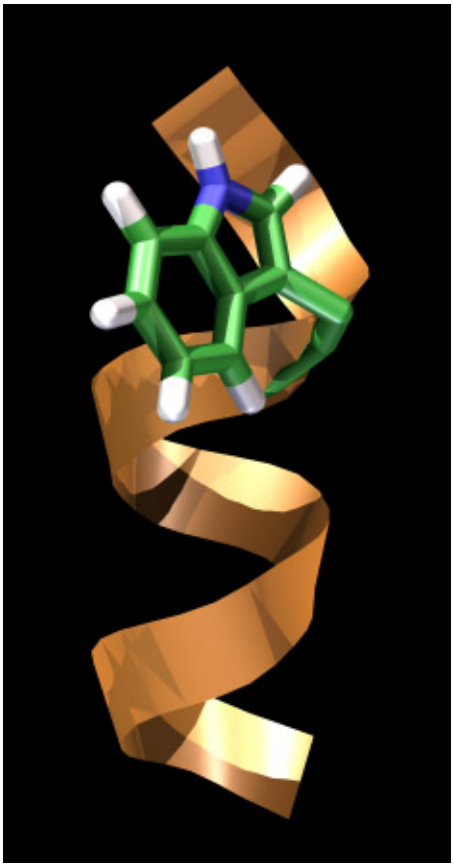
Hippo - Manual



Contents

Please note that we cannot guarantee the accuracy of this manual. Please report errors at: http://forums.biowerkzeug.org/

Summary

1. What is hippo?

- A molecular mechanics software to model, simulate and analyse bio-molecules
- Available online at : www.biowerkzeug.com

2. Features

- Molecular dynamics (NVT, NPT, NVE)
- Metropolis Monte Carlo (NVT, NPT)
- Replica exchange molecular dynamics and Monte Carlo
- Generalized Born implicit solvent (GB/SA)
- Generalized Born implicit membrane (GBIM)
- Multi-threading (openMP)
- Analysis tools
- Supported force fields: OPLS

3. Platforms

- Windows
- Linux (SUSE, UBUNTU, CENTOS)
- Mac OS X on Intel Macs (x86)

4. Files

Executable	hippo.exe (windows) or hippo (linux/mac)
Simulation settings	hippo_input.txt
Force field	oplsaa_forcefield.dat
Structure database	hippo_protein_database.dat
Temperatures	temperatures_remc.dat (for replica exchange only)

5. Input options

Sequence	SEQ: Single or triple letter code sequence
Protein databank	PDB: Structure file

6. Units

Length	Angstrom	1 Å = 10^{-10} m
Energy	kcal/mol	1 kcal/mol = 4.184 kJ/mol
Temperature	Kelvin	K
Pressure	bar	1 bar = 10^5 Pa = 10^5 N/m^2

1 Simulation

1.1 Setup

1. **Download hippo files from www.biowerkzeug.com**
   - hippo.exe (windows) or hippo (linux/mac)
   - hippo_input.txt
   - oplsaa_forcefield.dat
   - hippo_protein_database.dat
   - temperatures_remc.dat (for replica exchange only)

2. **Edit hippo_input.txt, e.g.**
   - mode: **MD simulation**
   - input type: **PDB** for protein databank
   - input file: Put the name of your PDB file here
   - max_mcsteps: Number of steps to run

   See below for a description of other settings (e.g. temperature, etc.)

3. **Execute hippo.exe (windows) or hippo (linux)**

1.2 Output

1. **Data files**
   - log.txt — logfile (text format)
   - movie.pdb — trajectory (PDB format)
   - movie.xyz — trajectory (binary format)
   - E.dat — energies (text format)
   - NPT.dat — thermodynamic variables (text format)
   - output.dat — simulation summary (text format)
   - av.dat — energy averages (text format)
   - av_vs_step.dat — energy averages (text format)
   - structure_initial.pdb — system coordinates after loading (PDB format)
   - structure_last.pdb — most recent system coordinates (PDB format)

2. **Check and debug**
   - out_nblist.dat — neighbour list (text format)
   - out_res_lists.dat — residue list (text format)
   - out_res_nb123.dat — covalent neighbour list (text format)
   - out_state_summary.dat — system load summary (text format)
   - errors_E.dat — energy errors (text format)

3. **Hippo internal**
   - state_restart — system snapshot for restarts (binary format)

2 Settings

1. hippo_input.txt

Name	Options	Description
mode	MD simulation MC simulation REMD simulation REMC simulation calc energy movie transrot scan	= Molecular Dynamics (MD) = Monte Carlo (MC) = Replica Exchange Molecular Dynamics = Replica Exchange Monte Carlo = Calculate energies of a PDB movie = Rigid body scan (for GBIM only)
input type	PDB seq	= Protein databank file = Sequence file
random number seed	integer < 0	seed for random number generator
set_helix_backbone	0/1	Set the backbone conformation to a perfect alpha helix
use_pdb_coordinates	0/1	Use coordinates from the PDB, if set to zero the structure will be build from the default in the hippo database (hippo_protein_database.dat)
match_hydrogens	0/1	Use the hydrogen atoms from the PDB file, if set to zero the hydrogen atoms will be added from the default in the hippo database (hippo_protein_database.dat)
align principal axis	0/±1/±2/±3	Align the protein along the x, y, z axis 0 = no align 1 = positive x axis -1 = negative x axis 2 = positive y axis -2 = negative y axis 3 = positive z axis -3 = negative z axis
align_backbone atoms only	0/1	Use only backbone atoms for aligning
max_mcsteps	integer	Length of the simulation (in steps) (both MD and MC)
T	double	Temperature (in Kelvin)
use_SSE	0/1	Use single or double precision
use_pbc	0/1	Use periodic boundary conditions
use_NPT	0/1	Use pressure coupling
pressure	double	Pressure (in bar)
isotropic pressure moves	0/1	Isotropic volume moves
fixed x,y box edge	0/1	Fix box dimensions in x and y
use_cutoff	0/1	Use a cut-off for non-bonded interactions (Coulomb and vdW)
cutoff	double	Cut-off distance (in Angstroms)

Name	Options	Description
cutoff_skin_depth	double	Skin beyond the cut-off for neighbour list update (in Angstroms) **Default: 2.0 Å**
use_feather	0/1	Feather the non-bonded potential smoothly to zero at the cut-off. **Default: 1**
cutoff_feather_depth	double	Smoothing length for feather (in Angstroms) **Default: 0.5 Å**
LJ cutoff correction	0/1	Energy cut-off correction for Lennard-Jones interactions **PBC only**: If use_pbc = 0, will be set to 0
use_grid	0/1	Use a grid for building the neighbour list **Default: 1** **PBC only**: If use_pbc = 0, will be set to 0
use_twinrange	0/1	Use a twin-range cut-off
cutoff_shortrange	double	Short range cut-off for twin-range
freq_pdb_out	integer	Output frequency for PDB movie (in steps)
freq_xyz_out	integer	Output frequency for binary movie (in steps)
freq_E_out	integer	Output frequency for energies (in steps)
freq_state_restart_out	integer	Output frequency for system restart files (in steps)
freq_check	integer	Frequency for system energy consistency checks (in steps)
freq_remc_swap	integer	Frequency for attempted Replica exchange swaps (in steps)
print_speed	0/1	Print the simulation speed in the shell
print_debug	0/1	Print debug output
conrot window_move_prob	0 < double < 1	Probability for attempting concerted rotation backbone move (**MC only**)
conrot a	integer	Concerted rotation size parameter
fixed chain ends	0/1	fix the chain ends for backbone moves
prob rigidbodymove/MC scan	0 < double < 1	Probability for attempting a rigid body move (**MC only**)
maximum rigidbody move size	double	Maximum move size (in Angstroms) (**MC only**)
use_gbsa	0/1	Use generalized Born implicit solvent
gb_mode	0/1/2	**Default: 2**
gb_cutoff	double	Cut-off for generalized Born correction sum (in Angstroms) (**MC only**)
r_probe	double	Radius of the solvent molecule (in Angstroms) **Default: 1.4 Å**
sigma	double	Solvent surface tension

Name	Options	Description
intra molecule epsilon	double	Solute interior dielectric constant: **generalized Born only** **Defaults: 1.0 for GB/SA, 2.0 for GBIM**
use_gb_flex123	0/1	Flexible 1-2 and 1-3 neighbours for the GB/SA asymptotic method (see below) **Defaults: 0 for GB/SA, 1 for GBIM**
use_gb_membrane	0/1	Use the generalized Born implicit membrane method (GBIM)
sigma_membrane	double	Membrane surface tension
born_mode	1/2	1 = Exponential switch 2 = Gaussian
sasa_mode	0/1/2/3	0 = Off 1 = Exponential switch 2 = Gaussian 3 = Linear
v(z)_mode	1/2/3	1 = Exponential switch 2 = Gaussian 3 = Linear
gamma_born	double	For Exponential or Gaussian
gamma_sasa	double	For Exponential or Gaussian
gamma_v(z)	double	For Exponential or Gaussian
L	double	Membrane half width (in Angstroms) the centre is always at z = 0
zmin	double	Translational-rotational scan minimum along membrane normal (in Angstroms)
zmax	double	Translational-rotational scan maximum along membrane normal (in Angstroms)
zinf	double	Translational-rotational scan bulk value along membrane normal (in Angstroms)
dz	double	Translational-rotational scan grid size along the membrane normal (in Angstroms)
dtheta (tilt)	double	Translational-rotational scan grid size for tilt angles (in degrees)
dphi (twist)	double	Translational-rotational scan grid size for rotation angles (in degrees)
MD timestep [fs]	double	Molecular dynamics integration timestep (in femtoseconds)
use rattle for H bonds	0/1	Constrain hydrogen bond lengths using RATTLE
freq_remove_com_motion	double	Frequency for removal of the system's centre of mass motion (in steps, 0 = off)
thermostat	1/2/3	1 = Andersen 2 = Lowe-Andersen 3 = Berendsen

Name	Options	Description
Collision frequency	integer	Collision frequency for Andersen and Lowe-Andersen thermostats (in steps)
Collision ratio	0 < double < 1	Percentage of atoms to couple in random collisions for Anderson and Lowe-Andersen thermostats
Berendsen coupling rate	double	Coupling constant for Berendsen thermostat (in 1/ps)
generate_velocities	0/1	Generate starting velocities from a Maxwell-Boltzmann distribution
barostat	1/2	1 = Monte Carlo pressure coupling 2 = Berendsen (**not implemented yet**)
freq_vol_move	integer	Frequency for volume moves (in steps)
Berendsen coupling rate	double	Coupling constant for Berendsen barostat (in 1/ps)
openMP numthreads	integer	Number of threads to use 0 = automatic (i.e. all available threads)
Auto-equilibrate	0/1	Equilibrate system at start of simulation
equilibration period	integer	Equilibration period (in steps)
reset time	integer	Reset time for equilibration (in steps)
Fmax	double	Force threshold to trigger equilibration (in kcal/mol/Å)
dt	double	integration timestep for equilibration period (in femtoseconds)
AE Collision frequency	integer	Collision frequency for Andersen thermostat during equilibration (in steps)
AE Collision ratio	0 < double < 1	Percentage of atoms to couple in random collisions for Anderson thermostat during equilibration

2. Replica exchange temperatures

- The temperatures are listed in: remc_temperature.dat
- For each replica add a line with the temperature

3. Command line options

- hippo in.pdb executes hippo with in.pdb as input PDB structure

4. Executables

Hippo	executable for Intel Core CPU (with SSE3)
Hippo_p3	executable for older processors (PIII, P4)
Hippo_mpi	executable for replica exchange jobs with MPI
Hippo_p3_mpi	executable for older processor MPI jobs

3 Analysis

3.1 Setup

The analyse program can be run in the same directory as the simulation. However, it is recommended to perform the analysis in separate sub-directories:

1. **The simulation is in the directory: /simulation**
2. **Create an analysis sub-directory: /simulation/analyse**
3. **Create further sub-directories for each analysis to be performed, e.g.**
   - /simulation/analyse/rmsd — root mean square deviation (rmsd)
   - /simulation/analyse/helix — helicity

   See below for a description of available analyses (e.g. rmsd, etc.)
4. **Run analyses**
   - analyse rmsd (options) — see below for details
   - analyse helix (options) — see below for details

3.2 Input options

Standard parameters passed to almost all commands:

- [pdb] — PDB file (usually to determine the topology)
- [xyz] — XYZ file (the binary movie trajectory)
- [freq] — frequency of frames in movie file ([freq] = x means use every x^{th} frame)
- [etc]

3.3 Descriptions

RMSD graph

Calculates the RMSD graph of the trajectory w.r.t. a structure in a reference PDB file

```
analyse rmsd [pdb] [xyz] [out] [ref] (backbone,all) [freq] (res1)
(res2)
```

pdb	PDB file (for topology), **default: structure_initial.pdb**
xyz	xyz file (binary movie), **default: movie.xyz**
out	output file containing the rmsd graph, **default: rmsd.dat**
ref	reference structure – must have same topology as system
backbone	= use only backbone atoms for RMSD fit
all	= use all atoms
freq	use every **freq** frame from xyz file, e.g. 1 = use all frames
res1, res2 (OPTIONAL)	include only atoms from residue number res1 – res2

Example:

```
analyse rmsd structure_initial.pdb movie.xyz rmsd.dat nativeNMR.PDB
backbone 10 14 32
```

will fit RMSD using every 10th frame to the backbone atoms from residues 14-32

Helicity graph

Calculates the helicity graph of the trajectory

```
analyse helicity [pdb] [xyz] [range] [startframe] [res1] [res2]
(chain_name)
```

pdb	PDB file (for topology), **default: structure_initial.pdb**
xyz	xyz file (binary movie), **default: movie.xyz**
range	range in degrees for helical dihedral, **default: ±30 deg**
startframe	first frame to be used in graph
res1, res2 (OPTIONAL)	include only atoms from residue number res1 – res2
chain_name	Char chain identifier, e.g. 'A'

Example:

```
analyse helicity structure_initial.pdb movie.xyz 30 0 2 17
```

will calculate the helicity with a range of ±30 degrees around perfect helical phi/psi dihdrals, starting from the frame 0, involving residues number 2 – 17

Note: always exclude terminal residues in helicity analysis.

Z, tilt, kink graph

Calculates center of mass, tilt angle, and kink angle of a peptide in a membrane as a function of simulation time. The membrane is in the xy plane, with z = 0 the membrane center. Kink angle is with respect to the membrane normal.

```
analyse z_tilt [pdb] [xyz] [out] (backbone,all) [freq] (res1) (res2)
chain_name
```

pdb	PDB file (for topology), **default: structure_initial.pdb**
xyz	xyz file (binary movie), **default: movie.xyz**
out	output file containing the rmsd graph, **default: z.dat**
backbone	= use only backbone atoms for fit
all	= use all atoms
freq	use every **freq** frame from xyz file, e.g. 1 = use all frames
res1, res2 (OPTIONAL)	include only atoms from residue number res1 – res2
chain_name	Char chain identifier, e.g. 'A'

Example:

```
analyse z_tilt structure_initial.pdb movie.xyz z.dat all 1 14 32 B
```

will use all frames to calculate z, tilt, kink of segment with residues 14-32 of chain 'B'

Cluster analysis

Performs a cluster analysis using the pairwise method by Daura et al.

```
analyse cluster [pdb] [xyz] (backbone,all) [freq] startframe cutoff
```

Fit to phase

Recenters the trajectory to the center of mass of a phase.

```
analyse fit_movie_phase [pdb] [xyz] [phase_name] [movie_out.pdb]
[movie_out.xyz] [freq]
```

pdb	PDB file (for topology), **default: structure_initial.pdb**
xyz	xyz file (binary movie), **default: movie.xyz**
phase_name	name of molecules to be fitted to, e.g. 'DPPC'
movie_out.pdb	PDB movie of fit
movie_out.xyz	xyz movie of fit
freq	use every **freq** frame from xyz file, e.g. 1 = use all frames freq acts only on the **movie_out.pdb** file, all frames will be used for the xyz file.

Example:

```
analyse fit_movie_phase structure_initial.pdb movie.xyz DPPC
movie_fit_to_DPPC.pdb movie_fit_to_DPPC.xyz 100
```

will refit the trajectory to the center of mass of all DPPC molecules, writing all frames to the xyz file and only every 100th frame to the PDB movie.

Fit solute to previous frames

Generate a PDB movie by RMSD fitting the solute of each frame to the previous frame.

```
analyse fit_movie [pdb] [xyz] [freq] (backbone,all) [movie_fit.pdb]
(res1) (res2)
```

pdb	PDB file (for topology), **default: structure_initial.pdb**
xyz	xyz file (binary movie), **default: movie.xyz**
movie_fit.pdb	PDB movie of fit
freq	use every **freq** frame from xyz file, e.g. 1 = use all frames
backbone	= use only backbone atoms for fit
all	= use all atoms
res1, res2 (OPTIONAL)	include only atoms from residue number res1 – res2

Example:

```
analyse fit_movie structure_initial.pdb movie.xyz 1 backbone
movie_fit.pdb 2 44
```

will give a RMSD fitted trajectory, where only the backbone atoms of residues 2 to 44 were used in the fit.