display the hydrogen bond, click on the Contacts button on the bottom of the screen:



Make sure the black square is visible to the right of the **H Bond** label, this button toggles hydrogen bond display. The number in the scroll box selects the interaction energy cut-off for labeling a hydrogen bond. Increase the cut-off to 1 kcal/mol. Zoom in to identify the hydrogen bond between the ligand and the protein. Have your instructor check your screen at this point.

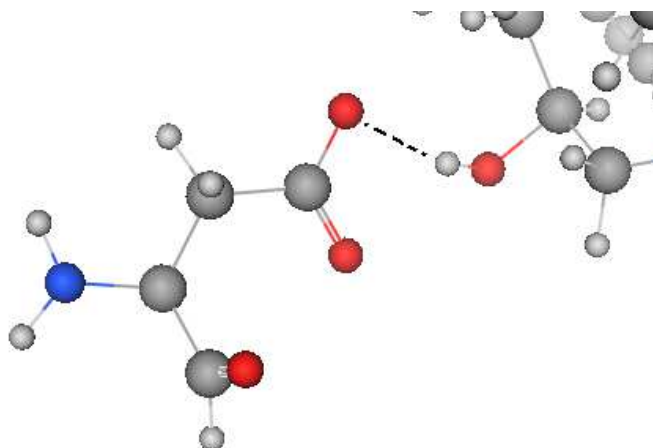


Figure 0.2. Hydrogen bond between Indinivir and aspartate 26 from HIV protease. The dashed line was added to help guide the eye to show the hydrogen bond to the COO⁻.

8. We now will calculate the dihedral energy plot for the –OH in Indinavir to help determine the degree to which the Indinavir O–H hydrogen is interacting with the ligand. Pull down the

Compute Menu, slide right on **Conformations**, and choose **Dihedral Energy Plot**. Now click on the four atoms on the ligand that define the dihedral angle in the order: H-O-C-H. The vertical axis units are kcal/mol and the red line indicates the current dihedral angle. The interaction with the COO⁻ from aspartate contributes strongly to the conformational preference for the -OH group. This dihedral plot is very different from the more symmetrical three-fold potential of a simple alcohol, because of the formation of a hydrogen bond with the COO⁻ of the aspartate. The strength of the interaction between the -OH on Indinavir and the COO⁻ from aspartate makes an important contribution to the efficacy of this drug.² When finished click Close on the Plot window.

9. Click on the **Close** button on the right-hand side of the molecule screen to finish.

References

1. Z.Chen, Y. Li, E.Chen, D. L.Hall,P.L.Darke,C.Culberson, J. A. Shafer, L. C. Kuo, "Crystal Structure At 1.9 Angstroms Resolution Of Human Immunodeficiency Virus (HIV) II Protease Complexed With L-735,524, An Orally Bioavailable Inhibitor Of The HIV Proteases," *J.Biol.Chem.*, **1994**, 269, 26344.
2. E. Rutenber, E. B. Fauman, R. J. Keenan, S. Fong, P. S. Furth, P. R. Ortiz de Montellano, E. Meng, I. D. Kuntz, D. L. Decamp, R. Salto, J. R. Rose, C. S. Craik, R. M. Stroud, " Structure of a Non-Peptide Inhibitor Complexed With HIV-1 Protease. Developing a Cycle of Structure-Based Drug Design," *J. Biol. Chem.*, **1993**, 268, 15343

Chapter 1. Building and Minimizing.

The following exercise will illustrate a few of the options available for structural input, minimization and display using MOE. We will begin with axial-methyl cyclohexane, Figure 1.1. We will use the Builder, where structures may be drawn on the screen. The minimum energy configuration will then be calculated using the MMFF94x force field.

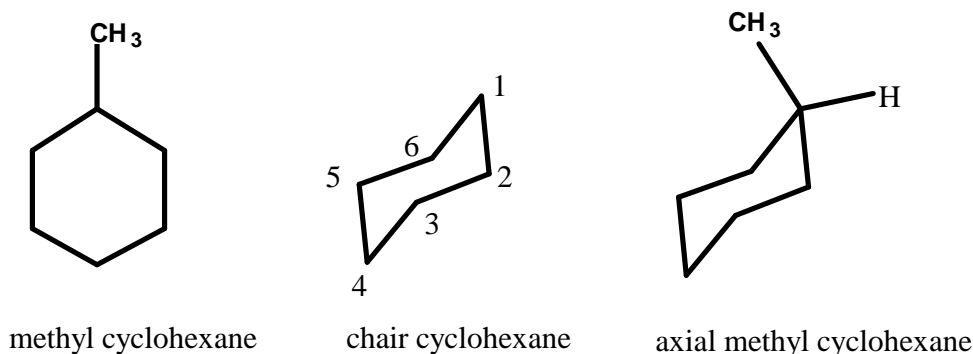


Figure 1.1. Axial-methylcyclohexane

Builder Click on the “Builder” button on the right of the screen. Drag the Builder tool palette to the left-side of the screen, so that you can see the main MOE window. Click on cyclohexane ring button. We now want to add the axial methyl group. Click on an axial hydrogen (you can reorient the molecule using the middle mouse button to see the orientation of the hydrogens). This position should now be marked with a pink square, showing that it is selected. Then click on the “C” button. A methyl group should then be added. The 3D structure is constructed using tabulated values of bond lengths and angles.

Building molecules always has the same flow:
 Start with a simple skeleton, like a ring, click on a hydrogen, and then add the functional group.
 Click on another hydrogen and then add the next functional group, etc.

Make sure no atoms are selected (no pink squares).

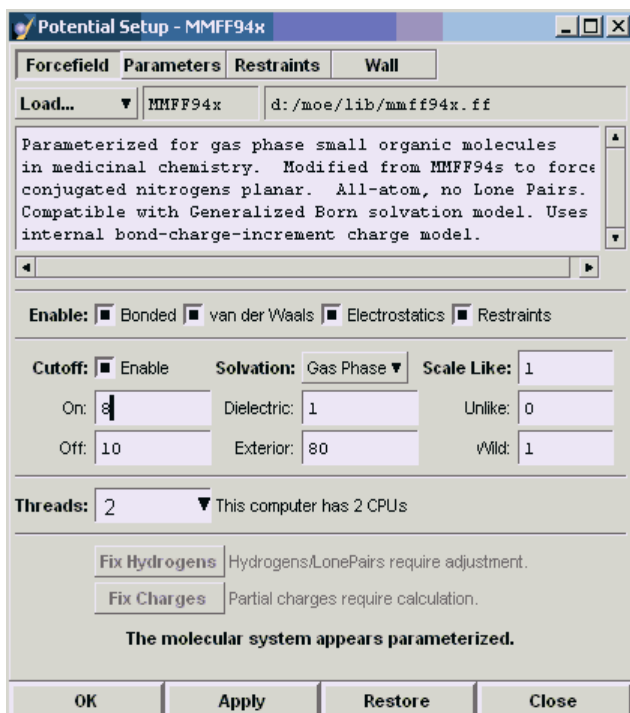
To save this molecule, pull down the **File** menu and choose **Save..** The file librarian dialog box will appear.

- On Windows systems type in “c:\Documents and Settings\All Users\Documents\moefiles” and press Enter to switch into the proper subdirectory.
- On OSX systems type in “Documents/moefiles” and press Enter to switch into the proper subdirectory.

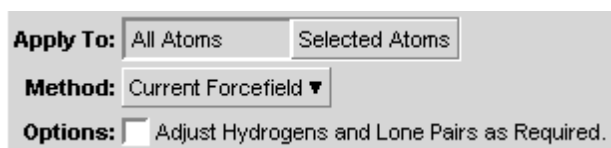
Then click on SetCWD. This button sets the current directory as the default directory. Any further access with the file librarian will default to this subdirectory for opening or saving files. Next double-click on the “smallmolecules” directory. Always save your small molecule work in this directory. Second, type in the name “amecyc6.moe”. Click on OK. Click the Close button in the Builder tools palette. Verify that you have constructed the axial isomer by reorienting the molecule on the screen.

MMFF We next need to find the minimum energy conformation and the steric energy of that structure. We must first specify the force field that we wish to use. Pull down the **Window** menu

and choose **Potential Setup**. If MMFF94x is not already selected, pull down the **Load...** menu and choose MMFF94x. For gas phase calculations make sure the “Solvation:” option is set to Gas Phase, as shown below. If your computer has multiple CPUs, you can select that number of Threads to help speed up calculations. If the Fix Hydrogens or Fix Charges button is highlighted, press that button. Press Apply, if the button is not grayed out.

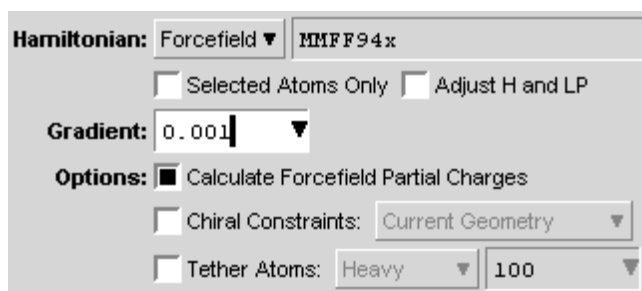


(The Distance Dependent dielectric option is used as a very approximate method for calculations in solvent. Unless otherwise stated we always do calculations in the gas phase, that is with a constant dielectric of 1.) Click on Close. Next the charges on the atoms need to be calculated using the force field standard values for each atom type. Pull down the Compute menu, and choose Partial Charges. Select the Method as Current Forcefield and click OK:



To see the charges that have been assigned, click the Atoms button at the bottom of the screen and click on the **Charge** button. Aliphatic C's and H's are neutral in the MMFF force field, so the charges should be 0.000. If you had O's, etc., then more interesting charges would be assigned. To clear the charge labels, repeat the process: click the Atoms button at the bottom of the screen and click on the **Charge** button, again. The Label buttons toggle. Click on the background to make sure no atoms are selected.

To prepare to minimize the structure, pull down the **GizMOE** menu and choose **Energy** to get the current energy displayed on the screen all the time, and then pull down the **Compute** menu and choose **Energy Minimize ...** Use the default settings for the parameters shown below, if they are not already shown. Except, for small molecules we should choose a smaller gradient of 0.001 for the final minimization steps to get a more accurate value of the energy:



Click on OK. The calculation may stop after a default maximum number of steps, but the energy won't necessarily be minimized. Repeat the minimization process repeatedly until the Giz-MOE energy listed in the upper-left of the main window no longer changes. A couple runs of Minimize will suffice. The final result should be 2.0828 kcal/mol (the equatorial isomer is 0.7115 kcal/mol). Make sure no atoms are selected (no pink squares). Pull down the **Compute** menu and choose **Potential Energy**. We often need to find the contribution to the total steric energy for each degree of freedom, i.e. bond stretching, bond angle bending, stretch-bend interaction (stb), dihedral torsions, out-of-plane bending (oop), electrostatic interactions, Lennard-Jones--Van der Waals energies (vdw), and solvation energies. The difference between the GizMOE energy and the total steric energy listed with the Potential Energy function is the solvation energy. For now we should have no solvation energy, as we set up the forcefield for gas phase calculations.

The Potential Energy results remain on the screen for a short period. To get a permanent listing, pull down the **Window** menu and choose the **Commands** option. A new dialog box will open. Make the Commands window a bit wider and then redo the **Potential Energy**. The results will be listed on the screen and in the Commands window.


Pull down the **File** menu and choose **Save...** If the current path is not your “/moefiles/smallmolecules” directory, you can position the cursor in the white file name field over the **V** symbol and press the right mouse button. This action opens the Recent Directories popup menu. Choose the “smallmolecules” directory where you saved your initial file. Then click on the amecyc.moe file name. Click OK. You will be instructed that the file already exists; click Yes.

After minimizing the molecule, you should always consider other conformations that might have lower energies. In other words, your first minimization might not have found the global energy minimum. One common approach to finding the global energy minimum is to change some dihedral angles in the molecule and then re-minimize. You can easily change the dihedral angles around non-ring bonds. The ability to change dihedral angles is also useful for finding the energy barriers for internal rotations. In other words, how easy is it for the molecule to change from one dihedral conformation to another?


For example, in methylcyclohexane, we can find the barrier to internal rotation for the $-CH_3$ group. Make sure methylcyclohexane is minimized. Record the energy. Click and then shift-click on the methyl C and the ring C to which the methyl C is attached. You can change the dihedral angle using the mouse or the dihedral thumbwheel that appears at the bottom of the screen. Hold down the “Alt” key and drag the mouse using the left mouse button to change the dihedral angle. Adjust the dihedral angle to give the eclipsed geometry. Carefully adjust the dihedral while observing the energy listed in the upper-left hand corner of the screen to find the dihedral angle that gives the maximum energy. Record this energy. The difference between this maximum and the minimized energy is called the “barrier height” for rotation around the methyl group C-C bond.

Starting from the eclipsed geometry, re-minimize the molecule. Did the molecule stay in the same conformation or did it return to the original conformation or a different low energy conformation?

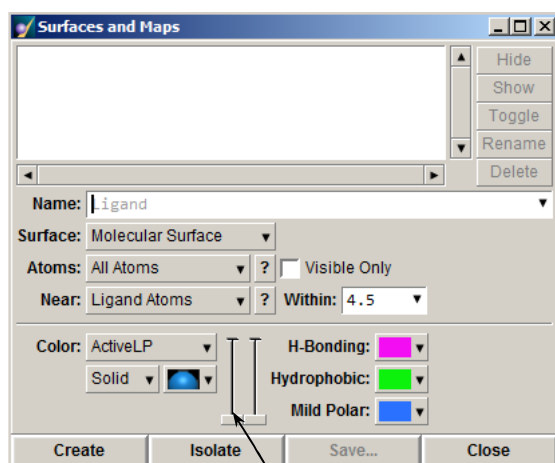
DISPLAY OPTIONS

Ball Stick Rendering Click on the **Atoms** button in the bar that runs across the bottom of the screen. Click on the  button for ball-stick rendering. You can reorient the molecule using the mouse by holding down the center mouse button. Is there a 'hole' in the middle of the ring? Ball Stick rendering is useful for creating graphics for papers for small molecules or for highlighting regions of large molecules.

Surface Rendering Stick structures of molecules are easy to visualize, but they present a very distorted view of molecular structure. Various techniques for displaying the surface of molecules are designed to present a more realistic model of what molecules "look like" to other molecules.

Make sure no atoms are selected by clicking in the black background. Click on the **Atoms** button and choose space filling, . You can reorient the molecule using the mouse by holding down the center mouse button. Is there a 'hole' in the middle of the ring?

An even more realistic view of the molecule is a Connolly surface model. Pull down the **Compute** menu and choose **Surfaces and Maps...** Choose Surface: Molecular Surface, Color By: ActiveLP, and the TF(left) and TB(right) transparency sliders at minimum, as shown below.



TF: front surface
transparency slider

Figure 1.3 Molecule Surface settings.

Click on Create. Reorient the molecule to observe the surface. This surface is equivalent to the Space Filling rendering, but more realistic. To see the original Van der Waals-spheres rendering, drag the TF slider to the middle of the range. (TF controls the transparency of the front surface and TB the back surface.) To delete the molecular surface, click on Delete in the Surfaces and Maps Dialog box. The Connolly surface gives a good idea of the size of a molecule, however, how close to the molecule can other molecules get? Are there some crevices or other hard to reach spots on the surface that are inaccessible to the solvent or other nearby molecules?

One way to answer the accessibility question is to use a spherical probe to find the surface of closest approach. In effect, a probe sphere with a fixed radius is "rolled" over the atoms of the

molecule, and the surface is drawn from the position of the center of the probe sphere. Water, for example, can be adequately approximated for this purpose as a sphere with a radius of 1.4 Å. In the Surfaces and Maps dialog box, choose the Surface: Interaction (VDW). This molecular surface option sets the probe radius at 1.4Å, to give the water accessible surface. The water accessible surface shows the distance of closest approach for water molecules and shows the portion of the molecule that is accessible to the solvent.

To delete the molecular surface, click on Delete in the Surfaces and Maps Dialog box. Click the Close button in the Surfaces and Maps dialog box.

Problem 1.1: What is the barrier height for rotation around the methyl group C-C bond? This difference corresponds to a very large energy change from a molecular mechanics perspective. Could you get the methyl group to stay in the eclipsed conformation?

Problem 1.2: The relative “size” of the various terms in the force field have no particular meaning; only differences in energy are important. You need to do a valid comparison and then look at energy differences. Record the forcefield terms for axial-methylcyclohexane in the table below. Click on Close on the button bar at the right of the MOE screen and then click OK. Start fresh in the Builder. Build 2-methylhexane by first clicking on the “C” button in the Elements button group six times. Next click on a H on the second C in the chain and add a methyl group by clicking on the “C” element button. Pull down the Compute menu, select Partial Charges and choose Current Forcefield. Minimize this molecule by pulling down the Compute menu, choosing Energy Minimize and setting the Gradient test criterion to 0.001 as you did before. Check the minimized energy and then click on the Minimize button a time or two more to make sure the energy doesn’t change on successive minimization runs. Record the energies in the table below. Record the various contributions to the steric energy in the table below. Only terms with significant values are listed for brevity. The total, however, should include all force field terms, as listed by the GizMOE Energy listing. Calculate the difference in energy for each contribution in the 4th column. In the 5th column record which molecule is favored by each contribution. Finally, from the difference column, decide which contribution dominates the preference. Which changes most, the ring strain (as measured by the sum of the bond stretch, bond angle, and dihedral torsion terms) or the through-space Van der Waals term?

Contribution	Methyl cyclohexane (kcal/mol)	2-methylhexane (kcal/mol)	difference (kcal/mol)	favored structure
bond <u>str</u> energy				
<u>angle</u> energy				
<u>torsional</u> energy				
<u>Van der Waals</u>				
<u>electrostatic</u>				
total				

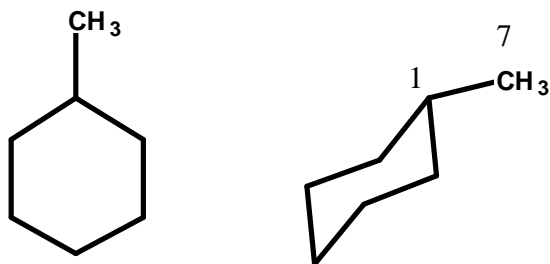
Use the following Check list as you complete the remaining exercises in this manual. You might want to make an extra copy of the checklist to tape in the front of your lab notebook so that you can always have it handy for homework and tests.

Checklist for Molecular Mechanics Calculations

1. Set the force field type. For gas phase calculations make sure the “Solvation:” option is set to gas phase. (Window | Potential Setup | MMFF94x | Solvation: Gas Phase | Apply [if active])
2. Apply atom charges to All Atoms. (Compute | Partial Charges | Method: Current Forcefield).
3. Display the current steric energy. (GizMOE | Energy).
4. Start initial minimization. Set the Hamiltonian option to Forcefield and set the “Calculate Forcefield Partial Charges” checkbox on. (Compute | Energy Minimize | RMS Gradient =0.001) (For proteins and other large systems use the default 0.05)
5. Check the energy on successive minimizations to make sure the calculation has converged. (click the Minimize button several times).
6. Get the total energy and force field terms. (Compute | Potential Energy)
7. Check the results for bad structures:
Look for unusually large force field terms.
Visually inspect the structure for high-energy conformations and bad contacts, or optionally: turn on the Clash option in **Contacts** (Contacts | Clash) to look for close contacts.
8. Search for the global minimum using various conformational search methods (see below). One good method is to adjust dihedral angles, as you did for the methyl-cyclohexane example, then re-minimize.
9. Save your file. The first time you use the file librarian, navigate to “C:\Documents and Settings\All Users\Documents\moefiles” (or “Documents/moefiles” on OS-X). Click on SetCWD if you haven’t already to establish the default directory.
10. When you are finished, pull down the Cancel menu in the upper right-hand corner of the screen and cancel out any instances of the Giz MOE tools. Press the Esc key repeatedly until the white command line is showing just beneath the menu bar.

Chapter 2: Conformational Preference of Methylcyclohexane

Does methylcyclohexane prefer the axial or equatorial conformation? Do Problem 1.2 first. If your axial structure is not on the screen, pull down the **File** menu and 'Open' your file amecyc6.moe. We now need to create the structure of the equatorial isomer (Figure 2.1). Click on the **Builder** button. Click on the C of the methyl group. Click the **Delete** button at the bottom of the Builder dialog box. Click on the equatorial hydrogen on the same C (Figure 1.1). Now click on the C button in the "Element" group. The equatorial isomer will be produced. Pull down the **Window** menu and choose **Potential Setup**. Click on the MMFF94x button, choose Gas Phase for the Solvation treatment, and then Apply if MMFF94x isn't the current force field. Pull down the **Compute** menu, select **Partial Charges**, choose the Method as **Current Forcefield**, and click OK. If the current energy is not displayed in the upper left of the MOE window, pull down the **GizMOE** menu and choose **Energy**, and then pull down the **Compute** menu and choose **Energy Minimize**.... Set the options as listed in Chapter 1. Check the minimized energy and then click on the Minimize button a time or two more to make sure the energy doesn't change on successive minimization runs. Pull down the **File** menu and choose **Save...** Change to the smallmolecules directory (if the correct directory is not listed, position the cursor over the file name field ∇ in the Save MOE panel and press the right mouse button). In the file name dialog box, type emecyc6.moe as the new file name.



After you minimize the structure, pull down the **Compute** menu and choose **Potential Energy**. Record the force field contributions, in the table below for Problem 2.1, to allow comparison with the results from Problem 1.2.

methyl cyclohexane equatorial methyl cyclohexane
Figure 2.1. Equatorial methyl cyclohexane

Problem 2.1 Compare the two conformers of methylcyclohexane. Record the various contributions to the steric energy in the table below. Calculate the difference in energy for each contribution in the 4th column. In the 5th column record which conformer is favored by each contribution. Finally, from the difference column, decide which contribution dominates the conformational preference. Which changes most, the ring strain (as measured by the sum of the bond stretch, bond angle, and dihedral torsion terms) or the through-space Van der Waals term?

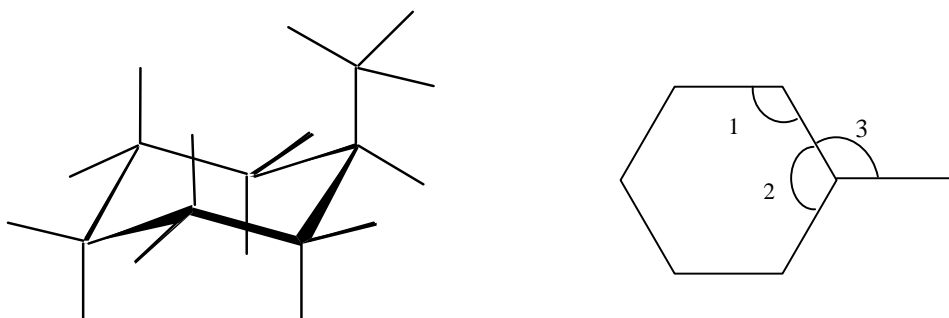
Contribution	equatorial (kcal/mol)	axial (kcal/mol)	difference (kcal/mol)	favored conformer
bond <u>str</u> energy				
<u>angle</u> energy				
<u>torsional</u> energy				
<u>Van der Waals</u>				
<u>electrostatic</u>				
total				

Chapter 3. Geometry (or How Does Molecular Mechanics Measure Up?)

In this chapter you will learn how to measure distances, bond lengths and angles from your minimized structures. We will make our measurements on axial- and equatorial-methylcyclohexane, so do Chapter 2 first. General Chemistry texts list the C-H bond length as 1.09 Å and the C-C bond length as 1.54 Å for sp^3 hybridized systems. The ideal bond angle around tetrahedral carbon is the tetrahedral angle, 109.5° . How close to these idealized values do real molecules come?

Make sure your axial-methylcyclohexane is on the screen. If it isn't, from the main MOE window, pull down the **File** menu and choose **Open...** If you didn't set the current working directory previously, use the pull down menu at the right of the file name dialog box to choose the "c:\Documents and Settings\All Users\Documents\moefiles\smallmolecules" directory or the directory you specified in Chapter 1. Select your axial file, click on OK. Click on the **Atoms** button in the bottom button bar and choose the thin line rendering option.

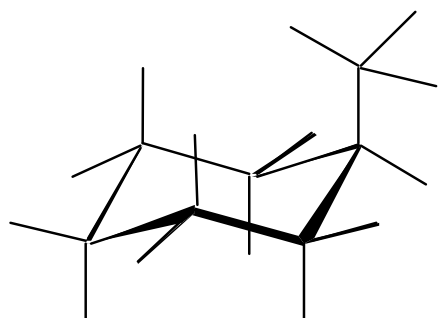
Bond distances: To find bond distances, first click on the **Measure** button at the right of the MOE window, slide right, and choose **Distances**. The white command line box below the menu bar should be replaced with "Measure: distance Pick first atom." Now whenever you click on any two atoms, the distance between those two atoms will be displayed. Measure the four unique C-C bond distances in your compound. Also measure one and only one of the ring C-H bond lengths. Record the values on the structure below. Don't measure every C-C bond length, only measure the ones that are not related by symmetry (see problem 3.2 below for the data table format for your observations). You can also measure the distances between atoms that are not attached. Find the shortest distance between a methyl hydrogen and a ring hydrogen. Include this distance on the structure below.



When you are finished press the ESC key. The command line box should reappear below the topmenu bar. To remove the distance labels slide right on the **Measure** button and choose **Remove Distances**.

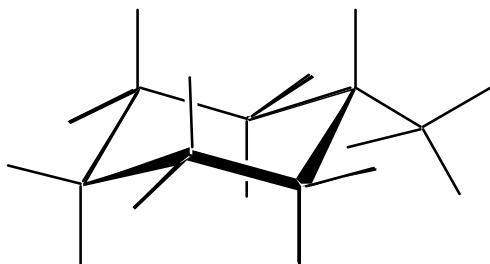
Bond Angles: To find bond angles, first slide right on the **Measure** button and choose **Angles**. The white command line box below the menu bar should be replaced with "Measure: angle." Now whenever you click on three atoms in a row, the bond angle will be displayed. Make sure that the central atom in the angle is the second atom that you click on. Measure the three angles shown in the structure above, at right. Record the angles in the structure above. When you are finished press the ESC key. The white command line box should reappear below the menu bar. To remove the angle labels slide right on the **Measure** button and choose **Remove Angles**.

Dihedral Angles: To find dihedral angles, first slide right on the “Measure” button and choose **Dihedrals**. The white command line box below the menu bar should be replaced with “Measure: dihedral.” Now whenever you click on four atoms in a row, the dihedral angle will be displayed. Make sure that you click on the four atoms in the order in which they are connected. For example, to find the ring dihedral angle for adjacent C-H bonds, click on the atoms in the order: ring-H, the attached ring-C, the adjacent ring-C, and finally the attached ring-H. Measure one of the anti H-C-C-H dihedral angles in your compound, and one of the ring C-C-C-C dihedrals. Record them in the structure below. Finish up as before by pressing the Esc key. Note that you can leave dihedral monitors on while you do other tasks, which include minimization, conformational searches, and dynamics.



Problem 3.1

Report the shortest distance between a methyl hydrogen and a ring hydrogen in equatorial-methylcyclohexane. Include this distance on the equatorial structure below. Do these shortest distances in the axial and equatorial conformers correlate with the change in Van der Waals energy that you found in Problem 2.1? The MMFF94 Van der Waals radius for H is 1.485 Å.



Problem 3.2

Compare the bond distances and angles in axial-methylcyclohexane to the “normal” values of the C-H bond length of $\sim 1.096 \text{ \AA}$, the C-C bond length of $\sim 1.527 \text{ \AA}$, and the angles listed in the Molecular Mechanics Tutorial Section 1 Table 4. Fill in the table below, using the values you have already recorded on the axial-structure, above, to help you answer this question. Deviations from the normal values cause bond strain. Which C-C bonds differ most from the normal values? Is it easier to deform the bond length or the bond angle; that is, do the bond lengths or bond angles deviate more from the normal values?

Bond Lengths			Bond Angles		
Bond	Distance Å	Expected Å	Angle	Angle °	Expected
C-C (Methyl)		1.53	C-C-C (1)		
C-C		1.53	C-C-C (2)		
C-C		1.53	C-C-C (3)		
C-C		1.53	H-C-C-H		
C-H		1.10	C-C-C-C		

Atom Charges: (see section 10 of the Molecular Mechanics Tutorial for more information on charge calculations) The partial charges on atoms and the resulting electrostatic interactions are an important factor in the energetics of polar molecules. How sensitive are steric energies to the charges that are assigned to the atoms in a molecule? The MMFF94 force field doesn't place charges on aliphatic carbons and hydrogens in non-polar environments¹. To see the charges that have been assigned, click the Atoms button and choose Charge. This approximation works well within the MMFF force field, but is really not realistic. One of the best ways to estimate partial atomic charges in molecules is called the Gasteiger method, or in MOE these are called PEOE charges². PEOE stands for Partial Equalization of Orbital Electronegativity. To calculate PEOE charges pull down the **Compute** menu, select **Partial Charges**, and for the Method select Gasteiger (PEOE). Record the charge on the methyl H's. Now calculate the potential energy with the PEOE charges by pulling down the **Compute** menu and choosing **Potential Energy**³. Hydrocarbons don't have large charges on the atoms, so the "electrostatic" contribution to the total steric energy is expected to be small. In compounds with heteroatoms, however, the electrostatic contribution can dominate the steric energy. To clear the charge labels, repeat the steps you took with the Atoms button. In the next problem we will build a polar molecule, where the electrostatic interactions play a more important role.

Problem 3.3: MMFF Forcefield Charges

Build ethylene glycol (1,2-dihydroxyethane: click the **Builder** "C" button twice then add the two OH's). Choose Current Forcefield for the Partial Charges. Minimize the molecule, find the lowest energy conformation, and then display the atom charges. Compare the electrostatic and Van der Waals terms; does the electrostatic term exceed the Van der Waals term in this molecule?

Problem 3.4: MMFF Forcefield Charges

Find the MMFF forcefield charges on the different hydrogens for ethylene glycol. Record these charges. We can use methylcyclohexane as a typical hydrocarbon for comparisons. Compare the MMFF charges for the hydrogens on ethylene glycol to the MMFF charges on the hydrogens for methylcyclohexane. Consider in particular the charges on the methylene hydrogens for ethylene glycol, how do they compare? The MMFF forcefield charges are derived from tabular values that don't include remote heteroatom influences.

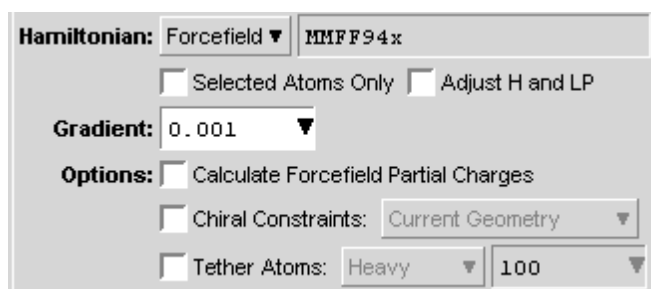
Problem 3.5: PEOE Charges

Calculate the PEOE charges on ethylene glycol. By what percentage difference does the PEOE charge on the methylene H's in ethylene glycol exceed the PEOE charge on the methyl H's in methylcyclohexane [$100\% (q_{EG} - q_{MCH})/q_{MCH}$]? This difference shows the remote effect of a

heteroatom. As you saw in the last problem, the MMFF charges on the methylene H's are zero for both methylcyclohexane and ethylene glycol.

Problem 3.6: PEOE Charges

Now minimize the energy using PEOE charges, using the following procedure, to see how much of a difference the details of the charge calculation make on the steric energy. Make sure PEOE charges have been assigned to the atoms (Problem 3.5). To use the PEOE charges in minimization pull down the **Compute** menu and choose **Energy Minimize...** Clear the Calculate Forcefield Partial Charges checkbox as shown below. (Otherwise the calculation will revert to the MMFF charges.)



Click on OK. Record the final total energy and the Potential Energy terms. Does the partial charge method make a significant difference in the overall steric energy? Does minimization with PEOE charges make significant changes in the individual forcefield energy terms? Compare the changes, are they in the same direction?

Contribution	ethylene glycol MMFF charges (kcal/mol)	ethylene glycol PEOE charges (kcal/mol)	difference (kcal/mol)
bond <u>str</u> energy			
<u>angle</u> energy			
<u>torsional</u> energy			
<u>Van der Waals</u>			
<u>electrostatic</u>			
total			

GizMOE (optional exercise)

It is sometimes useful to have a visual picture of forces in molecules rather than depending on changes in bond angles alone. The GizMOE options allow us to do that. Using the File menu, load your axial-methylcyclohexane structure. Change the rendering method to ball and stick. Then pull down the **GizMOE** menu and choose **Color Force**. If there are no strong forces acting on the molecule, the whole molecule should be colored blue. Record the energy. Now click on one of the methyl group hydrogens. While holding down the ALT and Shift keys, use the middle mouse button and drag the H atom away from its current position. (Alternatively, use the track

ball while holding down the Alt key). The atoms of the molecule that experience forces will be colored, with strong forces indicated by red. Now click the Minimize button and the molecule should relax back to an energy minimum position. Did the energy return to the minimized value?

You can also change the dihedral angle around non-ring bonds and see the change in forces. Make sure the molecule is minimized. Click and then shift-click on the methyl C and the ring C to which the methyl C is attached. Hold down the "Alt" key and drag the mouse using the left mouse button to change the dihedral angle. Adjust the dihedral angle to give the eclipsed geometry. Carefully adjust the dihedral while observing the energy listed in the upper-left corner of the screen to find the dihedral angle that gives the maximum energy, as you did in Chapter 1. Notice the colors of the atoms and the bonds. Do these colors show that the forces are higher than in the minimized structure and that some steric strain exists in the molecule? Which hydrogens on the methyl group experience the strongest forces?

The minimization process can be done continuously. Pull down the GizMOE menu and choose Minimizer. Now try pulling the H atom around. Try to get the methyl group to adopt an eclipsed position with respect to the C-C bonds in the ring. Can you make it stay in an eclipsed conformation? If the molecule moves too far off center, you can click on the button at the right of the MOE window to reset the view. The GizMOE minimizer is useful for generating low steric energy starting structures for small molecules and for ligand docking.

The GizMOE minimizer is not recommended for final minimization. Make sure to turn the GizMOE minimizer and force coloring off before continuing. Then minimize the molecule using the menu commands.

To turn off the GizMOE options, pull down the Cancel menu in the upper right-hand corner of the MOE window and choose each of the options in turn. Make sure no atoms are selected, click on the **Atoms** button, make sure the check box adjacent to C in the Color group is not checked, and choose Element in the Color group. The atoms in your molecule should now be colored by element.

Literature Cited:

1. Halgren, T. A., "Merck Molecular Force Field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interaction". *J. Comput. Chem.*, **1996**, *17*, 520-552.
2. J. Gasteiger, M. Marsili, "Iterative Partial Equalization of Orbital Electronegativity – A Rapid Access to Atomic Charges," *Tetrahedron*, **1980**, 3219-88.
3. Halgen¹ cautions that PEOE charges will tend to underestimate intermolecular interactions and in general PEOE charges don't work well for the MMFF force field.

Chapter 4: Building More Complex Structures

We will next illustrate how to build up complex structures from simple structures, by converting the axial methyl cyclohexane into 1-methyl-trans-decalin. Begin by Opening amecyc6: pull down the **File** menu and choose **Open...** If not already listed, use the pull down menu at the right of the file name dialog box to choose the “c:\Documents and Settings\All Users\Documents\moefiles\smallmolecules” directory on the PC or "Documents/moefiles/smallmolecules" on the MAC. If the axial-cyclohexane file is not available, see Chapter 1 if you need instructions to build the molecule. Select your axial file, click on OK. Click on the **Atoms** button in the bottom button bar and choose the thin line rendering option. Press the **Builder** button. Our final structure will look like Figure 4.2. We will append the ring from carbon 2 to carbon 3 (see Figure 1.1 and 2.1 for atom numbering). Click on C 2, then hold down the Shift key, and click on carbon 3. Now click on the cyclohexane button.

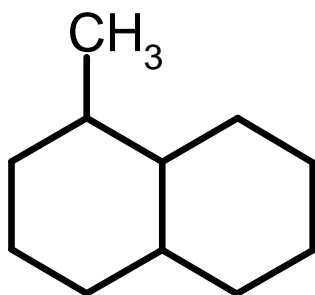


Figure 4.1. 1-methyl-trans-decalin

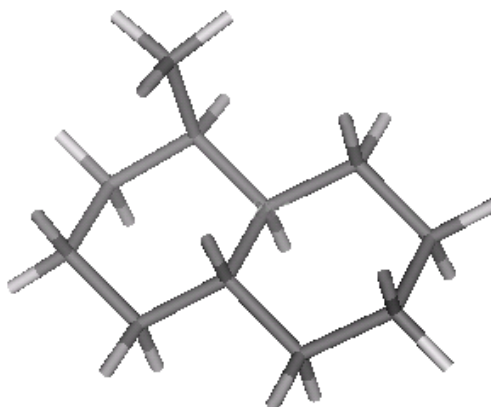


Figure 4.2: 1-methyl-trans-decalin.

The rings should be fused as shown in Figure 4.1. Use the Check List at the end of Chapter 1 to set up and minimize the molecule. If you don't get the conformer shown in Figure 4.2, do the following.

The results of minimization may not give the overall, or global minimum, structure. If you don't get the result in Figure 4.2 (both rings are in the boat form), rotate the molecule so that it is “side on,” but the atoms are not overlapped. Select the CH₂ group that you would like to move using click and shift click on each of the three atoms. Then:

Move the atoms by pressing the ALT key and Shift keys while dragging the mouse using the middle mouse button. Next, rotate the H's so that they point in the required direction by pressing the ALT key while dragging the mouse using the middle mouse button.

Reminimize and check the new conformation. Close the molecule.

Anatoxin

Ring systems are often the most difficult, and interesting, molecules to build. Anatoxin, Figure 4.3, is also known as Very Fast Death Factor¹. Anatoxin is a potent nicotinic acetylcholine receptor agonist. It acts as a potent neurotoxin, which often causes death by paralysis of the respiratory muscles. We will study anatoxin and similar ACE agonists in a later drug design study. Analogs of anatoxin may be useful in treatments for Alzheimer's disease.

Press the Builder button. Click on the cyclooctane button. To create the ring double bond, select two adjacent C's by clicking and shift-clicking. Then click on the double bond button. Next we'll attach the N atom by clicking on a H atom on the C adjacent to the double bond and then in the Element group clicking on the N button.

To create the five-membered ring, select the N atom and then the C atom in the ring as shown in Fig. 3. Click on the single bond button to make the new bond. To attach the side chain, click on the H on the C=C closest to the new ring. Then click on the -C=O. The terminal H should already be selected so that clicking on the C atom button add the methyl group. The molecule is now complete, except that the ring is highly strained. You can Close the Builder.

Use the Check List at the end of Chapt. 1 to run the minimization using MMFF94. The lowest energy conformation using MMFF94x at 20.1737 kcal/mol is shown in Figure 4.4. If you don't get this conformation try moving some atoms (select the atoms and use Alt-Shift-drag middle mouse and Alt-drag middle mouse) and reminimize. Also try changing the dihedral angle of the side chain (select two atoms then Alt-left mouse drag).

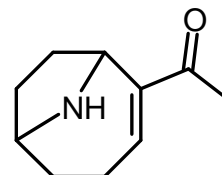


Figure 4.3 Anatoxin

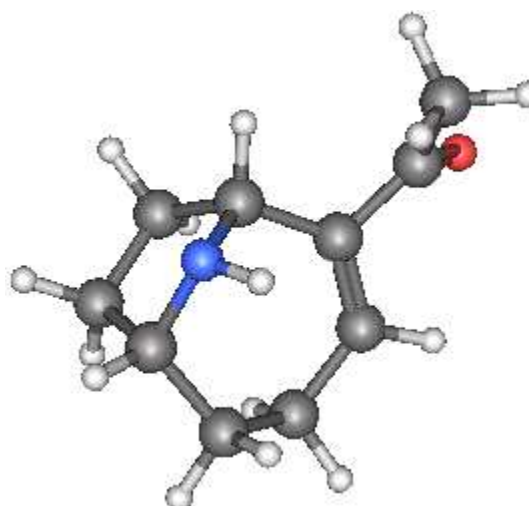


Figure 4.4. Lowest energy conformation of Anatoxin

Problem 4.1 Build the structure for camphor, Figure 4.5. Build the molecule starting with cyclohexane. Report the final steric energy and the various energy contributions. To find if camphor is highly strained, we need to make a fair comparison with an unstrained structure (consult Molecular Mechanics Tutorial Section 3). Decide which bond is best to break to generate a comparable unstrained structure, minimize the new molecule, and compare your results with camphor. Does this comparison confirm the expectation that camphor is a highly strained molecule? Which term in the force field shows the strain best? Make sure to draw your structure. Then, use a table similar to that in Problem 2.1.

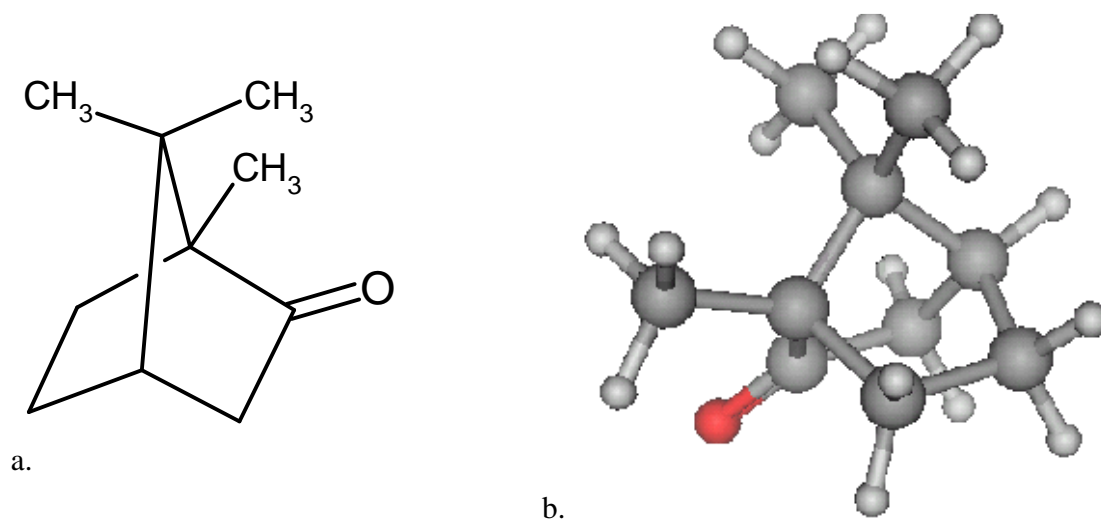


Figure 4.5. (a). Camphor. (b) Structure of camphor from molecular mechanics (rotated around the z-axis by 90° from a).

References:

1. Molecule of the Month: Anatoxin, <http://www.chm.bris.ac.uk/motm/antx/antx.htm>

Chapter 5. Conformational Preference for Butane

We will determine the conformational preference and corresponding equilibrium constant for butane, which is an important and experimentally well-studied system. We will also learn how to use the Conformational Search application.

First consider ethane. Two possible conformations of ethane are shown in Figure 5.1.

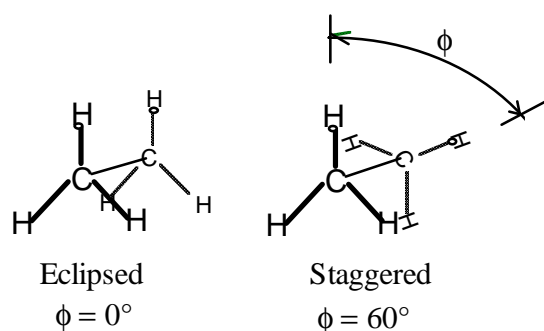


Figure 5.1. Eclipsed and staggered ethane.

The eclipsed conformer is higher in energy than the staggered form. The increase in **torsional energy** of the eclipsed form is caused by the repulsion of the electrons in the C-H bonds on different ends of the molecule. In the staggered form, the bonds are further apart thus reducing the electron-electron repulsion between the bonds. A plot of the dihedral energy of ethane is shown in Figure 5.2. The energy penalty of having eclipsed rather than staggered bonds is seen to be 2.7 kcal/mol (11.3 kJ/mol). The energy curve has three minima because the three atoms

attached to each end of the molecule are the same. Therefore, the conformations with $\phi = 0^\circ$, 120° , and 240° are all identical eclipsed conformations. The conformations with $\phi = 60^\circ$, 180° , and 300° are all identical with staggered, low energy conformations. Locate these energies in Figure 5.2.

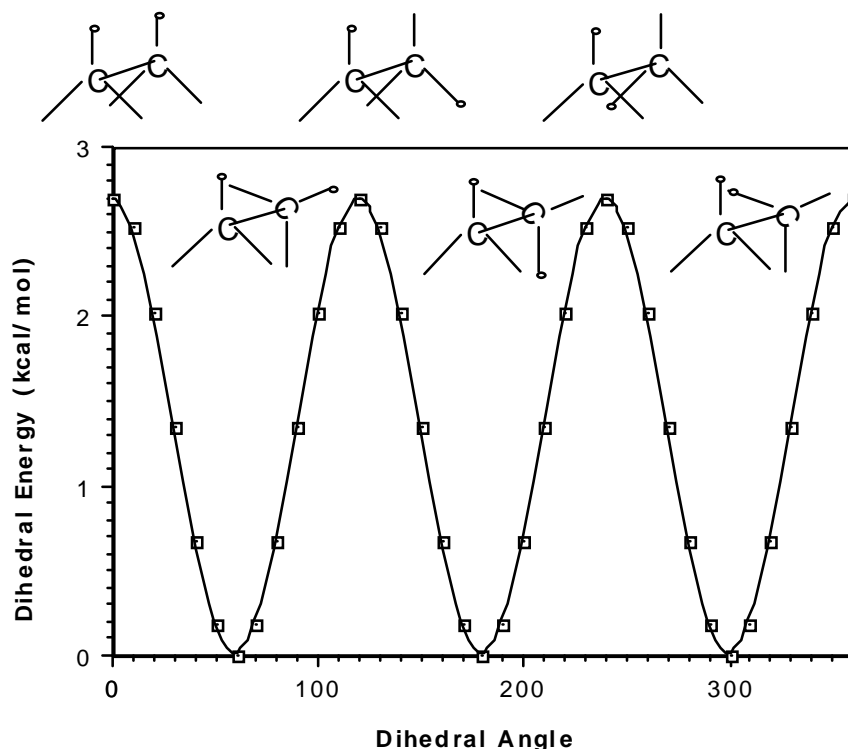
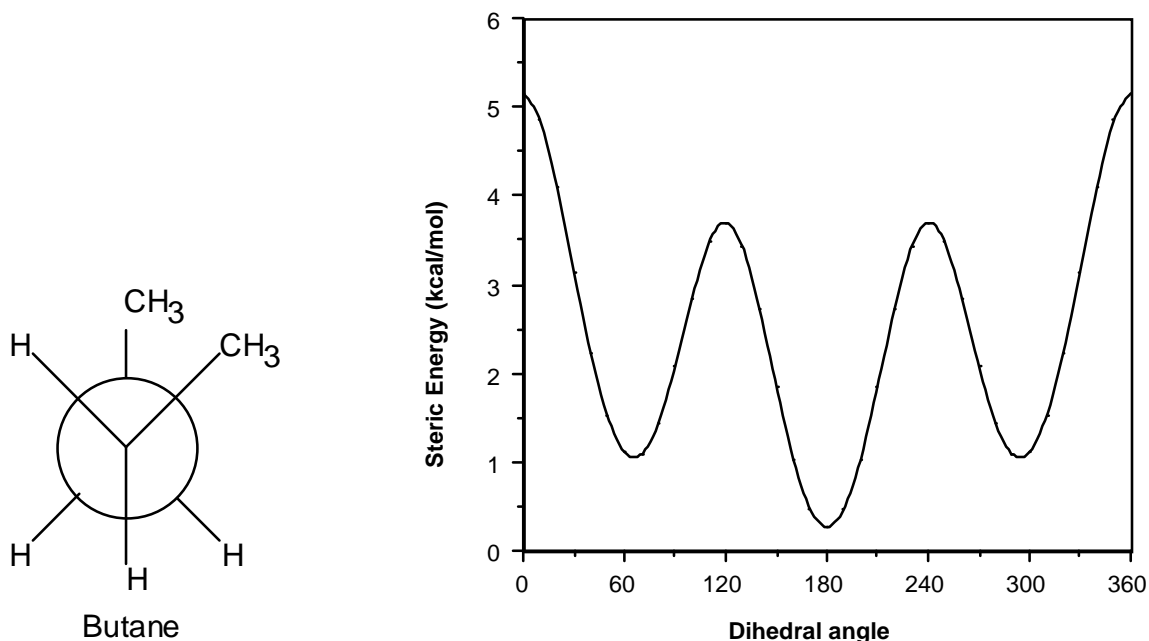


Figure 5.2. Dihedral energy in ethane. In the structures all hydrogens are equivalent, however one particular hydrogen on the front of the molecule and one on the back are shown with a dot so that you can follow the change in the dihedral angle over a full 360° .

Figure 5.2 is a plot of the dihedral, or torsional, potential energy for a 3ϕ , three-fold torsional barrier. Remember that the full torsional potential energy is given by:

$$E_{\text{tor}} = 1/2 k_{\text{tor},1} (1 - \cos \phi) + 1/2 k_{\text{tor},2} (1 - \cos 2 \phi) + 1/2 k_{\text{tor},3} (1 - \cos 3 \phi) \quad 1$$

Butane, Figure 5.3a, will also have a large term for the one-fold potential. The CHARMM steric energy as a function of dihedral angle is shown in Figure 5.3b.



a. Figure 5.3. (a.) Butane, in the gauche conformation. (b) Steric energy for butane.

In butane, the difference in energy between the anti and gauche forms is -0.8 kcal/mol. Also note that the minimum energy dihedral angle is 65° and not the ideal 60° . The equilibrium constant for the ratio of anti to gauche forms can be estimated from this energy difference. First, we will assume that there are no significant changes in vibrations between the two conformers. The steric energy difference is then ΔU . Remember $\Delta H = \Delta U + \Delta n_g RT$, where Δn_g is the change in the number of moles of gas. Since we are calculating the difference in energy between two conformers:



$\Delta n_g = 0$. Therefore, $\Delta U = \Delta H$. Next we need to calculate the change in entropy for the conformational change. Since there are two equivalent gauche conformers and only one anti conformer:

$$\Delta S (\text{anti-gauche}) = R \ln (1/2) = -1.38 \text{ cal/mol K} = -5.76 \text{ J/mol K} \quad 3$$

$$\text{Then } \Delta G (\text{anti-gauche}) = \Delta H - T\Delta S \quad 4$$

in calories:

$$\Delta G = -0.8 \text{ kcal/mol} - (298.2 \text{ K})(-1.38 \times 10^{-3} \text{ kcal/mol K}) = -0.39 \text{ kcal/mol} \quad 5$$

and in kJ:

$$\Delta G = -3.35 \text{ kJ/mol} - (298.2 \text{ K})(-5.76 \times 10^{-3} \text{ kJ/mol K}) = -1.63 \text{ kJ/mol} \quad 6$$

and the equilibrium constant can be obtained from:

$$\Delta_r G^\ominus = -RT \ln K \quad 7$$

Assuming the ΔG from molecular mechanics = $\Delta_r G^\ominus$ for a 1 bar standard state gives:

$$K = \frac{[\text{anti}]}{[\text{gauche}]} = 1.93 \quad 8$$

In other words, there are two molecules in the anti-conformation for every molecule in the gauche conformation at 25°C.

The following instructions will show you how to repeat the above calculations for the energy minimum structures for the anti and gauche forms of butane and also how to generate the energy plot in Figure 5.3b.

Conformational Preference for Butane

Build butane. Click on the Builder... button. Click on the "C" button four times. Close the Builder. Use the Checklist at the end of Chapter 1 to minimize the molecule using MMFF94x. Make sure the Solvation option is set to Gas Phase. Minimize with a RMS gradient tolerance of 0.0001 kcal. This conformation should be the anti-conformer. Record the force field term energies and the total steric energy.

To find the energy minimized structure for the gauche isomer: select the first atom defining the torsion by clicking on a -CH₂- carbon. Pick the second atom defining the torsion by shift-clicking on the second -CH₂- carbon atom. Hold down the Alt key and drag the mouse using the left mouse button. Adjust the molecule to give the gauche conformer. Minimize with a RMS gradient tolerance of 0.0001 kcal. Remember to click on the Minimize... button several times to make sure the structure is completely minimized. Pull down the Compute menu and choose Potential Energy. The contributions to the total steric energy will be listed in the MOE window. Record these energies and the total steric energy.

Problem 5.1

Record the various contributions to the steric energy in the table below. Calculate the difference in energy for each contribution in the 4th column. In the 5th column record which conformer is favored by each contribution. Finally, from the difference column, decide which contribution dominates the conformational preference in butane. Report the gauche dihedral angle.

Contribution	anti (kcal/mol)	gauche (kcal/mol)	difference (kcal/mol)	favored conformer
bond str energy				
angle energy				
torsional energy				
Van der Waals				
electrostatic				
total				

The Boltzmann Distribution: An Alternative Viewpoint

The Boltzmann distribution describes the probability of occurrence of a structure with energy E_i :

$$\text{probability of occurrence} = \frac{e^{-E_i/RT}}{q} \quad 9$$

where $e^{-E_i/RT}$ is called the Boltzmann weighting factor, R is the gas constant $8.314 \text{ J mol}^{-1}\text{K}^{-1}$, T is the temperature in degrees K, and q is the sum of the probabilities over all possible states. The q term, which is called the partition function, just assures that the probabilities sum to 1.0. The effect of a temperature increase is to increase the probability of high energy structures. For example, at a low temperature most molecules will be found in the lowest energy state, but as the temperature increases molecules gain energy through collisions and are promoted into higher energy states, Figure 5.4a. Alternatively, if the temperature is constant, systems with large energy differences have few molecules in high energy states. Systems with small energy differences between their levels have many molecules in upper energy states, Figure 5.4b.

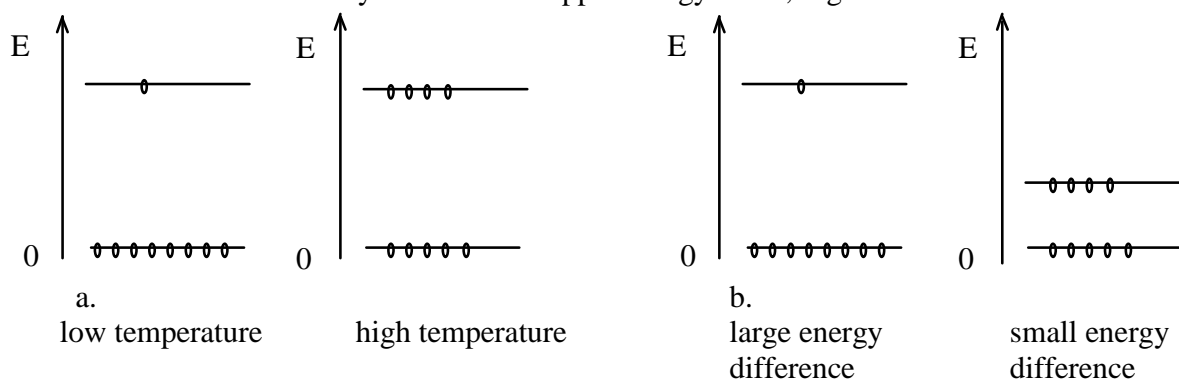


Figure 5.4 The Boltzmann distribution determines the probability of occurrence of a given energy state of a molecule. a. High temperatures favor higher energy states. b. Small energy differences favor higher energy states.

What determines the energy difference between energy states? A good example is the conformational energy of butane. The difference in energy between the gauche and anti forms is 0.8 kcal/mol . The Boltzmann distribution will tell us the relative numbers of molecules in the anti and in the higher energy gauche states. Another example is the conformational preference of axial and equatorial methylcyclohexane. The MMFF steric energy of axial-methylcyclohexane is 1.34 kcal/mol higher than the equatorial isomer (Chapter 2).

If there is more than one structure at a given energy, then we must multiply the probability by the number of structures at the same energy. The number of structures at the same energy is called the degeneracy and is given the symbol g . For example, butane has one anti-conformer, $g_{\text{anti}}=1$, and two gauche-conformers, $g_{\text{gauche}}=2$. The Boltzmann distribution with degeneracy is:

$$\text{probability of occurrence} = \frac{g e^{-E_i/RT}}{q} \quad 10$$

and

$$q = \sum_{\text{all states}} g e^{-E_i/RT} \quad 11$$

Take butane as an example. The anti-conformer has the lowest energy, which we can assign as $E_{\text{anti}} = 0$. Then the gauche-conformer has an energy $E_{\text{gauche}} = 0.8 \text{ kcal/mol} = 3.35 \text{ kJ/mol}$ above the anti-state. Table 5.1 shows how to calculate the probabilities from Eq. 10 and 11. The probabilities are in the last column.

Table 5.1. Calculation of the Boltzmann factors for gauche and anti-conformations of butane at 298.2K.

Conformation	Energy, E_i (kJ)	E_i / RT	$e^{-E_i/RT}$	$g e^{-E_i/RT}$	$g e^{-E_i/RT} / q$
gauche	3.35	1.35	.2589	0.5178	0.3411
anti	0	0	1	1	0.6588
sum=q=				q=1.5178	

To calculate q we sum the weighting factors in column 5. Then we use q to calculate the probabilities in column 6. Notice that if we take the ratio of the probabilities of the anti and gauche states we get the same result as Eq. 8, above, which was calculated from Gibb's Free Energy:

$$\frac{\text{probability for anti}}{\text{probability for gauche}} = \frac{0.6588}{0.3411} = 1.93 = K \quad 12$$

The Gibb's Free Energy and Boltzmann approach are equivalent but take slightly different points of view.

Dihedral Angle Conformation Searches

Pull down the Compute menu, slide right on Mechanics, and choose Dihedral Energy Plot. A prompt will then appear requesting you to pick the first atom that defines the dihedral. Click on the four carbons in the order they appear in the chain: in other words, click on the first CH_3 , then the $-\text{CH}_2-$, the second $-\text{CH}_2-$, and finally the last CH_3 . A grid scan search will change the dihedral angle in equal steps. A plot window will appear. If you want to change the axis ranges or the grid spacing so that you can zoom in on an area of the plot, click on Attributes... and use the pull down menu to adjust the x and y axes. The dihedral angle in the current structure is marked with the red line. This type of grid search does not do a full minimization at each new dihedral angle, so the other bond lengths and angles aren't allowed to adjust to the new dihedral. Nevertheless the plot will show you good conformations to try with full energy minimization as you search for the global minimum. If you want to allow a full minimization at each new dihedral then you should select the four atoms in the dihedral and then type

```
run '$MOE/sample/dihmplot.svl'
```

into the white command line at the top of the MOE window. For more complex molecules that have more than one rotatable bond you can use the Conformation applications from the Compute menu.

The Conformations applications work a little differently. They produce a database (group of files) to hold the lowest energy results from the conformations search. To do a Systematic Search with full minimization at each dihedral, pull down the Compute menu, slide right on

Conformations and choose Systematic Search. The tool tries to find all freely rotatable bonds for you. In butane there is only one. Double click on the line in the text box that reads something like “1 rot C(5) C(8) 120.” We can now change the Step size: click on 20 to set a finer step size.

Remember to set the c:\Documents and Settings\All Users\Documents\moefiles directory as your default directory (or Documents/moefiles on OS-X), otherwise the conformation search file would be saved in the Program Files directory on the C: drive (PC) or in the Applications folder (OS-X). In this and all other similar instances please make sure the moefiles directory is the default directory. Having done so, you can leave the Database entry to read “csearch.mdb.”

Click on the check box to “Open Database Viewer” at the end of the run. Also click on Energy Minimize Resulting Conformations. Set the RMS gradient to 0.001 for small molecules like butane. The other default settings should be fine. Click on OK. In short order the database viewer should appear with the low energy results from the conformational search. The structures will be sorted by energy. The lowest energy structure will be loaded into the MOE window. In the MOE window, press the Minimize button to finish the minimization. The results should be the same as you got for the trans conformer above.

To see the thumbnail structures better in the Database viewer, just click in any of the structure cells and drag the mouse downwards. The middle mouse button can be used to reorient each structure just as in the main MOE window. The second lowest structure should be a gauche conformer. Transfer this structure to the main MOE window by clicking the right mouse button in the second molecule cell. You should see a pop up menu; select Copy to MOE. Next you see a new dialog box; select the Clear Molecule Data option and then OK. The new molecule should appear in the MOE main window. Click on Minimize to make sure the molecule is completely minimized. You should get the same result as you got for the gauche conformer, above.

Click the go-away box in the Database viewer when you are finished.

Please note that for adjusting dihedral angles with the mouse, you select only two atoms. In other parts of MOE, for example in measuring dihedral angles and for Conformational Searches, you need to specify all four atoms of the dihedral.

Problem 5.2

Calculate the equilibrium constant for the anti to gauche conformers for dichloroethane from ΔG . Find the dihedral angle in the gauche conformer. Why is this angle different from butane? Also, use the Conformational Search application, dihplot.svl, to plot the steric energy as a function of dihedral angle. A rough sketch will do for your report or see Chapter 8 for printing.

Contribution	anti (kcal/mol)	gauche (kcal/mol)	difference (kcal/mol)	favored conformer
bond <u>str</u> energy				
<u>angle</u> energy				
<u>torsional</u> energy				
<u>Van der Waals</u>				
<u>electrostatic</u>				
total				

Problem 5.3

Using the energy difference from Problem 5.2, calculate the probabilities of occurrence of the gauche and anti forms for dichloroethane and the equilibrium constant.

Conformation	Energy, E_i (kJ)	E_i / RT	$e^{-E_i/RT}$	$g e^{-E_i/RT}$	$g e^{-E_i/RT} / q$
gauche					
anti	0	0	1	1	
sum=q=				q=	

K = _____

Problem 5.4

The dimer of methylvinylketone is shown in Figure 5.4. For this problem we will study just the axial conformer for the $-\text{CO}-\text{CH}_3$ side chain. An interesting question is which face of the carbonyl is more susceptible to nucleophilic attack? Nucleophilic attack will be perpendicular to the trigonal plane of the sp^2 hybridized carbon, Figure 5.4b. According to Cram's rule, the less hindered side is likely to be most susceptible. Make sure that you build the axial conformer. To begin this study we need to know the low energy conformers about the side-chain C-C bond to the ring. Do a conformational search around this bond. What are the low energy conformers? Draw these low energy conformers and note the less hindered side. Space filling models will be helpful in looking at steric influences.

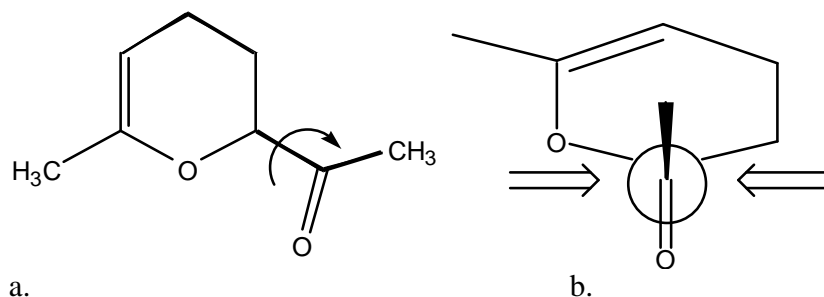


Figure 5.4 (a). Methylvinylketone dimer. The bond with free rotation is marked. (b) Newman projection. Which side of the carbonyl is attacked by nucleophiles? The favored direction of attack will change with conformation angle. Only one possible conformation is shown here.

Chapter 6: MM3

Energy minimization using MOE-MMFF and MM3 are very similar. The force fields are a little different, but the calculations do the same thing. One reason for using MM3 is to calculate enthalpies of formation, which most other molecular mechanics programs can't do. MM3 also treats conjugated pi-electron systems better than MOE-MMFF. You can't, on the other hand, use MM3 for large molecules. In this chapter you will calculate the enthalpy of formation of camphor, so do Problem 4.1 first. Please read the Introduction Section 2: Enthalpy of Formation.



MM3 Minimization

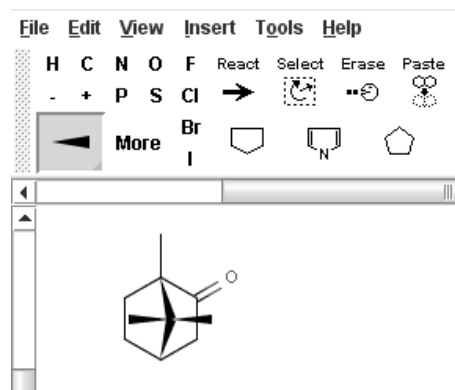
The easiest way to access MM3 is to use the "MM3 Molecular Mechanics" applet. This applet uses the Marvin 3D-Molecular Editor for structure input. You will find the link for this Web applet on the PChem home page or:

<http://schupflab.labs.keyes.colby.edu/pchemistry/webmm3/MarvinMM3d.html>

The Marvin Editor uses a very simplified molecular mechanics method for building 3-D structures, based on the Dreiding force field, to make approximate starting structures.

Building Camphor:

To build camphor, start by clicking on the cyclohexane ring button and then click in the molecule window. Next select the single bond button, . Drag the single bond for the methyl group and the ketone, as shown at right. Clicking on a bond will change it to a double bond. Finish by putting in the oxygen. Next pull down the bond menu and choose the "up wedge bond," . Complete the structure as shown. Don't add H's at this point. Use the Help menu for more information.



For complicated structures, it is often best to minimize first without hydrogens and then add H's and then reminimize. Molecular mechanics minimization is called "cleaning" in this applet. Accordingly, pull down the Edit menu and choose Clean > 3D > Clean in 3D (using Dreiding optimization). Inspect the structure for close contacts or unphysical bond angles. You can redraw the molecule if necessary to get a reasonable structure. Then pull down the Edit menu and choose ..H atoms > Add, and then reminimize using Clean > 3D > Clean in 3D. To see the starting 3D-structure, pull down the View menu and choose Open 3D-Viewer. Check the structure and then click the "go-away" box. Set the MM3 input file options:

Step 2: Generate the MM3 input file from the structure editor.

Using Delocalized-pi calculation: Using Automatic Type Assignments in MM3:

There are no delocalized pi-systems in this molecule. Normally we would leave the "Using Automatic Type Assignments" check box selected, but for this assignment we want to see the type assignments (they will be reassigned in the MM3 program in any case). Then click on Generate Structure from Structure Editor button.

Notice the MM3 input file that was generated. This input file highlights the information that is necessary for a molecular mechanics calculation. First, approximate coordinates are required. If the input coordinates are too unrealistic, the mechanics algorithm will not be able to converge on

a final low energy structure. The coordinates are listed as the decimal numbers in three columns, x, y, and z. The atom symbol is then listed, followed by the MM3 type. MM3 types are the same as MM2 types through type 57, see Table 1 in the Molecular Mechanics Tutorial for a listing. Type assignments are necessary to determine the force field constants. If the type assignment is incorrect, the calculation will be incorrect. For molecular orbital calculations, on the other hand, only the starting coordinates and the atomic number are required.

To run the MM3 calculation, click on the submit button. A new window will be opened with the output from the MM3 run. An example printout for cyclohexane is shown below.

```

FINAL STERIC ENERGY IS      8.0440 KCAL/MOL.
      COMPRESSION             0.2688
      BENDING                 0.1172
      BEND-BEND              -0.0130
      STRETCH-BEND           0.0375
      VANDERWAALS
      1,4 ENERGY            6.1713
      OTHER                  -0.3178
      H-BONDING               0.0000
      PI-COMPLX              0.0000

      TORSIONAL               1.7864
      TORSION-STRETCH        -0.0063
      DIPOLE-DIPOLE           0.0000
      CHARGE-DIPOLE           0.0000
      CHARGE-CHARGE           0.0000

```

bond stretching
bond bending
bend-bend interaction
stretch-bend interaction

{ *add these two for the*
total Van der Waals

{ *add these three for the*
total electrostatic
energy

```

HEAT OF FORMATION AND STRAIN ENERGY CALCULATIONS
      (UNIT = KCAL/MOLE)

      NOMAL (BE) AND STRAINLESS (SBE) ENTHALPY OF INCREMENTS
      (CONSTANTS AND SUMS OF INCREMENTS)

```

BOND OR STRUCTURE	NO	---NORMAL---		--STRAINLESS--	
C-C SP3-SP3	6	2.4470	14.6820	3.5060	21.0360
C-H ALIPHATIC	12	-4.5900	-55.0800	-4.5900	-55.0800
		BE =	-40.3980	SBE =	-34.0440

```

PARTITION FUNCTION CONTRIBUTION (PFC)
      CONFORMATIONAL POPULATION INCREMENT (POP)  0.00
      TORSIONAL CONTRIBUTION (TOR)                0.00
      TRANSLATION/ROTATION TERM (T/R)            2.40
      -----
      PFC = 2.40

HEAT OF FORMATION (HFO) = E + BE + PFC                -29.95
STRAINLESS HEAT OF FORMATION FOR SIGMA SYSTEM (HFS)
      HFS = SBE + T/R + ESCF - ECPI                -31.64
INHERENT SIGMA STRAIN (SI) = E + BE - SBE              1.69
SIGMA STRAIN ENERGY (S) = POP + TOR + SI              1.69

```

The different contributions for the energy force field are listed. The enthalpy of formation is listed on the line labeled HEAT OF FORMATION (HFO). The line labeled SIGMA STRAIN ENERGY is a very useful measure of the total strain in the molecule.

You can print this information by pulling down the File menu and choosing Print. You will also see your minimized structure in a MarvinView window. You can use all of the normal display options by clicking right at the lower-right hand side of the MarvinView window. The normal mode frequencies are listed below the enthalpy of formation calculation section. The careful statistical thermodynamic analysis of the contribution of the normal modes to the enthalpy and entropy is included. At the bottom of the page, the minimized structure is printed in the form of a new MM3 input file. Check the program assigned atom types in this file listing.

Problem 6.1: The Enthalpy of Formation of Camphor

Camphor is an interesting molecule because of its many uses and because it is a highly strained molecule. Because of the rings in the molecule, there are no torsional increments other than for methyl groups. There is also only one low energy conformation. As a very rough starting approximation for the bond enthalpy, calculate the enthalpy of formation from bond energies (Table I in the Introduction to Molecular Mechanics or from your General Chemistry or Physical Chemistry text). Contrast your calculation with the bond enthalpy from your MM3 calculation. Report your bond energy calculation, using Table I or data from your text, MM3 bond enthalpy, MM3 steric energy, the strain energy, and the enthalpy of formation of camphor.

Compare the calculated result for the enthalpy of formation with the literature by completing the following calculations. The enthalpy of combustion of camphor is -1411.0 kcal/mol. But we must also add the enthalpy of sublimation since our MM3 calculation is for the gas phase. The enthalpy of sublimation of camphor is 12.8 kcal/mol. From the enthalpy of combustion and the enthalpy of sublimation calculate the enthalpy of formation of gaseous camphor and compare with the MM3 value. How close are the MM3 and gas phase literature value?

Problem 6.2: Comparisons with Literature Values (or How Good is MM3?)

How well do MM3 enthalpies of formation match literature values? The monoterpenes are an important group of natural products, Figure 1. Determine the enthalpy of formation for each. The literature values are from Lange's Handbook or the CRC. Remember, you must add in the enthalpy of vaporization for liquids or the enthalpy of sublimation for solids, since molecular mechanics energies are for the gas phase. The enthalpy of vaporization or sublimation values in kJ/mol are: camphene, 43.9; α -pinene, 45.2; β -pinene, 46.4; limonene, 43.9; α -terpineol, 52.3; menthol, 56.5 kJ/mol. Even though you've been given the phase transition enthalpies, it is still a good idea to learn how to find them for yourself.

Enthalpies of Vaporization or Sublimation: The Clausius-Clapeyron equation describes the change in vapor or sublimation pressure with temperature:

$$\ln \frac{P_2}{P_1} = -\frac{\Delta_{tr}H}{R} \left(\frac{1}{T_2} - \frac{1}{T_1} \right) \text{ or equivalently } \ln P = -\frac{\Delta_{tr}H}{RT} + \text{cst} \quad (1)$$

where P_1 , and P_2 are the vapor pressures at temperatures T_1 and T_2 , respectively, and $\Delta_{tr}H$ is the enthalpy of vaporization or sublimation. Comparing with Eq. 1, if you use the CRC for enthalpies of vaporization from the vapor pressure versus temperature tables, the listed "a" parameters are equal to the enthalpy of vaporization in kJ/mol. Remember to change to kcal/mol (1 cal = 4.184 J). If you use the sublimation pressure versus temperature table from the CRC then the enthalpy of sublimation = 2.303 R B, as listed in the table caption. If you use R in J/mol K the result will be in J. If you use R=1.987 cal/mol K the results will be in cal. If the enthalpy of vaporization isn't available from tables directly, Eq. 6.1 shows that a plot of the ln (vapor pressure) versus 1/T gives a straight line with slope $-\Delta_{tr}H/R$. The CRC has tables of vapor pressure versus temperature for many organic compounds.

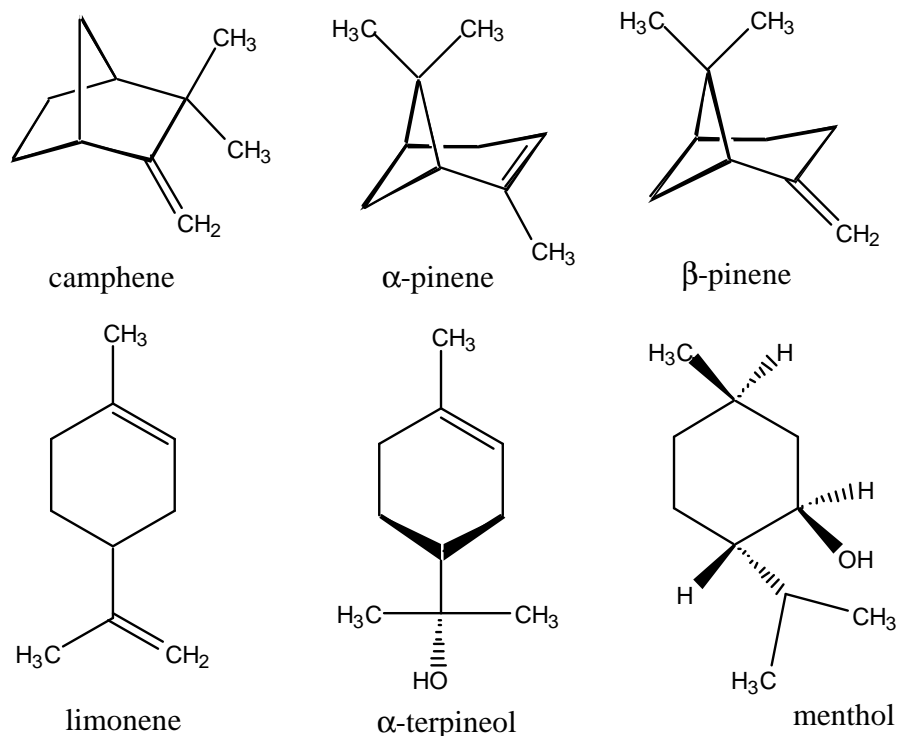


Figure 1 Some monoterpene natural products.

Report your results in the following table. The literature values are listed in kcal/mol. You can share data with other students, but make sure to do at least two of the molecules yourself. Label the two that you did. The Error = $\Delta_f H^\circ(\text{g}) - \Delta_f H^\circ \text{MM3}$. Are there any low frequency vibrations?

compound	Literature (kcal/mol)			Calculated (kcal/mol)			
	$\Delta_f H^\circ$ (l or s) from tables	$\Delta H_{\text{sublim.}}$ or vaporiz.	$\Delta_f H^\circ(\text{g})$	MM3 steric energy	MM3 bond enthalpy	$\Delta_f H^\circ$ MM3	Error (kcal/ mol)
Camphene	-18.22	10.49					
α -pinene	-4.04	10.80					
β -pinene	-1.84	11.09					
limonene	-23.51	10.49					
α -terpineol	-85.84	12.50					
menthol	-114.86	13.50					

Problem 6.3 Torsional Increments to the Enthalpy of Formation

If there are unconstrained or free bond rotations in a molecule, the MM3 $\Delta_f H^\circ$ should be low. In this problem we wish to determine if the addition of torsional increments (see Introduction Section 2) improve the agreement of the calculated values with literature values. Remember the torsional increment is estimated as 0.36 kcal/mol or 1.51 kJ mol⁻¹ for each internal rotation. For

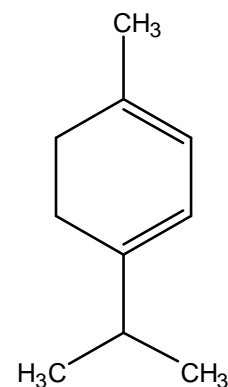
example, butane, $\text{CH}_3\text{-CH}_2\text{-CH}_2\text{-CH}_3$, has one additional internal rotation other than the methyl group rotations; so the torsional increment for butane would be 0.36 kcal/mol. Complete the following table (the data is from the CRC 46th Ed.):

compound	Literature $\Delta_f H^\circ(\text{g})$ kcal/mol	Calculated $\Delta_f H^\circ \text{MM3}$	Δ kcal/ mol	torsional increments		Calculated $\Delta_f H^\circ \text{MM3}$ + total increment	new Δ kcal/ mol	% improv- ment
				# free rotations	total increment kcal/mol			
n-pentane	-35.00							
1-pentene	-5.00							
<i>cis</i> -2-pentene	-6.71							
<i>trans</i> -2-pentene								
2-methyl-1-butene								

Hints: Try several starting geometries to ensure that you have found the lowest energy conformation.

Conjugated Pi-Electron Systems

α -Terpinene is an important mono-terpene (see Problem 6.2). However, the pi-electrons in the two double bonds are conjugated. MM3 in its simplest form does not do a good job on calculations of conjugated pi-electron systems. The MM3 $\Delta_f H^\circ(\text{g})$ as calculated in the same fashion as above is 27.7 kcal/mol whereas the experimental value is -4.89 kcal/mol. We must account for the extra stability of the conjugated pi-system and also the extra barrier to rotation about the bond between the two double bonds. This extra barrier to rotation is also caused by conjugation. MM3 accounts for these factors by doing a molecular orbital calculation on the conjugated pi-system. This molecular orbital calculation is called a self-consistent-field calculation, which is abbreviated SCF. The SCF calculation only covers the pi-electrons.



α -terpinene
Figure 2.

1,3-Butadiene, Figure 3, is a simple conjugated system that will serve as a good first example. The printout from the calculation on butadiene is shown in Figure 3.

The MO orbital diagrams and the energy diagram are not normally part of the printout, but they are included to help you learn how to interpret the molecular orbital portion of the results.

The MO diagrams are only shown for the lowest two orbitals, since only these two are filled with electrons. The molecular orbital coefficients are listed in columns. At the bottom of each column is the energy of the MO, in kcal/mol. For example, the coefficients for the lowest energy orbital are all positive; therefore all the p atomic orbitals have their positive lobes in the same direction. The energy diagram, at right, shows that the two filled orbitals have significantly lower energy than the empty orbitals. The bond order portion of the printout shows

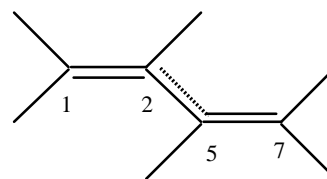


Figure 3. 1,3-butadiene. (The atom numbers correspond to the printout in Figure 4.)

that the end double bonds have a pi-bond order of 0.9662, which is less than a full double bond. However, the single bond between the two double bonds takes on some double bond character, with a pi-bond order of 0.2576. The bond energy in the pi-electron system is -118.06 kcal/mol and the total bond energy, sigma and pi, is -356.71 kcal/mol. The final $\Delta_f H^\circ$ with the pi-calculation included is calculated to be 25.09 kcal/mol. The experimental $\Delta_f H^\circ$ is 26.75 kcal/mol.

Problem 6.4: MM3 Calculations with SCF Pi Calculations

Calculate the MM3 enthalpy of formation of α -terpinene. The MM3 $\Delta_f H^\circ$ (g) as calculated without the SCF molecular orbital calculation is 27.7 kcal/mol; the experimental value is -4.89 kcal/mol (from the CRC).

Butadiene MM2 calculation with SCF calculation

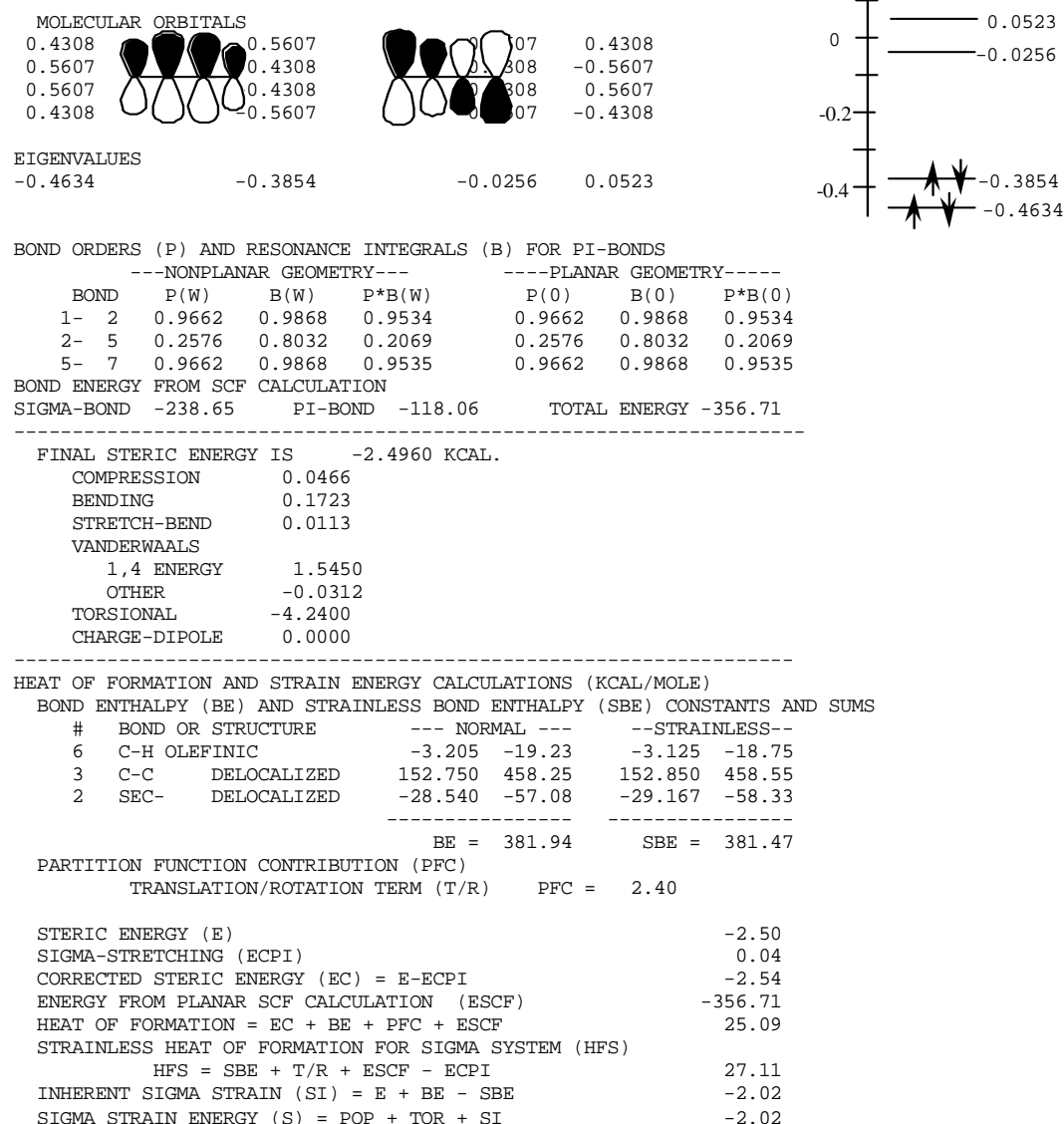


Figure 4: The MM3 printout for 1,3-butadiene. The calculation is done with the SCF pi-molecular orbital calculation.

Chapter 7: Comparing Structures

Changes in a molecule's structure not only affect the local environment, but also can have effects on the structure many bonds away. In this section you will compare the structures of axial- and equatorial- methylcyclohexane from Chapters 1 and 2. The "Interactive Superpose" application is used to produce an overlaid view of the two structures. Pull down the File menu and choose "Open..." If you did not previously set the current working directory, use the pull down menu at the right of the file name dialog box to choose the "c:\Documents and Settings\All Users\Documents\moefiles" directory (or Documents/moefiles/ on OS-X). Click on amecyc6.moe, and then click on "Open MOE file." Pull down Render and choose Ball and Stick. Now repeat the Open file process with emecyc6.moe. Rotate the two molecules so that the hexane ring is in the plane of the screen. We now need to move one of the molecules to the right so they are no longer overlapped. Click on an atom. To select the rest of the atoms in the same molecule, pull down the Selection menu, slide right on Extend, and choose Molecule. All the atoms in one of the molecules should now be selected. While holding down the Alt and Shift keys drag the mouse using the middle mouse button to separate the two molecules. Move the molecule far enough that the C atoms in the two molecules don't overlap. To make the two molecules easier to tell apart, pull down the Render menu, slide right on Color, slide right on Basic, and choose the color of your choice. Pull down the Edit menu and choose Interactive Superpose. The command line will be replaced with the Superposition prompt. To rotate the selected molecule, hold down the ALT key and drag the mouse using the middle mouse button. We must now choose atom pairs that we wish to superimpose in the two isomers. Click on carbon 1 (Figure 1.1 and 2.1) in each isomer. Click on the Set "2" button in the Superpose button group as shown below:



Now click on carbon 2 in each structure (carbon 2 is the secondary ring carbon adjacent to the tertiary carbon). You can reorient the molecules at anytime using the center mouse button as before. Click on the Set "3" button in the Superpose button group. Click on carbon 6 in each structure (carbon 6 is the other secondary carbon adjacent to the tertiary carbon). You should now have three sets of equivalent atoms. The Superpose button should now be active; click it.

The application will do a rigid body fit. No dihedral angles are changed, the algorithm simply does a least squares fit by adjusting the position of the center of mass and orientation of the molecules.

Notice first that the C-C bond to the methyls doesn't align with the C-H bond from the other isomer on the same tertiary carbon. The methyl groups are bent away from their respective ring to minimize repulsions. These bond angle changes are local differences. Also notice that the secondary carbon on the opposite side of the ring, carbon 4, and its attached hydrogen don't align. In other words, local changes can have an effect many bonds away. This may be caused by ring strain or through-space Van der Waals interactions. Also notice that in the axial conformer that the H's on the methyl face of the ring are pushed away from the methyl group by Van der Waals repulsions.

Problem 7: tert-butylcyclohexane

Compare axial and equatorial tert-butylcyclohexane. Which conformer is more stable this time? Is the ring more or less distorted than in the methylcyclohexane case?

Chapter 8: Printing Structures and Graphics

In the PC and OS-X versions of MOE you can't print directly from MOE, however it is still easy to print structures. You can print anything that is on the current screen, text and graphics.

Structures on PC Systems:

1. From the MOE window, pull down the File menu and choose Save Image... Change to the "c:\Documents and Settings\All Users\Documents\moefiles" directory if it is not the current working directory (position the cursor over the Path field in the Save MOE panel and press the right mouse button). In the file name dialog box, type "moe.bmp" or whatever file name you like with a .bmp extension. Press OK.
2. Start up the Paint program (click on the Start menu, Programs, Accessories, and then Paint). In the Paint application open your .bmp file. You can cut and paste this image into Word or PowerPoint. You can also print this image immediately; however, to save ink you might try:
3. To save ink on color printers it is best to print the structure with a white background. To do this first select any text on the screen (such as energy listings) using the rectangle selection tool. Pull down the Image menu and choose Invert Colors. The text should now be black with a white background.
4. In the lower left-hand corner change the foreground color to white by clicking on the top color square (which starts out set to black) and then clicking on the white square in the color palette. Then select the fill tool, and click on the background in the image.
5. You can move things around and add annotations, etc.

To Change the Image Background Color in MOE:

1. Pull down the Render menu and choose Set Up...
2. The Color: pull down menu should read Background, type in 255 for each of the R, G, and B entries. Click Apply
3. Pull down the Color: menu and choose text. Type in 0 for each of the R, G, and B entries. Click Apply.
4. Click Close.
5. Please do not change the defaults. Many people use these systems, so we need to show courtesy to our fellow chemists and leave the systems in a reliable state.

Printing Plots from MOE on a PC

Use the Start menu to run the Paint program. Make the window that you want to print the active window. Press Alt-Print Screen. Switch to the Paint program, pull down the Edit menu and choose Paste. For scatter plots, the background will be black. To switch to a white background, pull down the Image menu and choose Invert Colors. Print the graphics from the Paint program as you would normally.

Chapter 9: Conformational Preference of Small Peptides

The purpose of this lab is to determine the lowest energy conformation of alanylalanine and to compare this to the value found in the alpha helix in proteins. In particular, we wish to ask if the alpha-helix is the lowest energy conformation of the backbone, or is it a higher energy conformation that must be stabilized by hydrogen bond interactions in large systems? The backbone angles are defined in the Figure 9.1. ψ is defined by the N-C-C-N dihedral and ϕ is defined by the carbonyl carbons in the dihedral C-N-C(α)-C. The normal values in the alpha helix in a protein are $\psi = -47^\circ$ and $\phi = -57^\circ$. The structure of alanylalanine with the carboxyl terminus converted to a methylamide is shown in Figure 9.1. The N-terminus is shown in its non-ionized form. At neutral pH the N-terminus would be a $-\text{NH}_3^+$. However, in our current work the attraction of the charged end-group for the carbonyl oxygens would dominate the conformation. Since we want to study the conformational preference of the backbone, we will build the non-ionized form to avoid the charged end-group attraction, which does not play an important role in large proteins.

We also wish to include Born solvation so that we can generate the conformational preference in the aqueous, rather than gas, phase. Using a continuum solvation treatment won't greatly increase the calculation times. Using a solvation treatment will also decrease the importance of hydrogen-bonding between the terminal $-\text{NH}_2$ and the carbonyl oxygens. This hydrogen-bonding pattern is not important in the bulk of a protein structure.

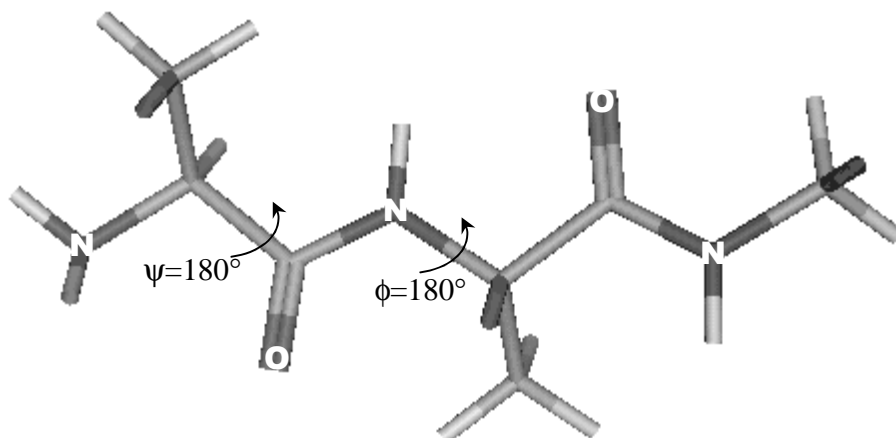
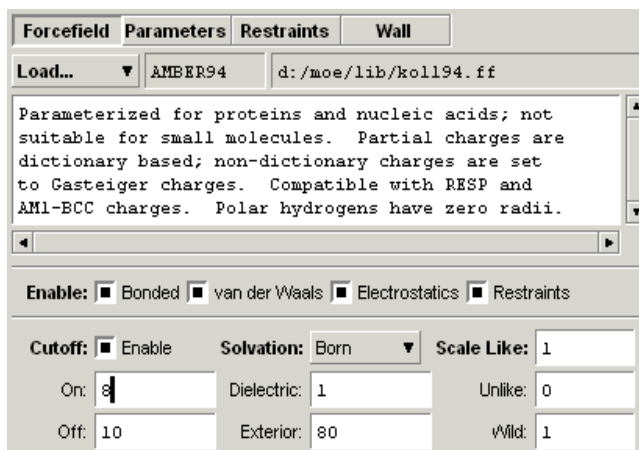


Figure 9.1. The backbone dihedral angles in alanylalanine. The peptide is shown in the all-*trans* conformation.

Procedure

First we need to set the force field type. Pull down the Window menu and choose Potential Setup. Pull down the Load... menu and choose AMBER'94. The AMBER force field was specifically designed for proteins. Peter Kollman's group in the Department of Biochemistry and Molecular Biology at UCSF¹ developed the force field, which is why the force field file is named koll94.ff. Make sure that the Born solvation option is chosen as shown below:



Next we need to build the dipeptide. Pull down the Edit menu, slide right on Build, and choose "Protein". Make sure the "extended" button is selected, so that our structure will be built in the all *trans* conformation. Click on "ALA" twice and Amidate C-term. Then click on Close. The peptide is built in the default "zwitter-ion" form, which we must now change. Click on the Builder... button. Select the NH_3^+ terminal N atom. In the upper right of the Builder panel click on the "0" charge button. An H atom should be removed from the N-terminus of your dipeptide. Close the Builder... Assign the atom charges by pulling down the Compute menu and choosing Partial Charges... The default "Current Forcefield charges" setting should be used. Click on OK. Check the charges that have been assigned to the atoms by clicking on the Atoms button at the bottom-right of the MOE window and choosing Label: Charge. The N-terminus N atom should have a small positive charge. Note the other charges, too. Remove the labels by clicking on the Atoms button again and choosing Label: Charge. Now we will check to see if the minimized structure changes much from this idealized -180° , -180° initial structure.

Use the Check list at the end of Chapt. 1 to minimize the molecule, using the Amber force field with solvation. Make sure to save your structure; use the moefiles/smallmolecules directory. What energy did you get? Use the following instructions to find the backbone dihedral angles.

To find the dihedral angles, first click on the "Measure" button at the right of the MOE window. Pull down the Measure menu in the upper left and choose Dihedrals. Now whenever you click on four atoms in a row, the dihedral angle will be displayed. Make sure that you click on the four atoms in the order in which they are connected. To select the ψ dihedral, start from the N-terminus and click on the backbone atoms: N-C-C-N in turn. To select the ϕ dihedral, start with the carbonyl-carbon on the N-terminus end, and then select the backbone atoms: N-C(α)-C(carboxyl) in turn. Compare these values to the "ideal" alpha helix values. Leave these dihedral monitors on. To quit the dihedral measurement mode press the "Esc" key repeatedly until the white command line reappears.

What hydrogen bonding exists for this conformation? Pull down the Render menu, slide right on Draw, and choose "Hydrogen bonds." If you find any hydrogen bonds, how do these hydrogen bonds stabilize the conformation? Are the hydrogen bonds that form similar to those in an alpha helix? (A file is available with 10 alanines in a right-handed helix, AAAAAAAAAA.moe, so you can compare to a regular helix. Alternately, you can build a short alanine polypeptide in the Protein builder to see what the normal hydrogen bonding pattern looks like. Just make sure the peptide is at least ten ALA's long, and choose the "Helix" secondary conformation option.)

Problem 9.1

Adjust the torsional angles in your dipeptide to give $\psi = -60^\circ$ and $\phi = -60^\circ$. To accomplish this do the following. Pick the first atom defining the ψ torsion by clicking on the C(α)- carbon at the N-terminus end. Pick the second atom defining the torsion by Shift-clicking on the adjacent carbonyl-carbon atom. Hold down the "Alt" or "Command" key and drag the mouse using the left mouse button to change the dihedral angle. Next repeat the above procedure for the ϕ angle, which should be set to -60° . Then click on Minimize button until the energy is minimized.

- Measure the new dihedral angles and record the energy.
- Which conformation is lowest in energy, the minimized structure starting from the all-*trans* conformation or this new one minimized starting from $\psi = -60^\circ$ and $\phi = -60^\circ$?
- Compare your results to the Ramachandran plot in the textbook Figure 8.7.11 or Atkins, section 22.3.² What regions of the Ramachandran plot correspond to your two structures?
- Are either of the structures stabilized by hydrogen bonds (they wouldn't be in an alpha helix)?
- Based on your results, is the alpha-helix like structure the lowest energy conformation of the backbone, or is it a higher energy conformation that must be stabilized by hydrogen bond interactions in large systems?

Please note that for adjusting dihedral angles with the mouse, you select only two atoms. In other parts of MOE, for example in measuring dihedral angles and for Conformational Searches, you need to specify all four atoms of the dihedral.

References:

- www.amber.ucsf.edu/amber/amber.html
- P.W. Atkins, J. de Paula, "Physical Chemistry, 7th Ed.," W. H. Freeman, New York, NY, 2002. Section 22.3.

Chapter 10: Dynamics in Small Peptides.

Purpose: The purpose of the chapter is to use molecular dynamics to find low energy conformations for the alanylalanine dipeptide. This chapter is a continuation of Chapter 9.

Introduction

Molecular dynamics is useful for visualizing the motions of macromolecules. Motional flexibility of enzymes plays a role in binding interactions and in catalytic events. In this Chapter we will study the alanylalanine dipeptide, which you built in Chapter 9. We choose such a small system so that the calculations will run quickly. However, the same procedures are used routinely for large enzymes and oligonucleotides. Molecular dynamics is also a good way to find low energy conformations. Often, energy minimization alone catches the molecule in conformations that are not the lowest energy conformation. Molecular dynamics helps the molecule explore other conformations that may be lower in energy. The take home message from dynamics simulations is that there is more motion than we expect from viewing static textbook models. The motion of molecules is exceedingly important in determining the energetics and the course of chemical reactions.

Molecular mechanics minimization corresponds to the structure the molecule would have at zero degrees K. Careful dynamics calculations in MOE are done in three steps. We first do a "heating" segment to warm the molecule to the chosen temperature. Then an "equilibration" period is used to equilibrate the molecule at the chosen temperature to ensure that all the degrees of freedom are at the same temperature. Next, we do a "simulation" segment that generates the trajectory of the molecule at the chosen temperature. The "equilibrium" run is used to answer questions about the motion of the molecule. Finally, a "cooling" segment is run to return the molecule to a low temperature. The cooling segment is used for "simulated annealing." If the final temperature in the "cooling" segment is ~ 0 K, then the final structure after the dynamics run will be comparable to a straight molecular mechanics minimization, however, hopefully in a very different conformation than the starting conformation.

In this exercise we will first do a "quick and dirty" dynamics run to familiarize you with the MOE interface for quick dynamics based conformation searches. Then we will do a careful dynamics based simulated annealing study with a heating, equilibration, simulation, and cooling period.

Procedure:

To begin, complete Chapter 9. We will see if molecular dynamics is successful in finding the lowest energy conformation of the dipeptide. Minimize the structure from Chapter 9. Make sure that hydrogen bonds are showing (Pull down the Render menu, slide right on Draw, and make sure Hydrogen Bonds is checked). Remove any dihedral angle monitors, if present. If the Commands window is not showing, pull down the Window menu and choose Commands.

Quick and Dirty Conformation Search Using Molecular Dynamics

Pull down the Compute menu, slide right on Simulations, and choose Dynamics. Set the time step to 0.0005 psec, which is 0.5×10^{-15} sec or 0.5 femtosecond!

In this and all other similar instances please make sure the current working directory is the "moefiles" directory on your system. The output database will then be stored in the proper directory. The NVT button should be selected giving a dynamics run at constant Number of

molecules, constant Volume, and constant Temperature. Change the Duration to 20 ps. Make sure the Open Database Viewer checkbox is selected.

Database: dynamics.mdb <input style="float:right" type="button" value="Browse..."/>		
<input type="checkbox"/> Resume Simulation <input checked="" type="checkbox"/> Open Database Viewer		
Save: <input type="checkbox"/> Position <input type="checkbox"/> Velocity <input type="checkbox"/> Acceleration Every: 0.5 ps		
Ensemble: NVE NVT NPT NPH		Algorithm: NPA NHA BER
T0: 300	T: 300	P: 101
Heat (ps): 0	Run: 20	Cool: 0
Constrain: Light Bonds	Water: Rigid	Tol: 1e-12
Time Step: 0.0005	QT: 0.2	QP: 5

Setting the Constrain: Light Bonds allows a larger time step to be used than would be used normally (for large molecules the time step could be set at 0.001 or 0.002 psec). Using these constraints the program holds the bond lengths to H atoms fixed. This option is called the "Shake" option in other programs. The "Shake" option should only be used in conformational studies and not for simulations that are designed to give thermodynamic parameters. At least a 100 psec duration is necessary to explore a significant portion of the "conformational space" of the molecule. In other words, trajectories require a lot of time to sample a wide range of conformational angles. However, for our purposes in this exercise 20 ps will be sufficient. Click on "OK." The structures are displayed as they are generated. During the run the Data Base Viewer will be opened.

The last structure in the trajectory is automatically loaded into the MOE window. Minimize this structure to see if you get a lower energy than you have obtained before. You can also sort the structures by energy: click on the "U" heading in the database table and then click right. Choose Sort Ascending. The structures are then sorted by energy. You can then load a few of the low energy structures into the MOE window and minimize to look for other low energy conformers. Close the Database Viewer window.

In general there is no reason to restrict the simulation to 300K. Often in dynamics studies temperatures as high as 900 K are common. It is often necessary to use temperatures in the 500-900K range to find new interesting conformers for small molecules.

Careful Conformational Study.

For the careful dynamics study, we will run a trajectory with a heating period of 50 ps to a temperature of 500 K, a 100 ps combined equilibration and simulation section at 500K, followed by a 20 ps cooling period. The cooling period will provide the "simulated annealing" result as the last structure in the trajectory. Again, pull down the Compute menu, slide right on Simulations, and choose Dynamics. Enter the parameters as shown below.

Database: dynamics.mdb <input style="float:right" type="button" value="Browse..."/>		
<input type="checkbox"/> Resume Simulation <input checked="" type="checkbox"/> Open Database Viewer		
Save: <input type="checkbox"/> Position <input type="checkbox"/> Velocity <input type="checkbox"/> Acceleration Every: 0.5 ps		
Ensemble: NVE NVT NPT NPH		Algorithm: NPA NHA BER
T0: 0	T: 500	P: 101
Heat (ps): 50	Run: 100	Cool: 20
Constrain: Light Bonds	Water: Rigid	Tol: 1e-12
Time Step: 0.0005	QT: 0.2	QP: 5

Run the trajectory as before. Overwrite the previous “dynamics.mdb” file. (If you get an error message about the time step being too large or the structure too strained, see Note 1 at the end of this section.)

When the dynamics run is complete a structure will be loaded into the MOE window. This structure is the result of the simulated annealing segment. Minimize this structure. Record the energy and the dihedrals. Is this structure similar to the results from Chapter 9?

To follow the temperature during the run, in the Database Viewer, pull down the Display menu and choose Plot. A new plot panel will appear. In the plot window pull down the Plot Field menu and choose T. A plot of the temperature will now appear. Note the heating and cooling segments. Click on (or drag the mouse around) one of the plot symbols about one-third of the way through the 500K constant temperature segment. This plot symbol will be circled and the corresponding row in the database will be selected. Scroll through the database to note the entry number for your selected point (Which should be around 140-180 or so). We will assign this data point as the end of the equilibration segment.

We now need to remove the structures from the heating and equilibration segments from the database. Click on the first row ID number at the far left of the database display. Then scroll down to your datapoint that corresponds to the end of the equilibration segment and shift click. The first though about 140 to 180 rows should now be highlighted. Pull down the Edit menu, slide right on Delete, and choose Selected Entries. Verify that you want to delete the first 170 (or so) entries. Now do the same deletion steps for the cooling period, that is for temperatures significantly less than 500K -- the last 15 psec or so. Clear the molecule in the main MOE window by clicking on Close at the right-side of the MOE window.

Now we want to see the simulation trajectories: in the Database Viewer, pull down the File menu in the Database Viewer and choose “Browse.” In the “Browser” tool, drag the Animation dial to the right. The trajectory will cycle through all the time steps and then repeat. If needed click on the View button in the main MOE window to center the molecule on the screen. To increase the speed of the animation, drag the dial further to the right. Click on “Close.” Send the first molecule in the database to the MOE window to prepare for the next step: right click on the first molecule in the Database Viewer and choose Send to MOE. Select Clear System before loading data.

Next we want to get the ψ - ϕ plot for the trajectory. The purpose of this plot is to determine the range of dihedral angles that were sampled in the trajectory. If a wide range was sampled, we can have some confidence that we have a good chance of finding the global minimum. If the range of sampled dihedral angles is small, we should redo the trajectory for a longer time or at a higher temperature. To generate the ψ - ϕ plot, in the Database Viewer, pull down the Compute menu, slide right on Molecule, and choose Conformation Geometry. In the main MOE window, trace out atoms in the ψ dihedral as you did before. The atom labels N-CA-C-N should be entered for you in the Dihedral Field in the Conformation Geometries dialog box. Click on Measure. The program will enter the value for this dihedral for each molecule in the Database in a new column. Now trace out the atoms in the ϕ dihedral. The atom labels C-N-CA-C should be entered for you in the Dihedral Field. Click on Measure to enter these dihedral measurements in the Database. Close the Conformation Geometry Dialog box. Now pull down the Compute menu, slide right on Analysis and choose Correlation Plot. Click on the N-CA-C-N column heading in the database to specify the x-axis variable, and the C-N-CA-C column heading to specify the y-axis variable. The ψ - ϕ correlation plot will appear. Click on Attributes... Pull down the top menu and choose Series Attributes. Choose Solid for the Line Style. Click Close. From the plot decide if the

trajectory has sampled conformations that are significantly different from the conformations studied in Chapter 9. Close the Correlation plot.

Check for low energy structures during the dynamics run by right clicking in the U column heading. In the pop-up menu choose Sort Ascending. The structures should then be sorted by energy. Transfer the lowest energy structure to the main MOE window by clicking the right mouse button in the first molecule cell. You should see a pop-up menu; select Send to MOE. Next you see a new dialog box; select the Clear System before loading data option and then OK. The new molecule should appear in the MOE main window. Click on Minimize several times to make sure the molecule is completely minimized. Record the ψ and ϕ angles and the final energy.

Problem 10.1

Does the dynamics trajectory find any new conformations that are significantly different from those you have already found in Chapter 9? Do any hydrogen bonds form to stabilize the structure? In the alpha-helix there would not be any such hydrogen-bonds, since the hydrogen bonds in the alpha-helix are between amino acids four residues apart along the backbone. If there are any hydrogen-bonds, are these hydrogen bonds different than in Chapter 9 (if any)? Is there more or less motion than you expected? Compare the energy of the minimized structure from the end of the dynamics simulation and the minimized lowest-energy structure with the structures that you obtained by minimizing starting from the initial $\psi \sim -60^\circ$, $\phi \sim -60^\circ$ and $\sim 180^\circ$, $\sim 180^\circ$ structures from Chapter 9.

Note:

1. To help set up the careful dynamics study, we can use a macro-function, an “svl”, to set the parameters for us. Pull down the File menu, choose Open... and click on the “md_heatcool.svl” entry in the file list. Then click on Run SVL button. The svl macro will run a trajectory with a heating period of 50 ps to a temperature of 500 K. Then a 100 ps combined equilibration and simulation section at 500 K, followed by a 20 ps cooling period.

Chapter 11: Solvation and β -Cyclodextrin

For small gas-phase molecules, advanced molecular orbital theory is more accurate and more useful than molecular mechanics. However, molecules with even as few as 20 atoms require large amounts of computer time for molecular orbital calculations. For large molecules or small molecules in a solvent, molecular mechanics is still the only practical computational technique. The real power of molecular mechanics calculations is in the ability to handle explicit solvent molecules. The conformation of a molecule can depend strongly on the presence of solvent. In addition, studies of molecular recognition and binding require careful consideration of solvent effects. In this chapter we study the interaction of water with the cyclic saccharide host, β -cyclodextrin, CD.

Cyclodextrins are often used as active site analogs for enzymes¹. Cyclodextrins are used to aid the absorption of drugs in the body. Other uses for cyclodextrins include the petroleum industry for separating aromatic hydrocarbons and in agriculture to reduce volatility of insecticides. Cyclodextrins are natural products produced by bacteria from starch. CD is made from seven D(+)-glucopyranose units linked through α -(1 \rightarrow 4) glycosidic bonds², Figure 1.

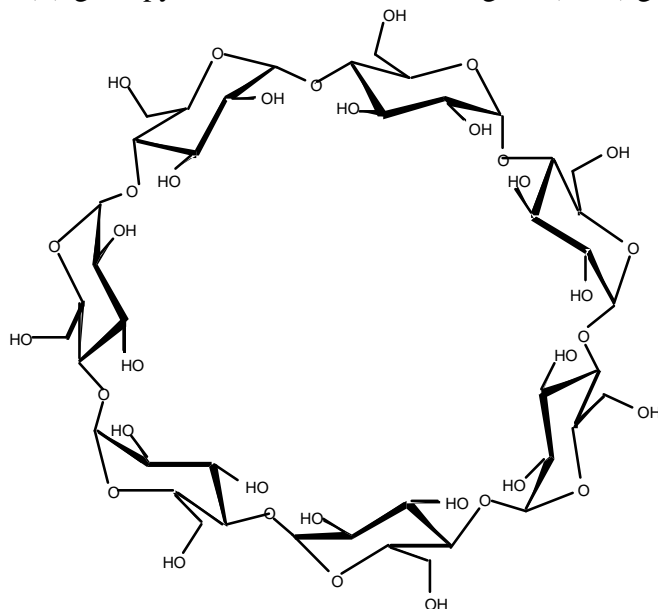


Figure 1. β -cyclodextrin (cycloheptaamylose).

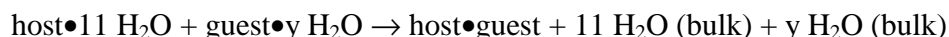
When dissolved in water, water molecules will fill the cavity of the host. Then when a guest interacts with the cavity of the host, water molecules are displaced. The binding affinity depends on the interactions of guest with the host and the difference in the interactions of bound water and water with the bulk solvent. The cavity of cyclodextrin holds around 11 water molecules. Most or all of these are excluded from the cavity upon binding with a guest. This process is called desolvation:



In addition, the guest interacts with water before it binds to the host, and these waters also must be released to the bulk of the waters in solution. The number of interacting waters is difficult to predict, so let's call the number y :



The change in Gibbs Free energy in this process is often unfavorable and is called the "desolvation" penalty. The net process is then:



In summary, solvation plays a very important role in molecular binding. In this chapter we will minimize β -cyclodextrin in aqueous solution and in vacuum to determine the number of waters in the cavity and any changes that occur upon solvation/desolvation.

When using explicit solvation it is best to use periodic boundary conditions. In this way we don't have to worry about what happens at the sides of the box of waters. Mirror images of the box will be stacked next to each other in all directions so that there are no surfaces to the solution. Periodic boundary conditions eliminate any surface tension effects.


Procedure:


Solvate β -Cyclodextrin: The β -cyclodextrin file is in the "c:\Documents and Settings\All Users\Documents\moefiles\Gh" directory on the PC and "Documents/moefiles/Gh" directory on OS-X and is called beta-cyclodextrin.moe. Open this file using the "File" menu and "Open." Please rename this structure so that the original file is not changed by pulling down the "File" menu and choosing "Save...", switch to the "moefiles/smallmolecules" directory, and give a new name in the file name dialog.

To add solvent, pull down the Edit menu and choose Build and slide right on Water Soak. Make sure the Soak Atoms Solute:All setting is selected. Click on the Box button to generate a cubic box of waters around your molecule. Change the box sizes to 17 Å on each side. To size the box to get complete layers of water click on the Adjust Box Size button. The box sides will be changed to 18.7 Å. Click on the Update Periodic Box checkbox to transfer the box sizes to the Crystal Parameters Dialog (which you will open next). Click OK. Pull down the Window menu and choose Crystal Parameters... The box dimensions should be entered for you as set in the Water Soak dialog. Check the Enable Periodic System checkbox to enable periodic boundary conditions in the energy minimization. Click on OK. The aquated structure now needs to be minimized. Use the Check List at the end of Chapter 1 to set up the forcefield. Since we are using explicit solvent molecules, use MMFF94x with the Solvation option selected as Gas Phase as usual. Minimize the molecule using the default Gradient of 0.05 and Hamiltonian option set to Forcefield. Pull down the File menu and Save your changes to the same file name. (Alternately, instead of doing a minimization, if time is short your instructor may allow you to open a pre-minimized structure: bcycludexaq.moe and you can skip to the next paragraph).

How Many Waters Are In the Cavity?

Count the number of waters in the cavity. This is most easily done by adjusting the clipping plane, or using Atom Selector tools to select waters in the cavity as described below. These steps can be done while the minimization is finishing.

Clipping . First make sure the system is centered in the MOE window by clicking on the View button. Then Zoom the window until the cyclodextrin almost fills the full window (Use the Zoom dial or Ctrl-middle mouse button). To help you see things better, try reducing the clipping plane to avoid displaying water molecules that are in the foreground and background. To reduce the clipping plane first click on the ZClip button at the very bottom-right. The Front setting determines where the clipping planes begins, in the Z direction. Pull the Z-Clip: Front slider into the middle of the range. Then pull the Width slider almost all the way to the bottom until as few as possible water molecules are displayed without any of the cyclodextrin atoms disappearing. You may need to reorient the cyclodextrin if the ring is tilted away from the plane of the window. To get an idea of the dense nature of solutions and the close interactions of the solvent with the solute make sure no atoms are selected, then click on the Atoms button and choose the space filling display, . To reset the clipping planes to their default distances, just return the Front slider all the way to the bottom and the Width slider all the way to the top. Use the Atom button to return to Line rendering mode.

Atom Selector Tools: Rotate the cyclodextrin to locate a water molecule that is in the center of the cavity. Click on that water molecule. Pull down the Selection menu and choose Atom Selector... Type in 5.5 for the Radius, which is a little larger than half the diameter of the ring. The units are Å. Click Proximity. To make sure that whole molecules are selected click on the Molecule button in the Extend button group. Several waters should now be colored pink, showing that they are selected along with the cyclodextrin. These waters should fill the cyclodextrin cavity. If things didn't go well, you can always click on the black background, and try again. When you have waters in the cavity selected, pull down the Render menu and slide right on Hide and choose All. Now pull down the Render menu, slide right on Show and choose Selected. Only the selected molecules should now be displayed on the screen. Click on the black background to remove the current selections. Close the Atom Selection window. Now only the waters in the cyclodextrin cavity should be visible. Count the displayed water molecules. Rotate the cyclodextrin and check to see if all the displayed waters are actually in the cavity; if they are not in the cavity don't include them in the sum. You can repeat the Proximity selection process again if the displayed waters are not well centered in the cavity. To get an idea of the dense nature of solutions and the close interactions of the solvent with the solute make sure no atoms are selected, then click on the Atoms button and choose the space filling display, . To finish, click on the Atoms button and choose the Line rendering mode. Then pull down the Render menu, slide right on Show and choose All.

Is the Diameter the Same in Solution and in the Gas Phase?

To make the distance measurements easier we will start by deleting the water molecules.

Delete the Solvent Molecules: To remove the solvent from the actively displayed molecule on the screen, click on one of the water molecules. Pull down the Selection menu, slide right on Extend, and choose Chain. Click on the Delete button at the right of the MOE window. In the Delete Atoms dialog box both options should be selected and then click OK.

Measure the diameter of the top of the cyclodextrin cavity, using the Measure button with the "Distances..." mode. Make sure to select H atoms that point in towards the cavity. See Chapter 3 for instructions. Choose three H to H atom distances and average your values. Now Minimize the molecule without solvent. Record the same diameters and average your values. To exit the

measurement mode, press the Esc key until the white command prompt box returns (just below the menu bar).

Problem 11.1: Report the number of water molecules in the cavity. Report the change in the average diameter of the cyclodextrin cavity. Also report the change in average diameter as a percentage, i.e. give a statement like “the cavity diameter enlarged by ~X% with waters present” or “the cavity diameter decreased by ~X% with waters present.”

Literature Cited

- (1) Furuki, T.; Hosokawa, F.; Sakurai, M.; Inoue, Y.; and Chûjô, R. *J. Am. Chem. Soc.* **1993**, *115*, 2903-2911.
- (2) Diaz, D.; Vargas-Baca, I.; and Garcia-Mora, J., *J. Chem. Ed.* , **1994**, *71*, 708.

Chapter 12: Docking: β -Cyclodextrin and β -Naphthol

In Chapter 11 we discussed some of the influences of the solvent on guest-host binding. In this chapter we will focus on the interactions of the guest and host. We want to find the conformation and energies of the interaction of the guest and host. β -Cyclodextrin and β -naphthol form a guest-host complex.

β -Cyclodextrin has a hydrophobic binding pocket and $-OH$ groups at both rims of the binding pocket. β -Naphthol, Figure 1, is expected to bind to CD because it has a hydrophobic group that fits into the cyclodextrin cavity, while the OH group participates in hydrogen bonds with the sugar OH groups. The reaction stoichiometry is 1:1.

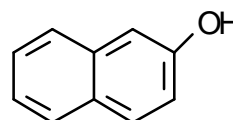


Figure 1. β -Naphthol

We will use the docking features of MOE to predict the conformation of the complex. Docking monitors the Van der Waals and hydrogen bonding interactions of the guest and host as you change the position and orientation of the guest in the cavity of the host. We will then do a full molecular mechanics minimization to estimate the binding energy for the guest-host interaction.

Procedure:

- β -Cyclodextrin:** The β -cyclodextrin(aq) file is in the "c:\Documents and Settings\All Users\Documents\moefiles\Gh" directory on the PC and "Documents/moefiles/Gh" directory on OS-X and is called "beta-cyclodextrin.moe." Open this file using the "File" menu and "Open."
- β -Naphthol:** The β -Naphthol file is also in the "c:\Documents and Settings\All Users\Documents\Gh" directory on the PC and "Documents/moefiles/Gh" directory on OS-X and is called b-naphthol.moe. Open this file using the "File" menu and "Open." Please rename this resulting combination structure so that the original file is not changed by pulling down the "File" menu and choosing "Save...", switch to the smallmolecules directory, and give a new name in the file name dialog.
- Pull down the Window menu and select Potential Setup and change the Solvation option to Born. The Born option uses a continuum solvation model to approximate solute solvent interactions. The Born approach is much faster than using explicit solvation. Make sure that the GizMOE energy is being displayed on the screen. Click on an atom in β -naphthol. Pull down the Selection menu, slide right on Extend, and choose Molecule. Move the guest around in the host cavity by holding down the Alt and Shift keys while dragging with the middle mouse button. The mouse keys are then:

<u>PC Mouse:</u>	Middle: Reorient molecule—xyz rotation
	Middle and drag in periphery of viewing area: rotate around z only
	Shift-Middle: xyz translation
	Ctrl-Middle: zoom in and out
	Shift-Alt-Middle: Translate selected atoms
	Alt-Middle: Rotate selected atoms
	Alt-Left: change dihedral angle between two selected atoms

Turn the guest around so that the OH group is close to the OH's that are directly attached to the rings at the wide end of the cavity. Center the guest in the cavity. You can switch between moving the guest only and moving both the host and guest together by either holding down the Alt key or not. Use the GizMOE energy to find low energy positions for the guest.

4. **Docking:** To turn on manual docking, pull down the Render menu, slide right on Draw, and choose Hydrogen Bonds. Note that as you move the guest around possible hydrogen bonds will be shown. Now pull down the Render menu, slide right on Draw, and choose VDW contacts. Now the yellow dashed lines will show you when two atoms are too close together. Position the guest to get as low an energy as possible. Pay close attention to the possibility of hydrogen-bonding between the guest and the host. Use the Check List at the end of Chapter 1 to Minimize and record the total steric energy. Also pull down the Compute menu and choose Potential Energy and record the energy contributions listed. You can use the Commands window to list the potential energy terms, if you like. You can get the Commands window by pulling down the Window menu and choosing Commands.

Now pull down the GizMOE menu and choose Minimizer. Now slowly move the guest around to see if you can lower the energy even further. If you can, record the total and various force field terms. When you are finished pull down the Cancel menu in the upper right-hand corner of the screen and cancel out any instances of the Minimizer and Giz MOE tools.

5. Load in just β -cyclodextrin and minimize and record the total steric energy. Also pull down the Compute menu, select Potential Energy, and record the energy contributions. Then load in just β -naphthol by itself, and repeat the minimization and energy recording steps. **PLEASE:** Don't "Save Changes" to avoid changing the original files.

Problem 12.1: Describe the conformation of the guest in the cavity of cyclodextrin. Find the binding energy for the complex by taking differences in the total steric energy of the reactants and product. Also, find the differences in each of the contributions to the steric energy. Which contributions to the steric energy dominate the binding interaction (eg. Bond stretch, bond bending, Van der Waals, electrostatic, etc.).

Auto-Docking

The ability to predict the conformation of a ligand in a protein or nucleic acid binding site is of central importance in pharmaceutical design. Detailed knowledge of the molecular recognition forces is necessary for designing new ligands. Knowing the binding site conformation helps to show the important interactions that stabilize the complex. As you have seen, determining the lowest energy conformation for even such a simple system as cyclodextrin can be difficult. Automated procedures are very helpful in exploring a wide variety of binding modes in complex systems. The advent of rapid parallel binding assays through combinatorial chemistry and high throughput screening also motivates a similar need for efficient, rapid computational screening procedures. The ability to computationally predict possible tight-binding ligands lowers the time and energy necessary to plan a major combinatorial study.

One approach to auto-docking is to use molecular mechanics techniques to minimize the energy of a ligand in a binding pocket starting from a series of random positions and orientations within the binding site. Finding the low energy binding conformations will then help to highlight the interactions that are responsible for tight binding. For example, the hydrogen bonds in the low energy complexes can be displayed. Electrostatic surfaces of the binding pocket and the ligand can also be displayed to observe how the distribution of charge in the ligand and binding pocket complement each other. Several methods have been developed to efficiently sample possible binding conformations.

Tabu Search: A Tabu search is a random searching algorithm that maintains a list of previously visited conformations^{1,2}. These conformations are forbidden (*tabu*) to future moves. This procedure insures that a wide range of starting conformations is tested. Many small changes in the internal conformation of the ligand and the position and orientation of the ligand relative to the binding pocket are then chosen. The lowest energy conformations are saved to a list.

Simulated Annealing: Simulated Annealing searches also start with a random set of conformations³. Many small changes in the internal conformation of the ligand and the conformation of the ligand relative to the binding pocket are then chosen. These small changes are chosen at random using the Monte Carlo technique⁴. In the Monte Carlo technique low energy conformations have a higher probability than high energy conformations, but because of the randomness both low and high energy conformations will be chosen. The low energy states are just chosen more often. The Monte Carlo technique starts by calculating the starting energy, E_0 , and the energy of the proposed conformation from the small change, E_1 . The change in energy for the proposed conformation is then $\Delta E = E_1 - E_0$. The Boltzmann weighting factor then gives the probability that the change will occur, given the set temperature:

$$\text{probability of occurrence} = e^{-\Delta E/RT} \quad 1$$

A random number is then generated in the range of 0 to 1. If this random number is greater than the Boltzmann probability the move is rejected, if the random number is less the move is accepted. The higher the temperature you set the more often high energy conformations will be chosen. The units for the temperature are degrees Kelvin. This process is complementary to and in some ways similar to molecular dynamics. The purpose is to efficiently sample many good starting geometries without the long run times that would be required in a dynamics study⁵. However, as a final result we seek the minimum energy configuration. Finding the minimum energy conformation is where simulated annealing is used.

The Monte Carlo technique is used to generate random changes in the structure. To find the energy minimum, as the calculation proceeds over hundreds of changes in conformation, the temperature is gradually lowered. The Boltzmann weighting factor thus becomes more selective for low energy conformations. As the temperature drops, the structure drops into a minimum energy state. The gradual lowering of the temperature is called simulated annealing. Annealing is the process in metallurgy and glass blowing of heating and then gradually cooling a sample to relieve internal stress. Simulated annealing can be applied to Monte Carlo simulations and molecular dynamics (with a Cool period after the equilibration phase).

The steps in docking are outlined below. The complete step-by-step instructions follow:

Docking Summary:

1. Select the ligand.
2. Build the Docking box to delineate the region in the binding pocket to place the ligand.
3. The program then does the calculations in two steps:
 - Step 1: Generate random starting orientations.
 - Step 2: Use random small moves to find a low energy conformation.
4. Extract the lowest energy conformer (or a few low energy conformers).
5. Minimize the resulting models using normal molecular mechanics.

Procedure:

1. Pull down the File menu, choose Open. If you have not set "c:\Documents and Settings\All Users\Documents\" on the PC or "Documents/moefiles/" for OS-X as the current working

directory please do so now. Switch to the \Gh subdirectory. Click on the “bnapbcycdex.moe” file and click on the Open MOE File button. Now save this file under a different name to avoid changes in the original file.

2. Use the Check List at the end of Chapter 1 to minimize the energy of the complex. You can use a Gradient of 0.01. Record the total steric energy. Pull down the Render menu, slide right on Draw, and choose Hydrogen Bonds. Note the hydrogen bond interaction between the guest and host. Which type of oxygen atom is involved?
3. Type

```
run '$MOE/sample/obsolete/dock.svl'
```

 into the white command line at the top of the MOE window.
4. Click on the Docking Box button. Drag the Docking Box and MOE-Dock dialog boxes to the side of the screen so that you can see the molecules.
5. In the Docking Box dialog, rotate and translate the docking box so that the box ends are parallel to the rims of the cyclodextrin cavity. You will need to rotate the ring to look at the molecule and box orientation from the x, y, and z directions. Use the Extent: x, y and z sliders to decrease the box size to just include the hydrogens that line the cyclodextrin cavity. (~24, ~24, ~24 works well, but different students should try different values). Click Apply and then Close.
6. In the Docking dialog box, click on Open Database Viewer and Change the Total Runs to 5 to help save time. The Random Start checkbox should be selected. You can choose to do a simulated annealing or a Tabu search.
7. If you choose to do a *Simulated Annealing search*: make sure the Simulated Annealing button is depressed. Set the following options:
 6 Cycles per run 500 Iteration Limit 1000 Initial Temperature (K)
 If you choose to do a *Tabu search*: make sure the Tabu button is depressed. Set the following options:
 500 Steps Per Run 5 Attempts Per Step 5 Tabu List Length
 These options are chosen to make the calculations run quickly (Tabu is shorter). A careful study would use the default settings or more.
8. Click on an atom in the beta-naphthol, pull down the Selection menu, slide right on Extend and choose molecule. The guest should now be highlighted. Click on OK.
9. In a few minutes the Database viewer will display the low energy conformations. To view the molecules better, click in any of the molecule boxes and drag the mouse down to enlarge the boxes.
10. In the main MOE window, the β -naphthol should still be highlighted. Click on the Delete button on the right side of the MOE window to remove the guest. Click OK with both Delete Options selected. You can also remove the yellow docking box by pulling down the Window menu, choosing Graphic Objects, clicking on the “docking box” line, and then clicking on Delete, and finally clicking on Close.
11. Find the lowest energy conformation in the Database viewer, which is the smallest U_{total}. Double click in the molecule box for this conformation. In the Copy Database Molecule to MOE window make sure that neither of the options are selected (we want to make sure that the cyclodextrin remains on the screen). Click OK. The docked conformation should now appear in the MOE window.
12. Minimize the complex (use the Check List for instructions). Record the energy.
 Note the hydrogen bond interaction between the guest and host. Which type of oxygen atom is involved?
13. Pull down the Database Viewer File menu and choose Quit.

Problem 12.2

Which conformation, the original or the MOE-Docked, has the lowest energy? Compare your results with others in the class. If some students used Simulated Annealing and some Tabu searches, which search strategy gave lower energies? Are the hydrogen bonding partners similar in the two structures?

References:

1. MOE Panel Index, Chemical Computing Group, Montreal Canada.
2. Baxter, C. A., Murray, C. W., Clark, D. E., Westhead, D. R. and Eldridge, M. D. Flexible Docking Using Tabu Search and an Empirical Estimate of Binding Affinity. *Proteins: Structure, Function and Genetics*. **33**, 367-382 (1998).
3. Hart, T. N. and Read, R. J. A Multiple-Start Monte Carlo Docking Method. *Proteins: Structure, Function and Genetics*. **13**, 206-222 (1992).
4. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. Equation of State Calculations by Fast Computing Machines. *Journal of Chem. Phys.* **21**, 1087-1092 (1953).
5. Morris, G.M., Goodsell, D.S., Huey, R., Olsen, A.J. Distributed Automated Docking of Flexible Ligands to Proteins: Parallel Applications of AutoDock 2.4. *Journal of Computer-Aided Molecular Design*. **10**, 293-304 (1996).

Chapter 13. Henry's Law Constants and Gibb's Free Energy of Solvation

Introduction

The fate of organic molecules in the environment is determined in part by their solubility in water. For example, an oil spill or leaking underground gasoline tanks introduce organics into surface and ground water. The long term damage done to the environment is determined by the solubility of the organic contaminants in the water^{1,2}. Soluble organics can travel long distances and allow the spread of the contamination over wide areas. Less soluble organics quickly evaporate and cause less of a problem. Henry's Law governs the solubility of compounds in dilute solution³:

$$P_B = x_B k_H \quad 1$$

where P_B is the partial pressure of dilute solute B above the solution, x_B is the mole fraction, and k_H is the Henry's Law constant for B. Compilations of k_H values are limited; many thousands of compounds are of concern in the environment and in the laboratory. The purpose of this Chapter is to calculate k_H values from continuum solvation studies. If these calculations are successful, considerable time, effort, and money can be saved in screening compounds for their environmental hazards.

We wish first to establish the connection of the Henry's Law constant to the Gibbs Free Energy of solvation. The equilibrium described by Eq. 1 can be written as:



The equilibrium constant for reaction 2 is:

$$K_{eq} = \frac{P_B}{x_B} \quad 3$$

if we measure the concentration of B in mole fraction. Comparing Eqs. 1 and 3 shows that the equilibrium constant K_{eq} is the same as k_H , the Henry's Law constant. The Gibb's Free Energy change for Eq. 2 is the Gibb's Free energy of desolvation, $\Delta_{desol}G_B$. Therefore, since k_H is the equilibrium constant for Eq. 2:

$$\Delta_{desol}G_B = -RT \ln k_H \quad 4$$

Therefore, calculations of $\Delta_{desol}G_B$ can be directly used to find Henry's Law constants.

The units of K as defined above are in atm. We will call this constant K_{PX} to help us remember the units. Environmental chemists often prefer to deal with unitless Henry's law constants, K_{CC} , where the gas phase pressure is replaced by the gas phase concentration and the solution mole fraction is replaced by the concentration:

$$k_{cc} = \frac{k_{PX}}{kRT} \quad 5$$

where k is the conversion factor from mole fraction to concentration in mol L⁻¹. Here, $k = 1000 \text{ mL } d_{H_2O} / \mathcal{M}_{H_2O}$, where d_{H_2O} is the density of water and \mathcal{M}_{H_2O} is the molar mass of water. At 25°C, $k = 55.35 \text{ mol L}^{-1}$. The unitless k_{cc} corresponds to 1 M ideal gas and 1 M aqueous solution

standard states. Also common in the literature is the Henry's Law constant with pressure for the gas phase and concentration for the aqueous phase:

$$k_{pc} = \frac{k_{px}}{k} \quad 6$$

Table 1 lists values for the various k_H 's and the $\Delta_{\text{desol}}G$ values derived from them using Eq. 4. The proper parameters for comparison with solvation calculations are k_{cc} and $\Delta_{\text{desol}}G_{cc}$.

Table 1. Henry's Law constants and Gibbs Energies of desolvation¹. The number in parenthesis is the source for that substance and following values. $\Delta_{\text{desol}}G = -RT \ln k_H$. The units are indicated as subscripts: p=pressure, x=mole fraction, and c=molarity.

substance	$k_{H,px}$ (atm)	k_{cc} (unitless)	k_{pc} atm L/mol	$\Delta_{\text{desol}}G_{px}$ (kJ/mol)	$\Delta_{\text{desol}}G_{cc}$ (kJ/mol)	$\Delta_{\text{desol}}G_{pc}$ (kJ/mol)	
Benzene(1)		294	0.216	5.32	-14.09	3.80	-4.14
toluene		358	0.263	6.47	-14.58	3.31	-4.63
ethylbenzene		433	0.318	7.83	-15.05	2.84	-5.10
m,p-xylene		406	0.298	7.34	-14.89	3.00	-4.94
o-xylene		278	0.204	5.02	-13.95	3.94	-4.00
1,1,1-trichloroethane		978	0.718	17.7	-17.07	0.82	-7.12
trichloroethylene		572	0.42	10.3	-15.74	2.15	-5.79
tetrachloroethylene		950	0.697	17.2	-17.00	0.89	-7.05
methyl- <i>t</i> -butyl ether		29.4	0.0216	0.532	-8.38	9.51	1.57
Tetrachloroethylene(2)		985	0.723	17.80	-17.1	0.80	-7.14
trichloroethylene		534	0.392	9.65	-15.6	2.32	-5.62
1,1-dichloroethylene		1457	1.069	26.32	-18.1	-0.17	-8.11
<i>cis</i> -1,2-dichloroethylene		228	0.167	4.11	-13.5	4.44	-3.50
<i>trans</i> -1,2-dichloroethylene		523	0.384	9.45	-15.5	2.37	-5.57
vinylchloride		1549	1.137	27.99	-18.2	-0.32	-8.26
1,1,1-trichloroethane		958	0.703	17.31	-17.0	0.87	-7.07
1,1-dichloroethane		313	0.23	5.66	-14.2	3.64	-4.30
chloroethane		621	0.456	11.23	-15.9	1.95	-5.99
carbon tetrachloride		1695	1.244	30.63	-18.4	-0.54	-8.48
chloroform		204	0.15	3.69	-13.2	4.70	-3.24
dichloromethane		122	0.0895	2.20	-11.9	5.98	-1.96
chloromethane		492	0.361	8.89	-15.4	2.53	-5.42
Methane(3)		413	0.303	7.46	-14.9	2.96	-4.98
O ₂	4.34x10 ⁴	31.849	784	-26.5	-8.58	-16.52	

A. Continuum Solvation GB/SA Calculation of K_{CC}

Read Section 8, Continuum Solvation, of the Introduction to Molecular Mechanics

The purpose of this section is to calculate K values from the Gibbs Free Energy of Solvation calculated using continuum solvation methods. The strength of continuum solvation techniques is the speed of the calculation.

Procedure

Use the Checklist at the end of Chapter 1 to build, minimize, and record the solvation energy for the four remaining molecules in Table 2. Use the MMFF94x forcefield with the Born solvation treatment⁴ as setup below:

The screenshot shows the 'Forcefield Parameters' dialog box in MOE. The 'Forcefield' tab is active, showing 'MMFF94x' selected. The description reads: 'Parameterized for gas phase small organic molecules in medicinal chemistry. Modified from MMFF94s to force conjugated nitrogens planar. All-atom, no Lone Pairs. Compatible with Generalized Born solvation model. Uses internal bond-charge-increment charge model.' The 'Enable' section has checkboxes for 'Bonded', 'van der Waals', 'Electrostatics', and 'Restrictions', all of which are checked. The 'Cutoff' section has 'Enable' checked. The 'Solvation' dropdown is set to 'Born'. The 'Scale Like' dropdown is set to '1'. The 'On' cutoff is 8, 'Off' is 10, 'Dielectric' is 1, 'Exterior' is 80, 'Unlike' is 0, and 'Wild' is 1. The 'Threads' dropdown is set to 2. At the bottom, there are buttons for 'Fix Hydrogens' (with a note: 'Hydrogens/LonePairs require adjustment.') and 'Fix Charges' (with a note: 'Partial charges require calculation.').

The solvation energy is found by pulling down the Compute menu and choosing Potential Energy and is listed as the “sol” contribution.

The molecules in the table below were chosen because the solvation energies are well known and they have roughly the same surface area. The MOE solvation energies do not include the surface tension term; the MOE results are just the $\Delta_{\text{sol}}G_{\text{elec}}$. Use $7.0 \text{ cal}/\text{\AA}^2$ and the solvent accessible surface area, ASA, to calculate $\Delta_{\text{sol}}G_{\text{vdw}} + \Delta_{\text{sol}}G_{\text{cav}}$. The ASA is calculated in the database portion of MOE as a QSAR descriptor; these values are listed below.⁴

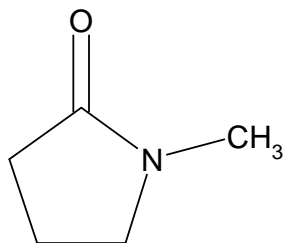
Table 2. Gibbs Free Energies of Solvation. The Free Energies of desolvation from Table 1 have been reversed to match the sign convention used in MOE.

Molecule	$\Delta_{\text{sol}}G$ (exp) kcal/mol	ASA (\AA^2)*	$\Delta_{\text{sol}}G_{\text{vdw}} + \Delta_{\text{sol}}G_{\text{cav}}$ kcal/mol ⁺	$\Delta_{\text{sol}}G_{\text{elec}}$ kcal/mol ⁺	$\Delta_{\text{sol}}G$ (GB/SA) kcal/mol
Butane	2.3	245.2	1.7164	1.1444	2.8608
1,1,1-trichloroethane	-0.208	254.9			
1,1-dichloroethane	-0.87	234.9			
benzene	-0.908	246.7	1.7269	-2.795	-1.0681
methyl- <i>tert</i> -butylether	-2.273	278.7			
acetone	-3.8	218.27	1.52789	-6.459	-4.93111
2-propanol	-4.8	224.90	1.57432	-5.0142	-3.43988
acetic acid	-6.7	201.09	1.40763	-7.762	-6.35437
acetamide	-9.7	239.25	1.67475	-9.0761	-7.40135
N-methylpyrrolidinone	-9.39	273.3			

*. Solvent accessible surface area

+. Assuming the surface tension term is $7.0 \text{ cal}/\text{\AA}^2$.

N-methylpyrrolidinone was included because the molecule is the subject of one of the lab experiments for this semester⁵. Methyl-*tert*-butylether, MTBE, is included because of the current interest in phasing out MTBE as a gasoline additive and the persistence of MTBE in groundwater.



N-methylpyrrolidinone

Problem 13.1

Make a plot of the GB/SA solvation energy versus the experimental solvation energy. Does the Generalized Born approach do a good job of reproducing the experimental trends? Does the GB/SA order reproduce the experimental order? Is the $\Delta_{\text{sol}}G_{\text{vdW}} + \Delta_{\text{sol}}G_{\text{cavity}}$ term a significant fraction (in magnitude) of the total $\Delta_{\text{sol}}G$ for these molecules? Calculate the Henry's Law constant for methyl-*tert*-butylether and compare with the experimental values.

Literature Cited:

1. G. A. Robbins, S. Wang, J. D. Stuart, *Anal. Chem.*, **1993**, *65*, 3113.
2. J.M. Gossett, *Environ. Sci. &Tech.*, **1987**, *21*, 202.
3. P. W. Atkins, *Physical Chemistry*, 5th. ed., W. H. Freeman, Co, New York, 1994. Table C18.
4. W. C. Still, A. Tempczyk, R.C. Hawley, and T. Hendrickson, "Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics," *J. Amer. Chem. Soc.*, **1990**, *112(16)*, 6127-6129.
5. T. W. Shattuck, "LogP and MR," Laboratory Manual, CH341, Physical Chemistry, Colby College, <http://www.colby.edu/chemistry/PChem/lab/LogP-MR.pdf>.

Chapter 14 Distance Geometry

Please read Section 6 “Distance Geometry” in the Introduction

The rings in progesterone can take many conformations. Building a model of progesterone with the proper ring conformations can be difficult. Applying constraints based on NMR data can greatly ease the construction of the proper conformation of the molecule. Distance geometry is an ideal technique for constructing molecules subject to constraints. The purpose of this exercise is to compare distance geometry with template based construction methods. We will use the EMBED program for distance geometry and the CORINA program for the template based approach, with access from the Web for both applications. The structure of progesterone with the conventional atom numbering is shown in Figure 1. The X-ray structure from the Cambridge Structure Database, CSD, is shown in Figure 2, to indicate the proper ring conformations.

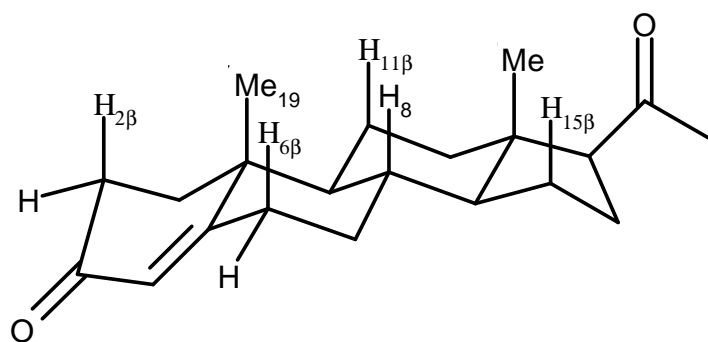


Figure 1. Progesterone

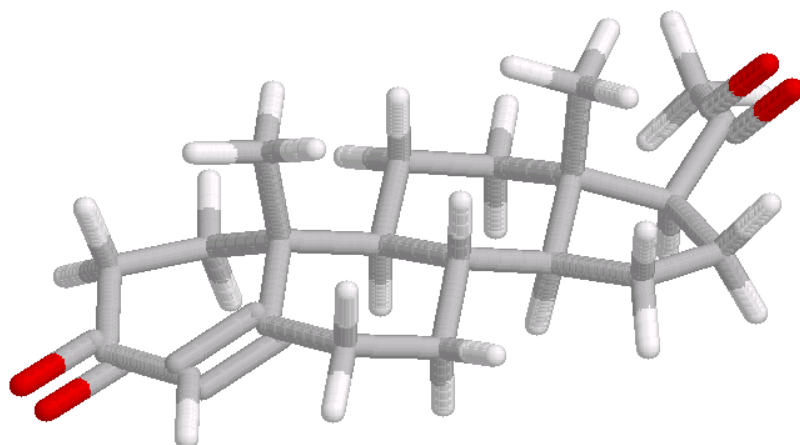


Figure 2. X-ray structure for progesterone from the CSD.

The 19-methyl of progesterone is reported to have short-range nOe couplings with four hydrogens on one face of the molecule¹, H_{2β}, H_{6β}, H_{11β}, and H₈. These nOe constraints are shown as input using the JME applet for EMBED, Figure 3a. A successful EMBED structure is shown in Figure 3b. Of course distance geometry is most useful when you don't have an X-ray structure. But for this beginning exercise comparison with the X-ray structure is very instructive.

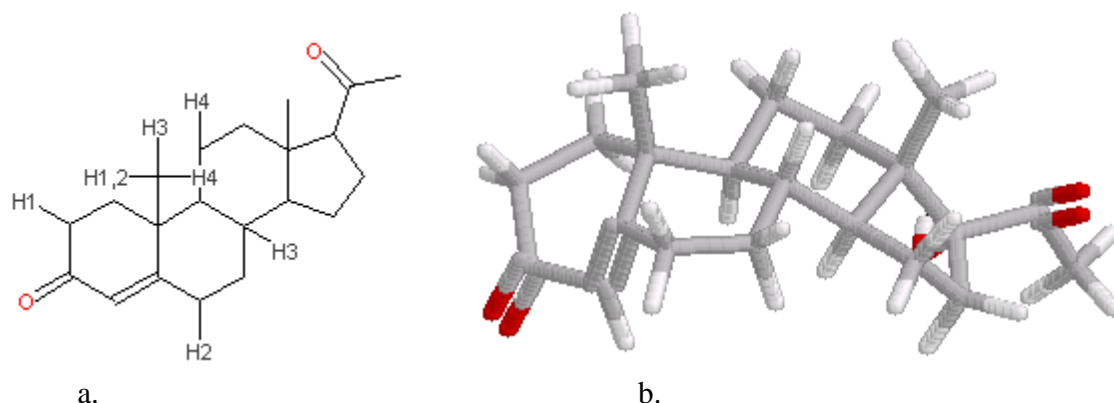



Figure 3: Progesterone with short nOe distance constraints. (a) Distance constraints as input using the JME applet. The distance ranges were specified as 2.5-3.5Å. (b) EMBED results. The Smiles notation is CC(=O)C3CCC4C2CCC1=CC(=O)CCC1(C)C2CCC34C.

Instructions

Distance Geometry: The distance geometry structure will be built using the “Distance Geometry” applet. This applet uses the Java Molecular Editor for structure input. You will find the link for this Web applet on the PChem home page.¹ Instructions for adding constraints are given below the JME editor window.

Building Progesterone:

To build progesterone, start by clicking on the cyclohexane ring button and then clicking in the molecule window to add the three fused cyclohexane rings. With the ring buttons, clicking on an existing ring bond adds a fused ring and clicking on a ring atom adds a spiro ring. If you make a mistake click on the UDO (undo) button. Click the cyclopentane ring and add it to the right-most cyclohexane. Next select the single bond button, . Drag the single bonds as shown at right. Clicking on a bond will change it to a double bond. Finish by putting in the oxygens.

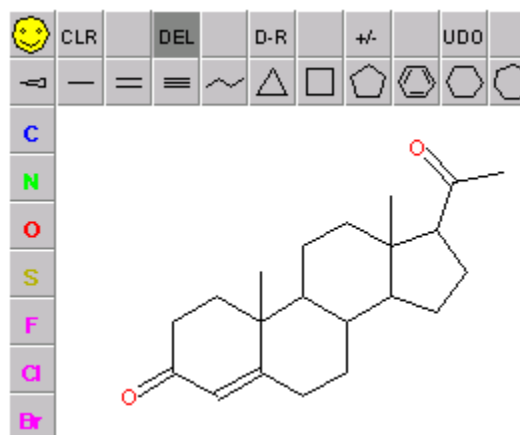





Figure 4. Progesterone in JME.

Click the Help button for more information on using the Java Molecular Editor. You would be finished at this point if you didn't have any experimental information. We now need to add the nOe based distance constraints. Use the single bond button,  to add bonds for the seven hydrogens as shown in Figure 3a. Use the  tool to set in the element names and constraint numbers. For example for H1, click on the  button, enter H1, and then click on the end of the bond at the H1 position. (For Macintosh systems make sure to highlight the entire dialog box when you change the label in the X tool, otherwise part of the old label may remain.)

When you have added all the atom labels, scroll down to the Distance Constraints section and click on the “Short” button for constraints 1-4. The typical distance range for a strong nOe will be entered automatically for you, as in Figure 5.



Figure 5. Select Short nOe distance constraints for all four distances.

Scroll down until you see the options button. The Add hydrogens button should be selected. The “SCF pi calculation” option won’t make any difference since progesterone has no conjugated double bonds. The “Set-up Gaussian 98” input button is used when the applet is run to prepare the input files for molecular orbital calculations, which we will ignore for now. Click Submit to start the calculation.

Write down the FINAL STERIC ENERGY and compare your results with Figure 2 and 3b. You will probably need to repeat the calculation several times to get the proper ring conformations. Just use the browser Back arrow and click on Submit again. Keep track of the FINAL STERIC ENERGY for each run. You should get the proper ring conformation for the rings around your distance constraints within three or five calculations. Getting the other ring conformations correctly requires a few more calculations. If you don't get all the ring conformations after you have tried 15 times, just stop with any reasonable conformation and believe us that sooner-or-later you would have gotten all the rings correctly.

In preparation for your CORINA calculations, return to the JME page. Click on the Delete button and remove the hydrogens from your structure. Your structure should now look again like Figure 4. Press the yellow “smiley face” button. A dialog box will appear with the Smiles string that corresponds to progesterone. Highlight and copy this Smiles string to the clipboard.

Template Based Builders: The CORINA program is available on-line through the University of Erlangen-Nürnberg at J. Gasteiger’s Group Web site. A link to this Web site is on the PChem Home Page. On this page scroll down to the JME applet. Paste your Smiles string into the Smiles string dialog box and click on the Generate 3D Structure button. (Alternatively you can draw in the structure using the JME editor.) Compare the CORINA generated structure to Figure 2. How many centers are inverted from the X-ray structure? Generating the structure a second time isn’t useful, since template based builders produce the same result every time. To get all the rings correctly requires the stereochemistry to be included in the Smiles string.⁴ (Don’t bother to do the calculation with the stereochemistry specified.)

MOE can also interpret Smiles strings. If you have time try inputting the progesterone Smiles string into the Builder Smiles dialog box.

Summary

The ring conformations of progesterone highlight the advantage of distance geometry. Template based model builders like CORINA give the same result every time. Distance geometry allows the various ring conformations of progesterone to be studied because each calculation has the potential to give a different result. Which is the better approach? CORINA has been highly optimized to give structures that are very close to X-ray structures. So if you can, using CORINA is a better approach. However, distance geometry is the best choice if you don’t know the proper conformation and need to generate several options, and distance geometry is especially useful if distance constraints are available from experiment.

Problem 14.1

- How many calculations were required to get the ring conformations for progesterone correctly around the nOe constraints?
- How many more calculations were necessary to get all the ring conformations correctly? (Remember you can stop at 15 tries.)
- Was the X-ray conformation (Figure 2) the lowest energy conformation?

- d. How many centers are inverted in the CORINA structure from the X-ray structure?
 e. Without looking at the structure of progesterone, use the Smiles string from the caption of Figure 3 to draw the structure of progesterone. Give your best effort, but don't cheat and compare to the Figures.

Problem 14.2 Build β -Ionone, Figure 8, without using constraints and do the MM2 calculation. Make sure the SCF pi calculation is not selected. What conformation and steric energy do you get? Now use your nOe constraints from your NOESY spectra. However, make sure to not do the MM2 or MMP2 calculation (both those check boxes should not be checked). Some of these possible nOe constraints are shown as input using the JME applet, Figure 8b.⁶ In Figure 8b, the constraints are shown to the carbon atoms since the nOe distances are average distances and the average position of a methyl group hydrogen is near the corresponding carbon atom. You can also give distance constraints to explicit hydrogens as we did above. You can also use constraints to the methyl groups attached to the top of the ring. Do several repeat calculations and compare the side chain conformations.

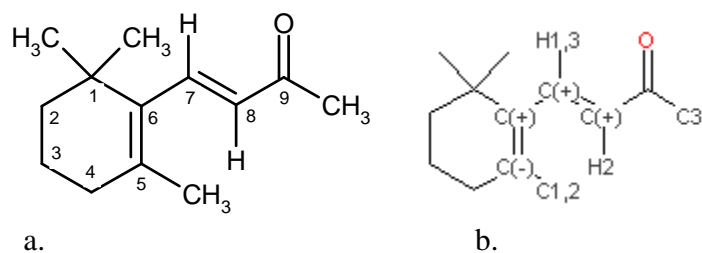


Figure 8: β -Ionone with short nOe distance constraints. (a) Atom numbering, (b) Distance constraints as input using the JME applet.

Literature Cited:

- Williams, D. H., Fleming, I., Spectroscopic Methods in Organic Chemistry, 4th Ed, McGraw-Hill, London, UK, 1987. p 118.
- Distance Geometry and MMP2: <http://iris12.colby.edu/~www/jme/dgmmp2.html> (accessed Jan 2003).
- Gasteiger, J. Research Group, *Corina*, http://www2.chemie.uni-erlangen.de/software/corina/free_struct.html (accessed Jan 2003).
- [H]C1CC2[C@]3([H])CC[C@H](C(C)=O)[C@@]3(C)CC([H])[C@]2([H])[C@@]4(C)CC([H])C(=O)C=C14
- Distance Geometry: <http://iris12.colby.edu/~www/jme/dg.html> (accessed Jan 2003).
- Honig, B.; Hudson, B.; Sykes, B. D.; Karplus, M., "Ring orientation in β -ionone and retinals," *Proc. Nat. Acad. Sci. U. S.*, **1971**, 68(6), 1289-93.

Chapter 15. Protein Structure and Gramicidin-S

One of the most active and interesting areas in biophysical chemistry is the study of protein structure. The problem is simply this: given the uncountable number of possible conformations for a protein, how can we determine the lowest energy structure? In this exercise we tackle a relatively simple problem, which retains the flavor of the more complicated problems under current study. We will model the structure of the antibiotic gramicidin-S.

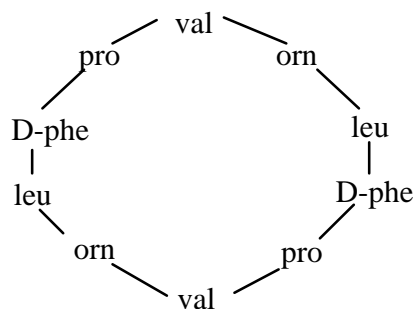


Figure 1 Gramicidin-S

Gramicidin-S is a cyclic decapeptide, Figure 1, produced by the soil fungus *Bacillus Brevis*. The protein is unusual for several reasons. First it includes D-phenylalanine, rather than the normal L-isomer. Secondly, the unusual amino acid ornithine is used. Thirdly, the protein is hydrophobic. Most proteins have a hydrophilic exterior, to enhance interaction with water, and a hydrophobic interior. Gramicidin-S has just the opposite. Its hydrophobic exterior suggests that the mode of action is through a strong membrane interaction¹. The linear gramicidins form ion channels in cell membranes.

Even a small peptide like gramicidin-S has too many possible conformations for each conformation to be exhaustively studied. To find the global minimum structure, we must rely on experimental information and some intuition. The NMR spectrum shows that gramicidin-S is symmetrical; the like-amino acid pairs have the same chemical shifts. Therefore, we really only need to worry about five amino acids, the other five are related by symmetry. In our modeling we must make sure that this symmetry is maintained. In lab you will be determining NMR constraints on the dihedral angles for some of the amino acids. The spin-spin J coupling between the α -CH and the backbone NH proton is about 4 Hz for alpha-helix type structures and 9Hz for beta-pleated sheet structures². The presence of alpha-helical or beta-pleated sheet type-regions will help to constrain our modeling. Of course just a few monomers with the proper dihedral angles aren't sufficient to establish a "real" alpha helix or beta-pleated sheet, but the NMR dihedral constraints can be used to point us in the right direction for molecular modeling.

We can also use some intuition. Prolines have a cyclic structure that is formed by the side chain and the backbone N, Figure 2a. Prolines often occur at turns, because of the kink caused by the cyclic structure, and proline can't assume the backbone dihedral angles necessary for alpha helix or beta-pleated sheet structures³.

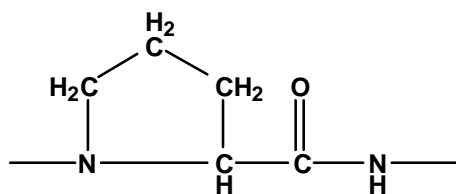
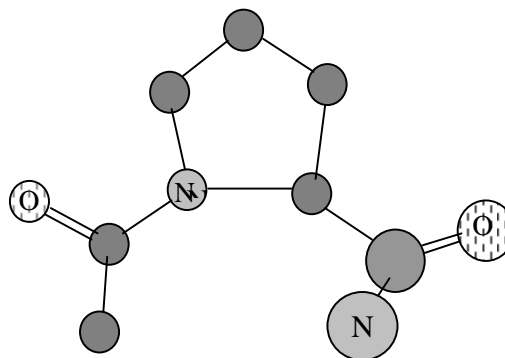


Figure 2 a. Proline in a protein



b. Proline in a proline-I structure based on *cis*-peptide linkages. (H's not shown.)

MOE has two types of proline based secondary structures, which can be used to establish proline turns. The proline-I secondary structure is based on the unusual formation of *cis*- peptide bonds, Figure 2b. In MOE you choose “poly pro” for the secondary structure and highlight the “cis” conformation button. On the other hand, proline-II turns are represented in MOE by “poly pro” dihedrals and the normal trans peptide conformation. Therefore, you first must decide whether to use “cis” proline or “trans” type turns. Since complete turns require two amino acids, only three residues remain in the “body” of the protein for us to worry about.

The average backbone angles for some regular secondary structures are shown in Table 1.

Table 1. Dihedral Angles for Regular Secondary Protein Structures³

	Bond Angle (degrees)		
	ϕ	ψ	ω
Antiparallel β -sheet	-139	+135	-178
Right-handed α -helix	-57	-47	180
Polyproline I	-83	+158	0
Polyproline II	-78	+149	180

Applying Constraints

Experimental constraints are necessary to help us narrow down the number of possible conformations for proteins. Constraints may be applied on dihedral angles or on distances. For example, an alpha-helix ϕ dihedral should be about -60° and a beta-pleated sheet ϕ dihedral should be about -140° . As mentioned above, we can measure these dihedral angles using spin-spin J coupling constants and then use these measured values as constraints.

Distance constraints can be determined from nOe measurements. The normal value for nOe based constraints is 3.0\AA . Distance constraints can also be inferred from secondary structure assignments. Examples for such inferences include N-H \cdots O=C hydrogen bond distances. For alpha-helices, strong hydrogen bonds form between residue *i* and residue *i*+4. For beta-pleated sheets, N-H \cdots O=C distances between strands can be constrained. Hydrogen bond lengths are in the range from $1.8\text{-}3.0\text{\AA}$, with 2\AA being normal for strong hydrogen bonds⁴. The hydrogen bond distance between residues *i* and *i*+4 in the alpha-helix is about 1.86\AA ⁴, and about 1.96\AA between beta-pleated sheet strands. Comparing dihedral and distance constraints, distance constraints limit the conformational flexibility of the molecule more, and are preferable if known.

Instructions

Using the Protein Builder to set the Sequence and Secondary Structure

In this section, you will find out how to specify the sequence for your protein. You will also find out how to change L-amino acids to D- amino acids and how to make cyclic proteins. You will also specify secondary structures like alpha-helical or beta-pleated sheet regions and specify turns. This exercise uses the program MOE. In MOE:

1. Pull down the Edit menu, and choose "Protein Builder."
2. In the Protein Builder, you select the secondary structure for the next amino acid and then click on the amino acids in your structure in the order that they appear in the protein. The builder doesn't have ornithine, so we will use lysine instead. At the end of the exercise use can use the Builder dialog to remove a CH₂ from each lysine if you like. Choose either helix (alpha helix) or anti-strand (beta-pleated sheet) secondary structure, as you have determined using your NMR spectra and then click on LYS. Continue around the ring until you come to PRO, Figure 1.

Before clicking on PRO, select the “poly pro” secondary structure, click on PRO, and then switch back to helix or anti-strand as you used before. (If you want proline-I turns also click the “cis” check box before adding the proline residue.) Continue around the ring until you get to PRO again, switch back to poly pro for PRO and then finish up in helix or anti-strand. Click on VIEW and then CLOSE.

3. Next we need to convert phenylalanine to the D-stereo isomer. Click on the Builder button on the right-hand side of the screen. Click on the C α carbon of phenylalanine. At the right-hand side of the Builder dialog you will see the Chirality section with the S button in blue, indicating the stereochemistry of the selected atom. Click on Invert to switch to the opposite chirality. You should see that the benzyl group switches positions with the C α -H. Repeat this process for the other phenylalanine.

4. We next need to form the ring peptide bond. Locate the carboxyl terminus and the singly bonded O. Click on this atom and then click on Delete at the bottom of the Builder window. Now click on the carbonyl carbon (it should turn pink). Next shift-click on the N atom in the terminal ammonium group at the other end of the chain. Both the N and the C should be highlighted. Now click on the Bond button at the bottom of the Builder window. A very long bond should now appear on the screen.

5. We now need to change the charge state on the N in the new peptide bond. You should note that this N atom still has two H atoms attached. Click on the black background of the MOE window to clear all selections. Now click on the N atom in the new peptide bond. This N atom should now be pink. In the upper-right corner of the Builder window is the atom charge section. The +1 button should be blue, indicating the charge on the selected atom is +1. Click on the adjacent “0” button. The N atom should now only have one H attached.

6. Use the checklist at the end of Chapter 2 to set up and run a minimization; however, use the CHARMM2.2 force field for this and all subsequent minimizations. The CHARMM force field was specifically (at least originally) designed for proteins, and works well for this exercise.

7. Observe the minimized structure. It should be fairly symmetrical. For example, the lysine side chains should both be on the same side of the molecule. The leucine side chains should both be on the opposite face from the lysine side chains. Also check all the peptide bonds. They should all be trans. Some may have switched because of the drastic conformational changes necessary to minimize the ring structure. If the molecule doesn't have both lysines on the same side of the ring, you should try building your molecule from scratch again. On a second try, use different stretches of extended backbone or helix or anti-strand secondary structures. Another way to get a good starting structure is to build the linear molecule as you did before, but adjust some of the dihedral angles in the backbone using the mouse before forming the ring bond. To change dihedral angles using the mouse, shift and then shift-click on two atoms in the bond you want to rotate and then press Alt and drag with the left mouse button.

Setting Dihedral Constraints

8. Pull down the Edit menu, slide right on Potential, and choose "Restrain...." A new dialog bar will appear at the top of the MOE model window. Use the pull down menu to select “dihedral” instead of distance. Click on the four consecutive atoms that define the dihedral you wish to constrain. You don't need to shift click in this mode. You need to set the allowable Target ranges for your dihedral. For the peptide bonds, if some are found to be cis instead of trans, use a narrow range of 178-182 degrees. The range for any ϕ or ψ constraints should be fairly broad ($\pm 15^\circ$) for this exercise, since we expect some strain in the ring to distort the angles from their ideal values. After you have entered the Target range, click on Create to enter the new constraint. Remember

to set the same constraints for both amino acids in the pair, since the structure should be symmetrical.

9. Repeat step 8 for each dihedral you wish to constrain.

10. Press the ESC key until the Restraints dialog boxes disappear and the white command line dialog box is displayed.

Setting Distance Constraints

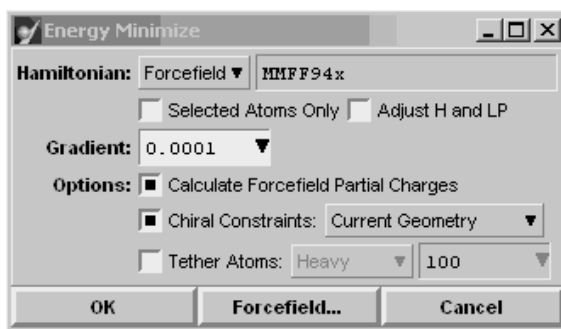
11. Pull down the Edit menu, slide right on Potential, and choose "Restrains..." Use the pull down menu to select "distance" if it is not already displayed. Click on the two atoms that define the distance you wish to constrain. You don't need to shift click in this mode. You need to set the allowable Target ranges for your distance. For hydrogen bonds try a range of 1.8Å to 2.5Å. After you have entered the Target range, click on Create to enter the new constraint. Remember to set the same constraints for both amino acids in the pair, since the structure should be symmetrical.

12. Repeat step 11 for each distance you wish to constrain.

13. Press the ESC key until the Restraints dialog boxes disappear and the white command line dialog box is displayed.

Maintaining the Proper Chirality

14. If during minimization, the chirality of your amino acids shifts, you can constrain the chirality during minimization when you pull down the Compute menu and choose Energy Minimize...



Minimization with Constraints

15. Any active constraints will be used during minimization or dynamics runs. To see the active constraints, pull down the Window menu and choose "Potential Setup." Click on the Restraints tab in the new window to see your restraints. You can edit the target values from this window also. If the restraints that you set aren't adhered to, you may need to increase the Weight of the constraint. Increased weights are often necessary for dihedrals. This window is also used to delete restraints; just click on the line for the restraint you wish to delete so that it is highlighted and then click Delete.

If minimization doesn't change your structure, you need to move some atoms using the mouse to move away from an energy minimum structure and then reminimize. To change dihedrals select two atoms and press Alt- and drag with the left mouse button. To translate a group of atoms, highlight the atoms you wish to move (shift and shift click) and then press the Shift and Alt- keys and drag with the middle mouse button. If your structure looks very strained, you might try a dynamics run to try to find a lower energy minimum. For dynamics use a high temperature, 600-900°, and at least 12000 heating and equilibration steps, otherwise you won't sample other possible conformations sufficiently (you don't need to use a Cool period).

15. Check to see that your final structure is symmetrical. If it is not, apply additional constraints, or adjust some of the dihedrals using the mouse. For example, you might need to use the mouse to rotate some side chains so they are symmetrical. Reminimize.
16. The constraints that you have applied are an artificial term in the potential energy function. For your **final minimization you should remove all constraints** (see step 14). Reminimize. Repeat minimization until the energy is minimized. This minimization may take 6000 steps or more. If the ring pops open, then you need to go back to using constraints.
17. Remember to save your data file and report your computer and data file name in your report.
18. Plot out several views of your structure. Measure some of the backbone dihedrals to see if the ϕ angles are close to the NMR determined values.

References

1. M. Waki, N. Izumiya, in *CRC Crit. Rev. Biotechnol.* , **1988**, 8, 206.
2. K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY, 1986, pp 166-7.
3. T. E. Creighton, *Proteins: Structures and Molecular Properties*, W. H. Freeman, New York, NY, 1983. pp171-5.
4. J. S. Richardson, D. C. Richardson, Principles and Patterns of Protein Conformation, in *Prediction of Protein Structure and the Principles of Protein Conformation*, G. D. Fasman, ed., Plenum Press, New York, NY, 1989. p9-10.