Poster I-77

LIGLIB: Generation of in silico analogous ligand libraries

Authors:
Wolf Cochrane (*University Pretoria*)
Fourie Joubert (*University Pretoria*)

Short Abstract: LIGLIB (LIGand LIbrary Builder) is an open source tool that builds in silico analogous ligand libraries. The user selects permutation positions on the structure, and for each a range of linkers and fragments. LIGLIB permutes over combinations of linkers and fragments at each position and creates a molecule for each.

Long Abstract:

There is an ever-increasing need for open source tools to aid scientist in the drug discovery process and in the exploration of chemical space. However, chemical space is vast and searching it in its entirety is impossible. Scientists need tools that will allow them to pick a starting position in chemical space and search the local space around it.

LIGLIB is an open source molecular enumeration tool that builds in silico, analogous chemical libraries, through a user-guided process. LIGLIB follows a Markush-like enumeration approach, where permutations are made on a scaffold structure. The permutations take place at user defined atom positions on the scaffold, and at each position, a range of linkers and fragments, also defined by the user, are permuted over. The fragments take part in interactions at interaction sites in the protein active site while the linkers vary the distance and orientation of the fragments to the interaction site. LIGLIB stores all the possible combinations of permutations at all of the selected permutation positions as a new molecule in an in silico analogous ligand library.

LIGLIB was incorporated into a vector application, Chimera (http://www.cgl.ucsf.edu/chimera/), a molecular visualization package designed to be extensible. LIGLIB exploits the visualization capabilities of Chimera and allows the user to enter the LIGLIB input in an intuitive, easy to use 3D environment.

The user would typically start by taking a hit molecule, know to bind the specific target or proteins in the same family, and removing some groups at interaction sites where permutations should take place. Such a scaffold would be a good position to start a search in chemical space since, the scaffold is known to fit into the target active site, the scaffold positions fragments at the correct ligand-target interaction sites and also because the scaffold might itself have interactions with the target active site. The scaffold is then used as the primary input of LIGLIB. Next, the user selects positions on the scaffold where permutations should be made, and for each position a range of linkers and fragments, which is used to make permutations. The linkers and fragments are selected from fragment libraries. LIGLIB has a library manager tool that the user can use to create new and manage existing fragment/linker libraries. The fragments/linkers are in SMILES (Simple Molecular Input Line Entry System) format, which means that the user can create a large set of

fragments/linkers by simply writing the SMILES string code of a fragment and entering it into the library.

The methodology employed by LIGLIB can be split up into four main sequential steps: Gather user input; convert input scaffold into SMILES format; Build the library and convert the library into 3D mol2 format. The user input includes a scaffold structure in mol2 format and linkers and fragments in SMILES format. The library generation takes place in SMILES format, which means that the input scaffold should be converted to SMILES. A graph isomorphism algorithm is implemented to map the same atom in the SMILES structure and mol2 structure to each other, which allows LIGLIB to find in the 2D SMILES structure the atom selected by the user on the 3D structure in Chimera. The isomorphism compares atoms by using tree data structures of which the selected atom is the root, and the leaves the atoms connected to each other in their order of a traversal from the root atom through the rest of the molecule graphs. The library is generated by recursively going through all the possible combinations of linkers and fragments at each position and creating a new molecule for each combination. LIGLIB generates the ligand library in 2D SMILES and 3D mol2 format of which the 3D mol2 library is generated by the 3D coordinate generator tool Corina (http://www.mol-net.com/software/category/gen3dcoord.html). Corina is however a proprietary tool and users who whishes to also have the library in 3D mol2 format will have to purchase Corina. LIGLIB was designed to be fast and allows the user to generate either large chemical libraries or small focused ones, by varying the number of permutation positions, linkers and fragment.

LIGLIB was developed with a Python front-end, which was used to embed the software into the Chimera vector application. The LIGLIB back-end was written in C++, which allows for extremely fast execution allowing the user to, if needed, also build large molecule libraries.

The library generated by LIGLIB can be used for further investigations such as QSARs, pharmacophore, ADMET and docking studies. LIGLIB can also be added to larger drug discovery systems to perform fragment-based molecular enumeration in conjunction with database searching, ab initio molecule building, ADMET refinement, in silico screening and lead molecule refinement.