THE UNIVERSITY of TEXAS

HEALTH SCIENCE CENTER AT HOUSTON

SCHOOL of HEALTH INFORMATION SCIENCES

Molecular Dynamics Simulation: Practice I

For students of HI 6327 “Biomolecular Modeling”

Willy Wriggers, Ph.D.

School of Health Information Sciences

http://biomachina.org/courses/modeling/04.html

Theoretical Calculations with X-PLOR

- Energy minimization
- MD calculations
- MD analysis
  (correlation functions)
- Free energy calculations
  (perturbation method, thermodynamic integration)
- Quantum–mechanical corrections (path integral)

X-PLOR Pros and Cons

Pros:

- free
- used in structural biology for refinement: widely accepted/known
- wealth of scripts and documentation available
- powerful scripting language
- great all-around package
- compatible with NAMD
- synthax similar to CHARMM

Cons:

- old (1998), no longer supported
- copyright ownership confusion, CNS
- doesn't have the latest MD and analysis tools (cf. NAMD, CHARMM, Amber)

X-PLOR Preparations

Make sure the following lines are in your .bashrc or .profile:

export PATH=$PATH:/home/public/Global/RH9/bin
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/home/public/Global/RH9/lib

Use as “xplor” at student account command prompt

Run jobs on machines:

flensted, emerson, bedia, pardo, borges, dacosta, carvalho

Interactive Use

- typing "xplor" generates the prompt

  ```
  X-PLOR>
  ```

- there is some online help

  ```
  X-PLOR> help
  ```

  produces a list of available commands

- X–PLOR has a powerful interpreted command language
  - variables (real and string)
  - if–statements
  - for– and while–loops
  - "vector" manupulations (e.g. coordinates)
  - data manipulations
  - mathematical functions

- example:

```
evaluate ($count = $count + 1)
if ($count > 5) then
  vector do (x = ran()*y^2)
end if
```

Commands

- commands are in general not case sensitive:

```
HELP = hElp
```

- commands can go over several lines
- some commands have fixed number of parameters:

```
ASSIgn (resid 1 and name ha)
       (resid 1 and name hn)
3.0 1.0 1.0
```

- others end with "end"

```
noe scale dist 50 end
```

- usually, 4 characters (sometimes 3 or 5) are sufficient:

```
ASSI = ASSIG = ASSIGN
```

Comments

3 types of comment lines:

1. exclamation mark !
   the rest of the line is ignored
   example

   ```
   coor   ! this is a comment...
   ```

2. curly brackets
   the contents of the { } is ignored
   example

   ```
   dynamics verlet init{ial}t = 1000 end
   ```

3. REMARK
   the line beginning with REMARK is ignored,
   but stored and written to the next output file
   (especially coordinate file)

Opening Files: @ and @@

- In general, files are opened with @ or @@.
- Both switch the "command stream" to the file.
- @–files are stored on internal command buffer
  (for *loops* or *if–statements*)
  are only opened once in a loop
- @@–files are only parsed
  cannot contain *loops* or *if–statements*
- "@ for command files – @@ for data files"
- Warning: some commands expect only the file name:

```
read trajectory
     input = coords.crd
     ...
end
```

File Names

Filenames are case sensitive (UNIX), and can be specified with absolute or relative path:

```
@@../../parallhdg.pro
@@/data4/Users/nilges/toppar2/parallhdg.pro
```

Environment variables

```
@@TOPPAR:parallhdg.pro
```

where TOPPAR has been defined by (e.g. UNIX tcsh:)

```
setenv TOPPAR /data4/Users/nilges/toppar2
```

Topology and PSF Files in X-PLOR

Topology file: residue library that defines standard molecule components (amino acids, nucleotides):

- atom definitions (masses)
- residue definitions (covalent topology)
- charges
- patches ("presidue") for modifications:
  - peptide bond
  - disulfide bridge
  - N– and C–termini
- no coordinates!
- no bond lengths, force constants etc!

The topology of a specific molecule is stored in the PSF:
Sequence + Topology $\rightarrow$ PSF

Example of a Residue in Topology File

```
residue ALA
  group
    atom N   type=NH1 charge=-0.36 end
    atom HN  type=H   charge= 0.26 end
    atom CA  type=CT  charge= 0.00 end
    atom HA  type=HA  charge= 0.10 end
    atom CB  type=CT  charge=-0.30 end
    atom HB1 type=HA  charge= 0.10 end
    atom HB2 type=HA  charge= 0.10 end
    atom HB3 type=HA  charge= 0.10 end
    atom C   type=C   charge= 0.48 end
    atom O   type=O   charge=-0.48 end

  bond N  HN    bond N  CA    bond CA  HA
  bond CA  CB   bond CB  HB1  bond CB  HB2
  bond CB  HB3  bond CA  C    bond C   O

  improper HA  N   C   CB  !chirality CA
  improper HB1 HB2 CA HB3  !methyl group CB
end
```

Example of a Patch

```
presidue PEPT
  add bond -C +N

  add angle -CA -C +N
  add angle -O  -C +N
  add angle -C  +N +CA
  add angle -C  +N +HN

  add improper -O  -C  +N  +CA
  add improper +HN +N  -C  -CA
  add improper -CA -C  +N  +CA
end
```

Input File Example: generate.inp

```
topology
    @@TOPPAR:topallhdg.pro
end

segment
    name="    "
    chain
       @@TOPPAR:toph19.pep
       sequence
           Ala Ala
       end
    end
end

REMARK ALA dipeptide
write structure
    output=INPUT:diala.psf
end

stop
```

Input File Example: generate.inp

- "segment" defines a new segment of the molecular structure.
- several segments possible (e.g. in complexes or multimers)
  - protein
  - DNA
  - water
- the name of the segment corresponds to the PDB coordinate file (last 4 characters before card number)

Input File Example: generate.inp

- "chain" concatenates residues
- "toph19.pep" determines standard patches to concatenate residues:

```
LINK PEPP  HEAD - *   TAIL + PRO   END
LINK PEPT  HEAD - *   TAIL + *     END
```

The "*" is a wildcard:

Before the patch PEPP (the "HEAD") is any residue
The HEAD is referred to in the patch as "-".

After the patch PEPP (the "TAIL") is PRO
The TAIL is referred to in the patch as "+".

Note: exceptions (like PRO) come first.

Input File Example: generate.inp

- and terminate the molecule

```
FIRSt PROP                    TAIL + PRO      END
FIRSt NTER                    TAIL + *        END

LAST  CTER   HEAD - *                         END
```

- "sequence" specifies the sequence
- "sequence ... end" can be replaced by
  "coor @@example.pdb"
  (note: "end" in coor file!)
- the REMARKs will be written to the PSF file
- PSF file is written by WRITE STRUCTURE ... END

Content of a PSF File

- Note: usually no need to look at the file – do not modify
- Header: REMARK records
  Filename,date etc are generated by WRITE

```
PSF

       3 !NTITLE
 REMARKS FILENAME="diala.psf"
 REMARKS ALA dipeptide
 REMARKS DATE:07-Sep-95  10:16:16
```

- list of all atoms

```
      23 !NATOM
       1       1    ALA  CA   CT      0.220000
       2       1    ALA  HA   HA      0.100000
...
      22       2    ALA  OT1  OC     -0.570000
      23       2    ALA  OT2  OC     -0.570000
```

Content of a PSF File

- list of all bonds

```
        22 !NBOND: bonds
         9       1       1       2       1       3
         3       5       3       6       1       7
...
        21      22      21      23
```

- same for
  - bond angles
  - dihedrals
  - impropers
  - hydrogen bond donors and acceptors
  - non–bonded groups

Resources and Further Reading

WWW:
http://xplor.csb.yale.edu

Books:
Chapters 1-3 in:
A.T. Brunger, X-PLOR Version 3.1 (Yale U Press, 1992)
online free at http://alpha2.bmc.uu.se/local_html/xplor_mirror.html

Acknowledgement:
Michael Nilges Primer
http://www.pasteur.fr/recherche/unites/Binfs/xplor/primer/
Lecture 1