Candock

Ver. 118

User’s Guide

1. Input / Output Files

```
obj ... directory with objective functions for KB energy minimization
data ... scoring function (csd...txt), gaff, amber FF
cando-ligands-3d.mol2 ... cando ligands converted to 3D
crystal_ligand.mol2 ... input ligand
receptor_abl1.pdb ... input receptor
centro_abl1.txt ... input/output file, predicted binding sites as centroids (n x
y z r) n...binding site number, x,y,z...center of centroid, r...centroid radius
bslibdb ... library of existing PDB binding sites /used by probis to predict
binding sites)
receptor_abl1.srf ... surface residues of receptor (intermediate file)
probis.json ... output of probis (intermediate file)
probis_with_ligands.json .... output of probis (intermediate file)
z_scores.pdb ... output of probis
ligand_clusters.pdb ... output of probis (used to determine binding sites)
small.mol2 ... small database of ligands
top_seeds ... rigidly docked seed fragments
cryst_cof_wat.txt ... ligand
centro_cof_wat.txt ... see above
receptor_cof_wat.pdb ... input file, example how to include cofactors and waters
for docking
probis.nosql ... intermediary file with aligned binding sites (output of probis)
gridpdb_hcp.pdb ... HCP grid of predicted binding site(s)
docked.pdb ... most important output file (minimized receptors and ligands,
energy - score see REMARK)
seeds.txt ... seeds, ie., rigid fragments hash table
prepared_ligands.pdb ... output file with typed ligands (idatm, gaff, rings,
seeds)
candock118 ... latest version of candock (118...commit number)
test_link118 ... last phase of docking (linking)
run.sh ... examples of candock usage (uncomment each example)
```

2. Options

```
USAGE:

   ./candock  [--ncpu <int>] [--clus_rad <double>] [--interatomic <double>]
              [--excluded <double>] [--conf_spin <double>] [--num_univec
              <int>] [--grid <double>] [--num_bsites <int>]
```

```
            [--max_clique_size <int>] [--top_percent <double>] [--pos_tol
            <double>] [--update_freq <int>] [--max_iter_final <int>]
            [--max_iter <int>] [--mini_tol <double>] [--fftype <kb|phy>]
            [--water_xml <string>] [--amber_xml <string>]
            [--max_num_possibles <int>] [--iterative] [--max_allow_energy
            <double>] [--docked_clus_rad <double>] [--docked_file
            <string>] [--link_iter <int>] [--max_possible_conf <int>]
            [--upper_tol_seed_dist <double>] [--lower_tol_seed_dist
            <double>] [--tol_seed_dist <double>] [--clash_coeff <double>]
            [--spin <5|10|15|20|30|60|90>] [--gaff_xml <string>]
            [--gaff_dat <string>] [--max_num_ligands <int>]
            [--topseedsfile <string>] [--clusterfile <string>] [--obj_dir
            <string>] [--potential_file <string>] [--scale <double>]
            [--step <double>] [--dist <string>] [--cutoff <4|5|6|7|8|9|10
            |11|12|13|14|15>] [--func <radial|normalized_frequency>]
            [--comp <reduced|complete>] [--ref <mean|cumulative>]
            [--max_frag_radius <double>] [--gridpdb_hcp <string>]
            [--centro_out <string>] [--centro_in <string>]
            [--centro_clus_rad <double>] [--z_scores_file <string>]
            [--lig_clus_file <string>] [--bio <string>]
            [--probis_min_z_score <double>] [--probis_min_pts <int>]
            [--probis_clus_rad <double>] [--neighb] [--names <string>]
            [--top_seeds_dir <string>] [--jsonwl <string>] [--json
            <string>] [--nosql <string>] [--bslib <string>] [--prep
            <string>] [--seeds <string>] --ligand <string> --receptor
            <string> [-q] [--] [--version] [-h]
```

Where:

```
   --ncpu <int>
     Number of CPUs to use concurrently (default is to use all CPUs)

   --clus_rad <double>
     Cluster radius for docked seeds (default is 2.0)

   --interatomic <double>
     Maximum interatomic distance (default is 8.0)

   --excluded <double>
     Excluded radius (default is 0.8)

   --conf_spin <double>
     Spin degrees for conformation generation (default is 10)

   --num_univec <int>
     Number of unit vectors evenly distributed on a sphere for conformation
     generation (default is 256)

   --grid <double>
     Grid spacing (default is 0.375)

   --num_bsites <int>
     Maximum number of predicted (or given) binding sites to consider for
     docking (default is 3)

   --max_clique_size <int>
     Maximum clique size for initial partial conformations generation
     (default is 3)

   --top_percent <double>
     Top percent of each docked seed to extend to full molecule (default is
     0.05)
```

```
--pos_tol <double>
  Minimization position tolerance in Angstroms - only for KB (default is
  0.00000000001)

--update_freq <int>
  Update non-bond frequency (default is 10)

--max_iter_final <int>
  Maximum iterations for final minimization (default is 100)

--max_iter <int>
  Maximum iterations for minimization during linking (default is 100)

--mini_tol <double>
  Minimization tolerance (default is 0.0001)

--fftype <kb|phy>
  Forcefield to use 'kb' (knowledge-based, default) or 'phy'
  (physics-based)

--water_xml <string>
  Water XML parameters (and topology) input file

--amber_xml <string>
  Receptor XML parameters (and topology) input file

--max_num_possibles <int>
  Maximum number of possibles conformations considered for clustering
  (default is 200000)

--iterative
  Enable iterative minimization during linking (not enabled by default)

--max_allow_energy <double>
  Maximum allowed energy for seed conformations (default is 0.0)

--docked_clus_rad <double>
  Cluster radius between docked ligand conformations (default is 2.0)

--docked_file <string>
  Docked ligands output filename (default is docked.pdb)

--link_iter <int>
  Maximum iterations for linking procedure (default is 1000)

--max_possible_conf <int>
  Maximum number of possible conformations to link (default is 20, -1
  means unlimited)

--upper_tol_seed_dist <double>
  Upper tolerance on seed distance for getting initial conformations of
  docked fragments (default is 2.0)

--lower_tol_seed_dist <double>
  Lower tolerance on seed distance for getting initial conformations of
  docked fragments (default is 2.0)

--tol_seed_dist <double>
  Tolerance on seed distance in-between linking (default is 2.0)

--clash_coeff <double>
  Clash coefficient for determining whether two atoms clash by eq.
  dist12 s< C * (vdw1 + vdw2) (default is 0.75)
```

```
   --spin <5|10|15|20|30|60|90>
     Spin degrees to rotate ligand (default 60)

   --gaff_xml <string>
     Gaff XML forcefield and ligand topology output file

   --gaff_dat <string>
     Gaff DAT forcefield input file

   --max_num_ligands <int>
     Maximum number of ligands to read in one chunk (default is 10)

   --topseedsfile <string>
     Top seeds                     output file

   --clusterfile <string>
     Clustered representative docked-seed conformations
      output file

   --obj_dir <string>
     Output directory for objective function and derivatives

   --potential_file <string>
     Output file for potentials and derivatives

   --scale <double>
     Scale non-bonded forces and energy for knowledge-based
      potential
     [0.0-1000.0] (default is 10.0)

   --step <double>
     Step for spline generation of non-bonded knowledge-based
      potential
     [0.0-1.0] (default is 0.01)

   --dist <string>
     Select one of the interatomic distance distribution
      file(s)
     provided with this script

   --cutoff <4|5|6|7|8|9|10|11|12|13|14|15>
     Cutoff length (default is 6)

   --func <radial|normalized_frequency>
     Function for calculating scores 'radial' (default) or
                        'normalized_frequency'

   --comp <reduced|complete>
     Atom types used in calculating reference state 'reduced' (default) or
     'complete'                     ('reduced' includes only those atom types
present in
     the specified receptor and small molecule, whereas    'complete'
                        includes all atom types)

   --ref <mean|cumulative>
     Normalization method for the reference state ('mean' (default) is
     averaged over all atom type pairs, whereas 'cumulative' is a summation
     for atom type pairs)

   --max_frag_radius <double>
     Maximum fragment radius for creating the initial rotamers (default is
     16.0)

   --gridpdb_hcp <string>
```

```
  Grid pdb hcp file for output (default is gridpdb_hcp.pdb)

--centro_out <string>
  Filename for outputting calculated centroids (default is centroid.txt)

--centro_in <string>
  Filename for reading centroids (default is centroid.txt)

--centro_clus_rad <double>
  Cluster radius for centroid centers (default is 3.0)

--z_scores_file <string>
  Binding site z-scores are outputted to this file (default is
  z_scores.pdb)

--lig_clus_file <string>
  Ligand clusters found by ProBiS are outputted to this file (default is
  ligand_clusters.pdb)

--bio <string>
  Directory with ProBiS-ligands bio database (default is
  bslibdb/data/bio)

--probis_min_z_score <double>
  Minimium z-score of ligands to be considered in clustering (default is
  2.5)

--probis_min_pts <int>
  The minimum number of points (for predicted ligands) required to form
  a cluster (default is 10)

--probis_clus_rad <double>
  Cluster radius for predicted ligands by probis (default is 3.0)

--neighb
  Allow only ligands that are in the similar regions according to
  REMARKs (default is disabled)

--names <string>
  Directory with ligand names (default is bslibdb/data/names)

--top_seeds_dir <string>
  Directory for saving top docked seeds (default is top_seeds)

--jsonwl <string>
  Json-formatted ProBiS alignments with transposed ligands output file
  (default is probis_with_ligands.json

--json <string>
  Json-formatted ProBiS alignments output file (default is probis.json

--nosql <string>
  NoSql-formatted ProBiS alignments output file (default is
  probis.nosql)

--bslib <string>
  Read binding sites library from this file (default is
  bslibdb/bslib.txt)

--prep <string>
  Prepared small molecule(s) are outputted to this filename (default is
  prepared_ligands.pdb)

--seeds <string>
```

```
  Read unique seeds from this file, if it exists, and append new
                    unique seeds if found

--ligand <string>
  (required)  Ligand filename

--receptor <string>
  (required)  Receptor filename

-q,  --quiet
  Quiet mode (default is verbose)

--,  --ignore_rest
  Ignores the rest of the labeled arguments following this flag.

--version
  Displays version information and exits.

-h,  --help
  Displays usage information and exits.


Command description message
```