Mumbo – Version January 2010 -

Principal authors: Martin Stiebritz and Yves Muller
Reference: Stiebritz, M. T. & Muller Y. A. (2006). MUMBO: a protein-design approach to crystallographic model building and refinement. Acta Cryst. D 62, 648-658.
Comments to ymuller@biologie.uni-erlangen.de

The following persons also contributed: Christof Jäger, Kay Diederichs, Yasmine Rudolph

General description

MUMBO finds the energetically most favourable combination of rotamers and amino acids to build up a given backbone conformation. Rotamer and amino acid diversity is generated using backbone-independent or backbone-dependent rotamer libraries. Selection of the energetically most favourable configuration is done using dead end elimination and variations thereof. In a final step an explicit brute force approach is used to determine the global minimum energy configuration (GMEC). Alternatively a Monte-Carlo Metropolis approach can also be used.

The total energy of a given configuration is estimated considering contributions from VDW-interactions, electrostatic interactions, rotamer solvation energies, rotamer probabilities, hydrogen bonding energies and if requested pseudo X-ray energies derived from electron density maps.

For crystallographic refinement and electron density fitting purposes, electron density maps can be read into MUMBO and pseudo X-ray energies derived from the electron density present at each rotamer position and subsequently used during the elimination process (see Stiebritz and Muller (2006), Acta Cryst. D 62, 648-658).

As an additional feature ligand binding can also be studied. Multiple ligand orientations can be generated and are processed the same way as the rotamer diversity.

The instructions to run MUMBO are specified in the input file **mumbo.inp**. The file mumbo.inp must be present in the working directory at the start of the program. An example file for mumbo.inp is provided further below.

Principal keywords

JOB, INPDB, OUTPDB
ENERGY_CALC, ENERGY_WEIGHTS, ENERGY_TUNING,
ENERGY_MC, ENERGY_BRUT, ENERGY_ANA, ENERGY_SPLIT
ROTAMER_TUNING, BACKRUB_TUNING, LIGAND_TUNING, XRAY_TUNING,
COMP_PICKNAT_TUNING
LOGFILE_TUNING
MUMBO_POS

Auxiliary keywords

INIT, MC, DEE, GOLD, DOUB, BRUT, ANA, MONSA, SPLIT, COMP, PICKNAT
VDW, ELEC, RPROB, SOLV, HBOND, XRAY
VDW_RADII, DIELEC, ESHIFT, ESWITCH, EDIST, CTONNB, CTOFNB, SOL_GREF, SOL_GFREE, SOL_LAM, VDW14, VDW13, ELE12, ELE13, ELE14, SOL12, SOL13, SOL14, HB_CTOF, HBGEOM, HBEMP, HB_DZ, HB_EZ
CUTOFF, KEEP, AUTO_RECOVER, MAXCOMB, MAXSORT, NCYC
PHIPSI_DEP, FINE_NDEG, FINE_NSTP, FINE_LIMIT, FINE_CHI_ALL, FINE_CHI_1_ONLY, BACK_NSTP, BACK_NDEG, REFPDB, NOCB, SWAP_QNH
LIGFILE, LIGNAME, LIGNGEN, LIGRMS, LIGSEED, NOSHIFT, NOROTATE, LIGFULLROT
MAPIN, MAPTYPE, HYDRO
ROTAMER_LIB_ALL, ROTAMER_LIB_PHIPSI, NONBONDED_LIB, PARAM_LIB, CONNEC_LIB
MC_SHORT, MC_LONG, ANA_LONG, ANA_SHORT, DBASE_SHORT, DBASE_LONG

Keyword description

JOB Specifies the different tasks to be executed by mumbo. The tasks listed after the JOB keyword will be executed stepwise. The various tasks are INIT, MC, DEE, GOLD, DOUB, SPLIT, BRUT, MONT, ANA, COMP and also PICKNAT

INIT Reads a coordinate file and generates the different amino acids together with their rotamers at the desired positions of the peptide chain. INIT also adds missing atoms to the coordinate file if need there is. At the end of INIT a set of PDB-files is written out that contain all information needed for the subsequent calculations (See Input – Output – files). The positions to be mutated are specified with the MUMBO_POS keyword. The keywords: ROTAMER_LIB_ALL, ROTAMER_LIB_PHIPSI and ROTAMER_TUNING allow to specify the way the rotamers are generated.

MC Calculates the energy of each rotamer with respect to the constant part of the structure. Those rotamers exceeding the energy threshold defined by the ENERGY_MC= CUTOFF= statement are eliminated at the end of the MC step. The different energy terms to be considered are specified with the ENERGY_CALC keyword. The additional keywords ENERGY_WEIGHTS and ENERGY_TUNING are available to fine tune the calculated energies.

DEE Invokes dead end elimination in order to reject rotamers that are not part of the global energy minimum. The same keywords as in MC apply to fine tune the energy calculations.

PDEE (Still being tested, multi processor dead end elimination)

GOLD Invokes the dead end elimination algorithm of Goldstein. Again the same keywords as in MC are used to fine tune the energy calculations.

PGOLD (Still being tested, multi processor dead end elimination according to Goldstein)

SPLIT Another DEE-Filter, namely conformational splitting.

DOUB Invokes dead end elimination while looking for dead ending pairs.

BRUT In this step, brute force is used to find the global minimum energy configuration. The energies of all possible combinations of rotamers and amino acids left over after dead end elimination are calculated step by step in order to determine the set of rotamers/amino acids that represent the global energy minimum. The percentage of combinations already searched is reported from time to time so that one can estimate the time needed to complete the search. The keyword ENERGY_BRUT allows to better monitor the output of BRUT. At the end of BRUT a file specified with the OUTPUT= keyword is written containing the GMEC structure (Input – Output – files).

MONSA Monte-Carlo Metropolis + Simulated annealing search for the configuration with the lowest energy. Here no guarantee is given that global minimum will be reached.

ANA Allows for an in depth analysis of the low energy configurations. For a certain number of low energy configurations, a lengthy analysis of individual interactions between side chains can be written out to the file mumbo.ene. ANA uses the same algorithm as BRUT, however because it runs on only one processor, it runs much slower than BRUT. See also ENERGY_ANA for fine tuning.

COMP Allows to compare the solution obtained in MUMBO with a reference coordinate file. Works only if solution obtained in MUMBO is unique, i.e. a single rotamer is left in the database at each position and if the sequence in reference file is identical to that obtained in MUMBO. See specific comments regarding the format of the reference coordinate file.

PICKNAT Allows to pick the ‘native’ rotamer, namely the rotamer corresponding the closest to the side chain orientation in the reference coordinate file. If followed by BRUT and ANA than PICKNAT allows obtaining a description of the side chain interaction energies as present in the reference coordinate file. Any design can thus be compared to the interactions in a wild-type structure using the same force field parameters and description.

INPDB Specifies the name of the file from which the starting coordinates are read. Please see 'specific comments' regarding potential trouble with the input coordinate file.

OUTPDB Specifies the name of the file into which the global minimum energy configuration is written to.

HYDRO If this keyword is present, then hydrogen atoms will not be deleted in the output-file. Hydrogen atoms are however only present if MUMBO is run with ENERGY_CALC HBOND specified.

ENERGY_CALC Specifies which energy contributions to consider during the calculation of the inter residue interaction energies. The various possibilities are:

VDW Van de Waals interactions

ELEC Electrostatic interactions

SOLV Atomic solvation energies according to Lazaridis and Karplus

RPROB Calculates a probability energy according to the rotamer probabilities listed either in ROTAMER_LIB_ALL or in ROTAMER_LIB_PHIPSI. The energy is calculated simply as the natural logarithm of the probability.

HBOND Hydrogen-bonding energies. Two options are available to derive hydrogen-bonding energies. Either a geometrical model (Gohlke et al., Proteins (2004), 56: 322-337) or an empirical model (Kortemme et al. JMB (2003), 326:1239-1259) can be used (See ENERGY_TUNING). Specifying HBOND automatically leads to an expansion of the number of rotamers generated in the INIT step for amino acids where the exact location of the hydrogen atom is ambiguous (Ser, Thr, Tyr).

XRAY A pseudo xray-energy is derived for each rotamer from electron density maps. E-xray equals minus the sum of the electron densities present at the positions of each atom in a rotamer. The electron density is calculated through linear extrapolation of the density at the grid points of the map (Rossman et al., J. Appl. Cryst. (1992), 25:166-180).

ENERGY_WEIGHTS The different energetic contributions can be weighted according to E(weighted) = weight * E(unweighted). The different weights are specified using subkeywords. The default values are: VDW = 1.0, ELEC = 1.0, RPROB = 3.0, SOLV = 1.0, HBOND = 1.0 and XRAY = 1.0.

ENERGY_BRUT Defines how the results of the BRUT step are listed. The subsequent keywords do not alter the way the global energy minimum is calculated but only how many rotamers will be kept at the end of the BRUT step. Subkeywords are:

KEEP Each rotamer that is part of a rotamer configuration with a total energy lower than that specified with the KEEP keyword is flagged and at the end of the BRUT step written out. These rotamers and their interactions can then be analysed in more detail during the ANA step.

ENERGY_MC

CUTOFF Energy cut-off applied during the MC step. Rotamers with energies higher than CUTOFF will be flagged and deleted at the end of MC. If all rotamers/amino acids at a given position have energies higher than CUTOFF then all rotamers/amino acids will be deleted and as a consequence the program **stops** unless the AUTO_RECOVER keyword is provided. Default value = 40.0 kcal/mol.

AUTO_RECOVER Allows MUMBO to recover in case all rotamers/amino acids at a given residue position have been flagged for deletion during the MC step because their energy exceeds the CUTOFF value. In case AUTO_RECOVER is specified than MUMBO returns to this position and retains the rotamer/amino acid with the lowest energy, irrespectively of the specified cut-off.

ENERGY_ANA

MAXCOMB Specifies the total number of possible rotamer/amino acid combinations accepted at the start of ANA for subsequent analysis. Because ANA is slow, this number should be kept low. Default value = 100000.

MAXSORT The lowest energy configurations are saved in ANA and sorted according to their energies. Because sorting is a slow process, MAXSORT specifies the number of configurations to be tracked in a sorting list. For each configuration present in the sorting list, a detailed account of the residue-residue interaction energies is written out to the file mumbo.ene. Default value = 100.

ENERGY_SPLIT

NCYC Number of cycles for the Split conformation dead end elimination step. Default value = 3.

ENERGY_BOUND

NCYC Number of cycles for the bound conformation dead end elimination step. Default value = 3.

ENERGY_MONSA

TEMP_START Starting temperature for Monte-Carlo metropolis simulated annealing search algorithm.

TEMP_END Ending temperature for Monte-Carlo metropolis simulated annealing search algorithm.

TEMP_STEP Number of temperature steps

TEMP_NUMIT Number of iterations during each temperature step. The total number of steps equals tem_step * temp_numit.

ENERGY_TUNING This keyword allows for the fine tuning of the energy terms:

VDW_RADII Allows to increase or to decrease the van der Waals radii of the various atom types. The atom radii are specified as sigma values in the file NONBONDED_LIB and are multiplied by the factor specified by VDW_RADII. Default value = 1.0.

VDW_SOFT If present, the repulsion part of the Van de Waals interaction is softened according to Pokala N, Handel TM, J Mol Biol. 2005, 347:203-227. Be aware that the well depth of E-min is also changed.

DIELEC Specifies the dielectric constant to be used during the calculation of the electrostatic interactions. Default value = 20.0.

ESHIFT, ESWITCH, EDIST and ECONT. These are mutually exclusive keywords that specify how to deal with long range electrostatic interactions. They are basically identical to those used in the programs CHARMM and X-plor.

ESHIFT Uses a distant dependent shift function so that the electrostatic interactions become zero at a distance equal to that specified by the CTOFNB keyword. The shift function has the form: (1 – (r(i,j) / CTOFNB) **2) **2

ESWITCH Uses a switch function in order to set the electrostatic interaction to zero at the distance specified by CTOFNB. CTONNB specifies the distance at which the switch function is switched on. The switch function has the form: ((r(i,j) - CTOFNB) **2) * (3 CTONNB -2 CTOFNB - r(i,j)) / (CTONNB - CTOFNB)**3.

EDIST Directly manipulates the dielectric constant. If EDIST is specified then the dielectric constant is multiplied by r(i,j) (unless r(i,j) is smaller than 1, in this case the dielectric constant remains unchanged).

ECONT In this case the dielectric constant is kept constant and neither shift or switch function is applied.

CTOFNB Specifies at which distance the shift function in ESHIFT and the switch function in ESWITCH cause the electrostatic interactions to be set to 0. CTOFNB has no effect on EDIST. Default value = 10.5.

CTONNB Specifies at which distance the switch function kicks in. It has no effect on either the shift function or on the behaviour of EDIST. Default value = 6.5.

SOL_GREF This factor can be used to manipulate the atomic reference solvation energy. In the Lazaridis Karplus model a reference solvation energy is postulated for each atom. The reference solvation energy is negative for hydrophilic atoms and positive for hydrophobic atoms. These reference energies are then reduced depending on the number of atoms surrounding the atom considered. SOL_GREF can be used to scale the reference solvation energy. Default value = 0.0.

SOL_GFREE This factor can be used to scale the free solvation energies. In the Lazaridis Karplus model a reference solvation energy is postulated for each atom. The reference solvation energy is negative for hydrophilic atoms and positive for hydrophobic atoms. The free energies are used to reduce the reference free energies (see Lazaridis and Karplus). Default value = 1.0.

SOL_LAM is a scaling factor for the atomic lamda values (solvation correlation lengths) . SOL_LAM > 1.0 will enhance the burial effect and helps to reduce the reference solvation energies. Default value = 1.0.

VDW13, VDW14 Van de Waals interactions between atoms are only calculated if the corresponding atoms are separated by at least 3 (VDW13) or 4 (VDW14) covalent bonds. Default = VDW14.

ELEC12, ELEC13, ELEC14 Electrostatic interactions between atoms only considered if atoms separated by at least 2, 3 or 4 covalent bonds. Default = ELEC14.

SOL12, SOL13, SOL14 the same as above but for atom solvation. Default = SOL13.

HBGEOM If specified then a geometrical hydrogen bonding model is used (see Gohlke et al., Proteins (2004)).

HBEMP If specified then an empirical hydrogen bonding model is used (see Kortemme et al). Either HBGEOM or HBEMP must be specified.

HB_CTOF Distance cut-off for hydrogen bonding to be considered. Distance between hydrogen atom and hydrogen bond acceptor atom (Default = 4 Angs)

HB_DZ Optimal donor acceptor distance (Default = 2.8 as in the Gohlke paper). Applies only to the HBGEOM hydrogen bonding model.

HB_ EZ Equilibrium hbond energy at optimal distance and geometry (Default = 8 as in the Gohlke paper). Applies only to the HBGEOM hydrogen bonding model.

ROTAMER_TUNING Allows to manipulate the way the rotamers are generated during the INIT step. Subsequent keywords are:

PHIPSI_DEP If this keyword is specified then backbone dependent rotamers will be generated. At each position phi and psi dihedral values will be calculated and depending on the result backbone dependent rotamers and their probabilities are retrieved from the file specified by the ROTAMER_LIB_PHIPSI keyword. In the absence of the PHIPSI_DEP keyword, the rotamers are build according to the instructions present in the file ROTAMER_LIB_ALL. In this case the rotamers are the same irrespectively of the conformation of the main chain.

FINE_NSTP, FINE_NDEG and FINE_LIMIT allow to drastically increase the number of rotamers through the generation of additional rotamers which should however only lightly deviate from those present in the input rotamer library file.

FINE_NDEG specifies the angle increment in degrees to be added to or subtracted from the dihedral values read from ROTAMER_LIB_PHIPSI or ROTAMER_LIB_ALL in order to generate additional rotamers. Default value = 5.0.

FINE_NSTP specifies the number of increments applied to each dihedral value. In case FINE_NSTP = 1 then each dihedral value is split into three new dihedral values corresponding to the initial value minus the increment specified in FINE_NDEG, the initial value itself and the initial value plus the increment. In case FINE_NSTEP = 2 each dihedral value is split into five new dihedral values, namely the initial value minus two increments, minus one increment etc.... By doing so, the number of rotamers can increase drastically. For example for the GLU side chain. Here three dihedral values are needed to describe its conformation. With FINE_NSTEP = 2 each dihedral value will be replaced by five new values, the combination of which generates 5 * 5 * 5 = 125 different combinations. If 15 different rotamers are present in the initial file than a total of 125 * 15 = 1875 rotamers would be generated with the FINE_NSTEP = 2 option.

The default FINE_NSTEP = 0 switches off the rotamer expanding mode. Default value = 0.
FINE_NSTP = 0 means no rotamer expansion through chi expansion.

FINE_LIMIT allows limiting the number of rotamers to be generated with the FINE_NSTP option. If the total number of rotamers for a given side chain exceeds FINE_LIMIT then FINE_NSTP will be automatically decreased by one. This is repeated until for a given side chain, the number of rotamers is lower than the value specified in FINE_LIMIT. Default value = 300.

FINE_CHI_ALL expand all dihedral angles present in the side chain. For example for a GLU side chain (three dihedral angles) and FINE_NSTEP = 2 each rotamer will be expanded to 5 * 5 * 5 = 125 new rotamers.

FINE_CHI_1_ONLY expand only chi1 dihedral angle. For example for a GLU side chain (three dihedral angles) and FINE_NSTEP = 2 each rotamer will be expanded to 5 new rotamers.

FINE_CHI_ALL and FINE_CHI_1_ONLY are mutually exclusive. Default is FINE_CHI_1_ONLY.

BACKRUB_TUNING Allows to introduce backbone flexibility through a backrub motion (See Davis et al. (2006), Structure **14**, 265-274). This leads to the generation of additional rotamers during the INIT step. BACKRUB motion is only performed in case BACK_NSTP > 0 (see below). Additional keywords are:

BACK_NSTP, BACK_NDEG allow to drastically increase the number of rotamers through the introduction of a backrub motion, namely a rotation of the Cα-Cβ vector of residue i around an axis that passes through the Cα atoms of residue i-1 and residue i+1.

BACK_NDEG specifies the rotation increment in degrees to be added or subtracted during the backrub motion (Default value = 3.0°).

BACK_NSTP specifies the number of increments applied during the backrub motion. In case BACK_NSTP = 1 then three rotations are generated during the backrub motion for residue i, namely –BACK_NDEG, 0 and + BACK_NDEG . In case BACK_NSTEP = 2 then five rotations are generated. This is highly similar to the FINE_NDEG and FINE_NSTEP key words in ROTAMER_TUNING.
BACK_NSTP= 0 means no backrub motion.

LIGAND_TUNING

LIGFILE Name of the file containing the starting ligand conformations/orientations.

LIGNAME Residue number and chain id of the ligand. If specified then this name replaces the ligand residue number present in the ligand file. Otherwise the residue number and the chain id is taken from the first set of ligand coordinates present in the ligand file.

LIGNGEN Number of ligand orientations to be generated applying random shifts and rotations. Default value = 30.

LIGRMS Target rms values for the shifted and rotated ligand coordinates. This is only a rough estimate of the diversity introduced by applying random shifts and rotations. The exact rms value is reported in the output after the random shifts have been applied. An increase of LIGRMS will increase the size of the shifts and rotations randomly applied to the ligand coordinations. Default value = 3.0.

LIGSEED Seed value for the random number generator.

NOSHIFT If specified then no shifts will be applied but only different ligand orientations are generated.

NOROTATE If specified then no reorientations take place only shifts will be applied.

LIGFULLROT If specified then random sampling of the full rotation space will be attempted. Makes only sense if LIGNGEN is very high. If LIGFULLROT is specified then LIGRMS is ignored.

XRAY_TUNING

MAPIN File from which the electron density is read in

MAPTYPE = CNS Map-type of the electron density file. At present only CNS-type (or XPLOR-ascii-type) maps can be read.

COMP_PICKNAT_TUNING. This keyword applies to both the COMP and the PICKNAT step.

REFPDB File containing the reference coordinates. See specific comments regarding the PDB format of the reference coordinate file.

NOCB Only side chain atoms beyond Cβ-atom are considered for the calculation of the root mean square deviation during comparison and/or for picking the 'native' rotamer.

SWAP_QNH. If this keyword is present, than the two alternative orientations for the amide group of Asn and Gln as well as the imidazole ring of His will be considered as being equivalent when matching any of their rotamers with the reference structure (This corresponds to a swapping of atoms N and O in Asn and Gln as well as a 180° rotation of the side chain of His). The possible swapping of identical but differently named atoms in Asp, Glu, Phe and Tyr (such as OD1 and OD2 in Asp, OE1 and OE2 in Glu) is always accounted for.

LOGFILE_TUNING Controls the level of details reported into the output files

MC_SHORT (default), MC_LONG Long/short output in MC step
ANA_SHORT (default), ANA_LONG Long/short output in ANA module

Only when ANA_LONG is specified, the files mumbo.ene2 and mumbo.lis will be generated DBASE_SHORT (default), DBASE_LONG Backing or not backing up all the files in the structure database after each elimination step.

MUMBO_POS Specifies the position to be substituted / mumboed. The keyword is followed by the chain id and the residue number that identifies the position to be 'mumboed'. The amino acids for which rotamers are to be built are subsequently listed using the three letter code. The chain id may or may not be present. Here two examples:

MUMBO_POS_6= B 28 ALA SER GLU ASP VAL THR
MUMBO_POS_25= 65 ALA PRO SER GLU ASP VAL THR TRP

ROTAMER_LIB_ALL Specifies the file from which the building instructions for the various amino acids are read. For more detail see the specific comments.

ROTAMER_LIB_PHIPSI Specifies the file from which backbone dependent dihedral values are read. These dihedral values then replace those provided in ROTAMER_LIB_ALL. Presently only the file bbdep02.May.lib can be read in here. This file is available from the Dunbrack lab home page at: http://dunbrack.fcc.edu/bbdep/.

PARAM_LIB Specifies a file in which various parameters for the individual amino acids are stored. Here all atoms belonging to a residue are listed together with atom type assignments, partial charges and hydrogen bonding parameters.

NONBONDED_LIB Specifies a file from which the parameters to calculate non-bonded interactions are read. The various atom types are listed together with sigmas and epsilons and atomic solvation parameters.

CONNEC_LIB Specifies the file from which atom connectivities are read.

Input – output - file description

mumbo.inp provides all input specifications, for example:

Input example: mumbo.inp
(Lines or keywords can be commented out using exclamation marks; the equal sign may or may not be present)

```
!  MUMBO_version January 2010
!
JOB=    INIT MC DEE GOLD SPLIT DOUB MONSA BRUT ANA COMP ! PICKNAT
INPDB=  lck_start.pdb
OUTPDB= lck_output.pdb
!
!  ENERGY TERMS:
!
ENERGY_CALC=    VDW ELEC RPROB SOLV !XRAY HBOND
ENERGY_WEIGHTS= VDW=1.0  ELEC=1.0 RPROB=3.0 SOLV=1.0
!ENERGY_WEIGHTS= XRAY=1.0 HBOND=1.0
!
```

```
!  ENEGRY FINE TUNING:
!
ENERGY_MC=     CUTOFF= 40  AUTO_RECOVER
ENERGY_SPLIT=  NCYC= 3
ENERGY_BRUT=   KEEP= 1000
ENERGY_ANA=    MAXCOMB=100000  MAXSORT=10
ENERGY_MONSA=  TEMP_START=5000 TEMP_END=300 TEMP_STEP=100 TEMP_NUMIT=1000
!
ENERGY_TUNING=   VDW_RADII=0.9 VDW_SOFT
ENERGY_TUNING=   DIELEC=20.0   ESHIFT
ENERGY_TUNING=   CTONNB= 5.5  CTOFNB= 12.0
ENERGY_TUNING=   SOL_GREF=0.0  SOL_LAM=1.2  SOL_GFREE=1.5
ENERGY_TUNING=   SOL13 VDW14
ENERGY_TUNING=   HB_CTOF=4.0
!ENERGY_TUNING=  HBGEOM    ! HB_DZ=2.8  HB_EZ=8
ENERGY_TUNING=   HBEMP
!
LOGFILE_TUNING= MC_SHORT ANA_LONG DBASE_SHORT !MC_LONG ANA_SHORT
!          D                                   BASE_LONG
!
!  ROTAMER FINE TUNING:
!
ROTAMER_TUNING=  PHIPSI_DEP
!ROTAMER_TUNING=  FINE_NSTP=1  FINE_NDEG=5.0  FINE_LIMIT=300
!ROTAMER_TUNING=  FINE_CHI_ALL ! FINE_CHI_1_ONLY
!
!  BACKBONE BACKRUB MOTION:
!
BACKRUB_TUNING=   BACK_NSTP=2  BACK_NDEG=5
!
!  LIGAND SPECS:
!
!LIGAND_TUNING= LIGFILE= pdb/est.pdb  LIGNAME= E 301
!LIGAND_TUNING= LIGNGEN= 10 LIGRMS= 1.0 LIGSEED= 12345 NOSHIFT ! NOROT
!
!  XRAY ELECTRON DENSITY:
!
!XRAY_TUNING=   MAPIN=2fofc.cns   MAPTYPE= CNS
!
!  COMPARISON AND ANALYSIS OF REF. STRUCTURE:
!
COMP_PICKNAT_TUNING= REFPDB= lck_ref.pdb NOCB  SWAP_QNH
!
!  LIBRARIES USED:
!
ROTAMER_LIB_ALL=     ../lib/rotamer_xpl.lib
ROTAMER_LIB_PHIPSI=  ../lib/bbdep02.May.lib
PARAM_LIB=           ../lib/parameter_xpl.lib
NONBONDED_LIB=       ../lib/nonbonded_xpl.lib
CONNEC_LIB=          ../lib/connectivity_xpl.lib
!
!  POSITIONS TO BE MUMBOED:
!
MUMBO_POS_1=    A   58 GLY ALA SER PRO
MUMBO_POS_2=    A   59 SER CYS PRO ALA GLU
MUMBO_POS_3=    A   60 PRO
MUMBO_POS_4=    A   61 LEU
MUMBO_POS_5=    A   62 GLN
MUMBO_POS_6=    A   63 ASP
MUMBO_POS_7=    A   64 ASN
MUMBO_POS_8=    A   65 LEU
MUMBO_POS_58=   A  116 VAL
```

```
MUMBO_POS_59=   A  118 LYS
MUMBO_POS_60=   A  119 ALA
…….
...
```

Additional files are provided using the keywords:
ROTAMER_LIB_ALL, ROTAMER_LIB_PHIPSI, NONBONDED_LIB, PARAM_LIB CONNEC_LIB and many others....

Input coordinates are specified using the INPDB statement. PDB format required. See also specific comments.

Output-files:

Output coordinates are written to the file specified using the OUTPDB statement

A database of all internally handled coordinates is written out after each step and reread into the memory at each subsequent step. The following pseudo PDB-type of files are written out.

0_atm_sum Contains the coordinates of the part of the structure which is held constant. At the positions to be mumboed the residues are replaced with Gly.

0_mch_sum Here the different positions to be mumboed are stored as Ala-residues.

1_atm_sum, 2_atm_sum, 3_atm_sum and so on Contain all the coordinates of the side-chains and rotamers which are considered at a certain position. 1_atm_sum contains those at position 1, 2_atm_sum those at position 2 and so forth …..

--1, --2, --3 and so on If older files exist in the working directory and DBASE_LONG is specified in LOGFILE_TUNING then these are backed up and the suffixes --1 , --2, --3 etc appended and incremented when a new file version is created.

mumbo.pdb, Files written at the end of the BRUT step. Mumbo.pdb contains the coordinates of the set of rotamers which correspond to the global energy minimum.

mumbo.ene, mumbo.ene2, mumbo.lis Files written at the end of ANA and containing a detailed analysis of the different rotamer energies. (mumbo.ene2, mumbo.lis only written out if LOGFILE_TUNING= ANA_LONG specified)

Specific comments

Problems with the input file: Most problems with running MUMBO are due to problems reading in the input coordinate files. Mumbo expects the coordinates being provided in PDB-file format. MUMBO only reads ATOM and HETATM records/strings. However, **MUMBO expects spaces to separate the various values provided in the input file**. Therefore it **will bomb** if B-factors are higher than 99.9 since no blanks will then be present between the occupancy and the B-factor. The same problems also occur when alternative conformations are present for a given residue in the input pdb file. Please not that B-factors and occupancies can be left out in the input PDB-file. For reasons unkown to the authors, MUMBO doesn't like residues numbered 0, -1, - 2 etc… **If MUMBO refuses to work chances are high that the INPDB file needs some attention.**

Description/specification of file paths: File paths can be specified in different ways. The following examples should all work:

/home/xray/alfons/mumbo_working_directory/lib/parameter.lis
lib/parameter.lis
../../eugene/lib/parameter.lis
or $MUMBO_LIB/parameter.lib

In the last example, the environment variable MUMBO_LIB has to be specified within the c-shell or t-shell environment before running mumbo. For example using the command:
Setenv MUMBO_LIB /home/xray/alfons/mumbo_working_directory/lib.
On some computers, some problems have been seen with the file paths. Usually these can be worked around by specifying the paths differently.

Parameter files: In order for mumbo to run, a consistent set of parameter files must be provided and specified using the ROTAMER_LIB_ALL, NONBONDED_LIB, PARAM_LIB CONNEC_LIB keywords. ROTAMER_LIB_PHIPSI is only needed if a backbone dependent rotamer library is requested. At present the following consistent set of files are provided, namely:

for ROTAMER_LIB_ALL the file rotamer_xpl.lib,
for NONBONDED_LIB the file nonbonded_xpl.lib
for PARAM_LIB the file param_xpl.lib
for CONNEC_LIB the file connec_xpl.lib
and for ROTAMER_LIB_PHIPSI the file bbdep02.May.lib

In these files the atom parameters are consistent with those used in Charm and X-plor, namely, toph19x.pro and param19x.pro. This also means that only polar hydrogens are generated during the INIT step and added to the structure. The geometrical parameters of the side chains and the fixed backbone rotamers in the file rotamer_xpl.lib were taken from Lovell and Richardson: Lovell SC, Word JM, Richardson JS, Richardson DC (2000). The penultimate rotamer library. Proteins 40, 389-408.
(http://kinemage.biochem.duke.edu/databases/rotamer.php). The backbone dependent library corresponds to the library published by Dunbrack:
http://dunbrack.fccc.edu/bbdep/bbdepdownload.php, namely bbdep02.May.lib.
The solvation parameters are included in the nonbonded_xpl.lib file and are identical to those published in Lazaridis T and Karplus M (1999). Effective Energy function for proteins in solution. Proteins 35, 133-152.

By now a consistent set of files with AMBER all atom parameters is also provided. See distribution.

Adding a new ligand, residue type: If a new ligand or a new residue type is to be added then a number of additions have to be made to the parameter files above, namely to ROTAMER_LIB_ALL, NONBONDED_LIB, PARAM_LIB CONNEC_LIB. **Even if no rotamers will be generated, the residue must be specified in ROTAMER_LIB_ALL**.