Molecular Modeling Cheat Sheets

September 2009 - Andrea Minoia

This document wants to be a first aid for beginners in molecular modeling. I will not discuss here the theory behind molecular modeling, I will rather focus on the practical aspects.

List of topics:

Tinker ‘s cheat sheet

Tinker is a free molecular modeling package for molecular mechanics and molecular dynamics simulations. Tinker 4.2 and Tinker 5 are available in the lab. In the table below, I have tried to summarize what I think the pros and the cons of Tinker are.

PROS	CONS
easy to use	serial code only
many “ready-to-use” force fields	slowly updated
results easy to be analyzed	poor support (community, manual, ...)
nice group function	poor rigid-body capabilities

Tinker 4.2
The main use of Tinker is to minimize molecular structures, search for the and run molecular dynamics (MD) simulation. Tinker is a collection of programs and the most important (i.e., most commonly used) are:

Energy minimization algorithms :
- minimize
- newton
- optimize (for small structures only)
- minrigid (minimize with rigid body)
- optrigid (optimize with rigid body)

Molecular dynamics algorithms:
- dynamic

Analysis algorithms:
- analyze (energetic analysis)
- radial (radial distribution function)

Structure manipulation algorithms:
- xyzedit

Most common force fields are:
- mm2/mm3
- oplsaa
- amber
- charmm

In general, to execute any of the Tinker programs you will need a valid structure and a key file.
Here an example of tinker input file:

4016

1 C -31.3176308 -36.2778015 -2.63730001 50 2 31 1771

2 C -29.8976307 -36.2778015 -2.63730001 50 1 32 1772

3 C -27.0576305 -36.2778015 -2.63730001 50 4 33 1773

4 C -25.6376305 -36.2778015 -2.63730001 50 3 34 1774

5 C -22.7976303 -36.2778015 -2.63730001 50 6 35 1775

6 C -21.3776302 -36.2778015 -2.63730001 50 5 36 1776

7 C -18.5376301 -36.2778015 -2.63730001 50 8 37 1777

…

The first line contains the total number of atoms N in the structure. This line is followed by N lines (one for each atom) containing:

•the progressive number of the atom in the structure,

•the atomic symbol, the X-, Y- and Z-coordinates,

•the atom type for the chosen force field,

•the connectivity, i.e. the progressive number for the atoms bounded to the current atom.

A key file is a file containing all the parameters for the simulation: force field, additional potentials, cutoffs, size of periodic cell, groups, etc. Below and example:

parameters ../params/mm2

a-axis 25.4553
b-axis 29.3946
c-axis 26.8688

cutoff 12.0
polymer-cutoff 5.5

bond 50 50 8.000 1.410
angle 50 50 50 0.500 120.0
torsion 50 50 50 50 -0.930 0.0 1 8.000 180.0 2

See the Tinker manual for more informations.

FAQs

How can I freeze part of my system?
Freeze part of the system in Tinker means that the frozen atoms cannot move and that they do NOT interact with other frozen atoms. If this is what you mean by “freeze”, then use the keyword INACTIVE in the key file.
E.g.

INACTIVE 1 5 10 20	this will freeze atoms 1, 5, 10 and 20
INACTIVE -1 100	this will freeze all the atoms from 1 to 100

Tip: Useful to freeze surfaces like graphite or structures like carbon nanotubes.

How can I perform a rigid body calculation?
A rigid body calculations means that all the atoms in the rigid body maintain a specific reciprocal orientation during the simulation. A rigid body calculation in Tinker implies divide the system in groups of atoms and ask for computing the inter-group interactions only.
E.g.:

GROUP 1 -1 10
GROUP 2 -11 20
GROUP-INTER

Indeed, usually a rigid body simulations is done considering as rigid body each molecule in the system. You can do easily this with Tinker as following:

GROUP-MOLECULE
GROUP INTER

Once the rigid bodies have been defined, you have to change the integration algorithm used by Tinker. To do this, in the key file you need to add the line

INTEGRATE RIGIDBODY

Note: mind that when you use groups, atoms cannot be in more than one group and then the number of atoms in all the groups must be equal to the number of atoms in the structure. The entire system has to be divided in groups. Also, mind to use the appropriate program for minimizations (minrigid, optrigid)

How can I perform a stochastic MD?
Just specify the proper integration algorithm in the key file:

INTEGRATE STOCHASTIC

Tip: have a look also to the keyword FRICTION, to modify the the frictional coefficient in ps^{-1} for use with stochastic dynamics.

Is there any algorithm similar to LINCS?
Have a look to the RATTLE keyword

How can I restrain some geometrical parameters?
Have a look to the following keywords:

RESTRAIN-DISTANCE
RESTRAIN-ANGLE
RESTRAIN-POSITION
RESTRAIN-GROUP
RESTRAIN-TORSION

Which Thermostat and Barostats are availble?
Berendsen and Andersen thermostats and barostats are available. The strength of the coupling can be modified via the keywords TAU-PRESSURE and TAU-TEMPERATURE.

Tinker is complaining about smaller parameters?
Tinker is written in Fortran without dynamic allocation of memory, that is that the size of all vectors and matrices is hard coded in the code. Most of the parameters you may need to increase are in the files sizes.i or in ewreg,i, both in the tinker source directory. Once you have increased the parameters by editing the files, you need to recompile tinker. In the virtual machine you have already the right scripts for the compiler used (g77), so that you only need to type at the prompt the following commands:

```
rm -rf *.o *.a *.x
./compile.make
./library.make
./link.make
./rename.make
```

That should to the job.

How can I run TINKER calculations in Namur (pbs queue) ?
You have create a script to submit to the queuing system. The script sounds like this:

```
#!/bin/bash
#$ -N at1-1
#$ -cwd
#$ -l virtual_free=1G
#$ -l h_cpu=168:00:00
#$ -j y
. /cluster/apps/modules/etc/profile.modules
module load common

dynamic cntatact1.xyz -k polycnt.key 500000 2 0.5 2 473 > dyn.out
```

Return to the list of topics

YASC scripts

YASC -Yet Another Script Collection- is a collection of some of the scripts I have written, in (too) many languages, during these years. Here a list of the main scripts in YASC:

yascgui simple gui in python for some of the scripts (currently under development):
molgeom polybuild buildCstruct
fixtnkatypes modtnkxyz fragpdb
vol2mol

molgeom calculate distances, angles and dihedrals angles in molecular systems.

polybuild builder for single polymeric chain. It use a library of repeating units to build the polymer chain with the proper tacticity.

builCstruct builder for sheets of graphite HOPG and for armchair and zigzag carbon nanotubes. Periodic structures can be built.

fixtnkatypes utility to assign atom types in tinker structures according to tinker templates or user specified atom types.

modtnkxyz utility to modify atomic coordinates in tinker files according to coordinates in xyz files.

fragpdb utility to make each molecule in the system a fragment for the pdb file

vol2mol utility to calculate the number of molecules in a given volume at a given density

zoa2tnk utility to convert zxf zoa file in tinker files. By using zoa group files, it can assign the proper atom types to the atoms in the system.

microtact tool to analyze the microtacticity of a polymer chain.

xplot plot on the fly a data distribution using gnuplot.

fan3dmap analyze 3D potential energy surface.

gro2tnk_oplsaa excel spreadsheet to convert Gromacs oplsaa parameters in Tinker oplsaa parameters.

Many other scripts way too messy to be written here, but feel free to ask me if I have already something to do “this” and “that” ...

Return to the list of topics

Some other software

Short list of useful and common (free) Software
There many software useful in molecular modeling. Some of these are:

Molecular Viewers:

VMD	one of the best viewer with some editing capabilities
Molden	viewer with nice and useful Z-matrix editor

Utilities for Chemistry:

babel (open babel)	translate between chemical formats with 3d coordinates generation capability

Plot

Gnuplot	2D/3D plot, fitting and more
Grace	2D plot, fitting and more
g3data	extract numerical values from a plot

Molecular Modeling Packages:

Tinker (4.2 and 5.0)	MM/MD molecular package
Gromacs	MM/MD molecular package
DLPOLY	MM/MD molecular package
Gamess	QC program like Gaussian

Programming:

Komodo edit	program editor
ipython	useful python shell

Short list of useful and common (commercial) software

Molecular Editors:

MStudio	3D editor, viewer and much more
Zoa	a bit of everything
Gaussview	viewer and 3D editor from Gaussian

Plot

Sigmaplot	2D/3D plot, fitting and more
Origin	2D/3D plot, fitting and more

Molecular Modeling Packages:

MStudio	QC/MM/MD molecular package
Gaussian03	QC program

Return to the list of topics

VMD's Cheat Sheet

VMD is one of the most versatile molecular viewer: it is free, it runs on Windows, Linux and Mac OS X, there is a huge community and a lot of plugins and scripts (the Tcl console allow you to write your own script) and, last but not least, it reads a lot (A LOT) of formats.

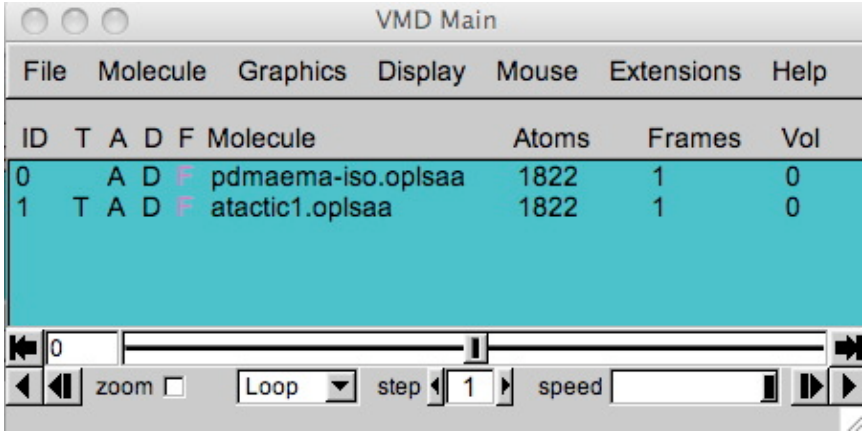
Open VMD from terminal:

vmd	run vmd
vmd file.xyz	open file with VMD. Mind that VMD tries to guess the file format from the file name. In this case VMD tries to open the file as xyz file.
vmd -tinker file.xyz	to open tinker files, specify the format with the option -tinker

Some shortcuts when focus on the view window:

1, 2, 3 or 4	select an atom, get the distance between two selected atoms, get the angle between three selected atoms or get the dihedral angles between four selected atoms, respectively.
7	allow to grab a fragment or molecule and displace it
=	reset view
s	zoom (also mouse wheel does this)
r	return in rotate mode
t	translate the entire system

The VMD main window



A number of controls are available directly from the main window. This windows can be divided in three main parts:

Menu Bar	Browse between different functions
Loaded molecules	display the molecules loaded in VMD
Trajectory controls	Control the trajectory

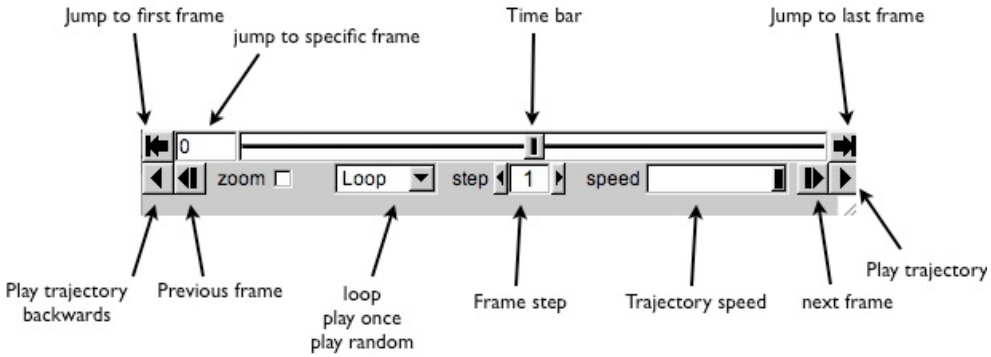
Useful info about Loaded Molecules part:

ID	id of the molecule. Useful to know when you make scripts
T	time line. Put the time for trajectory on a molecule
D	toggle on/off that molecule from the view windows
F	freeze the molecule: zoom, rotation, translation acts only on molecules that are not frozen
Molecule	name of the molecule

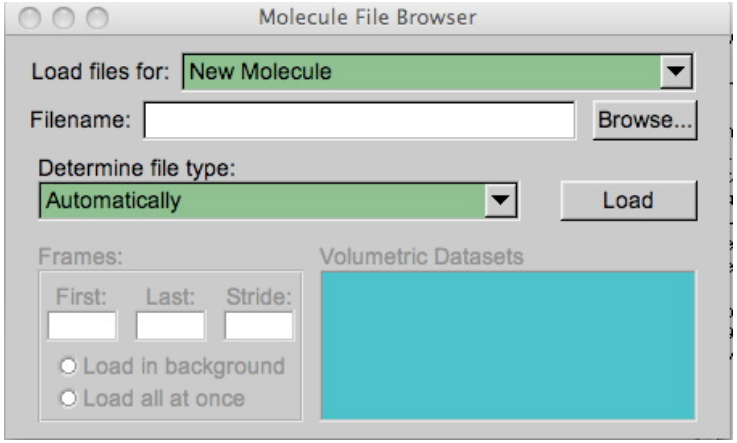
Atoms number of atoms in the molecule

Frames number of frames (a trajectory has more than one frame)

Trajectory controls



Load molecules in VMD (File -> New molecule):
To load a molecule in VMD you can also right-click on the main window



Load files for:

New Molecule	load a new molecule
Choose a loaded molecule	append new molecule to the selected one

Filename/**Browse**... select the file

Determine file type:

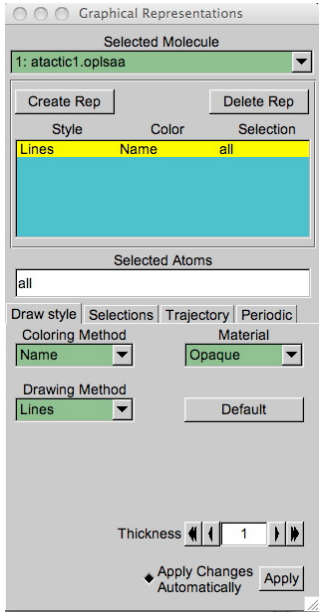
Automatically	let VMD try to guess the format or open the menu to choose manually the file format.

Frames:

load all at once	load a trajectory all in once, without showing all the frames
first/last/stride	load a part of trajectory

Load load the molecule

Graphical representations (Graphics -> Representations):
Another useful feature of VMD is the possibility to display only part of the molecules (e.g., specific fragments, atom types or atoms with specific atomic indices, ...) and to modify the style to display molecules or parts of it.



Use selected molecule menu to apply the changing to a specific molecule.

Create Rep/Delete Rep are use to create replica of the selected style or delete it

Selected atoms: choose the atom to be selected for the current representation:
E.g.:

all	all atoms
fragment x	only fragment x
not fragment x	all but fragment x
type C	all Carbon atoms C

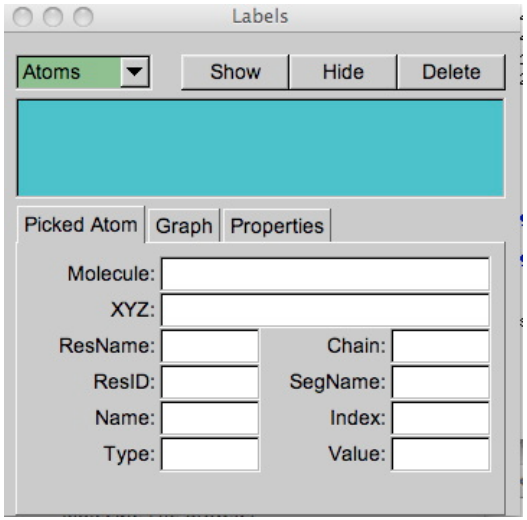
index 20 21 22 only atoms 21, 22 and 23 in the structure.
Mind that numbering in VMD starts from zero, not one!

More selection keywords can be combined using AND logical operator:
e.g.:
fragment 0 *and type* C

Some coloring methods:
Name based on the name
Color use only one color among those available

Drawing method: choose a style (VdW, lines,Bond, CPK, H-bonds, ...)
Depending on the style, some controls are displayed in order to enhance the graphical representation, e.g. sphere/bond resolution, line thickness, ..

Labels (Graphics -> Labels):
Use labels windows to delete/hide labels on the view and/or retrieve information about selected atoms, distances, angles and dihedral angles.



This is quite self explaining...

Interesting is also the Graph submenu, which allows to draw a plot for the values assumed by the selected distance, angle or dihedral angle along the trajectory. The values can be saved in a text file. **Mind that the distances, angles and dihedrals are calculated considering the system NOT periodic**, so be careful to pbc artifacts: you may want to consider to use molgeom instead.

FAQs

How can I display the periodic images?
Go to Graphics -> Representations and choose the Periodic tab. here you can check the directions where display the images and their number. Mind that not all formats can display the periodic images in this way. To be able to do so, you need to put the following script in the Tk Console (Extensions -> Tk Console):

```
#pick ID for current molecule
set molid [molinfo top]

#pick the number of frames
set n   [molinfo $molid get numframes]

#setting a, b, c in Angstrom and alpha, beta, gamma in degrees for
#your box
set a 72.4
set b 145.111
#   set c 57.0
set alpha 90.
set beta 90.
#   set gamma 90.

# Applying to all the frames
   for {set i 0} {$i < $n} {incr i} {
      molinfo $molid set frame $i
      molinfo $molid set a $a
      molinfo $molid set b $b
#       molinfo $molid set c $c
      molinfo $molid set alpha $alpha
      molinfo $molid set beta $beta
#       molinfo $molid set gamma $gamma

   }
#rewind trajectory
animate goto start
```

Lines started with # are comments. This is a really basic script and you can easily improve it. After executing the script, you should be able to display the periodic images from the Periodic tab in Graphical Representations

I have moved a fragment, how can I save the new system?
File -> Save State is not the way to go. This will save the VMD state (colors, styles, views orientations, ...). Instead right-click on the molecular name in the main window and choose Save Coordinates. You can save in a number of different formats, including PDB and XYZ.

How can I make an image out of my view?
File -> Render is the way to go. Choose a render mode (snapshot or pov3 are the most common for me), a name for the image and Start Rendering to create the image.

How can I make a movie of my trajectory?
Go to Extensions -> Visualize -> Movie maker. Choose the working directory, the name of the movie. Go to Movie Settings and uncheck the option 4 (Delete image file), and check Trajectory instead of the default Rock and Roll option. Click Make movie. Open a new terminal and go to the working directory of movie maker: you will find the movie and a bunch oh files. Delete the movie and leave the files. Follow the following two-steps procedure to generate the movie:

1)convert the images *.ppm into *.jpeg with imagemagick convert:
for f in peofront2..ppm ; do convert -quality 100 $f `basename $f ppm`jpg;done*

2) create a movie from the images with mencoder. Mind that the duration is related to the number of frames per seconds (here 2 s used):
mencoder "mf://peofront2..jpg" -mf fps=2 -o peofront.avi -ovc lavc -lavcopts vcodec=msmpeg4v2:vbitrate=800*

To do this you need to have installed the following packages: imagemagick, netpbm, mplayer/vlc, mencoder. If you are using the ubuntu virtual machine, those packages are already installed.

Return to the list of topics

Gnuplot's cheat sheet

Gnuplot is an extremely flexible and useful program to make 2D and 3D plot out of series of data stored in files. Non-linear fitting is also another interesting feature of gnuplot.

gnuplot	: start gnuplot
gnuplot -persist file	: use gnuplot with setting in file and produce a persistent plot (graphical windows remain after gnuplot quits)

plot - plot in 2D functions and series of data

plot f(x):	: plot f(x)
plot "file.dat"	: plot data in file.dat (min 1 column)
plot "file.dat" u 2:3	: plot 2nd col Vs. 3rd col in file.dat
plot "file.dat" u ($1*10):2	: plot 1st col Vs 2nd col in file.dat. Here the values in 1st col are multiply by 10
plot "file.dat" u ($0*10):2	: as before, but for file with one column only; $0 use the number of line as X in a 2D plot
plot "file.dat" w l	: plot data in file.dat using line style
plot "file.dat","file2.dat",f(x)	: plot data in file.dat and file2.dat and plot function f(x)
plot "file.dat" t "series"	: specify label for the plot

splot - plot 3D functions and datas

splot "surface.dat"	: 3D plot of data in surface.dat

set - set parameters for the plot

set title "Graph"	: set title for the plot
set xlabel/ylabel/zlabel "label"	: set label for X, Y and Z
set xrange [0:180]	: set range for X (yrange/ zrange)
set grid	: display grid
set style data line/point/ ...	: use line/point to display data
set angle deg/rad	: set angle in degrees/radiant (default)
set term png/eps/x11/ ...	: set the different terminals
set output "file"	: plot on file, not screen
set contour	: display contour plot
set contour base	: display contour plot on the base of a 3D plot
set pm3d	: 3D plot in colorscale

unset - unset parameters

unset title	: remove title for graph
unset xlabel	: remove label for X
unset grid	: remove grid

functions and parameters

f(x)=cos(x)	: define f(x)=cos(x)
g(x)=a+b*(1-cos(x))**2	: another function
a=2	: assign value 2 to parameter a

fit - non linear fitting

fit g(x) "file.dat" via a,b	: fit data in file.dat using the function g(x) by fitting the parameters a and b

rep - replot

rep	: update the last plot with last changes

help - display help

help set style data	: display the help for style data

save - save the session in a file

quit/exit - quit gnuplot

Use google to find out more, there is plenty of tutorials and examples out there

---

Return to the list of topics

Submit calculations to queuing system

Calculations are usually submitted to powerful servers with a queuing system to share the resources among all the users. There are different queuing systems out there; here two of them: pbs and loadlever.

pbs

qsub script.psb	: submit a job to the queue using a pbs script
qdel job_number	: remove the specified job from the queue
qstat	: display queue status
qstat -u user	: display status of the users's jobs in the queue

loadlever

llsubmit script	: submit a job to the queue using a script
llcancel job_number	: remove the specified job from the queue
llq	: display queue status
llq -u user	: display status of the users's jobs in the queue

To delete all your pbs job at once you can do:

```
qstat > q
awk '{print "qdel "$1}' < q > qdel
sh qdel
```

Return to the list of topics