

RMC ↔ GROMACS configuration converter

Orsolya Gereben, 2008-11-17

This program converts between the RMC version 3 **.cfg* file and the GROMACS **.gro* and **.top* file. The program needs a parameter file with freely chosen name, and it can be started either specifying the name of the parameter file as an command line argument (for example: `convert_gromacs my.dat`), or only with `convert_gromacs`, and then the program will ask for the name of the parameter file.

It has to be noted, that this manual is mainly written for RMS-users who now getting acquainted with GROMACS, so the concept and the file structure of RMC is supposed to be known, and some basic explanation of the GROMACS file structure concept is given without the aim of making the GROMACS concept fully understandable from this. Missing pieces can be looked up in the GROMACS manual.pdf and in the html format online.html, which are part of the GROMACS installation or can be downloaded from the www.gromacs.com website

Concept of the conversion, file structures

In case of RMC the structure of the **.cfg* file follows the simple rule, that the total number of atoms, the number of atom types, the half length of the simulation box and the number of atoms for each type are specified in the header of the **.cfg* file, then all the coordinates of the first type of atoms (each a separate line), then the coordinates of the second and so on atom types can be found consecutively. Because of this, if there are molecules consisting of different type of atoms, than the coordinates of the atoms belonging to the same molecule can be found in non-consecutive lines of the file. Even if there are more than one atom from the same type in the molecule, they not necessarily located consecutively, but sometimes in blocks, depending on the structure of the **.cfg* and the accompanying **.fnc* file. The only important thing is, that has to be a logic in the arrangements of the atoms, so if we know the atom indices of the atoms belonging to the first and second molecule, we can calculate the atom indices of the atoms in any molecule. The coordinates are reduced, going between -1 and $+1$ and the box half-length and other distance related properties are given in Ångstrom in RMC.

For example there are two different representations, two sets of **.cfg* with their matching **.fnc* file for SnI₄, see the atom indices in Table 1.

| Atoms | First representation | | Second representation | |
|-------|----------------------|-------------|-----------------------|-------------|
| | Atom indices | | Atom indices | |
| | first mol. | second mol. | first mol. | second mol. |
| Sn | 1 | 2 | 1 | 2 |
| I | 1001 | 1005 | 1001 | 1002 |
| I | 1002 | 1006 | 2001 | 2002 |
| I | 1003 | 1007 | 3001 | 3002 |
| I | 1004 | 1008 | 4001 | 4002 |

Table 1: The indices of the atoms in the two differently structured **.cfg* and **.fnc* file for SnI₄.

On the other hand in GROMACS the distances and coordinates are given in nm, and the simulation box goes from 0 → boxlength. In GROMACS the atoms of a residue have to follow each other consecutively in the text format **.gro* configuration file. The residue can be a molecule, a free ion, or it can be the part of a big molecule, like an amino acid in a protein. The structure of the **.gro* file is described in the GROMACS [online.html](#). It has a fixed format: first line is the title, second the total number of atoms, and then follows a line for each atom with fixed format:

- residue number (5 positions, integer)
- residue name (5 characters)
- atom name (5 characters)
- atom number (5 positions, integer)
- position (in nm, x y z in 3 columns, each n+5 positions with n decimal places)
- velocity (in nm/ps (or km/s), x y z in 3 columns, each 8 positions with 4 decimal places)

Normally n=3 in GROMACS, but we will use n=8 in the conversion program to achieve greater accuracy.

In GROMACS atoms have a name and a type. It has to be emphasized, that the GROMACS atom types should not be confused with the atom types in RMC. In GROMACS atoms of the same chemical type can have different GROMACS types, as the GROMACS type have to reflect the chemical role, bonds, hybridisation state of the atom causing it to have different Lennard-Jones, Coulomb parameters or charge. Each force field has its own defined different GROMACS atom types. You have to use one of the atom types of the force field you chose, so check the *ff*.atp* file of the force field for the available types in the top directory of the GROMACS installation. All atoms in GROMACS have to have a name, and in a residue each atom have to have different names. The names are the same for the atoms in the same position in different instances of the same residue (see the topology and **.gro* file of a molecule). The atom names can be freely chosen (max 5 character).

The topology of the residue is defined in the **.top* file, which can include topologies of residues defined separately in **.itp* files. Each molecule can have its **.itp* file, this way these can be included into any system's topology file, when it is necessary. Each atom type has to be defined first before it is used. If you want to define new types, you can do it at the beginning of the **.top* or **.itp* files with the `[atomtypes]` directive (see *methanol.itp* in the GROMACS' top directory). The concept of the topology file, the existing directives can be found in the GROMACS manual, there are much more, than it is shown in these examples (see Chapter 5 and especially Table 5.3 and 5.4 in the manual.pdf).

The topology file has a relatively free format, meaning it does not matter, how many space or tabs you are using to separate the data in a line, but the parameters of a line has to follow in a predefined order. It contains sections, each beginning with a directive enclosed in square bracket `[]`. The order of some of the directives are defined, others can follow in arbitrary order. The text after a semicolon is ignored by GROMACS, it can be used for comments.

In the line after the `[molecule type]` directive the name of the molecule and the number of exclusion (`nrexcl`) has to be defined. This means that for atoms `nrexcl` bonds or less apart non-bonded interactions will not be calculated (LJ and Coulomb). The GROMACS recommended value is 3, but you can specify anything here. Obviously, if you have a small molecule, where there are no atoms 3 bonds apart at all, you have to specify a smaller number (2 for example for water).

This will be followed by the listing of the atoms of a molecule after the `[atoms]` directive. Here all atoms to be simulated have to be listed. Atoms belonging to the same charge group have to follow each other consecutively. The parameters, which have to be given for each atom in the `[atoms]` section are the following:

```

nr           : serial number of the atom
type        : atom type
resnr       : residue number (index)
residue     : residue name
atom        : atom name
cngr        : charge group number
charge      : charge
mass        : mass in amu
qtot        : total charge so far

```

A charge group is a group of atom, where the total charge is zero, these will be kept together, when the cut-off for the calculation is set.

Obviously, these parameters cannot be deduced from the RMC configuration, so they have to be specified in the parameter file for the converter program.

The connectivity between the atoms can be defined by the [bonds] [angles] [dihedrals] or [constraints] directive. There are usually predefined bonds, angles, dihedrals between given type of atoms, check the force field's *ff*bon.itp* file for the available types, and the manual for details, in which format to add them manually to the *.top or *.itp file. If an accompanying *.fnc file is available for the RMC→GROMACS conversion, then the constraint from that can be automatically incorporated into the topology according to the parameters specified in the parameter file of the conversion. Both bonded and non-bonded distances can be constrained in GROAMCS.

An other important directive is the [pair] governing the calculation of the 1-4 interactions (between atoms 3 bonds apart), because of in certain force fields with nrexcl=3 non-bonded interaction with the normal set of parameters are not calculated for this pairs, but there exist a set of a bit weaker non-bonded interaction parameters, and for those 1-4 pairs listed after the [pair] directives non-bonded weaker interactions will be calculated. 1-4 interactions are never calculated for 1-4 hydrogen pairs. The [pair] section has to be added to the *.top or *.itp manually, if needed!

At the end of the topology file the [system] directive followed by the user-defined system name and then the [molecule] directive followed by as many lines as molecule types with the molecule name, and number of molecules has to be stated.

As an example the parameter file for converting the C₂Cl₄ RMC configuration consisting of 2000 C atoms (type 1) and 4000 Cl atoms (type 2) forming 1000 molecules according to the c2cl4.fnc file will be given, and the resulted *.gro and *.top and *.itp files as well. To demonstrate, how to give parameters, for more than one type of residue type, (and as there was no available *.cfg and *.fnc) the 1000 molecules will be split into two types of residues, 400 CIETH and 600 C₂Cl₄ molecules. It has to be noted, that the CR1 atom type used for the sp² C atoms of the tetrachloroethene theoretically not applicable, as it refers to aromatic carbon atoms, but there was no sp² C in the force field.

First some important sections of the c2cl4.fnc file are given below:

```

This is a .fnc file created by makecfg

No. of possible rmin-rmax pairs:
11
1.32  1.67  1.67  2.65  2.65  1.67  1.67  2.8  2.8  3.1
3.1

```

3.4 1.39 1.77 1.77 2.8 2.8 1.77 1.77 3 3 3.4

6000

| | | | | | | |
|------|------|------|------|------|------|------|
| | 1 | | 5 | | | |
| | 1001 | | 2001 | 3001 | 4001 | 5001 |
| 1 | 2 | 3 | 4 | 5 | | |
| | 2 | | 5 | | | |
| | 1002 | | 2002 | 3002 | 4002 | 5002 |
| 1 | 2 | 3 | 4 | 5 | | |
| | 3 | | 5 | | | |
| | 1003 | | 2003 | 3003 | 4003 | 5003 |
| 1 | 2 | 3 | 4 | 5 | | |
| | ⋮ | | | | | |
| | 1001 | 3 | | | | |
| | 1 | 4001 | 5001 | | | |
| 1 | 6 | 7 | | | | |
| | 1002 | 3 | | | | |
| | 2 | 4002 | 5002 | | | |
| 1 | 6 | 7 | | | | |
| | 1003 | 3 | | | | |
| | 3 | 4003 | 5003 | | | |
| 1 | 6 | 7 | | | | |
| | ⋮ | | | | | |
| | 2001 | 3 | | | | |
| | 1 | 3001 | 4001 | | | |
| 2 | 8 | 11 | | | | |
| | 2002 | 3 | | | | |
| | 2 | 3002 | 4002 | | | |
| 2 | 8 | 11 | | | | |
| | 2003 | 3 | | | | |
| | 3 | 3003 | 4003 | | | |
| 2 | 8 | 11 | | | | |
| | ⋮ | | | | | |
| | 3001 | 3 | | | | |
| | 1 | 2001 | 5001 | | | |
| 3 | 8 | 10 | | | | |
| | 3002 | 3 | | | | |
| | 2 | 2002 | 5002 | | | |
| 3 | 8 | 10 | | | | |
| | 3003 | 3 | | | | |
| | 3 | 2003 | 5003 | | | |
| 3 | 8 | 10 | | | | |
| | ⋮ | | | | | |
| 4001 | 4 | | | | | |
| | 1 | 1001 | 5001 | 2001 | | |
| 4 | 6 | 9 | 11 | | | |
| | 4002 | 4 | | | | |
| | 2 | 1002 | 5002 | 2002 | | |
| 4 | 6 | 9 | 11 | | | |

```

4003      4
      3 1003 5003 2003
4 6 9 11

      ⋮
5001      4
      1 1001 4001 3001
5 7 9 10
5002      4
      2 1002 4002 3002
5 7 9 10
5003      4
      3 1003 4003 3003
5 7 9 10

```

There are redundant distances, but this way the same FNC file could be used for structurally similar, but not symmetrical molecules.

The structure of the C₂Cl₄ molecule with all the bonds, angles, and constraints and GROMACS and RMC numbering is shown in Figure 1.

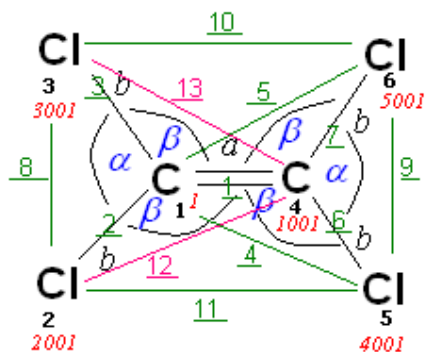


Figure 1: The C₂Cl₄ molecule. The black numbers (1-6) denote the numbering of the atoms in the topology file, the red italic numbers are the numbering of the first molecule in the *.fnc file, the green underlined numbers along the green lines are the constraint distances and their constraint index in the *.fnc file. α is the Cl-C-Cl and β is the Cl-C-C bond angle. Notice, that the constraints between the non-bonding C and Cl-s are not defined symmetrically in the FNC file, the missing 12 and 13 distances are shown with purple. The 2, 3, 6 and 7 distances are the same for C₂Cl₄, and the 4-5, the 8-9, 10-11 and 12-13 are pair-wise identical. The FNC constraint 1-3, 6-7 will be turned into bonds, the 4-5 into β the 8-9 into α angle constraint. The 10-11 can be turned into non-bonded constraint type 2.

The structure of the parameter file will be the same for both RMC \rightarrow GROMACS and GROMACS \rightarrow RMC conversion, although some of the parameters in the file will be ignored in one of the two cases. Parameters with italic and underlined fonts in the description field are used in both conversion direction, italic is used only in RMC \rightarrow GROMACS and underlined in GROMACS \rightarrow RMC conversion. If a parameter, which is not used in the conversion is followed by a used parameter, than some value (according to the type of the parameter) has to be given for the not used parameters as well!

The parameter file for the above described RMC \rightarrow GROMACS conversion should look the following way as in Table 2:

| Line index (not part of the file!) | Parameter | Description |
|---------------------------------------|----------------------------------|---|
| 1 | 0 3 | ! <u>conversion type (0) RMC->MD, (1) MD->RMC, number of configurations</u> |
| 2 | c2cl4.cfg 2 | ! <u>RMC configuration file, number of RMC types</u> |
| 3 | c2cl4.gro | ! <u>GROMACS *.gro file</u> |
| 4 | c2cl4.top | ! <u>GROMACS *.top file</u> |
| 5 | Tetrachloroethene | ! <u>Title</u> |
| 6 | ffG43a1.itp 2 2 | ! <u>force field parameter file; GROMACS bond type, angle type, D for force field default</u> |
| 7 | 0 0.020629 | ! <u>whether to rescale box (0: no, 1: yes), new number density in 1/(A3), if necessary</u> |
| 8 | 2 | ! <u>number of different residue types</u> |
| 9 | 1 2 | ! <u>character string (max 3 char) for each residue defining X in the bond angle and non-bonded constraint name definition Y_X_index.</u> |
| 10 | 400 6 CL4ETH 3 | ! <u>as many line as residue types containing first the number of this residue, number of atoms in the residue, the name (max 5 char) for the residue, number of bonds for exclusions (recommended=3)</u> |
| 11 | 600 6 C2Cl4 3 | - - |
| 12 | 1 2 C1 CR1 12.011 1 -0.06 1 | ! <u>as many lines, as atoms in all the residue types: the index of this atom in the RMC conf. in the first then in the second molecule; GROMACS atom name, atom type, mass; charge group, charge, RMC-type</u> |
| 13 | 001 2002 CL1 CL 1.008 1 0.03 2 | - - |
| 14 | 001 3002 CL2 CL 1.008 1 0.03 2 | - - |
| 15 | 001 1002 C2 CR1 12.011 2 -0.06 1 | - - |
| 16 | 001 4002 CL3 CL 1.008 2 0.03 2 | - - |
| 17 | 001 5002 CL4 CL 1.008 2 0.03 2 | - - |
| 18 | 401 402 C1 CR1 12.011 1 -0.06 1 | - - |

| | | |
|----|---|---|
| 19 | 401 2402 CL1 CL 1.008 1 0.03 2 | - - |
| 20 | 401 3402 CL2 CL 1.008 1 0.03 2 | - - |
| 21 | 401 1402 C2 CR1 12.011 2 -0.06 1 | - - |
| 22 | 401 4402 CL3 CL 1.008 2 0.03 2 | - - |
| 23 | 401 5402 CL4 CL 1.008 2 0.03 2 | - - |
| 24 | 3 c2cl4.fnc | ! whether to use FNC, name of the FNC file (0:no, 1:use FNC lower limit, 2: use FNC upper limit, 3:use FNC average limit) |
| 25 | 5 5 | ! number of bonds to create for each residue from FNC |
| 26 | 1 2 3 6 7 1 2 3 6 7 | ! constraint indices in the FNC file for each desired bond const. for each residue |
| 27 | 1.66e+7 8.12e+6 8.12e+6 8.12e+6 8.12e+6 1.68e+7 8.14e+6 8.14e+6 8.14e+6 8.14e+6 | ! force constants for each desired bond separated by space |
| 28 | 4 4 | ! number of angles to create for each residue from FNC |
| 29 | 4 5 8 9 4 5 8 9 | ! constraint indices in the FNC file for each desired angle const. for each residue |
| 30 | 500 500 632 632 550 550 632 632 | ! force constants for each desired angle separated by space |
| 31 | 4 4 1 1 4 4 1 1 | ! index of the central atom in the molecule (according to topology) for each desired angle const. for each residue |
| 32 | 119.8 119.8 120.4 120.4 119.9 119.9 120.2 120.2 | ! desired angle in degree for each desired angle const. of each residue |
| 33 | 2 2 | ! number of non-bonded constraints to create for each residue from FNC |
| 34 | 10 11 10 11 | ! constraint indices in the FNC file for each desired non-bonded const. for each residue |

Table 2: The parameter file for the converter program

Comments for some lines of the parameter files are given here, where necessary. The coloured blocks indicate the same type of line. As the number of lines from each group depend on the system type, the lines indices after the coloured blocks will not be the same for a different system!

- line 1: For GROMACS→RMC conversion nconf specify the number of configuration in the *.gro file, which has to appear consecutively. This type of file is created from the *.trr with trjconv (use -ndec X for trjconv, where X is the number of decimals, if larger than 3 is needed, -b start_time will specify the time value of the first configuration to extract). In case of more than 1 specified for nconf, each output RMC configuration will be saved into a separate file, and the time value extracted from the configuration will be added to the output RMC file names. For example if t=60.0000, t=80.0000 and t=100.0000 is the

time value in the *.gro configuration file for 3 configuration to convert and if c2cl4.cfg is specified for the RMC file name, then the files will be called c2cl4_60.0000.cfg, c2cl4_80.0000.cfg c2cl4_100.0000.cfg. If no .cfg extension can be found in the RMC file name, then the time value will be appended to the end of the file name.

- line 2: The number of RMC types needed only for GROMACS→RMC conversion.
-
- line 6: Used only for RMC→GROMACS, only the name of the force field file will be included into the topology file, the file itself is not needed for the conversion! The GROMACS bond and angle types can depend on the force field, and described in the GROMACS manual's Table 5.4. If nothing is given for the types, than the force field default will be used.
- line 7: Rescaling the box makes only sense with atomic systems, because the rescaling will change the intra-molecular distances as well! Rescaling occurs too, if not all the available particles are used from the original configuration, as the number density is kept constant.
- line 9: This name extension is used to differentiate between the names created in the define statements in the *.top (*.itp) files of different residues. Can be anything (max 3 character), only use different extension for each different residue.
- line 10-11: The residue name can be chosen freely for RMC → GROMACS conversion, but has to match the names in the *.gro file for GROMACS→ RMC conversion.
- line 12-23: The indices for an atom in the first and the second molecules can easily determined from the *.fnc file, if there is no *.fnc file, then have to be calculated based on the *.cfg file. Index start with 1!

After the green block start the parameters governing the building of bonds, angles and constraints based on the FNC file for RMC→GROMACS conversion. The constrained distances related to bonds can be converted to bonds, the 1-3 non-bonded distances to angles, and other distances more than two bonds away can be converted to type 2 constraints (see the GROMACS manual about bonds, angles and constraints). Bonds and angles can later be converted into GROMACS constraints by grompp, but as usually wider ranges are used in the *.fnc files for constraining a distance, the md run started with the use of constraints is usually crashes, as it is not able to adjust the atomic positions to fulfil the constraints, not even after energy minimalization done by bonds and angles. Using larger force constants for the energy minimalization sometimes help to pull the distances closer to the required constraint distance. Even after this it can happen, that there is trouble with the type 2 constraints created for atoms more than 2 bonds apart. Then take the constraint out of the topology, and introduce other ways to ensure the right molecule geometry, (for example dihedral angle constraint (type 1 [dihedrals]) to ensure the right conformation, improper dihedral angle constraint (type 2 [dihedrals]) to keep planar groups planar...

Only the bond and angle types has to be given for a molecule, not all the bonds and angles, so if there is more fnc pair in a molecule with the same fnc constraint, then only one f it has to nbe given, even in the case of the angles, specify one central atom only for an angle constraint. The program will determine all the others.

- line 24: If 1 is given, than the distances for the bonds, angles and the non-bonded constraints will be minimum distances from the *.fnc file. For 2 the maximum distances, for 3 the average of the minimum and maximum will be written into the topology file. You can edit it after the conversion easily, as these will be parameters defined at the beginning of the *.top or *.itp files, and appear only at one place.
- line 26,29 and 34: First specify the indices of the required FNC constraint (index starts with 1) for the first residue, than the second and so on.

- line 27 and 32: Search the force field's *ff*.bon.itp* file for the appropriate force constant, or at least to get an idea about the order of magnitude. It has to be noted, that the GROMOS-96 force field uses not the harmonic, but a different potential for bonds and angles (so called G96 bonds and angles, which are the type 2 [bonds] and [angles], while the other force fields mainly using the harmonic potential (type 1). The parameters in the *ff*.bon.itp* file are following the inherent type of the force field. You can use harmonic potential parameters with the GROMOS-96 force field, but then give type 1 in line 6, and give the force constant parameters according to the harmonic potential. See GROMACS manual.pdf 4.2.
- line 31: The index of the central atom of each angle has to be given here (first for all the angles of the first residue, then the second and so on...). The index has to be the index of the central atom in the molecule according to order the atoms following each other in the parameter file, and the numbering start with 1 for each molecule. In the example given above, the angle constraints are for the Cl-C-Cl, so the central atom with index 1 is Cl and with index 4 is C2 in both residues.

The files (or part of them created with RMC→ GROMACS conversion will be given now:

c2cl .gro file:

Tetrachloroethene

6000

| | | | | | |
|----------|-----|------|------------|------------|------------|
| 1ClETH | C1 | 1 | 0.29517188 | 0.37208103 | 0.22305788 |
| 1ClETH | CL1 | 2 | 0.41933430 | 0.33626471 | 0.10672363 |
| 1ClETH | CL2 | 3 | 0.25842666 | 0.22115058 | 0.30518070 |
| 1ClETH | C2 | 4 | 0.20488667 | 0.47513308 | 0.21004104 |
| 1ClETH | CL3 | 5 | 0.25170822 | 0.61074675 | 0.11398080 |
| 1ClETH | CL4 | 6 | 0.04941676 | 0.47393457 | 0.27674567 |
| 2ClETH | C1 | 7 | 0.61964957 | 0.27083498 | 0.65431289 |
| 2ClETH | CL1 | 8 | 0.55809396 | 0.42734495 | 0.69703021 |
| 2ClETH | CL2 | 9 | 0.51920196 | 0.19721709 | 0.53999672 |
| 2ClETH | C2 | 10 | 0.74074386 | 0.21372557 | 0.68942769 |
| 2ClETH | CL3 | 11 | 0.85775376 | 0.29146949 | 0.78145105 |
| 2ClETH | CL4 | 12 | 0.78412961 | 0.05466282 | 0.64826386 |
| ⋮ | | | | | |
| 401C2C14 | C1 | 2401 | 0.18167734 | 5.16664663 | 2.35673405 |
| 401C2C14 | CL1 | 2402 | 0.25787454 | 5.09328515 | 2.21504163 |
| 401C2C14 | CL2 | 2403 | 0.22838143 | 5.07776816 | 2.49361926 |
| 401C2C14 | C2 | 2404 | 0.09033629 | 5.26778450 | 2.35718321 |
| 401C2C14 | CL3 | 2405 | 0.03443982 | 5.34661955 | 2.21946546 |
| 401C2C14 | CL4 | 2406 | 0.00076152 | 5.30436269 | 2.49478321 |
| 402C2C14 | C1 | 2407 | 0.70644561 | 0.20074870 | 2.54129577 |
| 402C2C14 | CL1 | 2408 | 0.83395986 | 0.08600902 | 2.51513273 |
| 402C2C14 | CL2 | 2409 | 0.60631018 | 0.17258939 | 2.67522027 |
| 402C2C14 | C2 | 2410 | 0.66958453 | 0.29735569 | 2.45508897 |
| 402C2C14 | CL3 | 2411 | 0.74334637 | 0.32990204 | 2.29819076 |
| 402C2C14 | CL4 | 2412 | 0.59696872 | 0.44063159 | 2.50979799 |
| ⋮ | | | | | |
| ⋮ | | | | | |
| ⋮ | | | | | |

c2cl .top file:

```
#include "ffG43a1.itp"
```

```
#include "ClETH.itp"
#include "C2Cl4.itp"
```

```
[ system ]
; Name
Tetrachloroethene
```

```
[ molecules ]
      ClETH                400
      C2Cl4                600
```

***ClETH.itp* file:**

```
;Defining the parameters for bonds, angles, dihedrals and constraints
```

```
#define b_1_1      0.1355
#define b_1_2      0.172
#define b_1_3      0.172
#define b_1_4      0.172
#define b_1_5      0.172
#define nc_1_1     0.325
#define nc_1_2     0.325
#define a_1_1      119.8
#define a_1_2      119.8
#define a_1_3      120.4
#define a_1_4      120.4
#define fb_1_1     1.66e+007
#define fb_1_2     8.12e+006
#define fb_1_3     8.12e+006
#define fb_1_4     8.12e+006
#define fb_1_5     8.12e+006
#define fa_1_1     500
#define fa_1_2     500
#define fa_1_3     632
#define fa_1_4     632
```

```
[ moleculetype ]
;Name nrexcl
ClETH 3
```

```
[ atoms ]
;      nr      type  resnr residue  atom  cngr  charge      mass
qtot
-0.6      1      CR1      1  ClETH      C1      1      -0.6      12.011 ;
-0.3      2      CL      1  ClETH      CL1      1       0.3       1.008 ;
0         3      CL      1  ClETH      CL2      1       0.3       1.008 ;
-0.6      4      CR1      1  ClETH      C2      2      -0.6      12.011 ;
-0.3      5      CL      1  ClETH      CL3      2       0.3       1.008 ;
0         6      CL      1  ClETH      CL4      2       0.3       1.008 ;
```

```

[ bonds ]
;      ai      aj      funct      b0      kb
      1      2      2      b_1_2      fb_1_2
      1      3      2      b_1_3      fb_1_3
      1      4      2      b_1_1      fb_1_1
      4      5      2      b_1_4      fb_1_4
      4      6      2      b_1_5      fb_1_5

[ angles ]
;      ai      aj      ak      funct      theta      kb
      1      4      5      2      a_1_1      fa_1_1
      1      4      6      2      a_1_2      fa_1_2
      2      1      3      2      a_1_3      fa_1_3
      5      4      6      2      a_1_4      fa_1_4

[ constraints ]
;      ai      aj      type      b0
      2      5      2      nc_1_2
      3      6      2      nc_1_1

```

C2Cl .itp files:

```

;Defining the parameters for bonds, angles, dihedrals and constraints
#define b_2_1      0.1355
#define b_2_2      0.172
#define b_2_3      0.172
#define b_2_4      0.172
#define b_2_5      0.172
#define nc_2_1      0.325
#define nc_2_2      0.325
#define a_2_1      119.9
#define a_2_2      119.9
#define a_2_3      120.2
#define a_2_4      120.2
#define fb_2_1      1.68e+007
#define fb_2_2      8.14e+006
#define fb_2_3      8.14e+006
#define fb_2_4      8.14e+006
#define fb_2_5      8.14e+006
#define fa_2_1      550
#define fa_2_2      550
#define fa_2_3      632
#define fa_2_4      632

[ moleculetype ]
;Name nrexcl
C2Cl4 3

[ atoms ]
;      nr      type      resnr residue      atom      cngr      charge      mass
qtot
-0.6      1      CR1      1      C2Cl4      C1      1      -0.6      12.011 ;
-0.3      2      CL      1      C2Cl4      CL1      1      0.3      1.008 ;
0      3      CL      1      C2Cl4      CL2      1      0.3      1.008 ;
-0.6      4      CR1      1      C2Cl4      C2      2      -0.6      12.011 ;

```

```

-0.3      5      CL      1      C2Cl4      CL3      2      0.3      1.008 ;
0          6      CL      1      C2Cl4      CL4      2      0.3      1.008 ;

[ bonds ]
;      ai      aj      funct      b0      kb
      1      2      2      b_2_2      fb_2_2
      1      3      2      b_2_3      fb_2_3
      1      4      2      b_2_1      fb_2_1
      4      5      2      b_2_4      fb_2_4
      4      6      2      b_2_5      fb_2_5

[ angles ]
;      ai      aj      ak      funct      theta      kb
      1      4      5      2      a_2_1      fa_2_1
      1      4      6      2      a_2_2      fa_2_2
      2      1      3      2      a_2_3      fa_2_3
      5      4      6      2      a_2_4      fa_2_4

[ constraints ]
;      ai      aj      type      b0
      2      5      2      nc_2_2
      3      6      2      nc_2_1

```

The qtot in the *.itp file is the total charge so far. This is not strictly necessary, but usually is given .

Always check the file, and some additional parameters might be needed to make the md simulation work properly. In this example add the two missing angles for constraint 12 and 13 to the [angles] section:

```

      2      1      4      2      a_2_1      fa_2_1
      3      1      4      2      a_2_2      fa_2_2

```

After some experiments, it seems, that in this case to use the [constraints] to constrain the non-bonded 1-4 distances did not work, as the range of the constrained distances was too large, and the simulation crashed. Improper dihedral angle was used to keep the molecule planar:

```

[ dihedrals ]
;      ai      aj      ak      al      funct      angle      kb
      3      1      4      6      2      0.0      fi_1_1

```

with f1_1_1 = 167.5.

The [pairs] section also has to be added to calculate a modulated non-bonded interaction for the 1-4 Cl-Cl pairs, as due to the fact, that nrexcl=3 was given, no interaction would be calculated for the 1-4 Cl-Cl pairs at all:

```

[ pairs ]
;      ai      aj      funct      c0      c1
      2      5      1
      2      6      1
      3      5      1
      3      6      1

```

If only the 1-4 pair indices and 1 for the function type is given, than the parameters given in the *ff*nb.itp* file will be used, but if there is none there, or it should be overwritten, than the new value can be specified after the function type.

Virtual sites

The need to use water models like Tip4p and Tip5p where virtual sites are used made it necessary to include the possibility of handling virtual sites in the converter. For RMC → GROMACS conversion this makes it possible to convert a water molecule's 3 atoms into a water molecule containing 1 or more dummy atoms computed from the positions of the existing three atoms. All the possible virtual sites (2, 3, 3fd, 3fad, 3out and 4fd) can be calculated by the converter. This made it necessary to include additional parameters into the parameter file, if virtual sites has to be created. The unrealistic example of CaCl_2 in water using all the possible virtual sites will be given as an example. No fnc file is specified, in this example

| Line index (not part of the file!) | Parameter | Description |
|---------------------------------------|------------------------------------|---|
| 1 | 0 2 | ! <u>conversion type (0) RMC->MD, (1) MD->RMC, number of configurations</u> |
| 2 | cacl2_6p31_nx_1cc.cfg 4 | ! <u>RMC configuration file, number of RMC types</u> |
| 3 | cacl2_6p31_.gro | ! <u>GROMACS *.gro file</u> |
| 4 | cacl2.top | ! <u>GROMACS *.top file</u> |
| 5 | CaCl2 in water | ! <u>Title</u> |
| 6 | ffoplsaa.itp 1 1 | ! <u>force field parameter file; GROMACS bond type, angle type, D for force field default</u> |
| 7 | 0 0.020629 | ! <u>whether to rescale box (0: no, 1: yes), new number density in 1/(A3), if necessary</u> |
| 8 | 3 | ! <u>number of different residue types</u> |
| 9 | 1 2 3 | ! <u>character string (max 3 char) for each residue defining X in the bond angle and non-bonded constraint name definition Y_X_index.</u> |
| 10 | 510 1 CA2+ 1 | ! <u>as many line as residue types containing first the number of this residue, number of atoms in the residue, the name (max 5 char) for the residue, number of bonds for exclusions (recommended=3)</u> |
| 11 | 1020 1 CL- 1 | - - |
| 12 | 4490 10 SOL 2 | |
| 13 | 1 2 CA op1s_412 40.08 1 2.0 | ! <u>as many lines, as atoms in all the residue types: the index of this atom in the RMC conf. in the first then in the second molecule; GROMACS atom name, atom type, mass; charge group, charge, RMC-type, virtual type, indices from, virtual params</u> |
| 14 | 511 512 CL op1s_401 35.453 2 1.0 2 | - - |
| 15 | 531 1532 OW op1s_113 15.9994 3 | - - |

| | | |
|----|---|--|
| | 0.834 3 | |
| 16 | 021 6023 HW1 opl_s_114 1.008 3 .52 4 | - - |
| 17 | 022 6024 HW2 opl_s_114 1.008 3 .52 4 | - - |
| 18 | 0 0 MW opl_s_115 0.000 3 1.04 0 2 1 2 0.1 | - - |
| 19 | 0 0 MW2 opl_s_115 0.000 3 1.04 0 3 1 2 3 0.128 0.1 | - - |
| 20 | 0 0 MW3 opl_s_115 0.000 3 1.04 0 3fd 1 2 3 0.18 0.12 | - - |
| 21 | 0 0 MW4 opl_s_115 0.000 3 1.04 0 3fd 1 3 2 0.01 - .12 | - - |
| 22 | 0 0 MW5 opl_s_115 0.000 3 1.04 0 3fad 3 1 2 120 0.5 | - - |
| 23 | 0 0 MW6 opl_s_115 0.000 3 1.04 0 3out 2 1 3 0.2 0.1 .5 | - - |
| 24 | 0 0 MW7 opl_s_115 0.000 3 1.04 0 4fd 1 2 3 1 0.3 0.15 0.1 | - - |
| 25 | 0 | <i>! whether to use FNC, name of the FNC file (0:no, 1:use FNC lower limit, 2: use FNC upper limit, 3:use FNC average limit)</i> |

In the yellow block the virtual sites has to be counted among the atoms of a molecule, that is why there is 10 atoms for this water example with the 7 virtual sites.

In the green block the virtual sites has to be given similarly to the atoms, but the first and second index and mass has to be put to zero. This shows, that this is a virtual site, and extra information should be red for RMC→MD conversion after the RMC-type. The first of these is the virtual site type, which can be 2, 3, 3fd, 3fad, 3out and 4fd, see manual.pdf (4.7 and 5.2.2, maybe the index is different in different releases,) for details. Then follows the indices of the atoms (according to their number in the residue) building the virtual site (2-4 indices) and the parameters (1-3) related to the given virtual site type.

In case of MD→RMC conversion the virtual atoms will be left out from the RMC configuration.

The SOL.itp file created by the above example is given below. It has to be noted, that GROMACS before 4.0 used [dummiesX] instead of [virtual_sitesX], so change it if necessary.

```
[ moleculetype ]
;Name nrexcl
SOL 2

[ atoms ]
; nr type resnr residue atom cngr charge mass
qtot
0.166 1 opl_s_113 1 SOL OW 1 -0.834 15.9994 ;
-0.354 2 opl_s_114 1 SOL HW1 1 -0.52 1.008 ;
-0.874 3 opl_s_114 1 SOL HW2 1 -0.52 1.008 ;
```

```

-1.914      4  op1s_115      1  SOL      MW      1      -1.04      0  ;
-2.954      5  op1s_115      1  SOL      MW2     1      -1.04      0  ;
-3.994      6  op1s_115      1  SOL      MW3     1      -1.04      0  ;
-5.034      7  op1s_115      1  SOL      MW4     1      -1.04      0  ;
-6.074      8  op1s_115      1  SOL      MW5     1      -1.04      0  ;
-7.114      9  op1s_115      1  SOL      MW6     1      -1.04      0  ;
-8.154     10  op1s_115      1  SOL      MW7     1      -1.04      0  ;

[ virtual_sites2 ]
; Site from      funct      a
   4      1      2      1      0.1

[ virtual_sites3 ]
; Site from      funct      a      b
   5      1      2      3      1      0.128012065      0.1

[ virtual_sites3 ]
; Site from      funct      a      d
   6      1      2      3      2      0.18      0.12
   7      1      3      2      2      0.01      -0.12

[ virtual_sites3 ]
; Site from      funct      theta      d
   8      3      1      2      3      120      0.5

[ virtual_sites3 ]
; Site from      funct      a      b      c
   9      2      1      3      4      0.2      0.1      0.5

[ virtual_sites4 ]
; Site from      funct      a      b      d
  10      1      2      3      1      1      0.3      0.15      -
0.1

```

How to handle polarization with shell particles

Polarization can be included by attaching shell particles to certain atoms connected by a spring, see manual for the concept.

For conversion point of view this means, that extra shell particle has to be included into the GROMACS configuration. As no further information about the original position of the shell particle was given in the GROMACS manual, the initial position of the shell will coincide with the position of the atom, to which the shell is connected by the spring, as this seemed to work well. As an example the CsCl in SWM4-DP water will be given, where the water model consist of O, 2 H, a virtual site and a shell particle, connected to the oxygen.

| Line index (not part of the file!) | Parameter | Description |
|---------------------------------------|-----------------------------------|---|
| 1 | 0 2 | ! <u>conversion type (0) RMC->MD, (1) MD->RMC, number of configurations</u> |
| 2 | cscl.cfg 4 | ! <u>RMC configuration file, number of RMC types</u> |
| 3 | cscl.gro | ! <u>GROMACS *.gro file</u> |
| 4 | cscl.top | ! <u>GROMACS *.top file</u> |
| 5 | CsCl in water | ! <u>Title</u> |
| 6 | ffoplsaa.itp 1 1 | ! <u>force field parameter file; GROMACS bond type, angle type, D for force field default</u> |
| 7 | 0 0.2156 | ! <u>whether to rescale box (0: no, 1: yes), new number density in 1/(A3), if necessary</u> |
| 8 | 3 | ! <u>number of different residue types</u> |
| 9 | 1 2 3 | ! <u>character string (max 3 char) for each residue defining X in the bond angle and non-bonded constraint name definition Y_X_index.</u> |
| 10 | 50 1 Cs+ 1 | ! <u>as many line as residue types containing first the number of this residue, number of atoms in the residue, the name (max 5 char) for the residue, number of bonds for exclusions (recommended=3)</u> |
| 11 | 50 1 CL- 1 | - - |
| 12 | 3300 5 SOL 2 | |
| 13 | 1 2 Cs op1s_410 132.9054 1 .0 1 | ! <u>as many lines, as atoms in all the residue types: the index of this atom in the RMC conf. in the first then in the second molecule; GROMACS atom name, atom type, mass; charge group, charge, RMC-type, virtual type, indices from, virtual params</u> |
| 14 | 51 52 CL op1s_401 35.453 1 - .0 2 | - - |
| 15 | 01 102 OW1 op1s_113c 15.9994 1 | - - |

| | | |
|----|---|------|
| | 1.77185 3 | |
| 16 | 401 3402 HW2 opl_s_114 1.008 1 .55537 4 | - - |
| 17 | 701 6702 HW2 opl_s_114 1.008 1 .55537 4 | - - |
| 18 | 0 0 DW opl_s_115 0.000 1 1.1074 0 3 1 2 3 0.2031807494 0.2031807494 | - - |
| 19 | 01 102 SW WS 0.000 1 1.77185 -1 | - - |

For the shell -1 has to be specified as RMC type, so in case of MD→ RMC conversion the shell will not be converted back to the RMC file.