1 FFLD Documentation

Usage : flexSwarmDocking [OPTIONS] [LIGANDFILE]

With the introduction of the default values, *FFLD* works without giving any options. By typing *flexSwarmDocking*, *FFLD* prints out the possible command line options, which are also briefly described in the following subsection, as well as their default values. If a severe error is detected, for example if an essential file is missing, *FFLD* will exit immediately. If the error is not severe, *FFLD* will continue and eventually re-set the input values to the standard values.

Features in developpmental stage:

- *solvation term in scoring function*
- *coulombic term in scoring function*
- *flexible protein binding site residues*
- *monte carlo based optimization (MC,STUN,ELP,PT)*

In general *VdW*- and *H-bond*-terms should be used for scoring in productive runs (default). Neither the solvation nor the coulombic term of the scoring function have been exhaustively tested. Similarly the flexible protein binding site residues (*@<DAIM>BS*-section in *.info*-file) are in a developpmental stage.

1.1 Command Line Options

- *Optimizer (−o / −−optimizer)*
  Possible values are "GA" for genetic algorithm and "PS" for particle swarm. If no or a wrong option is given, the optimizer will be automatically set to "GA".
- *Seed (−s / −−seed)*
  Seed number for the initialization of the random number generator (default value is 1234).
- *Population Size (−p / −−popsize)*
  The population size is the number of individuals used in the "PS" and the "GA". The default value is 100 when not modified by the user.
- *Info-File (−−infofile)*
  This option allows to specify the info-file produced by *DAIM* (P.Kolb). If no name is given *FFLD* will look for an info-file with the same name as the ligand-mol2-file (for example ligand.mol2 → ligand.info).
- *Maximal Number Of Energy Evaluations (−e / −−maxeval)*
  FFLD will be stopped, when the total number of energy evaluations exceeds the specified value (default 300000).
- *Maximal Number Of Generations (−g / −−maxgen)*
  FFLD will be stopped, when the total number of cycles/generations exceeds the specified value (default 1000).
- *Parameter-File (−−paramfile)*
  The parameter-file can be specified with this option, if it's different from "./Inputs/FFLD_param".
- *VdW-Grids (−−vdwrepname and −−vdwattname)*
  Name of the repulsive and attractive VdW-grids (the default name being "prot_vdw_rep.ene" and "prot_vdw_att.ene" respectively).

- *VdW-Grids Access type (−−vdwaccess)*
  The possible values are “r” for reading or “w” for writing the grids, as well as “n” (“nothing”).

- *Coulombic-Grid (−−coulname)*
  Name of the coulombic grid (default “prot_coul_E1.ene”).

- *Coulombic Grid Access type (−−coulaccess)*
  Same as for the VdW-grids.

- *Solvation Grid (−−solvname)*
  Name of the solvation grid (default “prot_solv.ene”).

- *Solvation Grids Access type (−−solvaccess)*
  Same as for the VdW-grids.

- *Evaluation Mode (−−evalmode)*
  Specifies if *FFLD* is run in evaluation-mode. In this mode no optimizer is used and *FFLD* works as a scoring function, calculating the energy of the ligand provided by the user.

- *Developper's Mode (−−develop “off”,“on” or “chatty”)*
  Yields verbose output in standard output (printing of detailed energies at the end of the *FFLD*-run). An additional output file is written in “chatty” mode, including the energies of all individuals at every cycle (ga/ps_dev.dat). This option is not printed when FFLD is run without any options.

- *Autogeneration (−− autogeneration)*
  Sets the number of generations (iterations) of the genetic algorithm to = 20+2∗($numberofrotatablebonds$).

- *Name of Final PDB-Output (−−outputname)*
  With this option the name of the final PDB-output file can be chosen (standard “minimized.pdb”).

1.2 Parameter-File Options

- *VDWFactor*
  intra 1-4 VdW-scaling factor

- *ParentSelection*
  Parent selection (elitism). Possible values are “lin” (linear), “bst” (best windowing) and “wst” (worst windowing).

- *DynamicLinearization*
  Dynamic linearization of the softened part of the VdW-grid

- *Electrostatics*
  Electrostatics (bath-tub shaped H-bond function, including penalties for unfavorable polar-polar and polar-hydrophobic contacts between ligand and protein :“softHBond”; step function for H-Bonds : ”stepHBond” or coulombic electrostatics : “elec”).

- *LocalSearchFrequency*
  Local search frequency.

- *LocalSearchMaxEvaluation*
  Local search : maximal number of energy evaluations.

- *ProbabilityScaling*
  This option allows the user to modify the probabilities of operator selection and the probability of mutation (GA only !). The default is “CONST” where the operator probabilities and the mutation

probability remain constant over the entire run. If the user selects “ENERGY” or “SCORE” the operator probabilities and the mutation probability of the individuals are scaled according to energy or to the rank of the individual respectively. The operator probabilities are $mutationOperatorProbability = 1 - crossoverOperatorProbability$. The lower the energy, the lower the mutationOperatorProbability (unless the function is fliped, see “FlipFunction” below).

- *ProbabilityScalingFunction*
  This option defines with which function the probabilities are scaled according to energy or to the rank of the individual if “ENERGY” or “SCORE” are chosen in “ProbabilityScaling”. Possible functions are : exponential (“EXP”), linear (“LIN”) or sigmoidal (“SIG”).

- *ProbabilityOfCrossoverOperator*
  Defines the operator probability of crossover selection.

- *LowerProbabilityOfMutationOperator*
  This defines the lowest possible value for the mutation operator probability when the “ENERGY” or “SCORE” are chosen in “ProbabilityScaling”. If “CONST” is chosen this defines the constant mutation operator probability.

- *UpperProbabilityOfMutationOperator*
  This defines the highest possible value for the mutation operator probability when the “ENERGY” or “SCORE” are chosen in “ProbabilityScaling”.

- *LowerProbabilityOfMutation*
  This defines the lowest possible value for the mutation probability when the “ENERGY” or “SCORE” are chosen in “ProbabilityScaling”. If “CONST” is chosen this defines the constant mutation operator probability.

- *UpperProbabilityOfMutationOperator*
  This defines the lowest possible value for the mutation probability when the “ENERGY” or “SCORE” are chosen in “ProbabilityScaling”.

- *BoundaryMethod*
  This option allows the user to define with which method the scaling of the probabilities will be performed. Above and below, the probabilities will remain constant at the lowest or highest probability defined by the user above. The boundaries can either be defined as relative (“PERCENTAGE” of individuals) or absolute (“CUTOFF”) values.

- *LowerBoundary*
  Defines the lower boundary at which the scaling of the probabilties will start (below this value the probabilities are constant !). For both “CUTOFF” and “PERCENTAGE” this value can be negative.

- *UpperBoundary*
  As above for the upper boundary.

- *FlipFunction*
  If this flag is set from “NO” to “YES” the probability scaling function is flipped and the probability of mutation operator selection is increased for fitter inidividuals (see also “ProbabilityScaling”).

- *Exclusive*
  If this option is changed from the default “YES” to “NO” both crossover and mutation can be performed in one cycle. In the default mode “YES” these operators are mutually exclusive.

- *Dielectric*
  Dielectric constant.

1.3 Known Issues

- If the group name in the MOL2-file of the ligand exceeds 4 characters, FFLD will not crash, but might produce bogus output (atoms of the ligand appear as points on a single straight line, when watched in a molecular graphics programm).
- If the SEED geometrical centers are located outside of the grids, FFLD will crash with a segmentation fault.
- The same problem as before occurs when FFLD is run in evaluation mode and an atom of the ligand is lying outside of the grid.