

PROTOMOL Tutorial

Thierry Matthey, para//ab, UNIFOB, University of Bergen

- PROTO MOL & MD
- Getting Started & Running
- Design & Source
- Adding a New Force
- References

- Configuration file defining globals, integration and system input files
- CHARMM force field and compatibility with two versions (19 and 28a2, <http://www.lobos.nih.gov/Charmm>)
- Multiple time-stepping
- Fast electrostatic computations
- Comparison of force algorithms
- Hybrid Monte Carlo sampling
- Ability to interact with VMD – DCD trajectory files
- Ability to run on parallel machines

1. Construct initial configuration of x_0 , v_0 and F_0 ;

2. loop 1 to number of steps

(a) Update velocities

(b) Evaluate forces on each particle
($\vec{F}_i = -\Delta_i U$)

(c) Update positions

(d) Update velocities

3. Post-processing

$$U = U_{\text{bonded}} + U_{\text{non-bonded}} + U_{\text{external}}$$

$$U_{\text{bonded}} = U_{\text{bond}} + U_{\text{angle}} + U_{\text{dihedral}} + U_{\text{improper}}$$

$$U_{\text{non-bonded}} = U_{\text{electrostatic}} + U_{\text{Lennard-Jones}}$$

U_{bonded}

Angle, Bond, Dihedral and Improper

$U_{\text{non-bonded}}$

Lennard-Jones (van der Waals) and Coulomb using

cutoff, Ewald, PME or MG

U_{external}

Friction, Haptic, PaulTrap, Spherical, Gravitation, Magnetic Dipole and Magnetic Dipole Mirror

Getting Started & Running

- Binary: `http://www.nd.edu/~lcls/protomol/`

(Solaris, IRIX, Linux and AIX)

- Source: *protomol@cse.nd.edu*

- Examples: `http://www.nd.edu/~lcls/protomol/`

- User-guide: `http://www.nd.edu/~lcls/protomol/`

- Download: Binary & alanin example
- Run: `protomol alanin.conf --numsteps 1000`
- Help: `protomol -h`
- Forces available: `protomol -f`
- Visualize: `vmd` (Molecule → Load From Files → psf and dcd → select & load)

(<http://www.ks.uiuc.edu/Research/vmd>)

- Configuration file, e.g., alanin.conf
 - defines global parameters and the integration scheme
- Coordinates, e.g., alanin.out.pos.pdb
 - atom positions, either PDB or XYZ format
- Velocities or temperature, e.g., alanin.out.vel.pdb
 - atom velocities, either PDB or XYZ format
- Charmm parameter file, e.g., alanin.par
 - defines non-bonded parameter, Coulomb & Lennard-Jones
- PSF topology file, e.g., alanin.psf
 - defines the mass, charge, and the topology of each molecule

```
keyword1 value1  
keyword2 value2  
keyword3 value3
```

```
...  
} Integrator {
```

```
level N-1 <integrator type> {
```

```
<integrator arguments> (These will differ depending on the integrator type).  
<integrator forces> (These are all optional).
```

```
...
```

```
level 0 <integrator type> {
```

```
...
```

```
}
```

```
}
```

Note that the order of definition for each level is not strict, but PROTO MOL expects one definition for each level.

```

Integrator {
  level 0 Leapfrog {
    timestep 1
    force Improper
    force Dihedral
    force Bond
    force Angle
    force LennardJones
    -algorithm NonbondedCutoff
    -switchingFunction C2
    -switchon 1.0
    -cutoff 12
    force Coulomb
    -algorithm NonbondedCutoff
    -switchingFunction C1
    -cutoff 12
    force Haptic
    -port 2001
    -trate 1
    -timeout 1000
    -step_inc 100
    -wait 0
  }
}

```

```

temperature 300.0
firststep 0
numsteps 10000
cellsize 6.5
outputfreq 100
restartfreq 1000
seed 1234
posfile alanin.pdb
psfile alanin.psf
parfile alanin.par
usecharmm28parfile no
findbposfile alanin.out.pos.pdb
findbvelfile alanin.out.vel.pdb
dcfile alanin.out.dcd
allenergiesfile alanin.out.energy
restartfile alanin.out
boundaryConditions Normal
cellManager Cubic

```

Support of multiple time-stepping integration

```

Integrator {
  level 2 Impulse { # Long-range electrostatics
    cyclelength 4
  }
  level 1 Impulse { # Medium varying forces
    cyclelength 2
    force Improper, Dihedral
    force Coulomb -algorithm Full -switchingFunction ComplementSW1}
  level 0 Leapfrog { # Fast varying forces
    timestep 0.5 # [fs]
    force Bond, Angle}
}
Three-level Verlet/r-RESPA MTS

```

- Multiple Timestep Integrators

1. BSpineMOLLY

2. DMLeapfrog

3. EquilibriumMOLLY

4. HBondMOLLY

5. HybridMC (*Hybrid Monte Carlo Integrator*)

6. Impulse

- Single Timestep Integrators

1. BBK

2. LangevinImpulse

3. Leapfrog (*Velocity Leap-Frog Integrator*)

4. NoseNVTLeapfrog

5. PLLeapfrog (*Position Leap-Frog Integrator*)

Accuracy:

force compare <force-approximation type >
<force-approximation arguments >
force compare <force-exact type >
<force-exact arguments >

Benchmarking:

force time <force-to-benchmark type >

Accuracy & Benchmarking:

force compare time <force-approximation type >
<force-approximation arguments >
force compare time <force-exact type >
<force-exact arguments >

Sample of force comparison between PME and plain Ewald:

```
force compare Coulomb
-algorithm PMEwald
-real -reciprocal -correction
-gridsize 8 8 8
-order 4
-cutoff 6.5
force compare Coulomb
-algorithm FullEwald
-real -reciprocal -correction
```

Sample of force comparison and benchmarking PME with different accuracy:

```
force compare time force coulomb  
-algorithm PMEwald -real -reciprocal -correction  
-interpolation BSpline  
-cutoff 6.5 -gridsize 10 10 10  
force compare time force coulomb  
-algorithm PMEwald -real -reciprocal -correction  
-interpolation BSpline  
-cutoff 6.5 -gridsize 20 20 20  
force time LennardJones  
-algorithm NonbondedCutoff  
-switchingFunction c1  
-cutoff 8.0
```

- 1. line: Number of atoms

- 2. line: Blank or any comment.

- Rest: Atomic number or name and then the x, y, and z coordinates; separated by at least 1 space

8

Toluene from the Cambridge Crystallographic Database

```
C      1.58890      -1.44870      -0.47000
C      1.54300      -2.25990      0.77910
C      2.21440      -3.47410      0.88040
C      2.16940      -4.22350      2.03080
C      1.45120      -3.77740      3.11890
C      0.77320      -2.58240      3.03840
C      0.82440      -1.82660      1.89060
H      2.40210      -1.69480      -0.94010
```

Definition: <http://hackberry.chem.trinity.edu/IJC/Text/xmolxyz.html>

REMARK FILENAME="/usr/people/nonella/xplor/benchmark1/ALANIN.PDB"

...

```
REMARK DATE:16-Feb-89 11:21:32 created by user: nonella
ATOM 1 CA ACE -2.184 0.591 0.910 0.0 10.0 MAIN
ATOM 2 C ACE -0.665 0.627 0.966 100.0 0.0 MAIN
ATOM 3 O ACE -0.069 1.213 1.868 100.0 0.0 MAIN
ATOM 4 N ALA 0.000 0.000 0.000 100.0 0.0 MAIN
ATOM 5 H ALA -0.490 -0.462 -0.712 100.0 0.0 MAIN
...
ATOM 65 H CBX 8.050 8.324 9.225 100.0 0.0 MAIN
ATOM 66 CA CBX 9.223 8.571 11.014 0.0 10.0 MAIN
END
```

Definition: <http://www.rcsb.org/pdb/docs/format/pdbguide2.2/guide2.2.frame.html>
 or <http://www.rcsb.org/>

- More detail information

```

4 INTITLE
REMARKS FILENAME="pt1.psf"
REMARKS BPT1 COORDINATES TAKEN FROM CRYSTALLOGRAPHIC DATA W/O WATERS
REMARKS HYDROGEN POSITIONS GENERATED USING HBUILD ( 2 ITERATIONS)
REMARKS RMS FLUCTUATIONS (T, ICHEVE) FOR HEAVY PROTEIN ATOMS INCLUDED
REMARKS DATE:24-Apr-89 02:34:58 created by user: heller
568 INATOM
1 MAIN 1 ARG HT1 HC 0.260000 1.008000
2 MAIN 1 ARG HT2 HC 0.260000 1.008000
3 MAIN 1 ARG N NH3 0.000000E+00 14.0067
4 MAIN 1 ARG HT3 HC 0.260000 1.008000
582 INBOND: bonds
3 5 5 18 18
6 7 7 8 8
834 INTHTETA: angles
3 5 5 6 5
18 5 6 7 6
351 INPHI: dihedrals
3 5 6 7 7
6 7 8 9 9
259 INIMPHI: improper
11 12 15 18 9
10 18 15 18 6
9 12 13 12 9
114 INDON: donors
9 10 12 13 12
15 17 3 1 3
79 INACC: acceptors
19 18 26 32 25
35 34 47 46 46
24 INNB
45 44 43 97 96
208 224 223 222 236 235 234 210 209
222 0 INGRP
0 0 0 0 0 0 0 0 0

```

- Defines the topology of the molecules
 - Additionally some structural information
- Definition: <http://www.lrz-muenchen.de/~heller/ego/manual/node88.html> OR <http://www.rcsb.org/>

```

...
remark - parameter file PARAM19 -
...
bond c c 450.0 1.381 B. R. GELIN THEISIS AMIDE AND DIPEPTIDES
bond c CH1E 405.0 1.521 EXCEPT WHERE NOTED, CH1E,CH2E,CH3E, AND CT
bond c CH2E 405.0 1.521 ALL TREATED THE SAME. UREY BRADLEY TERMS ADDED
bond c CH3E 405.0 1.521
bond c CR1E 450.0 1.381
bond c CT 405.0 1.531
bond c N 471.0 1.331
bond c NC2 400.0 1.331 BOND LENGTH FROM PARAMFIX9 FORCE K APPROXIMATE
bond c NH1 471.0 1.331
bond c NH2 471.0 1.331
bond c NP 471.0 1.331
bond c NR 471.0 1.331
bond c O 580.0 1.231
bond c OC 580.0 1.231 FORCE DECREASE AND LENGTH INCREASE FROM C O
bond c OH1 450.0 1.381 FROM PARAMFIX9 (NO VALUE IN GELIN THEISIS)
bond c OS 292.0 1.431 FROM DEF NORMAL MODE FIT
bond CH1E CH1E 225.0 1.531
bond CH1E CH2E 225.0 1.521
bond CH1E CH3E 225.0 1.521
...
angle c c c 70.0 106.51 FROM B. R. GELIN THEISIS WITH HARMONIC
angle c c CH2E 65.0 126.51 PART OF F TERMS INCORPORATED. ATOMS
...
dihc CH1E c N CH1E 10.0 2 180.01 PRO ISOM. BARRIER 20 KCAL/MOL.
dihc CH2E c N CH1E 10.0 2 180.0
...
impr c c CR1E c CR1E CH2E 90.0 0 0.01USED HERE FOR TRP CG OUT OF PLANE
NONbonded H 0.0498 1.4254 0.0498 2.6157 0.0450 2.6157 i- charged group.
NONbonded HA 0.0450 2.6157
...

```

Definition: See user-guide

- Defines the parameters of bonded and non-bonded forces
- Atomic names correspond to PSF format

- Restart files
- Energy or Energy split
- Position, velocity, force and DCD trajectories
- Final positions, velocities and forces
- 3D graphic files, BSDL
- Volume, diffusion, temperature, PaulTrap, momentum, etc.

1. Run auto configuration

```
aclocal! autoheader! automake -a!
```

2. Chose platform and configuration

```
e.g., ./configure --with-gcc-debug
```

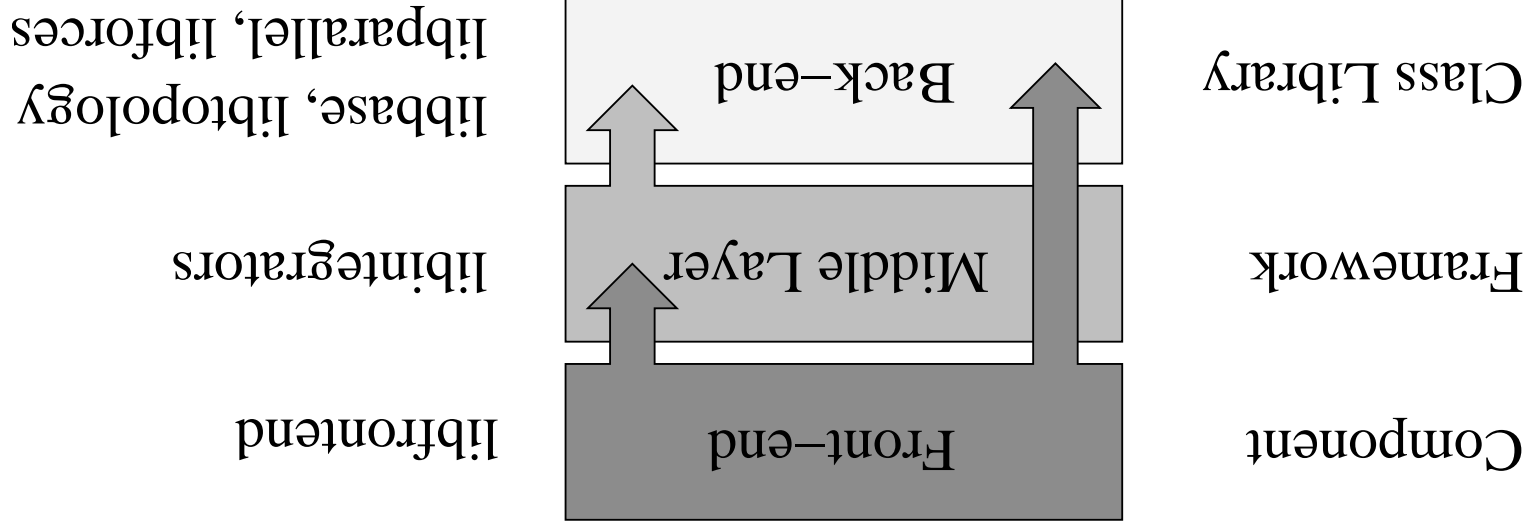
3. Compile

```
make depend! make clean! make
```

4. If any problems consult README and FAQ. compiler

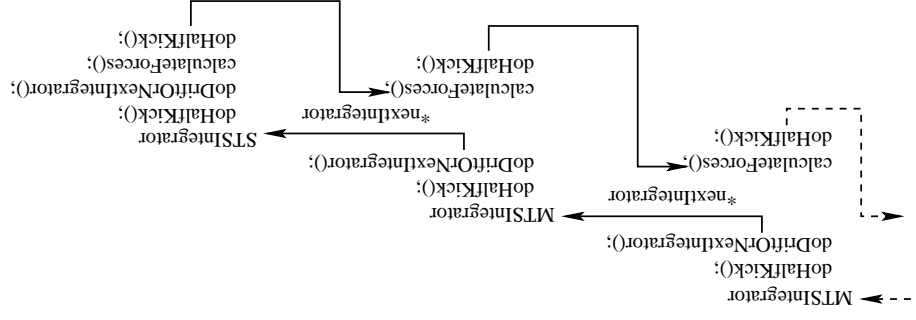
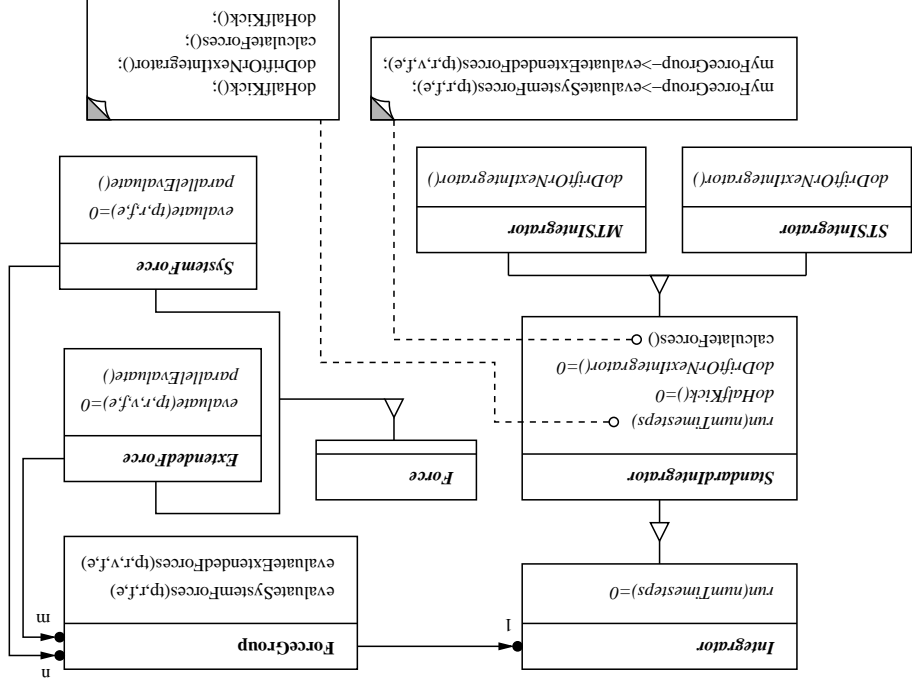
5. Binary: ./applications/protomol-app/protomol

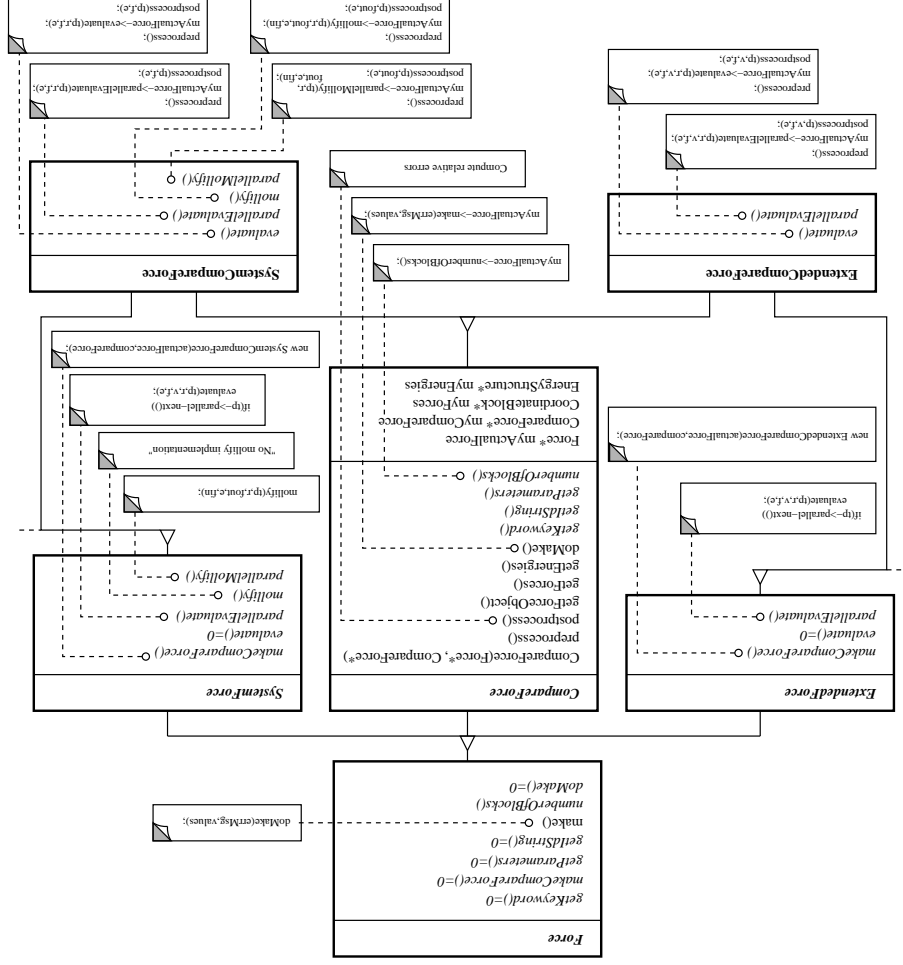
Design & Source



- framework/base : basic classes
- framework/forces : force computation
- framework/topology : topology, boundary conditions & switching functions

- framework/ind : networking, haptic device
- framework/parallel : parallelism (MPI)
- framework/integrators : Middle layer, integrators
- framework/frontend : I/O & user interface
- applications/protomol-app : protomol & main





R1 Algorithm to select an n -tuple of particles

R2 Boundary conditions

R3 Retrieve efficiently ($\mathcal{O}(1)$) the spatial information of each particle

R4 Potential defining the force and energy on the n -tuple

R5 Switching function to modify the potential

→ Policy, Abstract Factory and Prototypes

E.g.:

```
Force* aForce = new AnglesSystemForce<PBC>();
```

```
typedef OneAtomPair<PBC,CM,ClSwitchingFunction,CoulombForce,true> OneAtomPairType!  
class NonbondedCutoffSystemForce<PBC,CM,OneAtomPairType>!
```

Adding a New Force

- Find a similar force/ potential
- Take a copy and change it
- Add it to the Makefile
- Register your new force
- Reconfigure and recompile

- Copy from `BondSystemForce*. [Ch] or PaulTrapExtendedForce*. [Ch]`
- Replace `Bond` by `MyNewForceName`
- Change the implementation, remove `parallelEvaluate` and `numberOfBlocks` for pure sequential implementation
- Change the keyword, parameters, etc.
- Add `MyNewForceNameSystemForce*. [Ch]` to `Makefile.am`
- Register your new force in `frontend/registerOtherForceExamples.C`
- Reconfigure – `aclocal! autoheader! autocomf! automake -a!`
- Compile – `make depend! make clean! make`

- Copy from `gravitationForce.[ch]`
- Replace `gravitation` by `MyNewPotentialName`
- Change the implementation
- Change the keyword, parameters, etc.
- Add `MyNewPotentialName.[ch]` to `makefile.am`
- Register your new force in `frontend/registerOtherForceExamples.C` accordingly `gravitationForce` or `CoulombForce` – chose your requirements R2-R5
- Reconfigure – `aclocal! autoheader! autoconf! automake -a!`
- Compile – `make depend! make clean! make`

- <http://www.nd.edu/~lcls/protomol/>
- <http://www.ks.uiuc.edu/Research/vmd>
- <http://hackberry.chem.trinity.edu/IJC/Text/xmolxyz.html>
- http://www.rcsb.org/pdb/docs/format/pdbguide2.2/guide2.2_frame.html
- <http://www.rcsb.org/pdb/>
- <http://www.lrz-muenchen.de/~heller/ego/manual/node88.html>
- <http://www.it.uib.no/matthey/text/mattheyPhD.pdf>