Files Needed in MD Simulations: PDB, PSF, RTF, PRM, DCD, NAMD conf and out Files

- **PDB** – Protein Data Bank
- **PSF** – Protein Structure File
- **RTF** – Residue Topology File
- **PRM** – Parameter Files
- **DCD** – Coordinate/Velocity Trajectory Files
- **conf** – NAMD Configuration File
- **out** - NAMD Standard Output (Log) File

PDB file

- PDB files (downloaded from the Protein Data Bank) contain standard records for species, tissue, authorship, citations, sequence, secondary structure, etc.
- We only care about the ATOM records…
  - atom name (N, C, CA)
  - residue name (ALA, HIS)
  - residue id (integer)
  - coordinates (x, y, z)
  - occupancy (0.0 to 1.0)
  - temperature (B-) factor
  - segment id (6PTI)
- PDB files from the Protein Data Bank contain NO hydrogen atoms! (need to be added when building the system)

PDB file format (Appendix A NAMD Tutorial)

========== ATOM ==========			
COLUMNS	DATA TYPE	FIELD	DEFINITION
1 - 6	Record name	"ATOM "	
7 - 11	Integer	serial	Atom serial number.
13 - 16	Atom	name	Atom name.
17	Character	altLoc	Alternate location indicator.
18 - 20	Residue name	resName	Residue name.
22	Character	chainID	Chain identifier.
23 - 26	Integer	resSeq	Residue sequence number.
27	AChar	iCode	Code for insertion of residues.
31 - 38	Real(8.3)	x	Orthogonal coordinates for X in Angstroms.
39 - 46	Real(8.3)	y	Orthogonal coordinates for Y in Angstroms.
47 - 54	Real(8.3)	z	Orthogonal coordinates for Z in Angstroms.
55 - 60	Real(6.2)	occupancy	Occupancy.
61 - 66	Real(6.2)	tempFactor	Temperature factor.
73 - 76	LString(4)	segID	Segment identifier, L-justified.
77 - 78	LString(2)	element	Element symbol, right-justified.
79 - 80	LString(2)	charge	Charge on the atom.

From PDB:

```
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
```

Saved from VMD:

```
ATOM 1 N  MET X 1 27.340 24.430 2.614 1.00  9.67  1UBQ
ATOM 2 CA MET X 1 26.266 25.413 2.842 1.00 10.38  1UBQ
```

Protein Structure File (PSF)

needs to be built!!! ,e.g., by using psfgen

- **Every atom in the system is listed.**
- **Provides five main sections:**
  - atom
  - bonds
  - angles
  - dihedrals
  - impropers
- **What is not in the PSF file?**
  - coordinates (dynamic data, initially read from PDB file)
  - velocities (dynamic data, initially from Boltzmann distribution)
  - force field parameters (non-specific, used for many molecules)

PSF file format (Appendix B NAMD Tutorial)

exp: see the ***ubq.psf*** file you have generated

```
PSF
       1 !NTITLE
 REMARKS original generated structure x-plor psf file
    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
```

```
1237 !NBOND: bonds
    1    5    2    1    3    1    4    1
    5    6    7    5    7    8    7    9
```

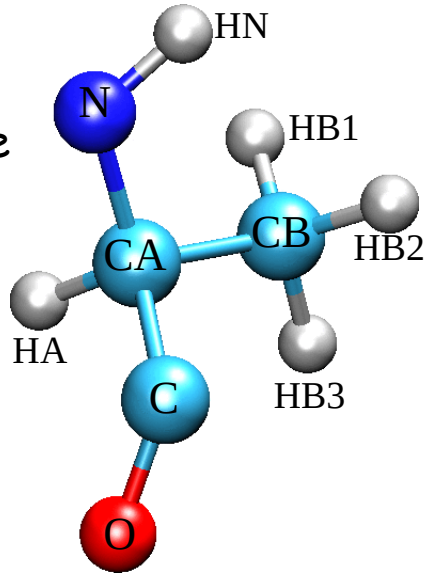
```
2257 !NTHETA: angles
    1    5    6    1    5   18    2    1    5
    2    1    4    2    1    3    3    1    5
```

```
3293 !NPHI: dihedrals
    1    5    7   10    1    5    7    8
    1    5    7    9    1    5   18   20
```

Topology File (RTF)

(Appendix C - NAMD Tutorial)

blueprints for building a PSF file



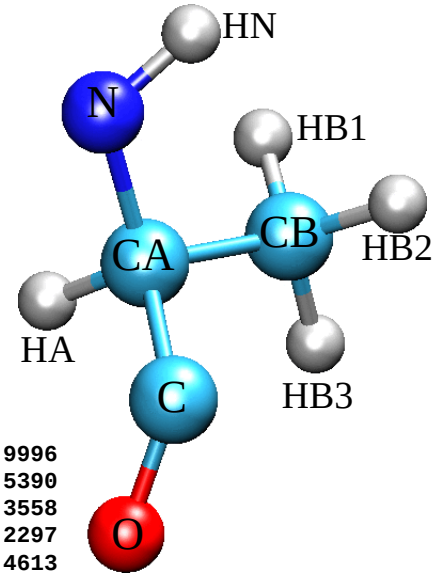
- For every type of residue known:
  - atom name, type, mass, and charge
  - bonds within the residue
  - bonds to other residues
  - any planar impropers (rare)
- Additional "patches" for:
  - terminating protein segments
  - joining protein segments
  - modifying protonation states
  - adding disulphide bonds

http://www.pharmacy.umaryland.edu/faculty/amackere/force_fields.htm

RTF File format

```
MASS    10 HS     1.00800 H ! thiol hydrogen
MASS    11 HE1    1.00800 H ! for alkene; RHC=CR
MASS    20 C     12.01100 C ! carbonyl C, peptide backbone
MASS    21 CA    12.01100 C ! aromatic C
MASS    22 CT1   12.01100 C ! aliphatic sp3 C for CH
------- (missing data here)
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
BOND CB CA  N  HN  N  CA
BOND C  CA  C  +N  CA HA  CB HB1  CB HB2  CB HB3
DOUBLE O  C
IMPR N -C CA HN  C CA +N O
DONOR HN N
ACCEPTOR O C
IC -C   CA   *N   HN    1.3551 126.4900  180.0000 115.4200  0.9996
IC -C   N    CA   C     1.3551 126.4900  180.0000 114.4400  1.5390
IC N    CA   C    +N    1.4592 114.4400  180.0000 116.8400  1.3558
IC +N   CA   *C   O     1.3558 116.8400  180.0000 122.5200  1.2297
IC CA   C    +N   +CA   1.5390 116.8400  180.0000 126.7700  1.4613
IC N    C    *CA  CB    1.4592 114.4400  123.2300 111.0900  1.5461
IC N    C    *CA  HA    1.4592 114.4400 -120.4500 106.3900  1.0840
IC C    CA   CB   HB1   1.5390 111.0900  177.2500 109.6000  1.1109
IC HB1  CA   *CB  HB2   1.1109 109.6000  119.1300 111.0500  1.1119
IC HB1  CA   *CB  HB3   1.1109 109.6000 -119.5800 111.6100  1.1114
```

Alanine



RTF File: H_2O

```
RESI TIP3          0.000
! tip3p water model
! generate using noangle nodihedral
GROUP
ATOM OH2  OT      -0.834
ATOM H1   HT       0.417
ATOM H2   HT       0.417
BOND OH2 H1 OH2 H2 H1 H2
! the last bond is needed for “shake”, i.e.,
  rigid solvent model
ANGLE H1 OH2 H2                ! required
ACCEPTOR OH2
```

RTF File

- Differences between RTF's for single molecules and residues
  - First residue converted to **N-terminus** NH_3^+
  - Last residue converted to **C-terminus** $CO\text{–}O^-$

- Special atoms
  - *-C, -O, +N, +H, +CA*
  - Refers to atoms in residues preceding (-) or following (+)

Parameter File (PRM)

Appendix E – NAMD Tutorial

- **Equilibrium value and spring constant** for
  - every pair of atom types that can form and bond
  - every triple of atom types that can form an angle
  - every quad of atom types that can form a dihedral or improper (many wildcard cases)
- **vdW radius and well depth for every atom type**
  - actually need these for every pair of atoms types!
  - pair radius calculated from arithmetic mean
  - pair well depth calculated from geometric mean
- **Closely tied to matching topology file!**

http://www.pharmacy.umaryland.edu/faculty/amackere/force_fields.htm

PRT File Format

```
BONDS
!V(bond) = Kb(b - b0)**2
!Kb: kcal/mole/A**2
!b0: A
!atom type Kb          b0
!
C    C     600.000     1.3350 ! ALLOW ARO HEM
               ! Heme vinyl substituent (KK, from propene (JCS))
CA   CA    305.000     1.3750 ! ALLOW   ARO
               ! benzene, JES 8/25/89
------------------------- (missing data here)--------------
ANGLES
!V(angle) = Ktheta(Theta - Theta0)**2
!V(Urey-Bradley) = Kub(S - S0)**2
!Ktheta: kcal/mole/rad**2
!Theta0: degrees
!Kub: kcal/mole/A**2 (Urey-Bradley)
!S0: A
!
!atom types     Ktheta    Theta0   Kub     S0
!
CA   CA   CA    40.000    120.00   35.00   2.41620 ! ALLOW   ARO
               ! JES 8/25/89
CE1  CE1  CT3    48.00    123.50
```

NAMD Configuration File

Appendix E – NAMD Tutorial

Example

```
# Minimization and Equilibration of
# Ubiquitin in a Water Sphere

#############################################################
#  ADJUSTABLE PARAMETERS
#############################################################
structure          ../common/ubq_ws.psf          # load system PSF file
coordinates        ../common/ubq_ws.pdb          # load system PDB file
set temperature    310                           # set temperature 310 K  (Tcl)
set outputname     ubq_ws_eq                     # base name of output files (Tcl)
firsttimestep      0                             # start simulation at t=0
```

SIMULATION PARAMETERS

```
# Input
paraTypeCharmm    on                          ;# use CHARMM forcefield parameters
parameters   ../common/par_all27_prot_lipid.inp   ;# load PRM file
temperature  $temperature                      ;# set temperature

# Force-Field Parameters    - specific to CHARMM
exclude          scaled1-4
1-4scaling       1.0
switching        on
cutoff           12.              # specified by user
switchdist       10.
pairlistdist     13.5

# Integrator Parameters
timestep         2.0               ;# 2fs/step
rigidBonds       all               ;# needed for 2fs steps
nonbondedFreq    1                 ;# non bonded forces calculated every steps
fullElectFrequency  2              ;# use PME only every other step
stepspercycle    10                ;# redo pairlists every ten steps
```

SIMULATION PARAMETERS (cont)

```
# Constant Temperature Control
langevin         on          ;# do langevin dynamics
langevinDamping    5     ;# damping coefficient (gamma) of 5/ps
langevinTemp    $temperature   ;# heat bath temperature
langevinHydrogen    off    ;# don't couple Langevin bath to hydrogens


# Output
outputName        $outputname    ;# set base name for output files

restartfreq       500         ;# write restart files w/ frequency 500steps = every 1ps
dcdfreq           500         ;# output coordinates in DCD trajectory file every 500steps
outputEnergies    100     ;# output energies/temperature,etc. every 100steps
outputPressure    100     ;# output pressure every 100steps
```

EXTRA PARAMETERS

```
# Spherical harmonic boundary conditions (BC)
sphericalBC        on      ;# turn on  spherical BC
# set location of sphere's center
sphericalBCcenter   30.3081743413, 28.8049907121, 15.353994423
# use only 1 boundary potential (max 2 can be used)
sphericalBCr1       26.0  ;# distance from center where 1st boundary potential sets in
sphericalBCk1       10    ;# force constant for 1st potential
sphericalBCexp1     2     ;# exponent for 1st potential
```

$$U(r) = k(r - r_0)^2, \quad if \quad r > r_0$$

```
#  EXECUTION SCRIPT
#####################################################################
# Minimization
minimize            100     ;# run structure minimization for 100 steps
reinitvels          $temperature  ;# reinitialize velocities corresponding to 310K

run 5000            ;# run MD simulation (equilibration) for 10ps
                    ;#  recall:  timeStep = 2fs!
```

NAMD Output File

Appendix F – NAMD Tutorial

- running NAMD2 jobs (simulations)

charmrun namd2 ++local my_job.conf > my_job.out &

```
ETITLE: TS BOND ANGLE DIHED IMPRP ELECT VDW BOUNDARY MISC KINETIC
        TOTAL TEMP TOTAL2 TOTAL3 TEMPAVG PRESSURE GPRESSURE
        VOLUME PRESSAVG GPRESSAVG

ENERGY: 1000 0.0000 0.0000 0.0000 0.0000 -97022.1848 9595.3175 0.0000 0.0000
          14319.5268 -73107.3405 300.2464 -73076.6148 -73084.1411 297.7598
          -626.5205 -636.6638 240716.1374 -616.5673 -616.6619
```

- use namdplot to graph these quantities vs time, e.g.,

namdplot TOTAL TEMP vs TS my_job.out &

NAMD Output File (cont.)

```
Info: Benchmark time: 47 CPUs 0.0475851 s/step 0.275377 days/ns 13540 kB memory
```

```
TIMING: 1000 CPU: 18.35, 0.01831/step Wall: 50.1581, 0.0499508/step,
        6.92374 hours remaining, 14244 kB of memory in use.
```

```
OPENING COORDINATE DCD FILE
WRITING COORDINATES TO DCD FILE AT STEP 1000
```

```
Warning: Pairlistdist is too small for 1 patches during timestep 17.
Warning: Pairlists partially disabled; reduced performance likely.
Warning: 20 pairlist warnings since previous energy output.
```

(too short pairlist distance AND/OR too long cycle length → reduced performance)

Do not ignore warnings you do not understand!