# MiCMoS <br> Milano Chemistry Molecular Simulation Description and User Manual 

## Tutorials

## T8. Monte Carlo: simulation of a crystal (acetanilide)

PURPOSE: In this tutorial, it is described how to prepare and equilibrate a periodic crystal of acetanilide with the Monte Carlo technique. The aim is also to get the user acquainted with the use of slave atoms for the Monte Carlo routine of MiCMoS.

REFERENCE to main manual: Sections 5.1 (Boxcry); 5.4 (Pretop); 6.2 (construction of molecular frameworks); 6.6 .1 (mcmain); 6.6.2 (instruction file); 6.6 .3 (topology and force field file); 6.6.4 (slave atom file).

FILES: You can find all the files you need to run this tutorial here:
https://sites.unimi.it/xtal chem group/images/MiCMoS package/T8/ACANIL01.cif https://sites.unimi.it/xtal chem group/images/MiCMoS package/T8/acan.mci https://sites.unimi.it/xtal chem group/images/MiCMoS package/T8/acan.oeh https://sites.unimi.it/xtal chem group/images/MiCMoS package/T8/acan.sla https://sites.unimi.it/xtal chem group/images/MiCMoS package/T8/acan.top
The crystal and molecular structure comes from the Cambridge Structural Database entry ACANIL01, giving acanil01.cif as starting point. Retcif, Retcor, Retcha and Crysaa may be run in sequence as described in Tutorials T 1 and T4. You will also need macros to run Pretop, Boxcry and mcmain. Place .oeh and pertinent (Windows/Unix) macros into your working directory.

If you call the run.cry macro according to

## ./run.cry acanylo1

it will perform the whole sequence Retcif, Retcor, Retcha and Crysaa, ending with files acanil01.oeh (the structure file) and acanyl01cry.pri, which contains information on static lattice energies. 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. Check the results of the static energy calculation from Crysaa (file acanyl01.pri) before going on. You should find that the calculation looks fine and reasonable (no errors in acanyl01.pri).

This example involves a full description of a semi-rigid molecular unit with core and slave atoms. The molecular diagram below shows that the benzene ring and the N atom can be treated as a rigid fragment, while the $-\mathrm{COCH}_{3}$ group is programmed to be flexible by rotation around three torsion angles (C1-N9, N9-C7 and C7-C8). This means that atoms C7, C8, 010 and their hydrogens, plus hydrogen H 19 , should be better defined as slaves.


Atoms are not in the optimum sequence, though, as slaves must follow backbone-defining core atoms. Thus, the first thing to do is to renumber the atoms so that the first 12 are those of the rigid unit, that is, the benzene ring, its 5 hydrogens and the formerly N9 atom. This is done by editing the acanil01.oeh file, changing atom numbers to produce the renumbered acan. oeh file, which looks like as follows.



Original numbering
Renumbered (1-6) and 16-19 unchanged)


The structural information for slave atoms is provided in the .s/a file that has the following content (see Manual, Section 6.6.4):
acan.sla slave atom file

| 5 |  |  |  |  |  |  |  |  |
| ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| 1.3550 | 127.62 | 17.58 | 0 | 0 | 3 | 13 | 0 | 0 |
| 1.5020 | 115.47 | 175.89 | 0 | 0 | 3 | 14 | 0 | 0 |
| 1.2340 | 121.38 | 179.31 | 0 | 0 | 0 | 15 | 0 | 0 |
| 1.0800 | 109.47 | 116.40 | 0 | 0 | 3 | 16 | 17 | 18 |
| 1.0000 | 117.20 | 0.00 | 0 | 4 | 3 | 19 | 0 | 0 |

For each $i, j, k$, $m$ sequence of atoms, each line has the $i j j$ distance, the $\dot{i j} k$ angle and the $i$ -$j-k-m$ torsion angle. The numbers to be typed in can be derived from any geometry check program on the original .cif file (e.g. Mercury, https://www.ccdc.cam.ac.uk/solutions/csdsystem/components/mercury/), or on the acan. dat file. The next three numbers in each line are the codes for the variation of each of the three parameters. Zero means fixed parameter,
and this is the case for all the bond lengths and for all bond angles except the 19-12-13 HNC angle, which is variable with maximum stepsize number 4 in the acan. $m$ ci control file ( 3.0 deg , see below). All torsion angles are variable, by stepsize number 3 ( 5.0 deg), except an improper torsion angle associated to atom 15 (carbonyl oxygen), which is kept constant to preserve the planarity of the carbonyl group (see below). The last numbers in each line are the atom sequence numbers.

Then the preliminary topology file can be prepared by the command

## ./run.pretop acan

Answer the dialog mode with unit factors (we do not want to rescale force constants):

```
give factors for ks,kb,kt force consts.
1 1 1
    normal end of operation
    Thank you for using MiCMoS
```

This produces an acantry.top file, that must be edited as usual.
acantry.top file. In red: lines to be erased; in green: lines to be edited.



As usual, "u" means "solute", "v" means "solvent". Make a copy of acantry.top and keep working on the new file, named acan.top.

The atom coordinates after atom 12 and all the unnecessary bond stretch and bend potentials must be deleted.

CAUTION: remember to update the number of entries where appropriate (for example, after deleting the lines highlighted in red, the number of explicit atom coordinates is 12 and no longer 19, while nstr-u and nbend $-u$ are all 0 ).

Slave atom connectivity must now be defined after nslav-u instructions:

| 5 nslav-u |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 13 | -1 | 0 | 12 | 1 | 6 | 10 | 1.0587 |
| 14 | -1 | 0 | 13 | 12 | 1 | 13 | -0.8973 |
| 15 | -1 | 0 | 13 | 14 | 12 | 27 | -1.2314 |
| 16 | 17 | 18 | 14 | 13 | 12 | 3 | 0.32897 |
| 19 | -1 | 0 | 12 | 13 | 14 | 7 | 0.4864 |

The meaning of that information is that atom 13 must be built by Z-matrix procedure, (13 -10 ) attached to atom 12 and with torsional chain made by atoms 12,1 and 6 (see Table 6.2 in the Manual). The atom is type 10 (carbonyl carbon, Table 1.1) and its point charge parameter is given as it was in the original .oeh file. A similar procedure holds for atom 14 (Zmatrix, chain $13,12,1$, type 13 , amide nitrogen). Atom 15, type 27 , carbonyl oxygen, is also built with the Z-matrix procedure, but it exploits an improper torsion (15, 13, 14, 12) that will be used to keep the carbonyl group planar. The three methyl hydrogens, type 3, are built over the chain $16,14,13,12$ using the $\mathbf{R X}_{3}$ procedure (Table 6.2) and the amide hydrogen, type 7 , by a $19,12,13,14$ chain with again the Z-matrix algorithm.

The rest of the top file has the data for the intramolecular force field. It is pointless to keep active all the 17 torsions, as the atomatic ring is rigid and we intend to focus just on the relevant ones through slave atoms. We decided that C1-N12, N12-C13 and C13-C14 bonds (new numbering) are rotatable, thus we now will define just 3 torsions and change "17" into " 3 ".

3 ntors-u torsional potentials for the 3 degrees of freedom

Then, you can safely erase all former lines and type the following ones in their place:

| 13 | 12 | 1 | 6 | 10.0 | 1.0 | 4.0 |
| ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| 14 | 13 | 12 | 19 | 17.0 | -1.0 | 2.0 |
| 16 | 14 | 13 | 15 | 2.0 | 1.0 | 3.0 |

The first four entries are torsion-defining atom id numbers (C13-N12-C1-C6; C14-C13-N12H19; H16-C14-C13-O15). Then, $K, f$ and $m$ parameters in $($ (tors $)=K\{1+\cos [f(m \cdot \tau)]\}$ (equation (6.4) in the manual) follow, $\tau$ being the torsion angle.


The first rotation occurs around the bond $\mathrm{N} 12-\mathrm{C} 1$ and can be modelled using parameters of acetanilide ( $\mathrm{n}^{0} 30$ in Table A7.5, $K=10, f=+1, m=4$ ).

The second rotation (C14-C13-N12-H19) should be discouraged, as reorientations that make O 15 antiperiplanar with respect to C 1 are difficult due to the $\pi$ electron system. Moreover, we
know that chemical conjugation keeps the amide group planar, i.e., that N12, C13, C14 and 015 atoms must be constrained in the same plane. Note that we have obtained these results by combining dihedral instructions and nslav-u ones: (i) we use the rather high barrier suggested by Table A7.5 for aminobenzene ( $K=17, f=-1, m=2$ ) (current instruction); (ii) C14 is frozen in the N12-C13-O15 plane as the slave-defining improper dihedral angle 015-C13-C14-N12 is kept fixed (see ns lav-u and .sla file instructions above).

The last dihedral $\mathrm{H} 16-\mathrm{C} 14-\mathrm{C} 13-\mathrm{O} 15$ describes the rotation of the terminal methyl group. The same parameters as in ethanol or methanol (Table A7.5, $\mathrm{n}^{0} 42$ and 43) are a reasonable choice: $K=2, f=+1, m=3$, as the potential barrier against the methyl rotation is very small.

Finally, you should include a couple of nlist-u intramolecular contacts to protect the intermolecular contacts of the carbonyl oxygen O15 to phenyl ring ortho-hydrogens H 7 and H11 with an intramolecular nonbonded potential (see manual, Sections 6.4.2, 6.6.2 and 6.6.3):

```
2 nlist-u intramolecular distances to prevent O atom collision
15 7 15 11
```

Eventually, the file acan.top should look like:
acan.top, after editing


The starting crystal box can now be prepared from .oeh and .sla files by calling Boxcry (manual, Section 5.1).

The program will ask for number of repetitions along the three crystallographic directions:
boxcry module 3.1 oct. 2018
cell $19.509 \quad 9.364 \quad 7.778 \quad 90.00 \quad 90.00 \quad 90.00$ give $n$. of cells on $a, b, c$
Give 245 to have an approximately cubic simulation box with 320 molecules (6080 atoms):


Boxcry has now prepared a file called acancry.bxi, which contains positions and relative orientations of all the molecules. This is what it looks like, with the slave atom information for each molecule (see also manual, Section 5.1.1 and Table 5.1):


Along with the .bxifile, Boxcry also generates a acancry. dat file for visualization of the crystal structure.

The run control acan.mci (manual, Section 6.6.2) is set up as follows:


Variability indices are non-zero for all six cell parameters; even if this is an orthorhombic structure, in principle the cell angles could be altered by the MC run. The stepsize lines specifies 0.30 and $20^{\circ}$ for the rigid-body translations and rotations, $5^{\circ}$ for the torsional parameters and $3^{\circ}$ for the variation of the bending angle. The ianis flag is now 4 for an anisotropic variation
of the box dimensions (see also Manual, Table 6.1 in Section 6.5). Note that, at variance with liquid cases (Tutotials T6 and T7), we can directly start the simulation from a high temperature (298 K), as our starting structure corresponds to the experimental X-ray one and we do not expect to have to dispose of hard contacts.

The MC run can be started by calling mcmain:

## ./run.mcmain acan acancry.bxi ac1

where "acan" is the name of the run control and topology files (.mci and .top), acancry.bxi is the starting simulation box produced by Boxcry and "ac1" will be the prefix of all the output files.

The results can be found in the usual file ac1mc.pri, whereas the box corresponding to the last frame are ac1mc.bxo and ac1mco. dat.

Now open the ASCII output, ac1mc.pri: it is very similar to that discussed in Tutorial T6 for liquid benzene. After regular printing of atomic parameters, rescaled CLP charges and box maxteps, we found the following information:

```
N.solutes-atoms, N.solvents-atoms
            320 19 0 0
    total parameters in solutes and solvents 6720 0
    actual variables, total,solute,solvent 3520 3520 0
    ivarib 0 number of particles for P control 320.
    initial nonzero intramolecular energies
        mol., Es,Eb,Et,Enb,Tot 
        mol., Es,Eb,Et,Enb,Tot 2 0.00 0.00 0.00 0.00 10.84 7.59
        mol., Es,Eb,Et,Enb,Tot 3 0.00 0.0.00 0.00 10.84 0.00.59
        mol., Es,Eb,Et,Enb,Tot 4 0.00 0.00 0.00 0.00 10.84 l
solute, solvent, tot intramolecular energy 2428.7 0.0 2428.7
```

We have 320 "solute" molecules, bearing 19 atoms each, for a total of 6720 variable parameters, of which only 3520 are active. Flag "ivarib" determines some details of how the isotermal-isobaric ensemble is handled by the Metropolis-Monte Carlo algorithm (see Section 6.6.2 and equation (6.6) in Section 6.5 in the main manual for detailed explanations). Then, for each molecule, the intramolecular contribution to the potential is displayed. Stertching ( $E_{s}$ ), bending ( $E_{\mathrm{b}}$ ), torsions ( $E_{\mathrm{t}}$ ) and nonbonded interactions ( $E_{\mathrm{nb}}$ ), as well as total intramolecular contributions (Tot), are shown. Note that Tot is rescaled by the 0.7 fudge factor FACTIN as given in the .mci input file (Sections 6.4.2 and 6.6.2 in the main manual). Eventually, the total intramolecular energy of the whole simulation box is printed.

Then, the simulation is started and box/energy values are printed as usual:


After 100 kstep, the program stops after printing relevant statistics (number of moves accepted and rejected, concerning both molecular and box degrees of freedom):

```
last step
    100000
general,tot,acc,rej, box,tot,acc,rej
    99889 69264 30625 111 87 24
    acetani acceptance ratio, total,box 69.34 78.38
final result at last MC step
    1942.6 -21304.1 -11592.5 0.0 -30954.1
    19.5095 9.3861 7.7773 89.83 90.44 90.11
Intermolecular energies
    LP energies, u,v,uv -21304.15 0.00 0.00
Coul energies, u,v,uv -11592.52 0.00 0.00
Total LP,Coul,total E -21304.15 -11592.52 -30954.12
    box dipole energy 0.00
    acetani final Ecoh,Et per mol.,vbox,dens -102.8 -96.7 56964.40 1.261
```

Note that the total sum of moves ( $99,889+111$, molecules+box) is 100,000 , the total number of steps. In other words, the moves ratio is 99.9:0.1 in favor of molecular degrees of freedom. Final "per mole" energies are printed, as well as the predicted box volume and density. The last information concerns the distribution of steps over parameters, that is, for each parameter is displayed the number of steps in which that parameter has been varied. The output ends with 3250 entries, which sum up to the total of 99,889 moves involving molecular degrees of freedom.


Please note that this is just an illustrative run, optimal to be ended a few minutes. A bit longer simulations ( $\geq 10^{6}$ steps) should be set up to ensure that the system is fully equilibrated.

## In summary:

1) Run Retcif, Retcor, Retcha, Crysaa on acanil01.cifto get acanil01.oeh;
2) Renumber atoms to have the first ones in the rigid part, obtain acan.oeh;
3) Prepare distances, angles and torsions for slave atoms by any geometry check program, for example by running Mercury (https://www.ccdc.cam.ac.uk/solutions/csdsystem/components/mercury/) on the original .ciffile;
4) Prepare the sla file (by hand, using templates), with actual values of the geometrical paramters and indication of variable parameters;
5) Launch Pretop to prepare a tentative .top file. Make a copy of this file and call it acan.top.
6) Edit acan.top to delete stretch /bend potentials and non-rigid atoms, and to insert codes
for the construction of slave atoms and torsion potential functions;
7) Run Boxcry on the acan.oeh file to obtain the starting box (acancry.bxi);
8) Prepare the acan.mci run control file by editing one of the many templates;
9) Run the MC module with acan.top, acan.mci and acancry.bxi.
