MiCMoS
Milano Chemistry Molecular Simulation
Description and User Manual

Tutorials

T10. Molecular Dynamics simulation of crystals

PURPOSE: In this tutorial, it is described how to prepare and equilibrate the $P2_12_12$ RT phase of pyridone with the Molecular Dynamics technique.

REFERENCE to main manual: Sections 5.4 (***Pretop***); 6.2 (construction of molecular frameworks); 7.6.1 (***mdmain***); 7.6.2 (instruction file, .*mdi*); 7.6.3 (input box file, .*dat*); 7.6.4 (topology and force field file, .*top*).

FILES: You can find all the files you need to run this tutorial here:
https://sites.unimi.it/xtal_chem_group/images/MiCMoS_package/T10/PYRIDO11.cif
https://sites.unimi.it/xtal_chem_group/images/MiCMoS_package/T10/pyrd.mdi
The starting structure file for pyridone (PYRIDO11.*cif*) has been downloaded from the Cambridge Structural Data Centre. You should produce a valid .*oeh* file, as well as a preliminary static energy analysis, as described in Tutorials T1 and T4. This means that you have to launch in sequence ***Retcif***, ***Retcor***, ***Retcha*** and ***Crysaa***: you can do this directly with the available macro **run.cry** (see below). You will also need macros to run ***Pretop***, ***Boxcry*** and ***mdmain***. Place .*cif* files and pertinent (Windows/Unix) macros into your working directory.

The procedure for the MD simulation of a crystal structure can be straightforward with minimal user intervention in most cases. The preparation of an .*oeh* file from a standard .*cif* file has been carried out several times in former tutorials (see T1, T2, T4, T5 and T8) and should be well known at this time. From the .*oeh* file, ***Pretop*** generates a topology file that can be applied usually with only minor human intervention, needed for special torsions (none if the molecule is rigid). Templates for the .*mdi* run control file and topologies are available and intervention may only concern the optimization of the relaxation parameters. The ***Boxcry*** module prepares a crystal box, namecry.*dat*, which can be input directly to the MD module ***mdmain*** without having to go through a preliminary Monte Carlo equilibration or minimization.

CAUTION: This is certainly true if we accept that starting experimental structures are more or less correct and correspond to well–behaved (meta)stable potential energy minima (at least). Indeed, we do not expect significant steric clashes, misplaced atoms or ill–positioned molecules from good–quality X-ray structures. However, always keep in mind the programmers' motto – *garbage in, garbage out*. Ensuring that your starting structure is correct is the first mandatory step of any computer simulation.

Since MD topology files are usually rather complex, the ***Pretop*** module produces a file called NAMEtry.*top* in the first place, which is then typically renamed ad NAME.*top*. This is cautionary to avoid an inadvertent overlay if corrections are to be inserted or, worse, by a successive run of the ***Pretop*** module that would destroy the previous .*top* file.

In MD simulations, a good choice of many parameters is essential to obtain physically meaningful results: force field constants, temperature and pressure control, cutoff, box dimensions, etc. (see manual, Section 7). The examples given in this tutorial give average and more or less adequate values, but cannot always be taken as absolute indications, since these choices may be quite different for different chemical systems. The MiCMoS environment provides two largely tested systems for the treatment of small organic molecules, the CLP scheme and a LJC scheme, but allows the introduction of user-defined parameters, both for intra- and intermolecular potentials.

First, make a copy of the starting PYRIDO11.cif entry in the CSD:

```
cp PYRIDO11.cif pyrid.cif
```

This molecule is strictly rigid so the MD run can be started almost immediately by the following command sequence. When the procedure is standard and there are no loopholes, the **run.cry** macro present in batch directory of the package combines ***Retcif***, ***Retcor***, ***Retcha*** and ***Crysaa*** deleting all unnecessary output files. Call the basic program sequence with the command:

```
./run.cry pyrd
```

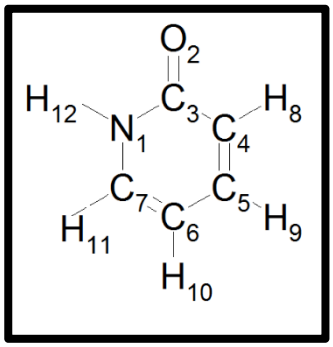
When prompted, answer 0 (normalize hydrogens), 0 (minimal output) and 0 (check the hydrogen count). See Tutorial T1 and Section 1.1 in the manual for detailed explanations. You will end with pyrid.*oeh* and pyrid.*dat* file bearing CLP (modified extended Hückel) charges. Check the results of the static energy calculation from ***Crysaa*** (file pyrdcry.*pri*) before going on. You should find that the calculation looks fine and reasonable (no errors in pyrdcry.*pri*).

Now prepare the topology file:

```
./run.pretop pyrd
```

For crystals with molecules constrained in a packing environment the force constants need not be upgraded, so use 1., 1., 1. rescaling factors when prompted. You have now a tentative pyridtry.*top* file, which is shown below.

File pyridtry.top, with the molecular scheme and atom numbering




```
#PYRIDO11 'P 21 21 2 topology
   12
    1   0.01285  -1.12334  -0.25939  21   -0.4151
    2  -0.01770   0.01349  -2.24689  27   -1.2925
    3   0.00355   0.04787  -0.98767  10    0.9400
    4   0.01763   1.24854  -0.21064  12   -0.3916
    5  -0.00280   1.19329   1.15056  12   -0.1739
    6  -0.01512  -0.03594   1.84148  12   -0.3637
    7   0.00272  -1.17359   1.10323  12    0.0499
    8   0.04471   2.20723  -0.70716   2    0.2994
    9  -0.00977   2.11419   1.71467   2    0.2847
   10  -0.03795  -0.07028   2.92073   2    0.2947
   11   0.00897  -2.13194   1.60120   2    0.2845
   12   0.02530  -1.97775  -0.77886   7    0.4835
    0  nslav-u
    0  ncore-v
    0  nslav-v
   88.1    0.0 volu-u,volu-v
   12  nstr-u
    1    3   1.379  4924.3   N- C
    1    7   1.364  5293.6   N- C
    1   12   1.000  5300.0   N- H
    2    3   1.260  7226.6   O- C
    3    4   1.430  4033.3   C- C
    4    5   1.362  5460.5   C- C
    4    8   1.080  3600.0   C- H
    5    6   1.410  4456.5   C- C
    5    9   1.080  3600.0   C- H
    6    7   1.356  5590.2   C- C
    6   10   1.080  3600.0   C- H
    7   11   1.080  3600.0   C- H
    0  nstr-v
   18  nbend-u
    1    3    2  120.00   702.0   N- C- O
    1    3    4  115.00   562.3   N- C- C
    1    7    6  121.00   618.1   N- C- C
    1    7   11  120.00   505.0   N- C- H
    2    3    4  124.00   745.7   O- C- C
    3    1    7  124.00   646.0   C- N- C
    3    1   12  117.00   460.0   C- N- H
    3    4    5  121.00   590.8   C- C- C
    3    4    8  120.00   505.0   C- C- H
    4    5    6  122.00   597.8   C- C- C
    4    5    9  119.00   517.5   C- C- H
    5    4    8  120.00   505.0   C- C- H
    5    6    7  118.00   569.7   C- C- C
    5    6   10  121.00   492.5   C- C- H
    6    5    9  119.00   517.5   C- C- H
    6    7   11  120.00   505.0   C- C- H
    7    1   12  119.00   460.0   C- N- H
    7    6   10  121.00   492.5   C- C- H
    0  nbend-v
   12  ntors-u
   7   1   3   2  50.00   -1.0    1.0   C- N- C- O
   3   1   7   6  50.00   -1.0    1.0   C- N- C- C
   1   3   7  12 100.00   -1.0    1.0   N- C- C- H
   1   3   4   5  50.00   -1.0    1.0   N- C- C- C
   3   1   2   4 100.00   -1.0    1.0   C- N- O- C
   3   4   5   6  50.00   -1.0    1.0   C- C- C- C
   4   3   5   8 100.00   -1.0    1.0   C- C- C- H
   4   5   6   7  50.00   -1.0    1.0   C- C- C- C
   5   4   6   9 100.00   -1.0    1.0   C- C- C- H
   5   6   7   1  50.00   -1.0    1.0   C- C- C- N
   6   5   7  10 100.00   -1.0    1.0   C- C- C- H
   7   1   6  11 100.00   -1.0    1.0   C- N- C- H
    0  ntors-v
    0  nlist-u
    0  nlist-v
   0.410   235.0   650.0   77000.0
    0  nextr
```

This .*top* file has all the stretching and bending force field parameters needed for the MD run. The labels at the right of each force field line (C-C, C-C-C, etc.) are just indicative but are not used.

Before proceeding, copy pyridtry.*top* into pyrid.*top* and keep editing this latter.

CAUTION: when dealing with flexible molecules, torsional potential must be updated according to Table A7.5 in the Appendix of the main manual. In any case, even in the case of rigid pyridone, a safe practice is to always inspect closely all assigned torsions. Actually, the automatic ***Pretop*** procedure assigns zero energy at zero torsion angle, but sometimes chooses the atom numbers corresponding to 180°.

This is exactly the case for torsion around the endocyclic N1-C3 bond (the first `ntors-u` instruction).

```
   7   1   3   2  50.00   -1.0    1.0   C- N- C- O
```

This corresponds to torsion C7–N1–C3–O2, which should be maximally stable when the torsion angle is 180º. That granted, recall that 50.0, –1.0 and +1 are *K*, *f* and *m* values in $E(tors) = K\{1 + \cos[f(m \cdot \tau)]\}$ (see manual, Section 7.4.1 and Figure A7.1b in Section A7.3, Appendix): the problem is that f = –1 and m = +1 stabilize torsions of 0º and destabilize those at 180º. Therefore, this instruction must be corrected. You have two choices. The simplest one is to re–define the torsion by choosing an endocyclic cis–chain, for which the potential stabilizing 0º conformers is correct:

```
   7   1   3   4  50.00   -1.0    1.0   C- N- C- C
```

Another equivalent possibility is to maintain the atom chain identities, and to correct the functional to have 0º conformations destabilized and 180º ones stabilized:

```
   7   1   3   2  50.00    1.0    1.0   C- N- C- O
```

Check by yourself all other torsions: you should agree in that all other parameters are correct. The pyrid.top file is now ready for entering Dynamics. Before that, we have to define the simulation box. We can call ***Boxcry*** to do that from the .*oeh* file:

```
./run.boxcry pyrid
```

The program will print:

```
boxcry module 3.1 oct.2018
   cell  13.657  5.912  5.701  90.00  90.00  90.00 give n. of cells on a,b,c
```

Give 2 5 5 to have an approximately cubic simulation box with 200 pyridone molecules in file pyridcry.*dat*.

```
#PYRIDO11 'P nTr,box dims,n.mols,n.atoms   2   5   5  27.3140  29.5600  28.5050  200  2400
 normal end of operation, mols. written  200
 Thank you for using MiCMoS
```

File pyridcry.*bxi* is also produced, but is for Monte Carlo but is not used in MD. Note that the crystal box is kept rather small in this short example.

Now prepare the run control file pyrid.*mdi* according to the following instructions. Have a look to the manual (Section 7.6.2) for a detailed description and meaning of all parameters.

```
pyridone PYRIDO11 clp potential minimal barostat
# n.steps  irvel  ipri  ibox idstr  timestep iengt ibias +Ebias Nbias tinon tinof
    10000     0      0    1    0     0.002     0     0
# cutoffu cutoffv cutoffuv factin  ipots   ianh   inano
   15.0    0.0     0.0       0.0      0      0      0
#  N(T)   Tset Tstart Trelax  0/1/2 weak/stiff/CSVR
    250    298   298    0.6      0
#  N(P)   Pset  comprs  0/1ianis ipr  ww   iextstr + stra11 22, 33, 12 13 23, GPa
    300    1.0   0.35      1      0   0.0    0
#  N(com)  nwbox   nwre  npri
     50     200    200    100
```

CAUTION: With one simulation box edge of 27.3 Å, the maximum cutoff allowed is 15 Å (must not exceed 0.55 times any of the three box lengths).

All other parameters should be obvious to the user who has gone through the preceding tutorials or has studied the manual; note that inano is zero, meaning that we are dealing with an unconfined periodic simulation box, as it should be. All the other values are more or less standard, with the only exceptions of N(T), Trelax, N(P) and comprs, that may need adjustment; as a rough guide, comprs is sensitive to the strength of crystal forces and may need smaller values in a hydrocarbon than in a strongly hydrogen-bound crystal that may tolerate a large pressure-volume jump (this is not a strict rule, though).

The comprs parameter is the first suspect in case of crash of the simulation. Then in order of frequency come a wrong assignment or rescaling of force constants (that may be evident from the starting values of intramolecular energies), and a badly prepared starting computational box. The latter case is more frequent for liquids than for crystals.

The actual MD run is started by the command

```
./run.mdmain pyrid pyridcry.dat py1
```

The prefix "pyrid" identifies input .*top* and .*mdi* files, pyridcry.*dat* is the starting simulation box and py1 will be the prefix of all the output files. These are: py1md.*pri* (regular printout), py1mdo.*dat* (last frame of the trajectory), py1mdc.*dat* (full trajectory) and py1md.*ene* (energu trajectory).

Now, open the py1md.*pri* printfile. You can check the total number of degrees of freedom (d.o.f.):

```
N.mols,N.atoms,(u-v),Mws   200   12    0    0   95.10    0.00
 n. of atomic degrees of freedom,     7200.        0.     7197.
```

We have 200 molecules consisting of 12 atoms each. Every molecule has 6 rigid body d.o.f. accounting for translations and rotations of the whole backbone, plus 3 x 12 – 6 internal d.o.f. due to intramolecular vibrations. Overall, the total number of d.o.f. is thus 200 x 12 x 3 =

7200. As we fix the centre of mass of the simulation box, the effective number of d.o.f. is 7197.

The last lines of the py1md.*pri* file summarize intermolecular energies, final unit cell parameters, and density, with corresponding percent changes from the starting estimates:

```
Intermolecular energies
   LP energies, u,v,uv,ubar,vbar  -9321.76      0.00      0.00      0.00      0.00
 Coul energies, u,v,uv,ubar,vbar  -6482.69      0.00      0.00      0.00      0.00
 final box edges, multiplicities  26.466  29.647  29.782  2  5  5
 final unit cell   13.233   5.929   5.956    90.0    90.0    90.0  1.352

   %edges, abs angles, %density variations
    -3.1   0.3   4.5    -0.0  -0.0  -0.0    -1.51
 final cohesive Ens, solute,solvent,solvation   -79.0     0.0     0.0
 cohesive energy variations   -0.7   0.0   0.0
```

You see that this short run changes only marginally the cell parameters. Temperature and pressure oscillate around the preset values. To have a look at them, you can resort to Unix grep command, as explained in the Tutorial T9:

```
grep 'nmove,Pset, Pcurrent' py1md.pri
grep 'old and new temperatures' py1md.pri
grep 'ens data' py1md.pri
```

for pressure, temperature and energies, respectively. Results are:

Pressures as a function of the simulation step:

```
       300 nmove,Pset, Pcurrent    1.000     13.3
       600 nmove,Pset, Pcurrent    1.000    399.0
       900 nmove,Pset, Pcurrent    1.000    492.1
....
      9300 nmove,Pset, Pcurrent    1.000   -109.0
      9600 nmove,Pset, Pcurrent    1.000      1.9
      9900 nmove,Pset, Pcurrent    1.000    145.7
```

Temperatures as a function of the simulation step:

```
       250 old and new temperatures   208.8  262.3    0.0    0.0  208.8  262.4
       500 old and new temperatures   241.4  275.3    0.0    0.0  241.5  275.5
       750 old and new temperatures   258.2  282.1    0.0    0.0  258.3  282.2
..............
      9500 old and new temperatures   303.8  300.3    0.0    0.0  304.0  300.5
      9750 old and new temperatures   286.9  293.6    0.0    0.0  287.0  293.7
     10000 old and new temperatures   302.7  299.9    0.0    0.0  302.9  300.0
```

Energies as a function of the simulation step:

```
       100 1816.4 1367.3  998.1    0.0    0.0    0.0    0.0    0.0  -9174.3  -7000.9 -11993.6   -80.9   1.372 ens data
       200 2054.6 1314.8 1194.1    0.0    0.0    0.0    0.0    0.0  -9605.7  -6866.2 -11908.4   -82.4   1.372 ens data
       300 2002.1 1401.0 1215.7    0.0    0.0    0.0    0.0    0.0  -9513.9  -6731.0 -11626.1   -81.2   1.372 ens data
.........
      9800 3147.5 2373.2 1565.0    0.0    0.0    0.0    0.0    0.0  -9245.5  -6536.2  -8696.1   -78.9   1.354 ens data
      9900 3185.4 2386.5 1612.2    0.0    0.0    0.0    0.0    0.0  -9266.6  -6561.9  -8644.5   -79.1   1.352 ens data
     10000 3030.9 2439.8 1652.0    0.0    0.0    0.0    0.0    0.0  -9321.8  -6482.7  -8681.7   -79.0   1.352 ens data
```

As for energies, from left to right: step number; stretching, bending and torsion energies of the solute; a dummy zero; stretching, bending and torsion energies of the solvent; a dummy zero; Lennard–Jones intermolecular energies; Coulomb intermolecular energies; total energies; cohesive energies per molecule per asymmetric unit; total density, in g/mL. All

energies are expressed in kJ/mol. From the *.top* file, we see that we have 12 stretching d.o.f. and 18 bending d.o.f., for a total of 200 x 12 = 2400 stretching d.o.f. and 200 x 18 = 3600 bending d.o.f.. Final internal stretching and bending energies are ~ 3000 kJ/mol and ~ 2400 kJ/mol, corresponding to 1.2 kJ/mol per d.o.f. (stretching) and 0.7 kJ/mol per d.o.f. (bending). These both comply well with the ½ RT estimate at 298 K (1.2 kJ/mol per d.o.f.).