Colby College Molecular Mechanics Tutorial

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

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.
Introduction to Molecular Mechanics

Summary

The goal of molecular mechanics is to predict the detailed structure and physical properties of molecules. Examples of physical properties that can be calculated include enthalpies of formation, entropies, dipole moments, and strain energies. Molecular mechanics calculates the energy of a molecule and then adjusts the energy through changes in bond lengths and angles to obtain the minimum energy structure.

Steric Energy

A molecule can possess different kinds of energy such as bond and thermal energy. Molecular mechanics calculates the steric energy of a molecule—the energy due to the geometry or conformation of a molecule. Energy is minimized in nature, and the conformation of a molecule that is favored is the lowest energy conformation. Knowledge of the conformation of a molecule is important because the structure of a molecule often has a great effect on its reactivity. The effect of structure on reactivity is important for large molecules like proteins. Studies of the conformation of proteins are difficult and therefore interesting, because their size makes many different conformations possible.

Molecular mechanics assumes the steric energy of a molecule to arise from a few, specific interactions within a molecule. These interactions include the stretching or compressing of bonds beyond their equilibrium lengths and angles, the Van der Waals attractions or repulsions of atoms that come close together, the electrostatic interactions between partial charges in a molecule due to polar bonds, and torsional effects of twisting about single bonds. To quantify the contribution of each, these interactions can be modeled by a potential function that gives the energy of the interaction as a function of distance, angle, or charge. The total steric energy of a molecule can be written as a sum of the energies of the interactions:

\[ E_{\text{steric energy}} = E_{\text{str}} + E_{\text{bend}} + E_{\text{str-bend}} + E_{\text{VdW}} + E_{\text{tor}} + E_{\text{qq}} \]  

The bond stretching, bending, and stretch-bend interactions are called bonded interactions because the atoms involved must be directly bonded or bonded to a common atom. The Van der Waals, torsional, and electrostatic (qq) interactions are between non-bonded atoms.

Bonded Interactions

\( E_{\text{str}} \) represents the energy required to stretch or compress a bond between two atoms, Figure 1.

![Figure 1. Bond Stretching](image-url)
A bond can be thought of as a spring having its own equilibrium length, $r_o$, and the energy required to stretch or compress it can be approximated by the Hookian potential for an ideal spring:

$$E_{str} = \frac{1}{2} k_{s,ij} (r_{ij} - r_o)^2$$

(2)

where $k_{s,ij}$ is the stretching force constant for the bond and $r_{ij}$ is the distance between the two atoms.

$E_{bend}$ is the energy required to bend a bond from its equilibrium angle, $\theta_o$. Again this system can be modeled by a spring, and the energy is given by the Hookian potential with respect to angle:

$$E_{bend} = \frac{1}{2} k_{b,ij} (\theta_{ij} - \theta_o)^2$$

(3)

where $k_{b,ij}$ is the bending force constant and $\theta_{ij}$ is the instantaneous bond angle (Figure 2).

Figure 2. Bond Bending

$E_{str-bend}$ is the stretch-bend interaction energy that takes into account the observation that when a bond is bent, the two associated bond lengths increase (Figure 3). The potential function that can model this interaction is:

$$E_{str-bend} = \frac{1}{2} k_{sb,ijk} (r_{ij} - r_o) (\theta_{ik} - \theta_o)$$

(4)

where $k_{sb,ijk}$ is the stretch-bend force constant for the bond between atoms $i$ and $j$ with the bend between atoms $i$, $j$, and $k$.

Figure 3. Stretch-Bend Interaction

Therefore, when intramolecular interactions stretch, compress, or bend a bond from its equilibrium length and angle, it resists these changes with an energy given by the above equations. When the bonds cannot relax back to their equilibrium positions, this energy raises the steric energy of the entire molecule.
**Non-bonded Interactions**

*Van der Waals interactions*, which are responsible for the liquefaction of non-polar gases like O₂ and N₂, also govern the energy of interaction of non-bonded atoms within a molecule. These interactions contribute to the steric interactions in molecules and are often the most important factors in determining the overall molecular conformation (shape). Such interactions are extremely important in determining the three-dimensional structure of many biomolecules, especially proteins.

A plot of the Van der Waals energy as a function of distance between two hydrogen atoms is shown in Figure 4. When two atoms are far apart, an attraction is felt. When two atoms are very close together, a strong repulsion is present. Although both attractive and repulsive forces exist, the repulsions are often the most important for determining the shapes of molecules. A measure of the size of an atom is its Van der Waals radius. The distance that gives the lowest, most favorable energy of interaction between two atoms is the sum of their Van der Waals radii. The lowest point on the curve in Figure 4 is this point. Interactions of two nuclei separated by more than the minimum energy distance are governed by the attractive forces between the atoms. At distances smaller than the minimum energy distance, repulsions dominate the interaction. The formula for the Van der Waals energy is:

\[
E_{VdW} = -\frac{A}{r_{ij}^6} + \frac{B}{r_{ij}^{12}}
\]

where A and B are constants dependent upon the identities of the two atoms involved and \( r_{ij} \) is the distance, in Angstroms, separating the two nuclei. This equation is also called the Lennard-Jones potential. Since, by definition, lower energy is more favorable, the \(-A/r^6\) part is the attractive part and the \(+B/r^{12}\) part is the repulsive part of the interaction. For two hydrogen atoms in a molecule:

\[
A = 70.38 \text{ kcal } \AA^6 \\
B = 6286. \text{ kcal } \AA^{12}
\]

**Figure 4:** Van der Waals interactions between two hydrogen atoms in a molecule, such as H₂O₂ or CH₃CH₃

*Torsional Interactions:* \( E_{\text{tor}} \) is the energy of torsion needed to rotate about bonds. Torsional energies are usually important only for single bonds because double and triple bonds are too rigid to permit rotation. Torsional interactions are modeled by the potential:

\[
E_{\text{tor}} = \frac{1}{2} k_{\text{tor},1} (1 - \cos \phi) + \frac{1}{2} k_{\text{tor},2} (1 - \cos 2 \phi) + \frac{1}{2} k_{\text{tor},3} (1 - \cos 3 \phi)
\]
The angle $\phi$ is the dihedral angle about the bond. The term in $3\phi$ is important for sp$^3$ hybridized systems (Figure 5a and b) and the term in just $\phi$ is useful for the central bond in molecules such as butane that have C-C-C-C frameworks (Figure 5c). The constants $k_{\text{tor},1}$, $k_{\text{tor},2}$, and $k_{\text{tor},3}$ are the torsional constants for one-fold, two-fold and three-fold rotational barriers, respectively.

![Dihedral Angle vs. Dihedral Energy](image)

**Figure 5.** Torsional Interactions, (a) dihedral angle in sp$^3$ systems. (b) three-fold, $3\phi$, rotational energy barrier in ethane. (c) butane, which also has a contribution of a one fold, $\phi$, barrier.

The origin of the torsional interaction is not well understood. Torsional energies are rationalized by some authors as a repulsion between the bonds of groups attached to a central, rotating bond (i.e., C-C-C-C frameworks). Torsional terms were originally used as a fudge factor to correct for the other energy terms when they did not accurately predict steric energies for bond twisting. For example, the interactions of the methyl groups and hydrogens on the "front" and "back" carbons in butane were thought to be Van der Waals in nature (Figure 5). However, the Van der Waals function alone gives an inaccurate value for the steric energy.

**Electrostatic Interactions:** If bonds in the molecule are polar, partial electrostatic charges will reside on the atoms. The electrostatic interactions are represented with a Coulombic potential function:

$$E_{qq,ij} = \frac{k Q_i Q_j}{4\pi \varepsilon r_{ij}}$$

(7)

The $Q_i$ and $Q_j$ are the partial atomic charges for atoms i and j separated by a distance $r_{ij}$. $\varepsilon$ is the molecular dielectric constant (normally 1.5), which approximates the dielectric effect of intervening solute or solvent atoms. $k$ is a units conversion constant. Like charges will raise the steric energy, while opposite charges will lower the energy. (The Del Re method is often used for estimating partial charges.)

The bond stretching, bond bending, stretch-bend, Van der Waals, torsion, and electrostatic interactions are said to make up a force field. Each interaction causes a steric force that the molecule must adjust to.

**Empirical Force Fields**

All the potential functions above involve some force constant or interaction constant. Theoretically, these constants should be available from quantum mechanical calculations. In practice, however, it is necessary to derive them empirically. That is, the constants are adjusted so that the detailed geometry is properly predicted for a number of well known compounds. These constants are then used to calculate the structures of new compounds. The accuracy of these constants is critical to molecular mechanics calculations. Unfortunately, no single best set of force constants is available because of the diversity of types of compounds. For example, the MM2 force field works best on
hydrocarbons because most of the known compounds used in deriving the force field were hydrocarbons. MM2 is less accurate for oxygen-containing compounds and even less reliable for nitrogen and sulfur species. This is because there aren’t as many hetero-atom containing compounds in the data base for MM2 and hydrocarbons are a more homogeneous class of compounds than substances with hetero-atoms. However, the MM2 force field is one of the best available and the most widely accepted force field for use with organic compounds. MM2 was specifically parameterized to reproduce experimental enthalpies of formation.

It is important to realize that the force field is not absolute, in that not all the interactions listed in Equation 1 may be necessary to accurately predict the steric energy of a molecule. On the other hand, many force fields use additional terms. For example, MM2 adds terms to the bonded interactions to better approximate the real potential function of a chemical bond. These additional terms take into account anharmonicity, which is a result of the fact that given enough vibrational energy, bonds will break. Purely quadratic potentials have steep "walls" that prevent bond dissociation (Figure 6a). Cubic terms are added to Equation 2 to adjust for this:

$$E_{str} = 1/2 k_{s,ij} (r_{ij} - r_0)^2 + 1/2 k_{s,ij} \chi_{ij} (r_{ij} - r_0)^3$$

where $\chi_{ij}$ is the anharmonicity constant. For example, for a C(sp$^3$)-C(sp$^3$) bond the anharmonicity is -2, see Figure 6b:

$$E_{str} = [4.40 \text{ mdynes/Å}] (r - 1.532 Å)^2 + [4.40 \text{ mdynes/Å}] [-2.00] (r - 1.532 Å)^3$$

The addition of the cubic term makes the small $r$ portion steeper or more repulsive. This is realistic for real bonds. At larger $r$ the curve is less steep, as desired. For $r$ very large ($r > 3$Å) the energy decreases, which is unphysical; the curve should approach a constant value. Even though the large $r$ behavior is incorrect, the bond length in compounds remains less than this value, so this region is unimportant under normal conditions.

The above potential functions represent the various non-bonded interactions that can occur between two atoms $i$ and $j$. A full force field determines the steric energy by summing these potentials over all pairs of atoms in the molecule.
**Enthalpy of Formation**

The steric energy of a molecule can be used to calculate the enthalpy of formation. First, a bond energy calculation is done using standard tabular values. It is customary to use bond increments rather than the bond energy calculations that you did in General Chemistry. However, the principle is the same. Thermal energy terms must then be added to account for the energy of translation and rotation of the molecule. The energy of translation (x, y, z motion of the center of mass of the molecule) is 3/2RT. The rotational energy of a non-linear molecule is also 3/2RT (1/2RT for each rotational axis).

The steric energy calculation in molecular mechanics corresponds to an internal energy calculation. Since $\Delta H=\Delta U+\Delta(PV)$, and $PV=nRT$ for an ideal gas, we must also add RT to convert from internal energy to enthalpy.

We have not yet considered molecular vibrations, especially internal rotations. In principle, every vibration, including internal rotations, contributes to the enthalpy. However, the contribution of vibrations is difficult to calculate. In practice the contributions are often small so they can be ignored. However, the internal rotation of the methyl group is always included; in fact the effect is included in the bond increment calculation. Extra terms must also be added for non-methyl free internal rotations. This contribution, which is called the torsional increment, is estimated as 0.36 kcal/mol or 1.51 kJ mol\(^{-1}\) for each internal rotation. For example, butane, CH\(_3\)-CH\(_2\)-CH\(_2\)-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. In summary the enthalpy of formation is then,

$$\Delta_f H^\circ = \frac{3}{2}RT + \frac{3}{2}RT + RT + \text{bond energy} + \text{steric energy} + \text{torsional increments} \quad (10)$$

This formula also assumes that there is only one low energy conformation of the molecule. If there are several low energy conformations, each must be accounted for in Equation 10.

**Bond Energy**

You are familiar with bond energy calculations from General Chemistry. The energy of a molecule is assumed to be an additive function of the energy of individual bonds (Table I). The $\Delta_r H$ for a reaction is given from $\Delta H^\circ(\text{bonds broken}) - \Delta H^\circ(\text{bonds formed})$.

<table>
<thead>
<tr>
<th>Bond</th>
<th>$\Delta H^\circ(\text{A-B})$ (kJ/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>436</td>
</tr>
<tr>
<td>C</td>
<td>412</td>
</tr>
<tr>
<td>O</td>
<td>463</td>
</tr>
</tbody>
</table>

C (graph) -> C (g) $\Delta H^\circ = 716.7$ kJ/mol

For example, the enthalpy of formation of acetaldehyde is calculated as:

$$2 \text{ C(graph)} + 2 \text{ H}_2 (g) + 1/2 \text{ O}_2 (g) \rightarrow \text{CH}_3\text{-CH}=\text{O} (g)$$

<table>
<thead>
<tr>
<th># Bonds Broken</th>
<th># Bonds Formed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 C (graph)</td>
<td>1 C=O</td>
</tr>
<tr>
<td>2 H-H</td>
<td>4 C-H</td>
</tr>
<tr>
<td>1/2 O=O</td>
<td>1 C-C</td>
</tr>
</tbody>
</table>

$\Delta H^\circ = 2553.9$ kJ/mol $- 2739$ kJ/mol = $-185.1$ kJ

The experimental value is $-166.19$ kJ, so the value derived from Table I is not very accurate.
The bond energy calculations in molecular mechanics are done slightly differently, using bond increments. Again the bond energies are assumed to be additive. The contributions are taken not only from each bond, but increments are added for certain structures, such as tertiary carbon linkages. The bond energy calculation for acetaldehyde from the MMX program is given below, with energies in kcal. MMX also calculates entropies, which are also listed for your interest.

<table>
<thead>
<tr>
<th>#</th>
<th>Bond or Structure</th>
<th>Each</th>
<th>Total</th>
<th>Tot S contrib.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>C-H ALIPHATIC</td>
<td>-3.205</td>
<td>-9.615</td>
<td>38.700</td>
</tr>
<tr>
<td>1</td>
<td>C=O</td>
<td>-25.00</td>
<td>-25.00</td>
<td>-2.300</td>
</tr>
<tr>
<td>1</td>
<td>C-H ALDEHYDE</td>
<td>-2.500</td>
<td>-2.500</td>
<td>26.800</td>
</tr>
<tr>
<td>1</td>
<td>C-C SP3-SP2 C=O</td>
<td>-3.000</td>
<td>-3.000</td>
<td>-0.600</td>
</tr>
<tr>
<td>1</td>
<td>ME-CARBONYL</td>
<td>-2.000</td>
<td>-2.000</td>
<td></td>
</tr>
</tbody>
</table>

bond energy = -42.115 kcal  \[S^\circ = 62.600 \text{ cal/K}\]

The bond energy is -42.115 kcal or -176.2 kJ. However, caution should be used since these calculations are designed to be used in conjunction with steric energies in a molecular mechanics calculation and not as general bond energy values. Using Equation 10, with the steric energy calculated by molecular mechanics gives the final \(\Delta f H^\circ = -169.33 \text{ kJ/mol}\), which is a significant improvement over the bond energy calculation from Table I of -185.1 kJ.
Introduction

QUANTA is a molecular modeling program, which is specifically designed to handle large biological molecules. CHARMm is a molecular mechanics and dynamics package that QUANTA uses for its mechanics and dynamics calculations. QUANTA can also be used to setup input files for MM2 molecular mechanics and MOPAC molecular orbital 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 QUANTA manuals. The QUANTA manuals have many interesting examples that extend well beyond the skills taught here.

The user interface is very similar to the Macintosh, with several notable exceptions. First there are three mouse buttons. Usually only the left one is used. In this manual, use the left button unless instructed otherwise. In order to type input to a dialog box, the mouse cursor must be in the same window. This is because, unlike the Macintosh, many windows can be active at the same time. Never use the "go-away" button in the upper left-hand corner of a palette, always use the Quit or Exit entry on the palette menu. To move a covered window into the foreground, click anywhere on the frame of the hidden window.

QUANTA 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 QUANTA on your own.

Chapter 1. Building and Minimizing.

The following will illustrate a few of the options available for structural input, minimization and display using QUANTA. We will begin with axial-methyl cyclohexane, Figure 1.1. We will use the ChemNote Application, where structures may be drawn on the screen in essentially the same way as they would in ChemDraw. The minimum energy configuration will then be calculated using CHARMm.

*Figure 1.1. Axial-methylcyclohexane*

**ChemNote Model Building** Pull down the Applications menu, slide right on "Builders" and choose "2-D sketcher." to start ChemNote. Click on the cyclohexane ring in the middle of the bonds palette, move the cursor to the middle of the sketch pad and click again. A green cyclohexane ring should appear. We now want to add the axial methyl group. Click on the icon for a single bond coming out of the plane: the solid triangle in the bonds palette. Move the cursor to the right most carbon on the cyclohexane. Hold the mouse button down and drag the bond away from the ring to the right. Don't worry about the hydrogens, ChemNote will add those automatically. To finish adding bonds, double click on the selection tool icon, ↑, in the Edit Icons palette in the upper left hand portion of the window.

To save this molecule, pull down the File menu and choose 'Return to Molecular Modeling.' Click Yes to the question 'Save Changes First.' The file librarian dialog box will appear; first choose the small_molecules/ directory with a single click, second type in the name
Rember-- don't use punctuation or spaces in file names. Click on 'Save.' Next the charge of the molecule is calculated using standard values for each atom type. This charge will be -0.180 for methyl cyclohexane, but we wish to have a neutral molecule. We need to smooth the charge over the atoms to yield a net charge of 0.000. Choose 'CT, CH1E, CH2E, CH3E, C5R, C6R, C5RE, C6RE, HA type', and then click OK. This choice is for all carbon atoms and non-polar hydrogens. All the carbons in our molecule are type CT, which stands for tetrahedral carbon.

Upon returning to QUANTA, hydrogen atoms are added automatically and the 3D structure is constructed using tabulated values of bond lengths and angles. The program then displays: Which molecule do you want to use? Choose the 'Use the new molecule only' option and click OK. Verify that you have constructed the axial isomer by reorienting the molecule on the screen using the following instructions.

Rotations, Translations, and Scaling  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 Dial palette. To use the mouse, position the cursor in the main window and hold down the center mouse button. Moving the mouse reorients the molecule. If you wish to rotate the molecule only around the axis perpendicular to the screen, hold down the right mouse button and move the cursor left and right. Alternatively, you can use the Dial palette. The Dial palette is in the lower right corner of the screen. If the Dials are not visible, pull down the Views menu, slide right on "Show Windows" and choose "Palettes." Clicking on the dial controls causes the listed action. If you hold the mouse button down, the action occurs smoothly, with the rate depending on the horizontal distance from the center of the control. Clicking on 'Reset' will allow you to start fresh with a centered molecule. The Dials allow you to rotate, translate, and scale (enlarge) the molecule. The Clipping control adjusts the size of the box in the z direction where atoms will be displayed. Atoms outside this box, both front and back, will not be displayed. Decreasing the Clipping control allows you to "drive" inside the molecule.

CHARMm Minimization  The structure made by ChemNote will not be the lowest energy structure. To prepare to minimize the structure, pull down the CHARMM menu and choose 'Minimization Options.' For small molecules that are close to the minimum geometry, choose the Conjugate Gradient Method. For rough starting geometries Steepest Decent is faster, less likely to fail, but less accurate. Use Adopted-Basis Newton Raphson for large molecules like proteins. Choose the Conjugate Gradient Method. Enter the default values for the parameters given below, if they are not already shown:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Minimization Steps</td>
<td>50</td>
</tr>
<tr>
<td>Coordinate Update Frequency</td>
<td>5</td>
</tr>
<tr>
<td>Energy Gradient Tolerance</td>
<td>0.0001</td>
</tr>
<tr>
<td>Energy Value Tolerance</td>
<td>0</td>
</tr>
<tr>
<td>Initial Step Size</td>
<td>0.02</td>
</tr>
<tr>
<td>Step Value Tolerance</td>
<td>0</td>
</tr>
</tbody>
</table>

Pull down the CHARMM menu, slide right on "CHARMM mode" and choose 'PSF's'. (The RTF options are for biopolymers.) You need only set the Minimization options and RTF options once. These choices will be used for all subsequent modeling, until changed. To actually do the minimization, choose 'CHARMm Minimization' in the Modeling palette. The calculation will stop after 50 steps, but the energy won't necessarily be minimized. Click on 'CHARMm Minimization' repeatedly until the energy listed in the main window no longer changes. You can monitor the progress of all CHARMM calculations in the TextPort at the bottom of the screen. The final result should be 6.3047 kcal/mol. Select 'Save Changes' in the Modeling palette to save your minimized structure to a file. The program asks you to 'Choose the MSF Saving Option.' A MSF is a molecular structure file, which QUANTA uses as the principal means of saving 3D information. Choose "Overwrite amecyc.msf" and click OK.
We often need to find the contribution to the total energy for each degree of freedom, i.e. bond stretching, bond angle bending, dihedral torsions, Lennard-Jones-Van der Waals energies, and electrostatic interactions. To find these contributions, select 'CHARMm Energy' in the Modeling palette. The results are listed in the TextPort by the keywords underlined above. Use the scroll bar to scan the results. The conformation of the molecule remains unchanged during this calculation. The "Improper" torsions entry is an additional term in the force field to get the proper conformation for small rings.

**DISPLAY OPTIONS**

Is there a 'hole' in the middle of a cyclohexane ring? We will construct solid models using literature values for atomic radii to answer this question. Van der Waals radii are usually used for this purpose, and refer to average covalent interactions. There are a wide variety of solid modeling options to use. Try them all. But, to get you started: reorient the molecule so that the methyl group points away from you as in Figure 1.2.

![Axial methyl cyclohexane](image)

**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.

Pull down the Draw menu, slide right on 'Solid Models' and choose "Van der Waal's". 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? When the solid model is displayed, a new window appears in the lower right portion of the screen. Click on the "No" box underneath the "Delete" column to remove the object. You should also try other options including "Ball and Stick" models in "Solid Models" and several of the "Raster Models." "Ray Trace" gives the best quality, but rotations aren't possible in ray trace mode. Click a mouse button to exit "Raster Models" or "Ray Trace" mode.

Another method of surface rendering that is especially popular for biomolecules is a dot surface. Pull down the Draw menu and choose 'Dot Surfaces.' Dot surfaces provide the fastest reorientation. Choose "Big Dots" and otherwise use the default settings, then click "OK." A Selection Palette is then displayed so that you can specify which atom's surfaces are to be modeled with dots. Make sure "Include" is highlighted, click on "All Atoms," and then click "Finish." To remove the dot surface pull down the Draw menu and choose "Dot Surfaces" again, select "Delete Dot Surface" , and click OK.

**Problem 1.1:** Using the printout in the TextPort after using the 'CHARMm Energy' option, decide which term in the force field dominates the steric energy of axial methyl cyclohexane. (e.g. bond stretch, bond bending, etc.) Compare absolute values for each term.
**Problem 1.2:** Start fresh in the ChemNote application, build axial methyl cyclohexane again (or Open your old ChemNote file) but this time minimize the structure using the Steepest Descents method. What happens? How do your results compare between Steepest Descents and the Conjugate Gradient minimization you did before? With rough beginning geometries, it is often best to minimize first with Steepest Descents, and then switch to Conjugate Gradient and minimize again. This two step process gives the best of both techniques. After you are finished with this problem, remember to switch back to Conjugate Gradient in the CHARMm "Minimization Options" dialog and then reminimize. Then "Save Changes" and once again choose the "Overwrite" option.
Chapter 2: Conformational Preference of Methylcyclohexane

Does methylcyclohexane prefer the axial or equatorial conformation? Do Problem 1.1 first. If your axial structure is not on the screen, pull down the File menu and 'Open' your file amecyc6.msf. We now need to create the structure of the equatorial isomer (Figure 2.1). You could do this by using ChemNote, but we will use the Molecular Editor to give you some practice with this powerful tool. Pull down the Edit menu and choose 'Molecular Editor.' Choose the Swap Bonds option in the Molecular Editor Palette. Notice that instructions are often printed in the lower left hand corner of the main window. To pick the first bond to swap, click on the bond to the equatorial hydrogen on carbon 1 (Figure 1.1). Now click on the C-C bond to the methyl group. The equatorial isomer will be produced. Click on 'Save and Exit' in the Molecular Editor palette. In the Save Options dialog box that follows, choose 'Reassign atom types,' 'Reassign atom charges,' and then OK. Adjust and smooth the charge using 'Carbon and non-polar hydrogen' types and click OK. In the MSF Saving Options choose "Save to New Filename" so that your old axial- file is not overwritten. Click OK to exit. In the file librarian dialog box, type emecyc6 as the new file name.

Figure 2.1. Equatorial methyl cyclohexane

To minimize the structure, check to make sure that the same options as in Chapter 1 are chosen. Click on "CHARMm Minimization" until the energy is minimized. The result should be 4.5006 kcal/mol. Select "Save Changes" and then "Overwrite emecyc6.msf." After you minimize the structure, use the 'CHARMm Energy' command can be used to list the contributions to the total steric energy of the molecule to allow comparison with the results from Problem 1.1.

**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 bond stretch, bond angle, and dihedral torsion terms) or the through-space Lennard-Jone terms?

<table>
<thead>
<tr>
<th>Contribution</th>
<th>equatorial (kcal/mol)</th>
<th>axial (kcal/mol)</th>
<th>difference (kcal/mol)</th>
<th>favored conformer</th>
</tr>
</thead>
<tbody>
<tr>
<td>bond energy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>angle energy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dihedral energy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lennard-Jones</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrostatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 1.09Å and the C-C bond length as 1.54Å for sp3 hybridized systems. The ideal bond angle around tetrahedral carbon is the tetrahedral angle, 109.5°. How close to these values do real molecules come?

Make sure your axial-methylcyclohexane is on the screen. If it isn't, from the main QUANTA screen, pull down the File menu and choose "Open." In the File Librarian select your axial file, click on the "Replace" button on the bottom of the screen, and then click on "Open." Bring the "Geometry" palette to the front. To bring a window forward, click on the border of its window (notice that the cursor changes to a >\ symbol when you are on the border of a window). Make sure "Show distance monitors," "Show angle monitors," and "Show dihedral monitors" are highlighted.

Bond distances: To find bond distances, first click on "Distance" in the Geometry palette. Now whenever you click on any two atoms, the distance between those two atoms will be displayed. Measure the bond distances in your compound. Record the values on the structure below. Don't measure every bond length, only the ones that are not related by symmetry. 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 click on "Distance" again to deselect it. To remove the atom labels click on "Clear ID" at the top of the Geometry palette. Finally, click on "Delete distance monitors."

Bond Angles: To find bond angles, first click on "Angles" in the Geometry palette. 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 bond angles in your compound. Record them in the structure above. When you are finished click on "Angles" again to deselect it. To remove the atom labels click on "Clear ID" at the top of the Geometry palette. Finally, click on "Delete angle monitors."

Dihedral Angles: To find dihedral angles, first click on "Dihedrals" in the Geometry palette. 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 the dihedral angles in your compound, including the ring C-C-C-C dihedral. Record them in the structure below. When you are finished click on "Dihedrals" again to deselect it. To remove the atom labels click on "Clear ID" at the top of the Geometry palette. Finally, click on "Delete dihedral monitors." Note that you can leave dihedral monitors on while you do other tasks, which include minimization, Conformational Searches, and dynamics.
Atom Charges: To display the charge on each atom, pull down the Draw menu, slide right on "Label Atoms," and select "Atomic Charge." 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 remove the atom charges from the screen, pull down the Draw menu, slide right on "Label Atoms," and select "Off."

The CHARMm Force Field
The CHARMm force field recognizes that bond lengths and angles change depending on hybridization and bonding partners even in normal-unstrained molecules. In Table 3.1 is listed the "normal" bond parameters that CHARMm uses in its force field for a few bond types. These parameters are the starting point for energy minimizations. Any deviations from these "normal" values will be reflected in increases in steric energy. These parameters are derived by finding the "best fit" to experimental data for a reference set of compounds. This reference set of compounds is often called the learning set. The experimental data is from electron and x-ray diffraction studies.

Table 3.1. CHARMm force field parameters.

<table>
<thead>
<tr>
<th>Bond</th>
<th>Distance(Å)</th>
<th>Angle</th>
<th>Angle (degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(sp^3)-C(sp^3)</td>
<td>1.529</td>
<td>C(sp^3)-C(sp^3)-C(sp^3)</td>
<td>112.70</td>
</tr>
<tr>
<td>C(sp^3)-O(alcohol)</td>
<td>1.405</td>
<td>C(sp^3)-C(sp^3)-O(alco)</td>
<td>110.5</td>
</tr>
<tr>
<td>C(sp^2 )*-C(sp^3)</td>
<td>1.502</td>
<td>C(sp^3)-C(sp^2 )-C(sp^3)</td>
<td>114.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C(sp^2 )-C(sp^3)-C(sp^3)</td>
<td>112.90</td>
</tr>
<tr>
<td>C(carbonyl)-C(sp^3)</td>
<td>1.530</td>
<td>C(sp^3)-C(=O)-C(sp^3)</td>
<td>117.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C(=O)-C(sp^3)-C(sp^3)</td>
<td>109.9</td>
</tr>
<tr>
<td>C(carbonyl)=O</td>
<td>1.207</td>
<td>C(sp^3)-C=O</td>
<td>124.80</td>
</tr>
<tr>
<td>H-C(sp^3)</td>
<td>1.090</td>
<td>H-C(sp^3)-H</td>
<td>107.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H-C(sp^3)-C(sp^3)</td>
<td>110.7</td>
</tr>
<tr>
<td>H-O(alcohol)</td>
<td>0.948</td>
<td>H-O(alcohol)-C(sp^3)</td>
<td>106.7</td>
</tr>
</tbody>
</table>

* sp^2 hybridized but not conjugated.
**Problem 3.1**
Measure the shortest distance between a methyl hydrogen and a ring hydrogen in equatorial-methylcyclohexane. Include this distance on the equatorial structure above. Do these shortest distances in the axial and equatorial conformers correlate with the change in Lennard-Jones energy that you found in Problem 2.1?

**Problem 3.2**
Compare the bond distances and angles in axial-methylcyclohexane to the "normal" CHARMm values of the C-H bond length of 1.090Å, the C-C bond length of 1.529Å, and the angles listed in Table 3.1. 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?

**Problem 3.3**
Build ethanol in ChemNote (pull down the Applications menu, slide right on "Builders, and select "2-D Sketcher). Minimize ethanol, and then display the atom charges. Compare the magnitude of these charges to the charges for methylcyclohexane. In which molecule will the Electrostatic contribution to the steric energy be greatest? Measure the C-C bond length in ethanol. By what % does this bond length differ from the C-C bonds in methylcyclohexane?
Chapter 4: Building More Complex Structures: 1-Methyl trans Decalin

We will next illustrate how to build up complex structures from simple structures in ChemNote, by converting the axial methyl cyclohexane into 1-methyl trans decalin. Start up ChemNote by pulling down the Applications menu, sliding right on "Builders," and choosing "2-D Sketcher." Begin by Opening amecyc6: pull down the File menu and choose "Open." Our final structure will look like Figure 4.4. We will append the ring from carbon 2 to carbon 3 (see Figure 1.1 and 2.1 for atom numbering). Click on the cyclohexane ring in the middle of the bonds palette. Position the cursor on carbon 2, hold down the mouse button, and move the cursor to carbon 3. The rings should be fused as shown in Figure 4.1. You can now pull down the File menu and choose 'Return to Molecular Modeling' to minimize the structure. Choose to 'Save' the changes and type in a unique name in the file librarian. Remember, don't use punctuation or spaces in file names. The minimized structure is shown in Figure 4.2.

![Figure 4.1. 1-methyl trans-decalin](image)

![Figure 4.2: 1-methyl-transdecalin. Left: the structure, Right: without hydrogens](image)

**Editing Molecules In Chem Note**

To illustrate how to edit atoms in ChemNote, we will now change axial methyl cyclohexane to chlorocyclohexane. Begin again by Opening amecyc6 in ChemNote. Click on the Cl icon in the atoms palette and then click on the methyl carbon in the sketch pad window. The methyl group should change to a Cl. Playing with ChemNote is the best way to learn how to build molecules. Try other changes to your structure. If you make a mistake, click on the selection tool icon, †, in the Edit Icons palette in the upper left hand portion of the window. Select the incorrect atoms, and then click on the eraser.

The edit tools palette is shown below with the action of each button:
Problem 4.1  Use ChemNote to build the structure for camphor, Figure 4.3. In ChemNote build the molecule starting with cyclohexane, and add the other bonds as if you were looking from above, Figure 4.4. Use Conjugate Gradient minimization to refine the structure. Report the final steric energy and the various energy contributions. Which term dominates the energy of camphor? Compare your results with methylcyclohexane, or better 2-methyl, 5-isopropylcyclohexanone. Does this comparison bare out the expectation that camphor is a highly strained molecule?

Figure 4.3. (a). Camphor. (b) Structure of camphor from molecular mechanics.

Figure 4.4. The appearance of camphor in ChemNote.
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.

The eclipsed conformer is higher in energy than the staggered form. The increase in **dihedral 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 bonds 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 $\phi = 0^\circ$, $120^\circ$, and $240^\circ$ are all identical eclipsed conformations. The conformations with $\phi = 60^\circ$, $180^\circ$, and $300^\circ$ are all identical with staggered, low energy conformations. Locate these energies in Figure 5.2.

Figure 5.2 is a plot of the dihedral, or torsional, potential energy for a $3\phi$, three-fold torsional barrier. Remember that the full torsional potential energy is given by:
\[ E_{\text{tor}} = \frac{1}{2} k_{\text{tor},1} (1 - \cos \phi) + \frac{1}{2} k_{\text{tor},2} (1 - \cos 2 \phi) + \frac{1}{2} k_{\text{tor},3} (1 - \cos 3 \phi) \]

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.

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 67° 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 \( \Delta U \). Remember \( \Delta H = \Delta U + \Delta n_g RT \), where \( \Delta n_g \) is the change in the number of moles of gas. Since we are calculating the difference in energy between two conformers:

\[ \text{butane (gauche)} \rightarrow \text{butane (anti)} \]

\( \Delta n_g = 0 \). 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:

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

Then

\[ \Delta G (\text{anti-gauche}) = \Delta H - T \Delta S \]

in calories:

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

and in kJ:

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

and the equilibrium constant can be obtained from:
\[ \Delta G = -RT \ln K \]
giving:
\[ K = \frac{\text{[anti]}}{\text{[gauche]}} = 1.93 \]

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 in ChemNote: pull down the Applications menu, slide right on "Builders", and choose "2-D sketcher" Click on the single bond icon in the bonds palette. Drag the three C-C bonds of butane out on in the ChemNote window. Don't worry about the hydrogens, ChemNote will add them automatically. Pull down the File menu and choose "Return to Molecular Modeling..." Proceed as you did in Chapter 1, using conjugate gradient minimization. After minimization, select "Save changes " from the Modeling palette, and select the "Overwrite.." option. This conformation should be the anti-conformer. Click on "CHARMm Energy..." in the Modeling palette. The contributions to the total steric energy will be listed in the window at the bottom left side of the screen. Record these energies and the total steric energy.

To find the energy minimized structure for the gauche isomer: select "Torsions..." from the Modeling palette. The "Torsions" palette will appear. Pick the first atom defining the torsion by clicking on a -CH2- carbon. Pick the second atom defining the torsion by clicking on the second -CH2- carbon atom. Select "Finish" in the torsions palette. The dials palette will now change to show only one dial, that for torsion 1. If the dials palette isn't completely visible, click on the border of its window (notice that the cursor changes to a >| symbol when you are on the border of a window). Click on the "torsion 1" dial until the dihedral angle is near 60°. Then select "CHARMm Minimization..." from the Modeling palette. Remember to click on "CHARMm Minimization..." repeatedly to make sure the structure is completely minimized. Click on "CHARMm Energy..." in the Modeling palette. The contributions to the total steric energy will be listed in the window at the bottom left side of the screen. Record these energies and the total steric energy. Select "Save changes " from the Modeling palette, and select the "Overwrite.." option.

**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.
<table>
<thead>
<tr>
<th>Contribution</th>
<th>anti (kcal/mol)</th>
<th>gauche (kcal/mol)</th>
<th>difference (kcal/mol)</th>
<th>favored conformer</th>
</tr>
</thead>
<tbody>
<tr>
<td>bond energy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>angle energy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dihedral energy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lennard-Jones</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrostatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**The Boltzman Distribution: An Alternative Viewpoint**

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

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

where $e^{-E_i/RT}$ is called the Boltzman weighting factor, $R$ is the gas constant 8.314 J mol$^{-1}$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](image_url)

Figure 5.4 The Boltzman 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 Boltzman 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 CHARMm steric energy of axial-methylcyclohexane is 1.8 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 Boltzman distribution with degeneracy is:

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

and

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

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 Boltzman factors for gauche and anti-conformations of butane at 298.2K.

<table>
<thead>
<tr>
<th>Conformation</th>
<th>Energy, $E_i$ (kJ)</th>
<th>$E_i / RT$</th>
<th>$e^{-E_i/RT}$</th>
<th>$g \ e^{-E_i/RT}$</th>
<th>$g \ e^{-E_i/RT} / q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>gauche</td>
<td>3.35</td>
<td>1.35</td>
<td>0.2589</td>
<td>0.5178</td>
<td>0.3411</td>
</tr>
<tr>
<td>anti</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.6588</td>
</tr>
<tr>
<td>sum=q=</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.5178</td>
</tr>
</tbody>
</table>

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:

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

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

### Dihedral Angle Conformation Searches

Pull down the Applications menu and choose "Conformational Search." The Conformation Search palette will then appear. Select "Torsions..." and then a new palette will appear. Make sure "Define all torsions." and "Use Default Names." is selected. Next click on "Pick torsions..." If a dialog box appears that says "Torsions already defined" then click on "Define New Torsion Angles" and then "OK." The "Pick Torsions" palette will appear. Click on the four carbons in the order they appear in the chain: in other words, click on the first CH$_3$, then the -CH$_2$, the second -CH$_2$, and finally the last CH$_3$. The "Torsion Name" window will then be displayed with "tor 1" as the default name and with "This is a backbone torsion" already selected; click "Apply." Next click on "Finish" in "Pick Torsions" palette, then "Exit torsions." Now click on "Setup search" and then select "Grid Scan." A grid scan search will change the dihedral angle in equal steps. In the "Define Ranges of Values" setup window choose "Absolute values", and change the settings to "From 0 to 350" and "Step size 10." Click on "OK" and then the "Grid Search Options" dialog will be displayed. Select "CHARMm minimization for each structure", "Constrain Grid torsions during minimization", "OK", and finally select "Do search" in the "Conformational Search" palette. The File Librarian will be displayed; type in a file name, for example "butanesrch." Click on "New." To see the results of the search, choose "ANALYSIS." In the "Plots" palette select "Scatter plot...". Choose the X-AXIS property as the "Torsion Angle," Click on "OK." Choose the Y-AXIS property as "Potential Energy," and then click "OK." The scatter plot will be displayed. To
set the scatter plot x-axis to 0 to 360°, pull down the Scatter tools menu and choose "Set 360 deg Scale."

To see the structure that corresponds to a given point in the scatter plot: Pull down the Scatter tools menu in the scatter plot window and choose "Select Structure." Now when you double click in the scatter plot window at various angles, the corresponding structure will be displayed. To exit the "Select Structures" mode Press the F1 key at the top of the keyboard. Pull down File in the scatter plot window and choose "Quit." Next click on "Exit Plots," "Exit Analysis" in the Analysis palette, and finally "Exit Conformational Search."

Please note that for the 'Torsions...' tool in the Modeling palette, you mark only two atoms. In other parts of Quanta, for example in the Geometry palette 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. Find the dihedral angle in the gauche conformer. Why is this angle different from butane? Also, use the Conformational Search application to plot the steric energy as a function of dihedral angle.

Problem 5.3

Using the energy difference from Problem 5.2, calculate the probabilities of occurrence of the gauche and anti forms for dichloroethane. Follow Table 5.1.

<table>
<thead>
<tr>
<th>Conformation</th>
<th>Energy, Ei (kJ)</th>
<th>Ei / RT</th>
<th>e^{-Ei/RT}</th>
<th>g e^{-Ei/RT}</th>
<th>g e^{-Ei/RT} / q</th>
</tr>
</thead>
<tbody>
<tr>
<td>gauche</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>sum=q=</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 -CO-CH₃ 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 sp² 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. Van der Waals solid 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: Working with MM2 from QUANTA

Energy minimization using CHARMM and MM2 are very similar. The force fields are a little different, but the calculations do the same thing. One reason for using MM2 is to calculate enthalpies of formation, which CHARMM can't do. MM2 also treats conjugated pi-electron systems better than CHARMM. You can't, on the other hand, use MM2 for large molecules. In this chapter you will find the enthalpy of formation of camphor, so do Problem 4.1 first.

MM2 Minimization  In the QUANTA screen, open your camphor file. Make sure the structure is energy minimized by clicking in "CHARMm minimization." Select "Save changes..." from the Modeling palette, and choose the "Overwrite..." option.

Pull down the Calculate menu, and choose MM2. Choose "Setup calculation..." from the new MM2 palette. Make sure the following three options are highlighted:

- Use MM2 Dipoles
- Do optimize
- Process Results automatically
- Cleanup Files automatically

Click "Finish" to return to the main MM2 palette. Choose "Run and Wait" to do the MM2 calculation. The File Manager dialog box will appear. Type in the name for your MM2 files and click "Save."

After the calculation is complete, wait for a red window outline to appear. Position this window at the very top of the screen and click the mouse button. The "jot" application window should appear, and in the window will be the output from the MM2 run. Scroll down to the last two pages of output to find the results. The enthalpy of formation is listed on the line labeled HEAT OF FORMATION (HFO)=. The line labeled SIGMA STRAIN ENERGY (S) = is also very useful as a measure of the total strain in the molecule.

You can print this information by pulling down the File menu and choosing Print. Select the "pin" printer then click "Print." You can edit the file before printing, to remove extraneous information to make the printout shorter. The extraneous information includes all of the updates for each iteration.

When you are finished with the output window, pull down File and choose Quit. You must then choose "Finish" in the MM2 palette before you can do anything else in QUANTA.

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. 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). The bond energy calculation for camphor from MMX is reproduced below for comparison. MMX uses a very similar force field to that in MM2.

<table>
<thead>
<tr>
<th>#</th>
<th>Bond or Structure</th>
<th>Each</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>C-C SP3-SP3</td>
<td>- 0.004</td>
<td>- 0.036</td>
</tr>
<tr>
<td>16</td>
<td>C-H ALIPHATIC</td>
<td>- 3.205</td>
<td>-51.280</td>
</tr>
<tr>
<td>1</td>
<td>C=O</td>
<td>-25.000</td>
<td>-25.000</td>
</tr>
<tr>
<td>2</td>
<td>C-C SP3-SP2 C=O</td>
<td>- 3.000</td>
<td>- 6.000</td>
</tr>
<tr>
<td>1</td>
<td>ISO (ALKANE)</td>
<td>0.078</td>
<td>0.078</td>
</tr>
<tr>
<td>1</td>
<td>NEO (ALKANE)</td>
<td>- 0.707</td>
<td>- 0.707</td>
</tr>
<tr>
<td>3</td>
<td>C(SP3)-METHYL</td>
<td>- 1.510</td>
<td>- 4.530</td>
</tr>
<tr>
<td>1</td>
<td>TERT-CARBONYL</td>
<td>- 1.300</td>
<td>- 1.300</td>
</tr>
</tbody>
</table>

bond energy = -88.775 kcal
Report your bond energy calculation, using Table I or data from your text, CHARMm steric energy, MM2 steric energy, MM2 bond energy, the strain energy, and the enthalpy of formation of camphor.

Compare the calculated results 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 MM2 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 MM2 value. How close did you come?

Problem 6.2. Comparisons with Literature Values (or How Good is MM2?)

How well do MM2 enthalpies of formation match literature values? Select a group from the following molecules and determine the enthalpy of formation for each. Look up the literature values; Lange's Handbook or the CRC are good places to look. 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.

Choose one of the following groups to work on.

Group 1: 1-chloropropane, 2-chloropropane, 1-chlorobutane, 2-chlorobutane, 3-chloro-1-propene
Group 2: methanol, ethanol, 1-propanol, 1-butanol, 2-butanol
Group 3: the monoterpenes are an important group of natural products. Some monoterpenes are shown below.

Report your results in a table with headings:
Enthalpies of Vaporization or Sublimation: The Clausius Clapeyron equation describes the change in vapor or sublimation pressure with temperature:

\[
\ln \frac{P_2}{P_1} = -\frac{\Delta f H}{R} \left(\frac{1}{T_2} - \frac{1}{T_1}\right) \quad \text{or equivalently} \quad \ln P = -\frac{\Delta f H}{RT} + \text{cst}
\]

where \(P_1\) and \(P_2\) are the vapor pressures at temperatures \(T_1\) and \(T_2\), respectively, and \(\Delta f H\) is the enthalpy of vaporization or sublimation. Comparing with Eq. 6.1, 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 kcal = 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 \(-\Delta f H/R\). The CRC has tables of vapor pressure versus temperature for many organic compounds.

Some of the values in kJ/mol are: for group 2: ethanol, 42.32; 1-propanol, 47.45; 1-butanol, 52.35; 2-butanol, 50.82. For group 3: camphene, 43.9; \(\alpha\)-pinene, 45.2; \(\beta\)-pinene, 46.4; limonene, 43.9; \(\alpha\)-terpineol, 52.3; menthol, 56.5 kJ/mol.

**Problem 6.3 Torsional Increments to the Enthalpy of Formation**

If there are unconstrained or free bond rotations in a molecule, the MM2 \(\Delta f H^g\) should be low. See if the addition of torsional increments improve the agreement of the calculated values with the literature values in Problem 6. 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. Make a table with the following headings: (some of the values for Group 3 are included)

<table>
<thead>
<tr>
<th>compound</th>
<th>Literature (\Delta f H^o) kcal/mol</th>
<th>Calculated (\Delta f H^o) MM2 kcal/mol</th>
<th>torsional # free rotations</th>
<th>increments total increment</th>
<th>Calculated (\Delta f H^o) MM2 + total increment kcal/mol</th>
<th>new (\Delta f H) kcal/mol</th>
<th>% improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>camphene</td>
<td>-7.73</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>(\alpha)-pinene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(\beta)-pinene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>limonene</td>
<td>-13.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>(\alpha)-terpineol</td>
<td>-73.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>menthol</td>
<td>-101.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Hints for Group 3: Many of the $\Delta_f H^\circ$ values for Group 3 are given, but make sure you know how to calculate them from values given in the literature (see problem 6). Camphene won’t have any free bond rotations, other than methyl groups. Use $\Delta_f H^\circ$ for camphor and camphene to judge the accuracy of our calculations when torsional increments don’t play a role. The other ring systems present a problem: how many free bond rotations should you add in? The rings hinder the motion in the ring, so perhaps no torsional increments should be added for ring bonds. On the other hand the rings do undergo conformational changes, so the ring bonds will contribute to $\Delta_f H^\circ$, but how much? The rings with double bonds will probably have less conformational flexibility than the saturated rings—why?

**Conjugated Pi-Electron Systems**

$\alpha$-Terpinene is an important mono-terpene (see Problem 6.2). However, the pi-electrons in the two double bonds are conjugated. MM2 in its simplest form does not do a good job on calculations of conjugated pi-electron systems. The MM2 $\Delta_f H^\circ$ as calculated in the same fashion as above is 22 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. MM2 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 calculation only covers the pi-electrons.

1,3-Butadiene, Figure 6.2, is a simple conjugated system that will serve as a good first example. The printout from the calculation on butadiene is shown in Figure 6.3. 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 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 listed as the "HEAT OF FORMATION" and is calculated to be 25.09 kcal/mol. The experimental $\Delta_f H^\circ$ is 26.75 kcal/mol. The default mode for the QUANTA interface to MM2 is to always do SCF pi-calculations for all pi systems, conjugated or not. The presence of a file called "noscf" in the QUANTA home directory is necessary to run MM2 without the SCF calculation.

**Problem 6.5 MM2 Calculations with SCF Pi Calculations**

Calculate the enthalpy of formation of $\alpha$-terpinene. The MM2 $\Delta_f H^\circ$ as calculated without the SCF molecular orbital calculation is 22 kcal/mol; the experimental value is -4.89 kcal/mol (from the CRC). Before you start the MM2 calculation, remove the file: "noscf" from the QUANTA home directory, if it is present.
Figure 6.3 The MM2 printout for 1,3-butadiene. The calculation is done with the SCF pi-molecular orbital calculation, which is the normal mode when running MM2 from QUANTA.

**Butadiene MM2 calculation with SCF calculation**

**MOLECULAR ORBITALS**

-0.4308  0.5607  0.5607  0.4308
0.5607  0.4308 -0.4308 -0.5607
-0.5607  0.4308  0.4308 -0.5607
0.4308 -0.5607  0.5607  0.4308

**EIGENVALUES**

-0.4634  -0.3854  -0.0256  0.0523

**BOND ORDERS (P) AND RESONANCE INTEGRALS (B) FOR PI-BONDS**

<table>
<thead>
<tr>
<th>BOND</th>
<th>P(W)</th>
<th>B(W)</th>
<th>P*B(W)</th>
<th>P(0)</th>
<th>B(0)</th>
<th>P*B(0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>0.9662</td>
<td>0.9868</td>
<td>0.9534</td>
<td>0.9662</td>
<td>0.9868</td>
<td>0.9534</td>
</tr>
<tr>
<td>2-5</td>
<td>0.2576</td>
<td>0.8032</td>
<td>0.2069</td>
<td>0.2576</td>
<td>0.8032</td>
<td>0.2069</td>
</tr>
<tr>
<td>5-7</td>
<td>0.9662</td>
<td>0.9868</td>
<td>0.9535</td>
<td>0.9662</td>
<td>0.9868</td>
<td>0.9535</td>
</tr>
</tbody>
</table>

**BOND ENERGY FROM SCF CALCULATION**

SIGMA-BOND  -238.65  PI-BOND  -118.06  TOTAL ENERGY -356.71

**FINAL STERIC ENERGY IS**  -2.4960 KCAL.

- COMPRESSION  0.0466
- BENDING  0.1723
- STRETCH-BEND  0.0113
- VANDERWAALS
  - 1,4 ENERGY  1.5450
  - OTHER  -0.0312
  - TORSIONAL  -4.2400
- CHARGE-DIPOLE  0.0000

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

**BOND ENTHALPY (BE) AND STRAINLESS BOND ENTHALPY (SBE) CONSTANTS AND SUMS**

<table>
<thead>
<tr>
<th>#</th>
<th>BOND OR STRUCTURE</th>
<th>--- NORMAL ---</th>
<th>--STRAINLESS--</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>C-H OLEFINIC</td>
<td>-3.205</td>
<td>-19.23</td>
</tr>
<tr>
<td>3</td>
<td>C-C DELOCALIZED</td>
<td>152.750</td>
<td>458.25</td>
</tr>
<tr>
<td>2</td>
<td>SEC- DELOCALIZED</td>
<td>-28.540</td>
<td>-57.08</td>
</tr>
</tbody>
</table>

BE = 381.94  SBE = 381.47

**PARTITION FUNCTION CONTRIBUTION (PFC)**

TRANSLATION/ROTATION TERM (T/R)  PFC = 2.40

**STERIC ENERGY (E)**  -2.50
**SIGMA-STRETCHING (ECPI)**  0.04
**CORRECTED STERIC ENERGY (EC) = E-ECPI**  -2.54
**ENERGY FROM PLANAR SCF CALCULATION (ESCF)**  -356.71
**HEAT OF FORMATION = EC + BE + PFC + ESCF**  25.09
**STRAINLESS HEAT OF FORMATION FOR SIGMA SYSTEM (HFS)**

HFS = SBE + T/R + ESCF - ECPI  27.11
**INHERENT SIGMA STRAIN (SI) = E + BE - SBE**  -2.02
**SIGMA STRAIN ENERGY (S) = POP + TOR + SI**  -2.02
Chapter 7: Comparing Structures

Changes in a molecule's structure not only affect the local environment, but 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 "Molecular Similarity" application is used to calculate the differences in two structures and to produce an overlaid view of the two structures. Pull down the File menu and choose 'Open'. Click on amecyc6.msf, at the bottom of the dialog box choose 'Append' (rather than 'Replace') so that both molecules will be displayed, and click on 'Open.' In the Molecule Management window, in the lower right portion of the screen, you can control which molecules are displayed by clicking in the 'Visible' column for the molecule of interest. Pull down the Applications menu and choose 'Molecular Similarity.' A new palette will appear. We now need to move one of the molecules to the right so they are no longer overlapped. Choose 'Move Molecule', click on an atom in the equatorial isomer, and move it to the right so that the isomers no longer overlap. To move the molecule use the Dials palette (lower right hand corner) or hold down the shift key and use the mouse. Click on "Move molecules" again to finish up. If the molecules aren't in orientations where you can see all the carbon atoms, choose 'Rotate Molecules in Place' and reorient the molecules. To rotate only one molecule, use the Molecule Management window to select the molecule you wish to rotate by clicking in the 'Active' column. Use the Dials palette to rotate the molecule. Make sure both molecules are active in the "Molecule Management" window, before proceeding.

We must now choose atom pairs that we wish to superimpose in the two isomers. Select 'Match Atoms.' A new palette will appear; choose 'Pick Equivalent Atoms'. Click on carbon 1 (Figure 1.1 and 2.1) in each isomer. A dotted line will be drawn between the two equivalent atoms. Do the same for carbon 2 in each structure (carbon 2 is the secondary ring carbon adjacent to the tertiary carbon). Also choose the equatorial hydrogens on carbon 2. If you make a mistake, choose 'undo last' and choose again. You can reorient the molecules at anytime using the center mouse button as before. You should now have three dotted lines between equivalent atoms. When the three pairs are choosen click on 'End Atom Picking' and then 'Exit Match Atoms.'

To do the comparison choose 'Rigid Body Fit to Target.' The target molecule is listed with an asterisk in the Molecule Management window. In this rigid body option, 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. The root mean square (rms) differences are displayed in the TextPort.

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 exactly 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 (Lennard-Jones) interactions. Choose 'Exit Molecular Similarity.'

Color Atoms

To make the two molecules easier to tell apart, use the 'Color Atoms' option. Pull down the Draw menu, slide right on "Color Atoms," and choose "By molecule." After you finish remember to return to normal colors by pulling down the Draw menu, sliding right on "Color Atoms," and choose "By element."

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

Structures can be plotted on the HP 870 printer. Orient your molecule in a good position on the QUANTA window, then follow the directions below.

1. Pull down the File menu, slide right on "Plot Molecules," and choose "Generate."
2. In the Plot Dialog box, choose "Artist Plot," and one of the styles listed. "Ball and stick" works well. The "Van der Waals" option is good for small molecules.
3. Enter a title for the plot in the "Title" edit box at the bottom of the screen. Choose the default option to "Plot with Titles and Border."
4. Click on "OK."
5. In the next dialog box select "Preview Plot," and click "OK."
6. A new window will appear with your plot. To continue, click the left mouse button in the Preview window.
7. The Plot Disposition window will be displayed. If the plot looked good, click on "Postscript Format," "Translate as color," "OK.," and go to step 11. If the plot wasn't sized properly, click on "Regenerate plot," and click on "OK."
8. The Plot dialog box will be displayed again. Select the same options as before. To change the plot scaling, change the number in the "Plot Scale" dialog box. The units are in mm/Å, so a bigger number increases the size of the molecule on the screen. Click on "OK" and continue at step 7.
9. The File Librarian window is displayed. Type in a file name. Click on "Save." The ".ps" postscript format file will then be generated. Next the "Plot Disposition" dialog box will be displayed, again. This time click on "Cancel."
10. To actually print the file, double click on the "quanta" folder icon on the desktop. Scroll the directory window until you find your file. The file should have the ".ps" suffix applied. Drag the file to the "HP" printer icon (on the desk top) for color printing or to the "Schupf Lab" printer icon for black and white. If there are no problems, the printer should begin printing within 10 sec. You are now finished. If there were problems go to the next step.
11. If the file didn't print: a) make sure the tray has paper loaded; b) make sure that the file name was correct. The ".ps" is added to your file name by QUANTA, so even though you didn't type it in, it is still necessary.

Fancier Plots

You can copy the current screen to the printer. This copy includes any solid surfaces. However, since these plots print with a black background, much ink is used. Therefore, please only use screen copies for special purposes like papers. To make a screen copy, in step 2 just choose "Color Screen Image."

Another possibility for good looking plots is to try the "stick plots" option. This type prints with a white background.
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 rather 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. $\psi$ is defined by the N-C-C-N dihedral and $\phi$ is defined by the carbonyl carbons in the dihedral C-N-C($\alpha$)-C. The normal values in the alpha helix are $\psi=-47^\circ$ and $\phi=-57^\circ$. The structure of alanylalanine is shown in Figure 9.1. The protein is shown in its non-ionized form. At neutral pH the N-terminus would be -NH$_3^+$ and the C-terminus would be -COO$. However, in our current work the attraction of the charged end-groups 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.

![Figure 9.1. The backbone dihedral angles in alanylalanine.](image)

First, Set up CHARMm for normal operation in RTF mode by: pulling down the CHARMm menu and sliding right on "CHARMm Mode" and choosing "RTF." RTF mode is the normal mode for working with polymers, which are composed of repeating units. Only use RTF mode with proteins and nucleic acids where QUANTA has RTF files that include all of your monomers. For small molecule work always use PSF mode.

Next pull down the CHARMm menu and select "Minimization options." and set the following options: Number of Minimization Steps 500
- Coordinate Update Frequency 5
- Energy Gradient Tolerance 0.0001
- Energy Value Tolerance 0
- Initial Step Size 0.02
- Step Value Tolerance 0
Next we need to build the dipeptide. Pull down Applications, slide right on "Builders," and choose "Sequence Builder." The File Librarian will be displayed for you to "Select a residue library." Choose the "AMINOH.RTF" line in the scroll box. Click on "Open." The "Sequence Builder" window will now be displayed with a list of available amino acids in the upper left corner. Click on "ALA" twice. The main window show now show "-ALA ALA." The peptide is built in the default zwitter ion form, which we must now changes. Changes to the sequence are made with "Patches." Pull down the Edit menu and choose "Apply Patches to Residues." The buttons in the upper left corner will change to the available patches. Click on "NH2" and then on the left hand "ALA." Next click on "COOH" and then the right hand "ALA." To set the initial conformation, we will choose the all-trans structure. Then we will check to see if the minimized structure changes much. Pull down the Conformation menu and select "Set Secondary Conformation." You are then instructed to "Pick the residue or range of residues." Click on the two "ALA" residues, and then click on "OK." A dialog box will be displayed, choose the "Extended Backbone (180.0)" option. Select the "OK" button. To exit the builder, pull down the Sequence Builder menu and choose "Return to Molecular Modeling." You will be asked: "Do you wish to save changes," click on "Yes." The File Librarian will then be displayed: type in a file name for your sequences and click on "Save." Next you will be asked "Which molecule do you want to use"; click on "Use new molecule only." The dipeptide is produced in the all-trans conformation. Now we can minimize the structure: choose "CHARMm minimization" repeatedly until the structure is at an energy minimum. What dihedral angles and energy did you get?

To measure the dihedral angles go to the "Geometry" palette. Make sure "Show dihedral monitors" is highlighted. Click on "Dihedrals" to begin selection of your angles. To select the \( \psi \) dihedral, start from the N-terminus and click on the backbone atoms: N-C-C-N in turn. To select the \( \phi \) dihedral, start with the carbonyl-carbon on the N-terminus end, and then select the backbone atoms: N-C(\( \alpha \))-C(carboxyl) in turn. Compare these values to the "ideal" alpha helix values. Leave these dihedral monitors on.

What hydrogen bonding exists for this conformation? Go back to the "Modeling Palette" and click on "Hydrogen bonds." How do these hydrogen bonds stabilize the conformation? Are the hydrogen bonds that form similar to those in an alpha helix? Select "Reject changes" to return to the all-trans structure. (You can build a short alanine polypeptide in the sequence builder to see what the normal hydrogen bonding pattern looks like. Just make sure the peptide is at least four ALA's long, and choose the "Right-Handed Alpha Helix" secondary conformation option.)

**Problem 9.1**

Adjust the torsional angles in your dipeptide to give \( \psi = -60 \) and \( \phi = -60 \). To accomplish this do the following. Select "Torsions..." from the Modeling palette. The "Torsions" palette will appear. Pick the first atom defining the \( \psi \) torsion by clicking on the C(\( \alpha \))- carbon at the N-terminus end. Pick the second atom defining the torsion by clicking on the adjacent carbonyl-carbon atom. Select "Finish" in the torsions palette. The dials palette will now change to show only one dial, that for torsion 1. If the dials palette isn't completely visible, click on the border of its window (notice that the cursor changes to a \( > \) symbol when you are on the border of a window). Click on the "torsion 1" dial until the dihedral angle is near -60\(^\circ\). Next repeat the above procedure for the \( \phi \) angle, which should be set to -60. Then select "CHARMm Minimization..." from the Modeling palette. Remember to click on "CHARMm Minimization..." repeatedly to make sure the structure is completely minimized. Measure the new dihedral angles and record the energy. After you are finished, select "Reject changes" in the "Modeling palette" before going on to Chapter 10. Which conformation is lowest in energy, the 180,180 or this one? Which structure is better stabilized by hydrogen bonds?

Please note that for the 'Torsions...' tool in the Modeling palette, you mark only two atoms. In other parts of Quanta, for example in the Geometry palette and for Conformational Searches, you need to specify all four atoms of the dihedral.
Chapter 10: Molecular Dynamics

Introduction
One of the most important developments in macromolecular chemistry is molecular dynamics. Molecular dynamics is the study of the motions of molecules. The time dependence of the motion of a molecule is called its trajectory. The trajectory is determined by integrating Newton's equations of motion for the bond stretching, angle bending, and dihedral torsions of the molecule. Molecules are always in motion. The motion of molecules is important in essentially all chemical interactions and are of particular interest in biochemistry. For example, the binding of substrates to enzymes, the binding of antigens to antibodies, the binding of regulatory proteins to DNA, and the mechanisms of enzyme catalysis are enhanced and sometimes completely determined by the conformational flexibility of the molecules. Different domains of an enzyme can have very different motional freedom. The problem of protein folding is the determination of the trajectory of the macromolecule as it assumes its active conformation after or during protein synthesis.

Most chemistry is done in solution. Molecular dynamics has proved to be an invaluable tool in studies of solvation energetics. Solute-solvent interactions are governed by the relative motions of the solute and solvent molecules and the motional-response of the solute to the presence of the solvent. Some of the earliest dynamics studies were to determine solvation Gibb's Free energies. In biochemistry, solute-solvent interactions play a particularly important role in determining the secondary and tertiary structure of biomolecules.

Another important use of dynamics is in the search for the global energy minimum in conformationally flexible molecules. Molecular mechanics find the energy minimum which is closest to the starting conformation of the molecule. This "local" energy minimum is rarely the lowest energy, or "global", minimum for the molecule. Finding the "global" minimum can be a very difficult task. In molecular mechanics a common procedure is to start with many different initial conformations and minimize them all looking for the lowest energy result. This kind of search can be very time consuming. Molecular dynamics, on the other hand, can help a molecule "explore" its conformation space more efficiently. The trajectory of the molecule is run at a high temperature, so that the atoms will move very far from their equilibrium positions. Such high temperature trajectories can overcome energy barriers that lead to more stable conformations. The trajectory often starts in one conformation and then ends up in another more stable conformation.

Molecular dynamics is an active area of research in biochemistry, molecular biology, and polymer chemistry. Current work is directed towards making molecular dynamics a reliable tool for the estimation of Gibb's free energies of solvation, conformational equilibria, and equilibrium constants for binding interactions. These thermodynamic parameters are determined by doing free energy perturbation studies using molecular dynamics trajectories; see Chapter 11 for more on free energy perturbation.

The difference between molecular mechanics and dynamics can be illustrated with a simple example. Let's direct our attention to a single bond in a molecule, a C-H bond for example. Assume that we start with the bond length too large, say 2 angstroms. If we were to run molecular mechanics, the bond length would decrease until the minimum in the potential energy was reached, Figure 10.1a. Further minimization would not change the bond length. If we were to run molecular dynamics on our stretched bond, the trajectory would decrease the bond length, but the bond length would continue decreasing past the equilibrium length until it was too short. Being too short, the bond length would then begin to increase. Over time the bond length will oscillate about its equilibrium value, never coming to rest, Figure 10.1b. In other words, in mechanics the potential energy is minimized, while the kinetic energy of the molecule is ignored. In a dynamics trajectory, both potential and kinetic energy are studied and the total energy is conserved by the motion.
Figure 10.1. The potential energy function for a bond. The initial bond length at 2 angstroms is too long. (a) Molecular mechanics finds the lowest energy state of the molecule. b. Molecular dynamics find the time dependent motion of the molecule. The vibration continues forever.

As chemists we often have too static a picture of molecules. Our mental images of molecular structure are derived from the printed page. Rather, molecules are always in motion. The results of molecular dynamics are very instructive, because dynamics trajectories show us how important motion is in chemical interactions. We should remember that chemical reactions, by their very nature, involve the motion of atoms as bonds are broken and made.

**Dynamics Trajectories: Integrating Newton's Laws**

Integrating Newton's Laws of motion is actually very straightforward. First, we use the molecular mechanics force field as the potential energy for our molecule. Therefore, the potential energy of our molecule involves bond stretching, angle bending, dihedral torsions, Van der Waals interactions, and electrostatic interactions. We then solve for the motion of each atom in the molecule as a function of time using this potential energy. However, as we begin to learn about dynamics, let's simplify our system to make things less complicated. Let's start with a diatomic molecule. The results of our work on a diatomic molecule will involve everything we need to know about more complicated systems. The molecular mechanics potential energy of a diatomic system has only one term, the potential energy for bond stretching:

\[ V = \frac{1}{2} k (r - r_0)^2 \]  

where \( r \) is the current bond length, \( r_0 \) is the equilibrium bond length, and \( k \) is the force constant for the bond. We can simplify Eq. 1 even further if we let \( x = r - r_0 \), then

\[ V = \frac{1}{2} k x^2 \]  

The force that acts on the system is the derivative of the potential:

\[ F = - \frac{dV}{dx} \]  

Taking the derivative of Eq. 2 gives:

\[ F = -k x \]
which is just the familiar Hooke's Law for a mass on a spring. Here the bond is the spring. Newton's Law tells us that \( F = m \cdot a \), where \( a \) is the acceleration. The acceleration is the rate of change of the velocity:

\[
F = -k \cdot x = m \frac{dv}{dt}
\]

The position of the system, \( x \), is determined by integrating the equation:

\[
\frac{dx}{dt} = v
\]

Integrating Eq. 5 gives the velocity as a function of time, starting from an initial velocity of \( v_1 \):

\[
\int_{v_1}^{v_2} dv = \int_{t_1}^{t_2} \frac{F}{m} \, dt
\]

giving

\[
v_2 = v_1 + \frac{F}{m} (t_2 - t_1)
\]

where \( m \) is the reduced mass for the vibrating bond. Integrating Eq. 6 gives the position as a function of time, starting from an initial position of \( x_1 \):

\[
\int_{x_1}^{x_2} dx = \int_{t_1}^{t_2} v_2 \, dt
\]

giving

\[
x_2 = x_1 + v_2 (t_2 - t_1)
\]

Since the velocity and position are both changing with time, Eqs 8 and 10 are solved repeatedly over short time steps, first updating the velocity and then updating the position. The value of \( x \) for each of these successive time intervals is then the trajectory of the system. In dynamics simulations the time step is very short, usually \( dt = t_2 - t_1 = 1 \times 10^{-15} \) sec or 1 femtosec.

All that remains is to determine the initial conditions. A common choice for the position is to choose \( x_1 = 0 \) at \( t = 0 \). But what about the velocity? The average velocity of a system is related to the temperature; the higher the temperature the larger amplitude the motions. At \( x = 0 \) all of the energy of an oscillating molecule is in kinetic energy. The kinetic energy is given as

\[
KE = \frac{1}{2} m v^2
\]

The Equipartition Principle of thermodynamics gives an estimate of the kinetic energy in a bond vibration as \( 1/2 \, RT \), where \( R \) is the gas constant; \( R = 8.314 \, \text{J mol}^{-1} \text{K}^{-1} \). Setting \( KE = 1/2 \, RT \) and solving for the velocity gives:

\[
v = \sqrt{\frac{RT}{m}}
\]

We therefore set \( v_1 = \sqrt{\frac{RT}{m}} \) at \( t = 0 \).

Eqs 8 and 10 are all that is meant by "integrating" Newton's Laws of motion. However, our example is a "one dimensional" system: there is only one motional variable. In more complicated molecules, equations 8 and 10 would be solved for the \( x, y, \) and \( z \) motion of each atom.
However, no new theory is needed; the problem just becomes more tedious. Computers are very good at solving simple, repetitive problems. In fact the advancement of molecular dynamics is very closely tied to the advancement of computer technology. The availability of fast computers means that molecular dynamics can now become one of the standard tools in computational chemistry.

**Problem 10.1: Dynamics trajectories**

Write a short EXCEL spreadsheet or BASIC program to determine the trajectory for a diatomic molecule. To make the problem more realistic, assume the bond is anharmonic, with potential energy function:

\[
V = \frac{1}{2} k x^2 + \frac{1}{2} k \text{anharm} x^3
\]

Please see the Chapter 1 for more information on anharmonic potentials for bond stretching. With your dynamics trajectory you will be able to see the time dependence of the vibration. You will also be able to determine the conditions for breaking a bond. For example, you can increase the anharmonicity to determine how anharmonic the bond must be to be broken at room temperature. Conversely, you can keep the anharmonicity constant and increase the temperature until the bond breaks, which is just what synthetic chemists do when they heat a reaction mixture. Differentiation of Eq. 13 gives:

\[
F = -k x + \frac{3}{2} k \text{anharm} x^2
\]

Display the results graphically as two asterisks separated by the distance \(x\). To make the graphics a little easier, you can use the program fragments below. Start with:

\[
\begin{align*}
R &= 8.314 \\
T &= 298.2 \\
k &= 200 \\
m &= 10 \\
\text{anharm} &= 0.05 \\
dt &= 0.1 \\
x &= 0
\end{align*}
\]

With these constants, increasing \(\text{anharm}\) to 0.1075 will cause the molecule to dissociate at 298.2K. Solve Eq. 12 for the initial velocity, \(v\). Because of the way that computer languages handle the "=" sign, you can drop the subscripts on \(v\) and \(x\), for example write:

\[
\begin{align*}
v &= v + F/m \ dt \\
\text{and} \\
x &= x + v \ dt.
\end{align*}
\]

After you get your spreadsheet or program to work, change the force constant \(k\), the anharmonicity, and the temperature to note the effect.

*The Spreadsheet Version:* Set up columns using the integrated Newton's equations 15 and 16 to calculate \(x\). Then to do the graphics, set up a column with values = \(x+10\). The 10 is an arbitrary offset to make the graphics look good. In the next column, put in statements similar to

\[
\text{=REPT(" ",15-D17/2)"*"&REPT(" ",D17)"
}
but, instead of "D17" use the cell address of the adjacent column with the x+10 values. The result should look something like:

<table>
<thead>
<tr>
<th>v</th>
<th>x</th>
<th>F</th>
<th>x+10</th>
<th>plot</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.74559</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>*</td>
</tr>
<tr>
<td>12.96835</td>
<td>1.574559</td>
<td>-277.72</td>
<td>11.5746</td>
<td>*</td>
</tr>
<tr>
<td>8.462302</td>
<td>2.871394</td>
<td>-450.61</td>
<td>12.8714</td>
<td>*</td>
</tr>
<tr>
<td>3.100163</td>
<td>3.717624</td>
<td>-536.21</td>
<td>13.7176</td>
<td>*</td>
</tr>
<tr>
<td>-2.52184</td>
<td>4.02764</td>
<td>-562.2</td>
<td>14.0276</td>
<td>*</td>
</tr>
<tr>
<td>-7.93464</td>
<td>3.775457</td>
<td>-541.28</td>
<td>13.7755</td>
<td>*</td>
</tr>
<tr>
<td>-12.5648</td>
<td>2.981993</td>
<td>-463.01</td>
<td>12.982</td>
<td>*</td>
</tr>
<tr>
<td>-15.5692</td>
<td>1.725515</td>
<td>-300.44</td>
<td>11.7255</td>
<td>*</td>
</tr>
<tr>
<td>-15.9021</td>
<td>0.168595</td>
<td>-33.293</td>
<td>10.1686</td>
<td>*</td>
</tr>
<tr>
<td>-12.7557</td>
<td>-1.42162</td>
<td>314.639</td>
<td>8.57838</td>
<td>*</td>
</tr>
<tr>
<td>-6.27013</td>
<td>-2.69719</td>
<td>648.561</td>
<td>7.30281</td>
<td>*</td>
</tr>
<tr>
<td>2.03583</td>
<td>-3.32421</td>
<td>830.596</td>
<td>6.67579</td>
<td>*</td>
</tr>
<tr>
<td>9.737817</td>
<td>-3.12062</td>
<td>770.199</td>
<td>6.87938</td>
<td>*</td>
</tr>
<tr>
<td>14.72284</td>
<td>-2.14684</td>
<td>498.502</td>
<td>7.85316</td>
<td>*</td>
</tr>
<tr>
<td>16.1402</td>
<td>-0.67456</td>
<td>141.737</td>
<td>9.32544</td>
<td>*</td>
</tr>
</tbody>
</table>

*The BASIC program*: The program listed below will then take care of the plotting. Just slip in your constants and initial conditions before the loop. Then put the integrated Newton's equations 15 and 16 inside the loop. The IF statement is put in to signal the dissociation of the bond. When the molecule dissociates the program will print out "rrrip." With these constants, increasing anharm to 0.1075 will cause the molecule to dissociate at 298.2K.

REM program to solve Hooke's Law dynamics
.
.
put constants and initial conditions in here
.
.
FOR i=1 TO 100
.
.
put Eq. 14, 15, and 16 in here
.
.
p=x+6
IF p>50 THEN LOCATE 1,1:PRINT"<rrrip>";GOTO qt
LOCATE 1,1
PRINT SPC(15-p);"*";SPC(INT(p+.5)+p);"*
LOCATE 1,1
PRINT SPC(15-p);";SPC(INT(p+.5)+p);" 
NEXT i
.
qt:
LOCATE 2,1
INPUT"type return to finish";a$
Chapter 11: 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 course of chemical reactions.

Molecular mechanics minimization corresponds to the structure the molecule would have at zero degrees K. Dynamics calculations are done in three steps. We first do a "heating" run to warm the molecule to room temperature. Next we "equilibrate" the molecule at room temperature to ensure that all the degrees of freedom are at the same temperature. Finally, we do a "simulation" run that generates the trajectory of the molecule at room temperature. The "simulation" run is used to answer questions about the motion of the molecule.

Procedure:

To begin, complete Chapter 9 and leave the alanylalanine molecule in the all-trans, extended backbone conformation. We will see if molecular dynamics is successful in finding the lowest energy conformation of the dipeptide. Make sure any dihedral monitors are off, but that hydrogen bonds are showing (click on Hydrogen Bonds in the Modeling palette if you haven't already done so).

Make sure the "Shake" option is on. This option keeps the C-H bonds in the molecule from gaining energy. By damping these high frequency vibrations, a longer time step may be chosen, thus decreasing the computation time. To turn "Shake" on, pull down the CHARMm menu and choose "SHAKE Options." Select the following options:

- SHAKE On
- BonH

Run dynamics with Shake ON.

- Shake tolerance: 1e-09
- Maximum number of iterations: 500
- Use Parameter-specified Geometry

Pull down the CHARMm menu and choose "Dynamics options." Select "Setup Heating," and click on "OK." In the heating setup dialog choose the following options:

Dynamics steps: 3000
Restart Read File: _____
Restart Write File: heat
Coordinate trajectory file: heat
Energy values file: heat
Output frequency: 10
Time step: 0.001
Initial temperature: 0
Final temperature: 300.0
- Start Heating From the Beginning
The time step is in picoseconds. Therefore, the time step of 0.001 psec is \(1 \times 10^{-15}\) sec or 1 femtosecond! Click on "OK." Next we set up the equilibration run: click on "Setup Equilibration." In the Equilibration setup dialog enter:

- Dynamics steps: 3000
- Restart Read File: heat
- Restart Write File: equil
- Coordinate trajectory file: equil
- Energy values file: equil
- Output frequency: 10
- Equilibration Frequency: 200
- Time step: 0.001
- Temperature: 300.0
- Restart Equilibration From the Restart File

Click on "OK." Finally, we setup the simulation run: click on "Setup Simulation." In the Simulation setup dialog enter:

- Dynamics steps: 3000
- Restart Read File: equil
- Restart Write File: simul
- Coordinate trajectory file: simul
- Energy values file: simul
- Output frequency: 10
- Time step: 0.001
- Temperature: 300.0
- Restart Simulation From the Restart File

Click on "OK." In the "Dynamics Setup" dialog select "Done." To run the dynamics trajectories select "CHARMm dynamics.." from the "Modeling" palette. The structures are displayed as they are generated by CHARMm. Notice that the total time of the dynamics run is 9 psec. After the run is complete, choose "Save changes.." (so that you can later minimize this structure).

Now we want to see the simulation trajectories: pull down the Applications menu and choose "Dynamics Animation." In the "Dynamics Animation" palette, click on "Select trajectories..." Click on the "Initialize Dynamics Files" check box and then click on "OK." The File Librarian will be displayed; click on the file "simul.DCD," and then "Open." Click on "Exit" to return to the "Dynamics Animation" palette. Next click on "Set Up Animation..." In the setup dialog verify the following settings:

- Use CHARMm Header
- Dataset range from: 6040 to: 9000
- Step size: 40
- Clock speed: 1
- Number Steps to Average Over: 0
- Regenerate hydrogen bonds
- Display geometry monitors
- Display trials
- Do not display dipole

Click on "OK." From the "Dynamics Animation" palette, select "Create Animation." To see the animation, click on "Cycle." The trajectory will cycle through all the time steps and then repeat. To change the speed of the animation, bring the "Dials" palette forward, and click on the vertical "speed" dial. Click on "Exit Dynamics Animation," to return to the main QUANTA palettes and window.
Minimize the current structure, to see the minimized structure that corresponds to the final dynamics conformation.

**Analysis of Dynamics Trajectories**

Does the dynamics trajectory find any new conformations that are significantly different from those you have already found? What do we mean by significantly different? We expect fluctuations about a local minimum energy conformation; these small changes in dihedral angles don't indicate a new conformation. The presence of a new conformation is shown by the trajectory traveling to and then fluctuating about a new local minimum, with distinctly different dihedral angles. These new angles, when energy minimized, give a new low energy structure. The presence of new conformations may be difficult to determine by watching the dynamics animation. The Analysis application can be used to produce a scatter plot of the ψ and φ angles during the trajectory calculation. From this plot you can determine if any significantly different conformations are part of the trajectory. You may already have used the Analysis application for looking at the results of conformational analysis for butane in an earlier chapter; the procedure here is very much the same as before.

1. Pull down the Analysis menu and choose "Analysis." A dialog box will appear asking, "Choose the Type of Input File." Choose "Dynamics File (.DCD)" and click on "OK." In the File Librarian, select "simul.DCD," and then "Open." The Analysis and Plots palette will be displayed.
2. Choose "Torsions..." in the Analysis palette. We need to define the ψ and φ dihedral angles for the Analysis plots. The "Torsions" palette will be displayed. Click on "Define Peptide Backbone Torsions." Click on "EXIT TORSIONS."
3. In the "Plots" palette select "Scatter plot...". Choose the X-AXIS property as a "Torsion Angle." Click on "OK." A dialog box will be displayed with the default torsion name "psi(1)" displayed. Click on "OK" to choose this as the first torsion. Choose the Y-AXIS property as "Torsion Angle," and then click "OK." A dialog box will be displayed with the default torsion name "psi(1)" displayed. You must change this to read "phi(2)." Click on "OK" to choose this as the second torsion. The scatter plot will be displayed.
4. To set the scatter plot x-axis to 0 to 360°, pull down the Scatter tools menu and choose "Set 360 deg Scale." Also pull down the Scatter tools menu and choose "Full Torsion Scale." To see the trajectory better, pull down the Scatter tools menu and choose "Show Trails."
5. To see the structure that corresponds to a given point in the scatter plot: Pull down the Scatter tools menu in the scatter plot window and choose "Select Structure." Now when you double click in the scatter plot window at various angles, the corresponding structure will be displayed. To exit the "Select Structures" mode **Press the F1 key** at the top of the keyboard. Pull down "File" in the scatter plot window and choose "Quit."

**Problem 11.1**

Does the dynamics trajectory find any new conformations that are significantly different from those you have already found? What hydrogen bonds form to stabilize the structure? Are these hydrogen bonds different than in Chapter 9? Is there more or less motion than you expected? Compare the energy of the minimized structure from the end of the simulation with the ψ=−60 and φ=−60 and 180,180 structure from Chapter 9.
Chapter 12: Free Energy Perturbation Theory

The greatest value in molecular dynamics is the ability to model the internal motions of a molecule. Internal energy, enthalpy, entropy, and Gibb's Free Energy all include contributions from the motion of a molecule. Therefore, molecular dynamics provides a way to estimate these important thermodynamic parameters. The current best method for practical calculations of Gibb's Free Energies is free energy perturbation theory, based on molecular dynamics. Free energy perturbation (FEP) theory is now in use in calculating $\Delta G$ for a wide variety of processes. For example, the Gibb's Free Energy of solution of hydrophobic molecules\(^1\), of binding of crown ethers to polar organics\(^2\), and the binding of NADP and NADPH to dihydrofolate reductase\(^3\) have been studied. In fact, the combined insights of x-ray crystal structure determination, NMR solution structure determination, and FEP studies have led to the consensus that the motions of proteins and nucleic acids play a major role in binding interactions. W. L. Jorgensen, in his article "Rusting of the Lock and Key Model for Protein-Ligand Binding," states simply that:

"These examples confirm the reasonable expectation that flexible molecules distort to form optimal interactions with binding partners."\(^4\)

Molecular mechanics calculates the steric energy of a molecule at absolute zero in temperature. What is the connection of the molecular mechanics steric energy to the thermodynamic internal energy and Gibb's Free energy of a substance? The hypothesis that makes the most sense is that the internal energy, $\Delta U$, is the time average of the total energy of the molecule. The total energy of the molecule is the kinetic plus potential energy:

$$E = \text{kinetic energy} + \text{potential energy} \quad 1$$

The potential energy is just the molecular mechanics steric energy. Molecular dynamics provides us with the time dependent energy of the molecule; all we need do to get $\Delta U$ is average the total energy during the trajectory calculation.

Now we turn to the relationship of the steric energy to the Gibb's Free Energy. In statistical mechanics, we find that the probability of a given state of a system occuring is proportional to the Boltzman weighting factor:

$$\text{probability of occurance } \propto e^{-E/RT} \quad 2$$

where $E$ is the total energy of the system, Eq. 1. In other words, states with low total energy are more likely to occur than states with high energy. A state of the system is determined by the conformation and motion of the molecule. The conformation determines the steric energy and the motions determine the kinetic energy.

In perturbation theory, we look at the effect of a small change in the structure of a molecule on its energy. To do the perturbation, the total energy is divided into two parts

$$E = E_0 + E_1 \quad 3$$

where $E_0$ is a reference structure and $E_1$ is a small perturbation from the reference structure. The perturbation is a small change that we place upon the system, say a small change in bond angle or a small change in the charge on an atom. The corresponding change in free energy of the system caused by the perturbation is given as\(^5,6\)

$$G - G_0 = -RT \ln \langle e^{-E_1/RT} \rangle_0 \quad 4$$

where $\langle >_0$ denotes the time average over the motion of the reference structure from a molecular dynamics run. The $e^{-E_1/RT}$ term is the probability of occurance for the small change in energy caused by the perturbation, from Eq. 2. The free energy then depends on the time average of the
probability of occurrence of the perturbed structure. In other words, if the perturbation produces a small change in energy, that change will contribute to the Gibb’s Free energy.

In our case however, we wish to find the change in free energy for large changes in a molecule. These changes, or mutations, include changing the conformations of bonds, or attaching a hydrogen ion, or changing a hydrogen to a methyl group or even a phenyl group. For example, we might like to mutate glycine into alanine for a study of site-specific mutagenesis of an enzyme. How do we apply Eq. 4 to such large changes? Assume that we wish to mutate molecule B into a different molecule A. First we define a total energy for mutating molecule B to A as

\[ E_\lambda = \lambda E_A + (1 - \lambda) E_B \]  

where \( E_A \) is the total energy for A and \( E_B \) is that for B, and \( \lambda \) is the coupling parameter. When \( \lambda = 1 \) the energy corresponds to molecule A, and when \( \lambda = 0 \) the energy corresponds to molecule B. When \( \lambda \) is at intermediate values, the system is a hypothetical superposition of A and B. It might seem quite strange to have such a combination of two molecules, in fact it is very unphysical; however, the theory is well-behaved and very useful none-the-less.

For the complete mutation to take place we vary \( \lambda \) from 0 to 1 over the course of the dynamics run. We divide this full range into short time slices, which are short enough that we can treat the change in each time slice as a perturbation. Then we apply Eq. 4 to each time slice and then add up the result for all the time slices. Let the \( \lambda \) value at each time slice be numbered \( \lambda_1, \lambda_2, \lambda_3, \text{ etc.} \). Then the difference in Eq. 4 is \( \Delta G(\lambda_i) \) for each time slice, \( i=1,2,3...n \), for n total time slices. Then the total change in \( \Delta G \) for the perturbation is

\[ \Delta G_{B\rightarrow A} = \sum_{i=1}^{n} \Delta G(\lambda_i) \]  

Since each time slice in the mutation is a small change, we can simplify Eqs. 4 and 6. We do the mutation in small steps; therefore \( E_1 \ll RT \) for each time slice in the perturbation. Remembering that \( e^{-x} \approx 1-x \), we can expand the exponential in the Boltzman distribution:

\[ e^{-E_1/RT} \approx 1 - E_1/RT \]  

Then Eq. 4 simplifies to:

\[ G - G_0 = -RT \ln < 1 - E_1/RT >_o = -RT \ln (1 - < E_1 >_o /RT) \]  

Next remember that \( \ln(1-x) \approx -x \), when \( x \) is small. This approximation on Eq. 8 gives:

\[ G - G_0 = -RT (1 - < E_1 >_o/RT) = < E_1 >_o \]  

In words, this simple result means that the change in Gibb’s Free Energy for a perturbation is just the time average of the total perturbation energy. Now applying Eq. 9 to each time slice in the total mutations simplifies Eq. 6 to:

\[ \Delta G_{B\rightarrow A} = \sum_{i=1}^{n} < E(\lambda_i) >_o \]  

where \( E(\lambda_i) \) is the total energy for the time slice in the mutation from Eq. 5. This very simple result makes FEP studies easy to do. The time average in Eq. 9 is automatically calculated during trajectory calculations. All we need do is to change \( \lambda \) in small steps during the trajectory. This approach to FEP simulations is called the slow-growth method.
Our initial efforts to use molecular dynamics are frustrated, however, because molecular dynamics is a classical theory, which gives too high a weight to high frequency vibrations. We must be careful to account for the difference between classical theories and the true distribution of vibrational energies in molecules. We can do this by always calculating the difference between our system and a reference system. In calculating differences, errors tend to cancel, and in so doing, classical molecular dynamics is a surprisingly useful tool for understanding complex systems. The success of classical dynamics is due in part to the observation that the major contributions to \( \Delta G \) for solvation and binding interactions are low frequency vibrations, especially torsions, which are handled adequately by classical theory. In addition, these low frequency vibrations tend to change the most in systems of interest; high frequency vibrations change little, therefore the high frequency vibrations cancel out in comparisons.

For example, to study the Gibb's Free Energy of solvation of molecule B, \( \Delta_{\text{sol}} G_B \), we will choose molecule A as the reference structure. The mutation will then be from B to A. To determine the difference in Free Energy of solvation between B and A, we will construct the following thermodynamic cycle:

\[
\begin{align*}
\text{A (aq)} & \quad \Delta_{\text{sol}} G_A \quad \text{---->} \quad \text{A (g)} \\
\uparrow \quad \Delta G_{\text{aq}}^{\text{B->A}} \quad | \quad -\Delta G_{g}^{\text{B->A}} \\
| \quad \Delta_{\text{sol}} G_B \quad \downarrow \\
\text{B (aq)} & \quad \text{---->} \quad \text{B (g)}
\end{align*}
\]

where \( \Delta G_{\text{aq}}^{\text{B->A}} \) is the Free Energy of perturbation of B to A in the solution phase, and \( \Delta G_{g}^{\text{B->A}} \) is the Free Energy of perturbation in the gas phase. Adding contributions around the cycle gives:

\[
\Delta_{\text{sol}} G_B = \Delta G_{\text{aq}}^{\text{B->A}} + \Delta_{\text{sol}} G_A - \Delta G_{g}^{\text{B->A}} \tag{12}
\]

We then determine the difference

\[
\Delta_{\text{sol}} G_B - \Delta_{\text{sol}} G_A = \Delta G_{\text{aq}}^{\text{B->A}} - \Delta G_{g}^{\text{B->A}} \tag{13}
\]

These kinds of differences are often called \( \Delta \Delta G \) values:

\[
\Delta \Delta G = \Delta_{\text{sol}} G_B - \Delta_{\text{sol}} G_A = \Delta G_{\text{aq}}^{\text{B->A}} - \Delta G_{g}^{\text{B->A}} \tag{14}
\]

We choose a reference system, A, where \( \Delta_{\text{sol}} G_A \) is known from experiment. We can then predict our final result:

\[
\Delta_{\text{sol}} G_B = \Delta_{\text{sol}} G_A(\text{experimental}) + \Delta \Delta G \tag{15}
\]

**Perturbation Setup Files**

There are currently three setup files for FEP studies. The first: "pert1to2" manages the perturbation between two structures produced by QUANTA. This file can be used for atom changes, for example Cl to H, or for dihedral angle changes, for example axial to equatorial or cis to trans. In general, "pert1to2" can be used whenever both structures have the same number of atoms and the mutation is not too large. The second setup file, "pertCltoH," mutates the chloro group to a H atom. To use "pertCltoH" you must build your structure in ChemNote starting with the chlorine to be mutated as the first atom. The reverse process is handled by "pertHtoCl." Here the hydrogen to be mutated must be the first atom.
Neither of these FEP methods use the SHAKE option. Recent studies have shown that results are more accurate without SHAKE. You should always start dynamics runs with a minimized structure, so don't forget to minimize before calling these files.

**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 a leaking underground gasoline tank 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. 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 solution:

\[ P_B = X_B K \]

where \( P_B \) is the partial pressure of dilute solute B above the solution, \( X_B \) is the mole fraction, and \( K \) is the Henry's Law constant for B. Compilations of \( K \) 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 \( K \) values from Free Energy perturbation studies (FEP). If FEP 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 Gibb's Free Energy of solvation. The equilibrium described by Eq. 16 can be written as:

\[ B (aq) \rightarrow B (g) \]

The equilibrium constant for reaction 17 is:

\[ K_{eq} = \frac{P_B}{X_B} \]

if we measure the concentration of \( B \) in mole fraction. Comparing Eqs. 16 and 18 shows that the equilibrium constant \( K_{eq} \) is the same as \( K \), the Henry's Law constant. The Gibb's Free Energy change for Eq. 17 is the Gibb's Free energy of solvation, \( \Delta_{sol}G_B \). Therefore, since \( K \) is the equilibrium constant for Eq. 17:

\[ \Delta_{sol}G_B = -RT \ln K \]

Therefore, FEP calculations of \( \Delta_{sol}G_B \) 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:

\[ K_{CC} = \frac{K_{PX}}{kRT} \]

where \( k \) is the conversion factor from mole fraction to concentration in mol L\(^{-1}\). Here, \( k = 1000 \text{ mL} \times d_{H_2O} / M_{H_2O} \), where \( d_{H_2O} \) is the density of water and \( M_{H_2O} \) is the molar mass of water. At 25°C, \( k = 55.35 \text{ mol L}^{-1} \). Also common in the literature is the Henry's Law constant with pressure for the gas phase and concentration for the aqueous phase:
$K_{PC} = \frac{K_{PX}}{RT}$

Table 12.1 list values for the various $K$'s and the $\Delta_{sol}G$ values derived from them using Eq. 19. The proper parameters for comparison with FEP calculations are $K_{CC}$ and $\Delta_{sol}G_{CC}$.

Table 12.1. Henry's Law constants and Free Energies of solvation. The number in parenthesis is the source for that substance and following values. $\Delta_{sol}G = -RT\ln K$. The units are indicated as subscripts: p=pressure, x=mole fraction, and c=molarity.

<table>
<thead>
<tr>
<th>Substance</th>
<th>$K_{px}$</th>
<th>$K_{cc}$</th>
<th>$K_{pc}$</th>
<th>$\Delta_{sol}G_{px}$</th>
<th>$\Delta_{sol}G_{cc}$</th>
<th>$\Delta_{sol}G_{pc}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>benzene(8)</td>
<td>2.96E+04</td>
<td>0.216</td>
<td>535.51</td>
<td>-25.52</td>
<td>3.80</td>
<td>-15.57</td>
</tr>
<tr>
<td>toluene</td>
<td>3.61E+04</td>
<td>0.263</td>
<td>652.04</td>
<td>-26.01</td>
<td>3.31</td>
<td>-16.06</td>
</tr>
<tr>
<td>ethylbenzene</td>
<td>4.36E+04</td>
<td>0.318</td>
<td>788.40</td>
<td>-26.48</td>
<td>2.84</td>
<td>-16.53</td>
</tr>
<tr>
<td>m,p-xylene</td>
<td>4.09E+04</td>
<td>0.298</td>
<td>738.81</td>
<td>-26.32</td>
<td>3.00</td>
<td>-16.37</td>
</tr>
<tr>
<td>o-xylene</td>
<td>2.80E+04</td>
<td>0.204</td>
<td>505.76</td>
<td>-25.38</td>
<td>3.94</td>
<td>-15.43</td>
</tr>
<tr>
<td>1,1,1-trichloroethane</td>
<td>9.85E+04</td>
<td>0.718</td>
<td>1780.09</td>
<td>-28.50</td>
<td>0.82</td>
<td>-18.55</td>
</tr>
<tr>
<td>trichloroethylene</td>
<td>5.76E+04</td>
<td>0.42</td>
<td>1041.28</td>
<td>-27.17</td>
<td>2.15</td>
<td>-17.22</td>
</tr>
<tr>
<td>tetrachloroethylene</td>
<td>9.56E+04</td>
<td>0.697</td>
<td>1728.03</td>
<td>-28.43</td>
<td>0.89</td>
<td>-18.48</td>
</tr>
<tr>
<td>methyl-t-butyl ether</td>
<td>2.96E+03</td>
<td>0.0216</td>
<td>53.55</td>
<td>-19.82</td>
<td>9.51</td>
<td>-9.87</td>
</tr>
<tr>
<td>tetrachloroethylene(9)</td>
<td>9.92E+04</td>
<td>0.723</td>
<td>1792.49</td>
<td>-28.52</td>
<td>0.80</td>
<td>-18.57</td>
</tr>
<tr>
<td>trichloroethylene</td>
<td>5.38E+04</td>
<td>0.392</td>
<td>971.86</td>
<td>-27.00</td>
<td>2.32</td>
<td>-17.05</td>
</tr>
<tr>
<td>1,1-dichloroethylene</td>
<td>1.47E+05</td>
<td>1.069</td>
<td>2650.30</td>
<td>-29.49</td>
<td>-0.17</td>
<td>-19.54</td>
</tr>
<tr>
<td>cis-1,2-dichloroethylene</td>
<td>2.29E+04</td>
<td>0.167</td>
<td>414.03</td>
<td>-24.89</td>
<td>4.44</td>
<td>-14.94</td>
</tr>
<tr>
<td>trans-1,2-dichloroethylene</td>
<td>5.27E+04</td>
<td>0.384</td>
<td>952.03</td>
<td>-26.95</td>
<td>2.37</td>
<td>-17.00</td>
</tr>
<tr>
<td>vinyl chloride</td>
<td>1.56E+05</td>
<td>1.137</td>
<td>2818.89</td>
<td>-29.64</td>
<td>-0.32</td>
<td>-19.69</td>
</tr>
<tr>
<td>1,1,1-trichloroethane</td>
<td>9.65E+04</td>
<td>0.703</td>
<td>1742.90</td>
<td>-28.45</td>
<td>0.87</td>
<td>-18.50</td>
</tr>
<tr>
<td>1,1-dichloroethane</td>
<td>3.16E+04</td>
<td>0.23</td>
<td>570.22</td>
<td>-25.68</td>
<td>3.64</td>
<td>-15.73</td>
</tr>
<tr>
<td>chloroethane</td>
<td>6.26E+04</td>
<td>0.456</td>
<td>1130.53</td>
<td>-27.38</td>
<td>1.95</td>
<td>-17.43</td>
</tr>
<tr>
<td>carbon tetrachloride</td>
<td>1.71E+05</td>
<td>1.244</td>
<td>3084.17</td>
<td>-29.86</td>
<td>-0.54</td>
<td>-19.91</td>
</tr>
<tr>
<td>chloroform</td>
<td>2.06E+04</td>
<td>0.15</td>
<td>371.89</td>
<td>-24.62</td>
<td>4.70</td>
<td>-14.67</td>
</tr>
<tr>
<td>dichloromethane</td>
<td>1.23E+04</td>
<td>0.0895</td>
<td>221.89</td>
<td>-23.34</td>
<td>5.98</td>
<td>-13.39</td>
</tr>
<tr>
<td>chloromethane</td>
<td>4.95E+04</td>
<td>0.361</td>
<td>895.00</td>
<td>-26.80</td>
<td>2.53</td>
<td>-16.85</td>
</tr>
<tr>
<td>methane (10)</td>
<td>4.13E+02</td>
<td>3.01E-03</td>
<td>7.46</td>
<td>-14.93</td>
<td>14.39</td>
<td>-4.98</td>
</tr>
<tr>
<td>oxygen</td>
<td>4.34E+04</td>
<td>0.316</td>
<td>784.10</td>
<td>-26.47</td>
<td>2.85</td>
<td>-16.52</td>
</tr>
</tbody>
</table>

**Free Energy Perturbation Calculation of $K_{CC}$**

Before we use FEP on a new system, we should determine the accuracy of FEP methods by running some known compounds and comparing with literature values in Table 12.1. Let's work through an example of a FEP study. The comparison we will discuss is the mutation of 1,1,1-trichloroethane to 1,1-dichloroethane. This mutation is a relatively simple one where a Cl atom is mutated to a H atom. However, we also need to adjust the charges on the attached C atom. We choose 1,1-dichloroethane as our reference structure, and from that reference calculate $\Delta_{sol}G$ of 1,1,1-trichloroethane. From Table 12.1, we expect $\Delta_{sol}G = \Delta_{sol}G(1,1,1\text{-trichloroethane}) - \Delta_{sol}G(1,1\text{-dichloroethane}) = 0.87 - 3.64 \text{ kJ/mol} = -2.77 \text{ kJ/mol}$ or equivalently $-0.66 \text{ kcal mol}$. To calculate $\Delta_{sol}G$ we need to run two FEP studies, one Cl -> H mutation in the gas phase to determine
\[ \Delta G_{B \rightarrow A}^g \] and the same Cl \rightarrow \text{H} mutation in the aqueous phase for \[ \Delta G_{B \rightarrow A}^{aq} \]. The difference in Eq. 14 gives \[ \Delta \Delta G \].

**Procedure**

In this example, we run the mutation discussed above. However, these instructions will work for any Cl \rightarrow \text{H} mutation. The procedure we use follows the outline:

1. Build and minimize 1,1,1-trichloroethane.
2. Copy the structure files to new files named PERT1.
3. Run the gas phase FEP.
4. Solvate 1,1,1-trichloroethane and minimize.
5. Copy the solvated structure files to new files named PERT1.
6. Run the aqueous phase FEP.

The detailed instructions follow.

1. **Build and minimize 1,1,1-trichloroethane:** Pull down the Applications menu, slide right on "Builders" and choose "2D Sketcher." Build 1,1,1-trichloroethane. Pull down the File menu and choose "Return to Molecular Modeling." "Save changes" and choose the default smoothing option, i.e. all carbon atoms and non-polar hydrogens. In the QUANTA window, pull down the CHARMm menu, slide right on "CHARMm Mode," and choose "PSF." Pull down the CHARMm menu and choose "Minimization Options." Choose "Conjugate Gradient," "800 steps," and energy tolerance "0.001." Click on "OK." Minimize the structure: remember to click on "CHARMm Minimization" repeatedly until the energy no longer changes. In the Modeling palette, choose "Save Changes," and select "Overwrite ____.”

2. **Copy the structure files to new files named PERT1:** Pull down the File menu and choose "Open System Window." In the system window type: mkpert1. At this point you can close the system window by typing: exit, or you can minimize the window using the minimize button on the title bar of the system window.

3. **Run the gas phase FEP:** Type: pertCltoH to begin the FEP dynamics run. At the conclusion, about four minutes later, type: p.out; the "jot" application will open and display the final results. The output file has many details, however at this point you only need record the final total \[ \Delta G_{B \rightarrow A} \] value. Start at the end of the file and scroll backwards. Find the heading "PERTURBATION>TOTALS since last reset:". The result is in the third column, which is labeled "HFCTote" several lines above, and in the row labeled "PTOT PROP>". Pull down the File menu in "jot" and choose "Quit." At this point you can close the system window by typing: exit, or you can minimize the window using the minimize button on the title bar of the system window.

4. **Return to the 1,1,1-trichloroethane and solvate and minimize:** Pull down the CHARMm menu, slide right on "Solvate," and choose "15Å length box water." Next you will be asked how you want to center the water solvation box, choose "Centroid." Back in the QUANTA screen choose "Save Changes" from the Modeling palette. Choose "Overwrite ___," and then "OK." Pull down the CHARMm menu and choose "Minimization Options." Change the number of minimization steps to 800, then click "OK." Minimize your structure. The minimization will require about 5 minutes. We set the number of steps to 800, so that you won't have to repeatedly click on "CHARMm Minimization." Make sure your structure is minimized to better than 0.1 kcal. Choose "Save Changes" from the Modeling palette, "Overwrite______," and click "OK."

5. **Copy the solvated structure files to new files named PERT1:** If you closed the system window in step 5, pull down the File menu and choose "Open System Window." In the system window type: mkpert1.

6. **Run the aqueous FEP:** Because the aqueous run takes much longer we want to do the CHARMm run in the background. This way others can use the computer if need be. To run a job
in the background we just follow the command with an "&." Type: "pertCltoH &" to begin the FEP dynamics run. At the conclusion, about three hours later, type: jot p.out; the "jot" application will open and display the final results. Record the final total $\Delta G_{B>A}$ value. Pull down the File menu in "jot" and choose "Quit." Close the system window by typing: exit. You can view the p.out file at any time, as long as no one else has done a perturbation calculation in the mean time.

**Problem 12.1**

Calculate $\Delta \Delta G$, using Eq. 14, and compare with the literature value. Since these numbers are expected to be small, the % error will be large; instead just report the difference. Is the sign correct on your value? Using Eqs. 15 and 19, calculate the K for your compound using the $\Delta_{sol} G$ literature value for 1,1-dichloroethane (or whatever reference compound you chose). The % error is expected to be large in this final value. But at least your result should be in the correct direction. To shorten the length of the simulation, we chose too small a number of dynamics steps for each time slice (150). Better results should be obtained with longer simulation times at each time slice.

**Literature Cited**

Chapter 13. Protein Structure and Gramicidin-S

One of the most active and interesting areas in bio-physical 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.

Gramicidin-S is a cyclic decapeptide, Figure 13.1, 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 interaction. 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 \( \alpha \)-CH and the backbone NH proton is about 4 Hz for a alpha-helix type structures and 9 Hz for beta-pleated sheet structures. 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 13.2a. 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 structures. QUANTA/CHARMm has two types of proline based secondary structures, which can be used to establish proline turns. The proline-1 secondary structure is based on the unusual formation of cis-peptide bonds, Figure 13.2b.

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 13.1.
### Table 13.1. Dihedral Angles for Regular Secondary Protein Structures

<table>
<thead>
<tr>
<th>Bond Angle (degrees)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Antiparallel β-sheet</td>
<td>-139</td>
<td>+135</td>
</tr>
<tr>
<td>Right-handed α-helix</td>
<td>-57</td>
<td>-47</td>
</tr>
<tr>
<td>Polyproline I</td>
<td>-83</td>
<td>+158</td>
</tr>
<tr>
<td>Polyproline II</td>
<td>-78</td>
<td>+149</td>
</tr>
</tbody>
</table>

### 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 \( \phi \) dihedral should be about \(-60^\circ\) and a beta-pleated sheet \( \phi \) dihedral should be about \(-140^\circ\). 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 3.0Å. QUANTA uses 3.0Å as the default distance constraint. Distance constraints can also be inferred from secondary structure assignments. Examples for such inferences include N-H...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...O=C distances between strands can be constrained. Hydrogen bond lengths are in the range from 1.8-3.0Å, with 2Å being normal for strong hydrogen bonds. The hydrogen bond distance between residues \( i \) and \( i+4 \) in the alpha-helix is about 1.86Å, and about 1.96Å 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 Sequence 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 by applying patches. You will also specify secondary structures like alpha-helical or beta-pleated sheet regions and specify turns.

1. Pull down the Applications menu, slide right on "Builders," and choose "Sequence Builder." In the File Librarian you must choose the "RTF" file for the type of polymer you want to build. Click the "AMINOH.RTF" entry and then "OK." (RTF stands for residue topology file.)
2. In the Sequence Builder, you click on the amino acids in your structure in the order that they appear in the protein. Start with ORN and continue around the ring, Figure 13.1.
3. Next we need to apply "patches" to phenylalanine to convert to the D-stereo isomer. Pull down the Edit menu and choose "Apply Patches to Residues."
4. Click on the "LTOD" button in the left hand portion of the screen, and then click on one of the PHE's in your structure. Then click on "LTOD" again, and then click on the other PHE.
5. Next we need to set the secondary structure, if you know it. If you don't have guesses for the secondary structure proceed to step 6. From your NMR spectra, determine the type of secondary structure that you have. Hint: two of the amino acids will have the same dihedral angles. Use these two only. Leave the third to be determined by the minimization. Pull down the Conformation menu and choose "Set Secondary Conformation." Select the two amino acids that you have determined and then on "OK." A scroll box will appear, choose the secondary structure for these two residues, and click "OK." Repeat this process for the corresponding residues on the other side of the ring (remember that the structure must be symmetric).
6. We need to specify turns for the prolines. Select the first PHE and PRO and click "OK." In the scroll box choose PROLINE 1, and click on "OK." Now select the second PHE and PRO pair, click "OK," and the select PROLINE 1 again. Click on "DONE" to return to the main screen.
7. Now we need to create the cyclic structure. Pull down the **Edit** menu and choose "Apply Patches to Residues." This time click on "LINK" and then the first ORN and then the last VAL. The sequence should now read:

```
7-ORN [ LINK 10 ] - LEU
3-PHE [ LTOD ] - PRO - VAL
6-ORN - LEU - PHE [ LTOD ]
10-PRO - VAL [ 1 LINK ]
```

8. Pull down the **Sequence Builder** menu and choose "Return to Molecular Modeling." When the system asks if you want to save your changes click on "YES." In the File Librarian type in a file name, remembering not to use any punctuation in the file name and **eight characters or less.** Click on "Save." In the next dialog box click on "Use the new molecule _____.msf only," and then "OK."

**Minimizing Your Structure**

After you build your sequence, you will note that one of the bonds is very long. This long bond was caused by specifying a cyclic structure. You must minimize your structure to get a reasonable starting conformation. You then will specify either dihedral or distance constraints and reminimize using conjugate gradient techniques to attempt to find a minimum energy structure that is consistent with the experimental evidence.

9. Pull down the **CHARMm** menu, slide right on "CHARMm mode," and choose "RTF." Pull down the **CHARMm** menu and choose "Minimization Options." Select "Steepest Descents," and then if not already shown:

- Number of Minimization Steps: 200
- Coordinate Update Frequency: 5
- Energy Gradient Tolerance: 0.0001
- Energy Value Tolerance: 0
- Initial Step Size: 0.02
- Step Value Tolerance: 0

Click on "OK."

10. Click on "CHARMm minimization" in the "Modeling" palette. The energy of this structure is still high, however. To speed subsequent minimization steps, you can apply dihedral or distance constraints. If you wish to apply dihedral constraints go to step 11. If you wish to apply distance constraints, which are better, go to step 15.

**Setting Dihedral Constraints**

11. Pull down the **CHARMm** menu, slide right on "Constraints Options" and choose "Dihedral/Distance." Click on "Define Dihedral Constraint(s)." Click on the four consecutive atoms that define the dihedral you wish to constrain. A new window will appear, labeled "Constraints_Database.con." In this window, you specify the constraints. QUANTA takes the current value of the dihedral as the target value. You also need to set the allowable ranges for your dihedral. This range should be fairly broad for this exercise, since we expect some strain in the ring to distort the angles from their ideal values. 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 dihedral you wish to constrain.

13. Click on "Save Constraints" in the "Edit Constraints" palette. Then click on "Exit Edit Constraints" to return to modeling.

14. Pull down the **CHARMm** menu, slide right on "Constraints Options" and choose "Dihedral On." Minimization will now include your constraints. Continue with step 21.

**Setting Distance Constraints**

15. Pull down the **CHARMm** menu, slide right on "Constraints Options" and choose "Dihedral/Distance." Click on "Define Distance Constraint(s)." Make sure "Individual Atoms" is checked.
16. Click on the two atoms that you wish to constrain. A new window will appear, labeled "Constraints_Database.con." In this window, you specify the constraints. You also need to set the allowable ranges for your distance. This range should be fairly broad for this exercise, since we expect some strain in the ring to distort the angles from their ideal values. For hydrogen bonds try a target of 2.0Å with a range of 3.0Å to 1.9Å. Remember to set the same constraints for both amino acids in the pair, since the structure should be symmetrical.

17. Repeat step 16 for each distance you wish to constrain.
18. Click on "Save Constraints" in the "Edit Constraints" palette. Then click on "Exit Edit Constraints" to return to modeling.
19. Pull down the CHARMm menu, slide right on "Constraints Options" and choose "Distance On." Minimization will now include your constraints.
20. A dialog box will appear that allows you to choose an overall scale factor; just use the default value of 1.000, by clicking "OK."

Minimization with Constraints
20. The bond lengths should now be close enough to normal values that we can use conjugate gradient minimization. Pull down the CHARMm menu and choose "Minimization Options." Select "Conjugate Gradient," and then if not already shown:

- Number of Minimization Steps 1000
- Coordinate Update Frequency 5
- Energy Gradient Tolerance 0.0001
- Energy Value Tolerance 0
- Initial Step Size 0.02
- Step Value Tolerance 0

Click on "OK."
21. Click on "CHARMm minimization" in the "Modeling" palette.
22. Check to see that your final structure is symmetrical. If it is not, apply additional constraints, use the "Torsions..." option, or use the "Move Atoms" function in the "Modeling" palette to produce a symmetrical structure. For example, you might need to use "Torsions..." in the "Modeling" palette to rotate some side chains so they are symmetrical.
23. The constraints that you have applied are an artificial term in the potential energy function. For your **final minimization you should remove all constraints** by completing the following. Pull down the CHARMm menu, slide right on "Constraints Options" and choose "Dihedral off" or "Distance Off," depending on the constraint type(s) you used. Reminimize. Repeat minimization until the energy is minimized. This minimization may take 6000 steps or more.
24. Remember to click on "Save Changes" in the "Modeling" palette. Overwrite the data file.
25. Plot out several views of your structure. Measure some of the backbone dihedrals to see if the \( \phi \) angles are close to the NMR determined values.

References