SUMMARY OF DATA FORMATS

Modelling: formats of atomic coordinates

XYZ format

Extension: .xyz (in most programs) .ani (in SIESTA)

Atomic xyz coordinates measured in Angstrom

```
[ comment line                ] !! NOT IMPLEMENTED !! DO NOT INCLUDE
[ N                           ] # of atoms, required by this xyz reader plugin  line 1
[ molecule name               ] name of molecule (can be blank)                 line 2
atom1 x y z [optional data]     atom name followed by xyz coords                line 3
atom2 x y z [ ...             ] and (optionally) other data.
...
atomN x y z [ ...             ]                                                 line N+2
```

see

http://www.ks.uiuc.edu/Research/vmd/plugins/molfile/xyzplugin.html

Modelling: formats of atomic coordinates

PDB format (originally from Protein Data Bank)

Extension: .pdb

<record type> <atom ID> <atom name> <residue name> <residue ID>
< x, y, and z coordinates in Angstrom > <occupancy> <beta factor> < segment name> <line>

```
ATOM      1  N   MET     1      27.340  24.430   2.614  1.00  9.67      1UBQ  71
ATOM      2  CA  MET     1      26.266  25.413   2.842  1.00 10.38      1UBQ  72
ATOM      3  C   MET     1      26.913  26.639   3.531  1.00  9.62      1UBQ  73
ATOM      4  O   MET     1      27.886  26.463   4.263  1.00  9.62      1UBQ  74
ATOM      5  CB  MET     1      25.112  24.880   3.649  1.00 13.77      1UBQ  75
ATOM      6  CG  MET     1      25.353  24.860   5.134  1.00 16.29      1UBQ  76
ATOM      7  SD  MET     1      23.930  23.959   5.904  1.00 17.17      1UBQ  77
ATOM      8  CE  MET     1      24.447  23.984   7.620  1.00 16.11      1UBQ  78
ATOM      9  N   GLN     2      26.335  27.770   3.258  1.00  9.27      1UBQ  79
ATOM     10  CA  GLN     2      26.850  29.021   3.898  1.00  9.07      1UBQ  80
ATOM     11  C   GLN     2      26.100  29.253   5.202  1.00  8.72      1UBQ  81
ATOM     12  O   GLN     2      24.865  29.024   5.330  1.00  8.22      1UBQ  82
ATOM     13  CB  GLN     2      26.733  30.148   2.905  1.00 14.46      1UBQ  83
ATOM     14  CG  GLN     2      26.882  31.546   3.409  1.00 17.01      1UBQ  84
ATOM     15  CD  GLN     2      26.786  32.562   2.270  1.00 20.10      1UBQ  85
ATOM     16  OE1 GLN     2      27.783  33.160   1.870  1.00 21.89      1UBQ  86
ATOM     17  NE2 GLN     2      25.562  32.733   1.806  1.00 19.49      1UBQ  87
```

see
https://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/node22.html

Modelling: formats of atomic data

General Topology Data for molecules

Extension: .inp .rtf (usually the name starts by top)

- Atom type and mass, connectivity (bonds, angles, dihedrals)
- Other atomic info can be included such as atom charge (obtained by DFT or semi-empirically)

```
MASS     1 H      1.00800 H ! polar H
MASS     2 HC     1.00800 H ! N-ter H
MASS     3 HA     1.00800 H ! nonpolar H
```

```
RESI ALA          0.00
GROUP
ATOM N    NH1    -0.47  !      |
ATOM HN   H       0.31  !  HN-N
ATOM CA   CT1     0.07  !      |      HB1
ATOM HA   HB      0.09  !      |     /
GROUP                   !  HA-CA--CB-HB2
ATOM CB   CT3    -0.27  !      |     \
ATOM HB1  HA      0.09  !      |      HB3
ATOM HB2  HA      0.09  !    O=C
ATOM HB3  HA      0.09  !      |
GROUP                   !
ATOM C    C       0.51
ATOM O    O      -0.51
```

see

https://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/node24.html

Modelling: formats of atomic data

Structure data File **Extension:** .psf

A first section with: atom ID, segment name, residue ID, residue name, atom name, type, charge and mass, and an unused field with 0.

Second section with molecular topology: bonds, angles, dihedrals

```
PSF CMAP

      6 !NTITLE
 REMARKS original generated structure x-plor psf file
 REMARKS 2 patches were applied to the molecule.
 REMARKS topology top_all27_prot_lipid.inp
 REMARKS segment U { first NTER; last CTER; auto angles dihedrals }
 REMARKS defaultpatch NTER U:1
 REMARKS defaultpatch CTER U:76

   1231 !NATOM
      1 U    1    MET  N    NH3   -0.300000       14.0070           0
      2 U    1    MET  HT1  HC     0.330000        1.0080           0
      3 U    1    MET  HT2  HC     0.330000        1.0080           0
      4 U    1    MET  HT3  HC     0.330000        1.0080           0
      5 U    1    MET  CA   CT1    0.210000       12.0110           0
      6 U    1    MET  HA   HB     0.100000        1.0080           0
      7 U    1    MET  CB   CT2   -0.180000       12.0110           0
```

```
1237 !NBOND: bonds
   1       5       2       1       3       1       4       1
   5       6       7       5       7       8       7       9
  10       7      10      11      10      12      13      10
```

see

https://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/node23.html