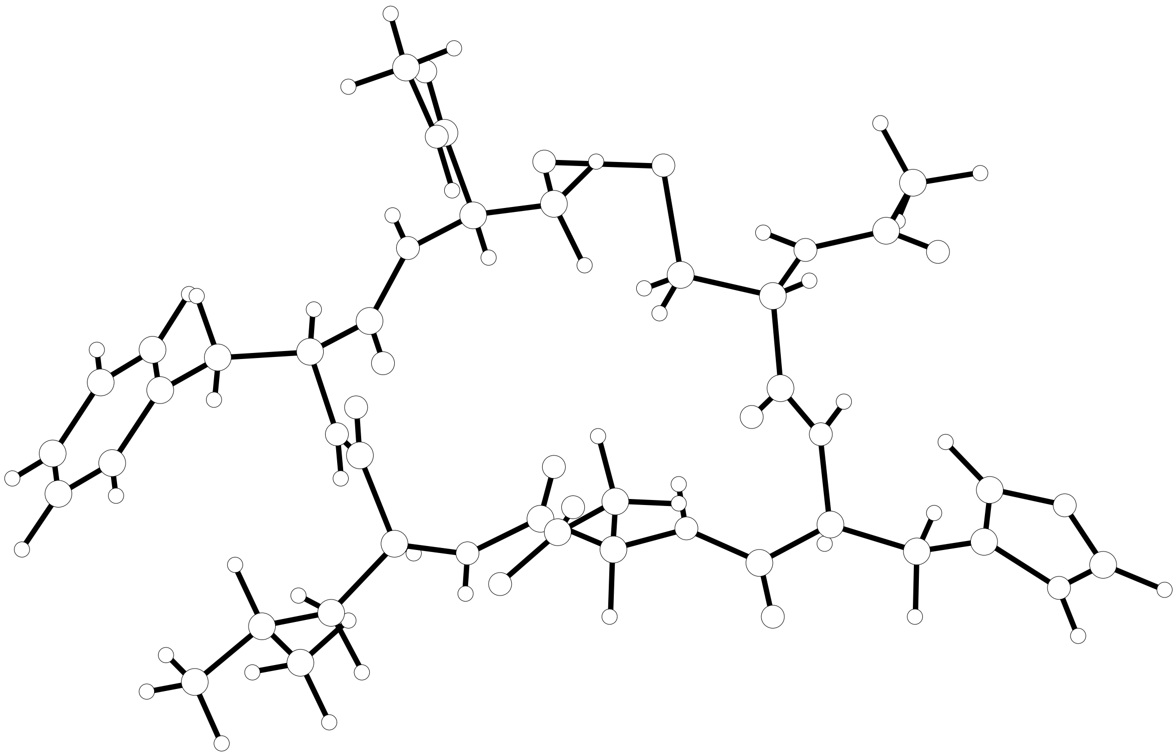
MOIL–View



Documentation for version 8.0, May 22, 1995

Written by Carlos Simmerling, while in the groups of

Ron Elber, at the University of Illinois at Chicago,
Department of Chemistry

Peter Kollman, at the University of California, San Fransisco,
Department of Pharmaceutical Chemistry

MOIL−View V8.0

MOIL−View [1] (**MO**lecules at **IL**linois) is a program for Silicon Graphics and IBM RS/6000 workstations designed to allow the user to view and analyze molecular structures and dynamics. Many of the options have been implemented with the user of MOIL [2] in mind, and some of the routines come directly from the MOIL program. Use of MOIL is not required.

GENERAL FEATURES

mouse and menu interface
read/write coordinates in CHARMm CRD, PDB and AMBER formats
OPTIONAL use of MOIL or AMBER topology files for bonding information
allows rotation, translation etc. of structures
display objects as lines, ribbons and/or spheres, sphere quality variable
two molecules may be loaded and displayed simultaneously
overlap function to minimize the RMSD of selected particles of the two molecules
user definable colors and particle selection
text labels, fixed to particles or screen location
show distances, angles and torsions
find and display hydrogen bonds
postscript output of structures, including ball−and−stick and space filling images
print monomer−based contact maps of molecules
trajectory files may be read and a movie can be shown
print 2D−RMS postscript plots for trajectories, with user−definable overlap selection
cluster analysis of trajectories, postscript plots of clusters, clustering information
trajectory modification− format conversion, frame averaging, particle removal and more
plot trajectory data such as distances, angles, torsions, RMS, Ramachandran plots,
plotting on−screen with options to save postscript plots and raw data
compression algorithm for faster playback of complicated dynamics scenes
support for NTSC/genlock video
solvent molecules can be added to a structure (shell or box)
full FORTRAN source code provided for ease of modification

OBTAINING MOIL−View

The current anonymous ftp site for MOIL−View is munin.ucsf.edu (128.218.14.87). Log in as anonymous, type 'cd pub', and 'get MOIL−View.src.V8.SGI.tar.Z'.

CREATING AN EXECUTABLE FILE

Uncompress the file ('uncompress MOIL−View.src.V8.SGI.tar.Z') , and use 'tar MOIL−View.src.V8.SGI.tar' to extract the source code. A 'MOIL−View8' directory will be created to hold the program files. Use 'cd' to move into this directory. There is a makefile, and all source code (in FORTRAN) is in the 'source' directory. Type 'make' and the compiler will create 'moil−view' in the exe/ directory. To run the program, type moil−view from the exe directory, or move the program to a convenient place (or simply use an alias).

The program was developed and tested most recently on an SGI Indigo2 with XZ graphics. It should function on a variety of graphics configurations, including 8 bitplane systems (Z−buffer highly recommended). IRIX 4.0.1 or newer is required.The makefile has flags that must be set to match your system configuration. If you have an R4x00 machine, change the makefile to include the −mips2 flag. A library flag must be changed for 5.x IRIX. A separate version is available for GL−compatible IBM RS/6000 machines. Read the file README for last minute corrections.

If you do not have the Developer Option with your SGI machine, binary code may be available through ftp. The best performance will be obtained by compiling on your own machine, however.

MINI MANUAL (for those who can't wait)

LEFT button for the tool choices and controlling the molecule's orientation

MIDDLE button to select a particle

RIGHT button for the menus, or for stopping a movie

* HOW TO USE MOIL−VIEW *

MOIL−View is a menu−driven program. Almost all user options are obtained by making choices from the program menu.

Press and hold the **right** mouse button at any time during the main program to bring up the menu. The options are described below. Menu options that are currently selected will have a check mark next to the menu item (for example, a check will appear next to 'Display 2nd molecule' when the 2nd molecule is being displayed). Selecting this option AGAIN will remove the check mark and also reverse the status of the function. When a menu option is unavailable (for example, 'overlap molecules' when no molecules are loaded) it will be gray, and you will be unable to select it until it becomes appropriate.

CONTROLLING THE MOLECULE'S ORIENTATION:

There are two methods of manipulating objects− using the toolbox and using the 'virtual trackball'. When the LEFT mouse button is pressed, the program looks to see if the mouse cursor is inside the toolbox, and performs the appropriate function. If not, the molecule will directly follow the cursor motion based on the status of the 'mouse mode' option. Both of these options are discussed in more detail in later sections.

THE FILE BROWSER

The file browser allows you to move through the directory structure using left mouse button clicks. Click on a directory name to show the files contained in it, click on <parent dir> to move up one directory. If more files are in a directory than can be shown at once, <more files> will be shown at the top or bottom of the list. Click on this line to scroll through the list of filenames.

You may also type directly on the Path: and Name: lines. A white underscore bar is shown at the end of the currently active line. Backspace to delete characters, or use the 'CLEAR' button to clear the text (name only). Press 'OK' when finished, 'Cancel' to quit without making a selection. The 'CLEAR PATH' and "CLEAR FILE' buttons function as the names imply.

Pressing the 'LIST DIRS" button will open a window showing you a list of other directories that you have visited. Press the left button on one of them to replace the current path with the selected on. The directory list is saved in the session file.

The program will warn you if you ask it to write over an existing file, asking for confirmation.

THE TOOLBOX

The toolbox is a small movable window labelled 'Toolbox'. It is set to always bring itself to the front of the main window, even if you try to hide it. This is done so that it will not accidentally get 'lost'. If you wish to remove it, choose the menu option **options/show toolbox.** Choose the same option to turn it back on. A checkmark next to the option will be displayed when the toolbox is on.

The toolbox is activated any time the LEFT mouse button is pressed. If the pointer is over the toolbox at the time of pressing the LEFT button, the structure will be transformed according to the current selection. The rotations and translations are sensitive to the distance of the pointer from the center– the farther away from the center you press, the faster the structure will move.

Size will zoom in and out, lines changes the number of pixels used to draw the bond lines, and reset view returns to the original settings for size, rotation, and translation. It does not affect linewidth. Note that rotations are done in the **screen−fixed** coordinate system, and the axes in the upper corner reflect the original axes in the first molecule.

The toolbox can be moved to wherever is most convenient by simply grabbing the window at the top title bar (using the left button) and then moving and dropping it.

The letters 'R', 'T' and 'S' above the 'Mouse Mode' label allow you to quickly change the mouse mode, as described above. A red box appears behind the currently selected mode.

Keep in mind that the toolbox is active even during a dynamics movie.

******IMPORTANT******

You can control which of the two structures is affected by the toolbox by clicking the left button on the '1' or '2' located just under the 'Molecule #' label. If the molecule is active, a red box will be drawn behind the number. Click again to turn it off. Either one or both of the molecules may be transformed at one time. Reset View, Linewidth, and Size all affect BOTH structures, regardless of the status of the red boxes.

PICKING PARTICLES– THE 'PICK' SYNTAX

The 'pick' commmand is used for particle selection. This syntax is used in MOIL, and also here in MOIL−View for selecting colors, mixed line/spheres, ribbons, hbonds, contact maps and overlap of structures.

Do not confuse this 'pick' with the *selection* of particles using the mouse, which displays information on the screen! That type of picking is described below.

An explanation of the 'pick' syntax from MOIL follows:

example: pick #prt 1 60 | chem mono ALA & chem prtc CA != #mon 2 3 done

The command line includes a series of selections interconnected by logical instructions. It is read from left to right until "done" is detected. No brackets are allowed, and hopefully this will not be too severe a restriction.

The above line stands for (in brackets we wrote the corresponding expression)

pick particles 1 to 60 (#prt 1 60), or (|) like a good chemist (chem) pick all particles of the monomer (mono) alanine (ALA) anywhere this monome is found, but you should pick this under the .and. selection (&) of chemical (chem) particles (prtc) that are called (CA), then you should remove (!=) from the selection monomer numbers (#mono) from 2 (2) to 3 (3). and this is all (done).

The pick line is of the format
pick A lexp B lexp C lexp ... done

<u>*don't forget the 'pick' and 'done'!*</u>

where A−C are selection commands and lexp are logical expressions
note that the line must begin with 'pick' and end with 'done'

logical expressions
A | B – standing for A U B (A union B)
A & B – standing for A /\ B (common parts of A and B are kept)
A != B – only parts of A that are not in B are kept.

selection commands
#prt – defines range of particles to be picked, e.g. #prt 1 6
#mon – defines range of monomers to be picked, e.g. #mon 1 3
chem – An indicator for "chemical" notation will be used.

prtc – select by chemical name of particle
mono – select by chemical name of monomer
examples: "chem prtc CH3E" "chem mono HEME"

Modifications added in MOIL–View include:

new selection commands useful for proteins:
#sid – picks side chain atoms by mono # – anything BUT: C, O, N, H, CA, NH, HN, HA
#bac – picks backbone atoms by mono # – C, O, N, H, CA, NH, HN, HA
#ba2 – picks backbone atoms by mono # – C, CA, N, NH

You may also use '*' as a completion wildcard, eg. CH* would pick CH2, CH, and CH3. You are not allowed to have any characters after the wildcard. Note that *H will **NOT** select 1H and 2H (but will instead act like * and select everything), and C*N will act like C*.

ON–SCREEN PARTICLE SELECTION

The following applies if you have chosen '**Display selection info**' from the options submenu (described below).

Pressing the MIDDLE mouse button while near a particle will place a circle around it. The screen will display the monomer and particle information. Once a particle is selected, the structure can be centered with this particle at the origin (see options submenu above). This centering will remain in effect EVEN DURING A MOVIE. This can be helpful for tracking a particular particle during a dynamics run.

Selecting a second particle will give the distance to the previous selection, and a dotted line will be drawn between the two. Both the line and the displayed distance will be updated each dynamics frame. This can be useful for monitoring distances during dynamics. Note that if the two particles are in different molecules, a line will NOT be drawn, since the molecules are kept in different coordinate systems. Selecting a third and fourth particle will display the current angle and torsion, along with the particle names for the particles that were selected.

SESSION FILES

When you quit the program, your current file names are saved, and used next time you run the program to restore your current session. Also saved are the current bond algorithm, color and sphere choices, centering choices and more. These ’session’ files can also be loaded and saved at any time. The default session file created when you quit is ’save.current’. Remove this file before running the program again if you don’t want to restore the session.

<u>THE MENUS:</u>

Pressing and holding the RIGHT mouse button at any time will bring up the program menu. Note that the menu may not available during the execution of many of the program functions. In the following description of the features, each main menu option is <u>underlined</u>, with submenus in **bold**. Program options are described in the order in which they appear in the menu structure.

<u>FILE:</u>

This has submenu options to read the coordinate and connectivity files. You need to load the coordinate file *before* loading the connectivity file. Currently, the structure can be a CHARMm CRD file, AMBER RESTART file, or a PDB file. Note that AMBER coordinate files have no particle information, and a topology file will be necessary to assign atom types and names. Without this information, dummy atom types will be assigned to all particles (possibly resulting in incorrect bonds, due to an inability to properly determine the covalent radii).

There is also a second load coordinate command, used to load a second structure for comparison. The second molecule does not have to be related to the first. Note that many program functions (dynamics, spheres, hydrogen bonds) function only on the first molecule. The second molecule is mostly useful for overlap and comparison (see below).

When a second molecule is loaded, many menu choices can be applied to either molecule. A window will open and ask you which molecule to change, (1 or 2), and some options will ask (1, 2 or both).

Connectivity (topology) files are *optional* and can be in one of several formats: MOIL topology, AMBER topology, or a simple list of bonded particle pairs preceded by two lines containing the number of particles and number of bonds. For example:

4	(4 particles)
3	(3 bonds)
1 2	(bond 1 to 2)

2 3	(bond 2 to 3)
3 4	(bond 3 to 4)

You can also choose to write the current coordinates (CRD or PDB format) to a file from this menu. *This will apply any rotations or centering that have been made, but not scaling or translation*. The coordinates are also changed after overlap and dynamics. Read the warning below concerning writing coordinates after changing the default file format type.

You may also load and save 'session' files (described above) in this menu.

ABOUT FILE FORMATS: The FILE menu contains a submenu headed 'set file formats'. If you have previously selected a file format, or loaded a session file containing format information, ALL reading/writing of files of that type (eg. coordinates, trajectories, etc.) will be in the selected format, regardless of the original format of the file. This allows maximum flexibility in file format conversion. To change the current format, select the desired option from the menu. The first time you attempt to load a file of a given type, the program will automatically ask about your desired default format. This format will remain in effect until you change it. The program will display an error message if you try to load a file that does not match the currently selected format.

If you are unsure of the current format, check the menu. A check mark is placed next to the current format for each type. In addition, each file requester will inform you of the format it expects to load.

Coordinates and topology for each of the two molecules can be in different formats.

Be careful when writing coordinate files− the write function always uses the CURRENT format, which may be different from that of the original coordinate file (if you have changed it since loading the original file).

OPTIONS:

This menu contains several options that you may change. Many are on/off choices, some are choices between algorithms.

The **2nd molecule** submenu contains options for controlling the second structure. **Display** turns on and off the display of the second structure described above. **Delete** will actually remove the second molecule from memory. **Flash** will toggle the flash mode where the program alternates between the display of the first and second molecule every few seconds. This can be useful for comparison of two structures.The rate of flashing can be set under the *set parameters* menu described below.

Show Toolbox controls the display of the toolbox (for rotations, etc). The default is ON. See 'about the toolbox' below for more information.

Bond Algorithm creates your structure based on 1) a **distance** algorithm in which any two atoms closer than the sum of their covalent radii are bonded, or 2) based on **connectivity** information from the connectivity file. When a coordinate file is loaded the program will ask if you want to use the distance algorithm. If the current connectivity file does not match (in number of particles) a newly loaded coordinate file, the program will also ask if you want to switch to a distance algorithm. If not, you will need to load a new connectivity file. Option 3), **consecutive atoms**, creates bonds between each set of consecutive atoms. This can be useful in some circumstances, such as displaying a coordinate file that contains only the C_a atoms of a protein.

The bond algorithm can be chosen separately for each of the two molecules, and the information will be saved in the session file.

The bond list is assumed constant during dynamics movies, and is NOT recalculated for each frame. You may need to modify the code if you are changing bonds during the dynamics.

Center Object allows you to center the structure at an origin calculated at either the **center of mass (CM)** of the structure, or at a **selected atom**. You must select the particle (see below) and THEN choose 'center on selection' for this to work properly. The centering will remain effective even during a dynamics movie. The centering can be removed using '**don't center**'. This option will work with either molecule.

Note that the modifications made for centering are applied to the coordinates when they are written to a new file. AMBER users must be careful when using this option for periodic systems, which are NOT centered at the origin. It is easier to view the periodic box using centering, but make sure to remove centering and reset the rotation before saving the coordinates to a new file (if you plan to use them for input to AMBER).

Selection Info turns on and off the display of information about the selected particles, such as molecule #, monomer and particle name. It also shows the distance to the last particle that was selected, and the name of the last particle. Selecting three particles will show the angle value, and four selections will return a torsion angle value. The <u>center</u> mouse button is used to actually select the particle. All information is displayed in the lower left corner, and a circle is drawn around the current selection, with a dotted line drawn to the previous selection. Particles from either molecule can be selected.

These particles are NOT necessarily those used to **plot trajectory data**! This is different from the behavior of previous versions of MOIL−View. The particles selected *at the time the plot is made* are used, and even if the selection is changed, the particles used for the plot do not change. This allows the user to select different particles for distances and angles, for example. To change the particle definition for a plot, simply re−open the plot as described below in the **plot** sub−menu.

Objects Are allows display of your structure as **lines** (default), **ribbons** or **spheres** (or any combination). Tool control is still available, but may be slow for a large number of spheres. Try changing the sphere parameters (see below) for faster spheres if you need them. Expect a short delay when switching between object types, especially for large structures. No lines will be shown if 'spheres' is selected.

You may combine spheres and lines by using the '**mix lines/spheres**' command. A requester will ask for a 'pick' command (see below for syntax) and any particles matching your choice will be ADDED to the sphere list. Note that you cannot turn OFF individual spheres, you must turn them all off and add new ones.

Ribbons are created through the '**protein ribbons−add**' command. First you are asked to give a pick command to specify which part of the molecule will have the ribbon. See 'pick' syntax below. Next a requester will ask for the RGB color components of this ribbon. You are allowed to make different ribbons for different parts of the molecule, and they can be different colors (see color options for information on mixing RGB values to form colors). Select the '**remove**' option to remove all ribbons.

The '**spheres are**' menu allows you to choose how spheres are displayed. '**Polygons**' creates standard filled spheres. **Lines** create open spheres and can be useful to see what is within a certain distance (changing the VDW parameter) of a particle. '**Points**' creates a dot pattern for the sphere outline. Finally, '**opaque**' and '**transparent**' control whether the particle transparency (or alpha) values are used. This option only functions on those machines that support transparency. Individual alpha values for particles can be set under the '**color options**' menu. Experiment until you are satisfied with the results.

'**Change Parameters**' allows the user to adjust several parameters, such as the sphere quality level, shading and lighting of objects, and the distance criterion for the hydrogen bond option described below. Each of these choices will be described next.

Sphere quality level can be from 3 to 30, and specifies the tesselation level of the octahedron used to approximate the sphere. Higher number give better quality but slower spheres.

’**Hbond map cutoff**’ changes the cutoff distance for calculating hydrogen bonds. This is the distance between the hydrogen atom and the hydrogen acceptor atom. The A−H...B angle must be between 113 and 180 degrees, and is not currently user definable. Change the numbers in findhb.f if you need to change this value, and then re−compile the program.

’**Contact map cutoff**’ changes the distance used to determine what is a ’contact’ between two point particles. Any monomer pair with any pair of atoms within this distance will be considered to be in contact.

Use shading and **Use lighting** simply toggle these functions, try it and see the difference.

’**Use z−buffer**’ turns on and off the use of the z−buffer. This buffer is essential to the proper display of space−filling objects, but on some machines (such as the Indy) there is a **significant** performance penalty for its use. The default z−buffer is ON with machines that support z−buffering. While important for spheres, it is not as necessary for lines, and you may decide the performance penalty of a software Z−buffer is not worthwhile. Try this option and find out.

’**Show unbonded particles**’ allows the user to choose whether unbonded atoms should be displayed with a ’+’.

’**Change VDW radius**’ allows you to give a ’pick’ command to select a group of particles, then change the VDW radius for the group. This allows you to customize the sphere size to your potential function parameters. Currently these numbers are NOT saved in the session file. If you want a permanent change, modify the numbers in the subroutine ’setparam.f’ and recompile the program.

’**Change flash rate**’ adjust how rapidly the two molecules are interchanged when the flash feature described above is used. This may need to be adjusted to suit your tastes and machine speed.

’**Use depth−cue**’ determines whether depth−cueing is used. If this is selected, the object colors will appear to merge with the background color as the distance from the viewer increases. Change the behavior here if you do not like this effect.

’**Hydrogen Bonds**’ allows hydrogen bonds to be found and displayed using dashed lines. Again you will be allowed to ’pick’ which section of the molecule should be searched for h−bonds. You are also allowed to choose the color to use for the dotted lines (**Pick HB color**). The criterion for valid hbonds can be adjusted: distance through the option described above, but currently the angle must be changed through the code in findhb.f. ’**Turn Off**’ will remove the hydrogen bonds.

The ’**Solvate**’ menu allows you to add solvation to your molecule. This option requires that you already have a file (in CRD, or PDB format) containing the coordinates of the water molecules. This coordinate set will be used to add either a **solvation shell** (of user defined thickness) or a **cubic box** (suitable for periodic boundary conditions). A box outline can be drawn to aid in determining the proper box dimension for your system. After the parameters are provided, the coordinate sets are overlapped and any water molecules that overlap the original molecule are removed. The program also removes any water molecules outside the shell or box boundaries. At this point you should <u>save</u> <u>the</u> <u>coordinates</u> to a new file.

Note that the ’**Draw box outline**’ will provide a box that will be updated during dynamics, if the trajectory file contains this information (AMBER PBC).

’**Movies’** has two options. They are described below.

Generate: Here you can make a movie, just like in **Dynamics** described below, except that each frame will be compressed and saved to your choice of disk file. This file can get VERY large.... This is mostly useful if you have a very complicated scene and not very expensive graphics hardware! The delay time to render each frame will be replaced with the time to decompress and display the file, which is usually much shorter.

In the current version, only the lower left portion (640 by 480 pixels) of the physical screen is saved. This is done to maintain compatibility with early SGI video output devices. Current harware such as Galileo Video can convert the entire screen to NTSC format, but the program does not yet support compression for these devices due to memory constraints. Send me e−mail if you want to try full screen movies on your system.

Watch: This will retrieve and uncompress frames from a disk file made with the **generate**. The playback speed using this method is essentially independent of the complexity of the structure, unlike the actual sphere rendering. This is mostly useful to make videos of large numbers of highly detailed spheres, to be played back at reasonable speeds. If you are lucky enough to have <u>very</u> fast SGI graphics, this may not be necessary. Due to the large size of the movie file, it is most efficient to use a local disk for playback files.

Video Timing: This option is for users of NTSC monitors (genlock or videotape). Choosing this option toggles between hires and NTSC timing. In NTSC mode, only the lower left quadrant of the screen is shown on your NTSC monitor. This mode is useful for recording a dynamics movie to videotape to impress your friends, colleagues, and funding sources. If you do not have a genlock, this can be very annoying if chosen accidentally. Comment the line in source/main.f that controls this command (a call to the toggle() subroutine).

<u>COLOR</u> <u>OPTIONS</u>:

This is where you can choose what colors are used for the particles, and which particles are displayed.

'**User Defined**': using a MOIL−style 'pick' command (described below), you can sssign new colors to any particles. The first three lines in the requester are for the red, green and blue values. You can mix these like paint to form millions of colors (depending on the number of bitplanes in your system). Choose a red, green and blue and then give a pick command on the fourth line. Any particles 'picked' will be given the color you chose. **Choose black (0,0,0) to turn OFF display of a particle with a black background, and white (255,255,255) on a white background.** Lines and spheres for background−color particles are not drawn. This can speed up sphere rendering dramatically. It is sometimes useful to turn everything OFF first (0,0,0) and then turn ON the particles you want to show.

In addition to typing in a specific red, green or blue value, you can adjust the sliding box over the color bars to choose the proper value. A preview of the current color mix is shown underneath the color bars.

Select 'HELP ON' from the color selection window if you need a reminder on the 'pick' syntax. Select 'LIST PICKS' to open a window listing previous pick commands that you have used. Click on one with the left mouse button to transfer it to the current pick: line.

'**Backbone Only**': this will show only the alpha−carbon, carbonyl carbon, and nitrogen of the protein backbone. Currently, it turns off everything else, and leaves the backbone with its current colors. This may give unexpected results if the backbone is not already displayed.

Currently, the atom names that define the backbone are hard−coded in the program as

C, CA, N, H, and O. Modify the code in main.f to change this default behavior.

’**Restore Default**’: this returns you to the default colors. Change the file ’setparam.f’ if you want to change the default values.

’**Transparency**’ allows you to set the alpha values (0−255) which control particle opaqueness (on machines that support this feature). Higher values are more opaque. The transparency feature does not work perfectly. If you don’t have alpha support on your system, dot spheres give a similar effect (see below).

’**Background Color**’: This allows you to choose white or black as the background color. If the background is black, any black (color 0,0,0) particles will NOT be displayed, and if white, color (255,255,255) will not be displayed. *Use these color choices to keep certain portions of the molecule from being displayed, or printed in the postscript option.* Switching between background colors automatically changes the color of particles that are not displayed**.** While the postscript DRAWING option will always be on a plain page background (white), the postscript BITMAP image WILL save the background color.

’**Close Waters Only**’: This turns off all waters not within the hbond cutoff distance from non−water molecules. The cutoff criterion can be adjusted in the options/change parameters menu. Waters are monomers that have WAT or TIP as the first three characters. Updating which waters are ’close’ can be done optionally during dynamics.

’No waters’: This turns off all waters.

’Pick HB color’ performs the same function as described above, and is also placed here simply for convenience.

All color information (for both molecules) is saved in the session files.

OVERLAP MOLECULES:

This allows you to issue a ’pick’ command to select particles to be used to overlap the two molecules. The structures will be overlapped (using the Kabsch[3] algorithm) to minimize the root mean square deviation (RMSD) of the picked particles. Any errors, such as zero determinant of the transformation matrix, will be written to standard output. After overlap, both structures will be active in the toolbox (see below). The RMS deviation between the picked particles in the overlapped structures will be displayed.

You can use the same 'pick' command for both molecules, which will apply the same pick to each molecule. Note that this does NOT mean that you will pick the same particle numbers, or even the same number of particles. An example may be if you pick all alpha carbons, and the two molecules have different sequences.

You may also choose DIFFERENT pick commands for each molecule. An example of this might be to overlap the final segment of two proteins that differ in initial segments (eg. overlap monomers 24−25 of one structure with monomers 28−29 of the other structure).

Select the 'HELP ON' box from the overlap window if you need a reminder of the pick syntax. Select 'LAST PICKS' for a list of previous picks, and use the left mouse button to selct one of them as the current pick command.

This same basic procedure also applies to defining the overlap for the trajectory overlap plot, 2D−RMSD plot and cluster analysis.

DYNAMICS:

This option allows you to use a CHARMm (or MOIL) DCD type, MOIL PATH, or AMBER trajectory file compatible with the current structure to view dynamics (for molecule 1 only). After selecting the file using the file browser, a window will request you to give the delay between frames, start and end frame number, number of frames to skip between each frame displayed. DCD files contain information such as total number of frames in the trajectory, but AMBER files do not. Only formatted AMBER trajectory files are supported, while DCD and PATH files are unformatted (binary).

You may also choose whether to 'cycle' the movie, which returns to the first frame after the last, continuing until the user interrupts.

To pause the movie, press "P" (capital p) on the keyboard. During pause, the molecule orientation may be changed. Press 'Q' to quit the dynamics or 'C' to continue. In addition, pressing the RIGHT mouse button during dynamics will bring up a window asking if you want to cancel or continue the movie.

When you discontinue the movie, the current structure is no longer the INITIAL coordinate file, but the LAST SHOWN from the dynamics file. Re−load the coordinate file if you want the original coordinates. During the movie, the current frame number is displayed in the bottom right corner.

Any centering choices will be repeated on each dynamics frame.

If a distance line is shown, it (and the calculated distance shown) will be updated for each frame. Angle and Torsion values will also be updated for each frame.

Ribbon positions will be updated during dynamics, but since hydrogen bonds can be slow to calculate, the user can decide whether to update these bonds during dynamics. If hbonds are active, a requester will open and allow the user to make this choice.

The user can also decide whether to update the list of 'close' waters for each dynamics frame if **color options/close waters** was previously selected. This update can be slow for large systems.

If a second molecule exists and an overlap has been performed, the program will ask if the user wants to repeat the overlap procedure for each dynamics frame. This will be done automatically if the user is plotting overlap data for the trajectory (see **plot trajectory data** below).

It is important to note that, during a movie, coordinates for all frames are not being saved, but the movie is shown WHILE the dynamics file is being read. This may change in the future. Some performance deterioration may be noticed for large structures, ESPECIALLY IF THE TRAJECTORY FILE IS ACCESSED OVER THE NETWORK.

See the sections below on plotting trajectory data, clustering and modifying trajectories for more information..

POSTSCRIPT DRAWING:

There are two basic options, **normal** or **ball and stick**. Normal drawings are described below. Ball and stick drawings are similar except that each particle will be drawn as a circle 1/10 the size of its VDW radius. This circle is drawn in addition to the lines connecting bonded particles.

This option generates postscript output for a simple hardcopy of the display. The structure will be in the current view, using sticks. Circles are used for any particles that were spheres. You will be asked whether you want to depth−cue the image. If you choose this option, the lines for particles farther from the screen will appear lighter, fading into the page. You can also choose to use the particle colors, or use black for all particles. Note that, even if particle colors are NOT used, particles that are the background color WILL NOT be printed. Color choice here means that particle colors will be taken into account, otherwise all particles are the same color. Whether you actually get color depends on your printer. Colors will show as gray scales on B&W printers, and full color if your printer supports it.

The thickness of the sticks is related to the screeen line width chosen in the toolbox. Ribbons, spheres and hydrogen bonds will be printed if they are displayed on the screen. Text labels will also be printed (see below).

The current postscript plotting is not very sophisticated...

After creating the postscript file, you must send it to your printer using 'lpr' or whatever command is appropriate on your system.

POSTSCRIPT BITMAP IMAGE:

This option will take the entire displayed image, and convert it to a postscript bitmap file. These files are VERY large− 2 megabytes for B&W and up to 6mb for color. In addition, they tend to take a very long time to print. The advantage to this method is the the image will be exactly like your screen image (within printer limitations), with proper sphere shading and interpenetration. It makes little difference for line objects. Again, you are responsible for sending the postscript file to your printer (and deleting it when you are finished!).

Note that the color/gray option here determines whether full color information is saved in the postscript file, or only gray scales. This is different from the postscript drawing option, where the 'color' choice determined whether particle colors were used at all, and a single file was appropriate for color or black−and−white printers. Here, particle colors are always used. Full color bitmap files are 6+ megabytes in size, and gray scale only 2+ megabytes, so choose the type that is appropriate for your printer.

CONTACT MAP:

This will generate a postscript file for a simple contact map based on particle distances. You issue a 'pick' command for the range of monomers to use for the map, and the program calculates distances between particle pairs in this pick. A postscript file is generated which shows the pattern of MONOMER contacts. Any monomer pair with at least one particle pair touching will be considered to be in contact. The distance used for contact can be changed in the 'change parameters' submenu (see above). As with the postscript files described above, you need to send the file to your printer using the command appropriate on your system.

TRAJECTORY 2D−RMS:

This option will call the standard dynamics routine described above, but data will be collected for the RMSD of each frame in the trajectory to every other frame. If you do not have a currently defined overlap from **overlap molecules**, the program will prompt you for a pick command to select the particles to overlap. A second molecule is not required for this option− if it does not exist, the program will ask if you want to load it or simply use the first coordinate set as the reference

It is important to note that the program must save the coordinates of each of the 'picked' atoms (not ALL atoms) for each frame of the trajectory. The limits on the number of atoms and number of frames for this option are defined in the source code in LENGTH.BLOCK, and recompiling is necessary to change the default sizes. The parameters are named MAXRMSPT and MAXRMSFR.

The program will use the standard file browser to obtain the trajectory file name, and then present the user with the standard dynamics options described previously. Skipping frames results in a smaller plot that covers the same time period, but with reduced resolution− a quick fix for trajectories that exceed MAXRMSFR.

After the trajectory has been read and displayed, the file browser will request a name for the new postscript file. Next, you must choose whether to print larger RMSD values as increasingly darker or lighter shades of gray. You may also choose 'smooth' shading where the RMSD values are directly mapped to gray values, or discrete shading where four ranges of RMSD values are used. In the case of smooth shading, you must choose a cutoff value for the high end of the scale− all values above this will be considered the same, and those below it will be smoothly shaded. For discrete shading, you must choose three cutoff values, which result in the ranges: 0−LOW, LOW−MIDDLE, MIDDLE−HIGH, and HIGH+. In either case, a legend is printed on the page for reference, along with the appropriate file names and overlap information.

Once again, you are responsible for sending the file to your printer/previewer. The program will ask if you want another plot (changing shading or cutoff options), and allow you to create a new plot without needing to re−calculate the RMSD matrix.

Finally, you are asked if you want to save the RMSD data directly to a file. This data may be useful as input for other programs. The data is saved one value to a line, starting with frame 1 and continuing to the last frame, printing the RMSD to each frame (including itself). This provides (number of frames)^2 data points.

CLUSTER TRAJECTORY:

The program provides a simple cluster analysis function, based on a 2D−RMSD matrix generated exactly as described in the preceding section. A pick command must be provided, and the trajectory file is selected. The standard dynamics options are requested. After the dynamics are completed, the RMSD values are calculated. You may then save the entire data set to a file as described previously.

Each frame is initially placed in a cluster by itself. The actual clustering then proceeds as follows: The program prompts the user for a cutoff value to be used in the clustering. For each pair of clusters, the average RMSD between all pairs of structures in the two clusters is calculated from the RMSD matrix data. We then focus on the pair with the lowest average RMSD (the most similar pair). If the average RMSD of this pair is less than the specified cutoff value, the two clusters are combined. This process of calculating average RMSD and combining clusters is repeated until all cluster pairs have an average RMSD greater than the cutoff value. Then, a 'representative' structure is found for each cluster. This structure is the one that has the lowest average RMSD to all other members of the cluster to which it belongs.

The information about the number of clusters found and the number of structures in each cluster is written to the <u>standard</u> <u>output</u> in the present version. The user may then decide to try a different cutoff value to see how the clustering is affected. If another cutoff is not desired, selecting 'NO' will open another window presenting four options: PS1, PS2, TRAJ and NONE. Each will be described in detail below. While this window is displayed, waiting for the user choice, the molecule may be rotated, scaled, etc. just as it normally can be from inside the main program. This is important for re−orienting the molecule for postscript plots, as will be described below

The first option, 'PS1', will save a single cluster to a postscript plot, identical to that obtained using **postscript drawing**. The program will ask whether the user wants to plot ALL of the structures in the cluster, or only the single 'representative' structure. The program then asks which cluster number to print (from the clustering info written to the standard output), a name for the postscript file, and the standard postscript plot options. The program will then go back through the original trajectory file and extract the full coordinate set for each of the desired structures, and then write the plot to the file.

It is important to note that this postscript plot will use the current view of the molecule, and also the current color selections, just as the standard **postscript drawing** would use. It is not currently possible to change the color selection while inside the clustering routine, but the view can be changed *while the option window is open.*

The second option after clustering is labelled 'PS6'. This will create a postscript drawing similarly to that just described, with the exception that instead of a single cluster, the 6 most populated clusters will be drawn separately on a single page. Each cluster will be labelled with cluster number and number of members in the cluster. Again the current colors and view of the molecule are used.

The next option 'TRAJ', saves all of the structures in a single cluster to a new trajectory file. The program will request the cluster number and a new trajectory filename and format. This trajectory file can then be used as input to re−cluster or sub−cluster the original cluster(s).

After any of the options has been performed (except 'NONE') the program will again present the same options. In this manner the user can create multiple postscript plots from different views, plots of several clusters, and so on.

If the last option, 'NONE', is selected, the program will ask if the user wants to save some of the clustering information to a file. This information includes cutoff values, RMSD values between representative structures of cluster pairs, the cluster number to which each of the structures belongs, the RMSD of each structure to the representative structure of its cluster and the RMSD value between each pair of clusters that were combined. Examination of these RMSD values can be useful in determining an 'natural' cutoff value.

At this point the program will ask if the user wants to try a different cutoff value. If so, the program will re−cluster and present the same options for saving postscript files, trajectories and cluster information. Selecting 'NO' will return the user to the main program.

Be patient− due to the large number of calculations that are necessary, clustering can be quite slow. Try clustering a smaller number of frames.

MODIFY TRAJECTORY:

This option allows you to change your trajectory file in many ways. The options described below can also be combined to modify several aspects of the file in a single pass.

The program will first ask if you want to save the modified trajectory to a new file. If you select 'no', the new trajectory will be displayed but not saved.

Next, the program will ask if you want to save the trajectory to a postscript file. This will create a printable file with all of the structures of the trajectory superimposed on a single page. Output is similar to that obtained for a single structure in the postscript drawing option described above. You may also specifiy that the program should skip a certain number of frames between each saving to the postscript file. This can be useful for paring down very large trajectories that contain too many frames to reasonably view on a single page.

Next the program will ask if you want the trajectory overlapped. This functions just like the single−frame **overlap trajectory** function described previously. If a second coordinate set is loaded, it will be used as the reference coordinates for the overlap. If one is not present, you may load a second molecule OR you may simply use the first coordinate set as the reference (copying it to the second). The program will prompt you to make a choice. As usual, pick commands are required to specify the particles to overlap as was described in the section on simple overlaps above.

Next, a cutoff value can be specified such that only structures with an RMSD less than the cutoff will be saved (in the postscript file and in the new trajectory file).

The next option is to decide which atoms should have coordinates saved in the new trajectory file. One possible use is to save only non−water atoms to reduce the size of a solvated trajectory. Keep in mind that such a trajectory will no longer match the coordinate and topology files that are used for the original trajectory, and you may need to create a new coordinate and/or topology file for the program to be able to use this new trajectory. A standard pick command is used to choose which atoms to SAVE.

The file browser now opens to allow you to choose the OLD trajectory file. Select it as you would any file.

The next option is whether to average frames for the trajectory. A running average is used, and the user may specify how many frames should be combined to produce the running average. For example, if the user chooses a value of 5, frame #1 in the new trajectory will be an average of frames 1 through 5, frame #2 will be 2 through 6, etc.

The standard dynamics option window opens next, allowing you to choose which part of the original trajectory should be used. Skipping frames will result in a smaller trajectory file that still covers the same time period as the original but with reduced resolution.

The output format of the new trajectory file can be chosen next. Note that writing an AMBER PBC format file when the original trajectory did not contain box size information will *not* place correct box sizes in the file. In some cases, however, you may need to have a file in this format so the option is still provided.

Finally, the filename of the new trajectory file is specified using the file browser. Be careful not to overwrite the original file, since the old file is still required during the modification process.

If you asked for a postscript file to be created, a filename will now be requested. You will also be prompted for the usual options for postscript files, such as color and depth−cue.

You are NOT able to change the molecule's orientation during **modify trajectory**, since changing the relative orientation mid−trajectory would adversely affect the postscript file and overlap options.

PLOT TRAJECTORY DATA:

This menu allows you to **create** or **remove** windows that plot data from a dynamics trajectory. **Distances, angles** and **torsions** (such as are already displayed in the lower left corner of the screen) can be plotted (for molecule # 1 only, however). After opening a plot, the program will ask whether points or lines should be used to draw the data. Points may more useful than lines for torsion values where, for example, a line between + 180 and −180 degrees may not be drawn in the direction of the true transition.

If desired, the overlap **RMS** can also be plotted as a function of the trajectory time. In this case a second coordinate set (for overlap) should be loaded, and the 'overlap molecules' menu choice should have been used to 'pick' the particles in each molecule for overlap. If a second coordinate or an overlap choice have not been defined, the program will allow the user to do so at this point. The user may also use the current first molecule as the reference if a second molecule is not loaded.

A Ramachandran−style plot can be made (phi and psi angles for a single protein residue). The user has the option of retaining all data points for the full trajectory (providing a plot of the cumulative conformational space sampled), or displaying only the current point, to show the instantaneous sampling. When the plot is opened, the program looks for consecutive bonds of the type C−N−CA−C (phi) and N−CA−C−N (psi) in which the CA atom is *in the same monomer as the currently selected particle*. If there are problems in the monomer numbering or particle names, the program may not be able to find these torsions, and attempting to create the plot will be unsuccessful.

It is important to note that the data in the window represents the particles that were chosen *at the time the plot was created*. These particles are displayed in the title bar of each plot window. In this manner, the on screen particle selection can be changed without affecting a current plot. If the user decides to plot a new set of data, the plot should be re−opened after selecting the new particles, and the trajectory should be viewed again to make the data represent the new particle selection.

In addition to data calculated from the trajectory, the user can load an **external data** file to plot. This can be useful to monitor the trajectory while comparing to data that can not be calculated by MOIL−View. The data file must have a single data value per line.

After drawing a plot of the data, data can be saved to disk using the **save raw data** menu option. These files can then be used for further analysis or plotting outside of MOIL−View. The data file will contain the frame number and data value for each point, with one point per line.

Simple postscript plots can also be created to print a hardcopy of the plot using the **save postscript plot** menu option. As with previous drawings, you need to send the postscript file to the printer by yourself.

Only one plot of each type can be active at a time− selecting ’create plot’ when the plot is already active will simply close the plot and re−open it (possibly using a new particle selection). This does allow you to change the plot options, such as points vs. lines and the particles that are used to define the plotted value.

ADD TEXT:

Text labels can be placed anywhere on the MOIL−View screen. There are two types of labels− those that are attached to a particle (and follow its movements), and those that are in a fixed position on the screen. Color can be selected for each label, but currently *the text size* *is fixed*. The size that is used in the postscript output can be chosen, but this will not affect the size of the screen label. This may be changed in a future version of the program.

'**Fix to picked atom**' will allow the user to select a particle using the middle mouse button, and then enter a text line for the label. '**Fixed to screen position**' will allow the user to type text and then move it around the screen until the desired position is determined. '**Remove one label**' allows the user to select a label using the middle mouse button, and then remove it. '**Remove all labels**' will do exactly that. Text labels are NOT saved in session files.

Follow the on−screen prompts to obtain detailed instructions for creating/removing labels.

MOUSE MODE:

This menu allows you to choose what function to perform when the left mouse button is held down and the mouse is moved. The default function is to perform a **rotation** on a 'virtual sphere' centered at the origin. This allows for a more intuitive and accurate rotation of the molecule than is possible using separate x, y and z rotations from the toolbox.

Other possible choice for this function are **translation**, **scaling** or **no function** (like earlier versions of MOIL−View). When translation is selected, the structure will follow the mouse pointer when the left button is held down. When scaling is selected, motion toward the center of the window will reduce the size of the structure, motion toward the edges of the window increases the size.

This mode choice can also be made directly from the toolbox, by pressing the left button over the 'R' (rotation, 'T' (translation) or 'S' (scaling) letters at the top of the toolbox. A red box will appear under the current function. Pressing the button over a letter already selected will remove the red box and return the mouse mode to 'no function'.

If the left mouse button is pressed while the cursor is over the toolbox, the toolbox function will be performed regardless of the status of the mouse mode.

CLEAR ALL:

This option is useful if things gets confusing, or if you have trouble displaying a new structure after using something else previously. It resets all particle numbers to zero, and erases all display objects. It can be better than quitting and re−starting, since quitting will save your session. Quit and remove the 'save.current' file before restarting if things get really bad. E−mail me if you have to, but keep in mind how much you paid for the software when deciding if the response was fast enough!

QUIT:

When you quit, a file is made containing your session information as described above. This will be restored next time you run the program. See 'loading saved session' above. Session files currently save particle colors, sphere and ribbon information, centering, coordinate and connectivity filenames and formats, dynamics information, toolbox status and more. See savesess.f for details.

Enjoy the program (bugs and all!) and please feel free to send me any suggestions, bug reports or improvements that you have made.

Please use reference [1] below if you include any images or data generated by MOIL−View in a publication.

Carlos Simmerling, Ph.D.
University of California at San Fransisco
Department of Pharmaceutical Chemistry
Box 0446
San Fransisco, CA 94143

carlos@hugin.ucsf.edu

REFERENCES:

[1] Simmerling, C., Elber, R. and Zhang, J., MOIL−View − A Program for Visualization of Structure and Dynamics of Biomolecules and STO− A Program for Computing Stochastic Paths, in *Modelling of Biomolecular Structure and Mechanisms*, A. Pullman et al. (eds.) 241−265, 1995, Kluwer, Netherlands

[2] Elber, R., Roitberg, A., Simmerling, C., Goldstein, R., Verkhivker, G., Li, H. and

Ulitsky, A., MOIL: A Molecular Dynamics Program with Emphasis on Conformational Searches and Reaction Path Calculations, in *Statistical Mechanics, Protein Structure, and Protein Substrate Interactions*, S. Doniach (ed.), 165−191, 1994, Plenum, NY.

[3] W. Kabsch, Acta. Cryst., A34, 827 (1978).