User Guide for LeFrag

Email: htzhao@lephar.com

Website: www.lephar.com

1. Usage

$LeFrag

Get "Help" information.

1.1 Fragmentation of a compounds library

$LeFrag –flib [Library of Tripos Mol2 molecules]

This will automatically decompose all small molecules in a library into fragments. The fragments are written into a file named "flib.mol2". A file named "flib.dat" records some basic physicochemical properties of the generated fragments, particularly the frequency of a fragment in the library. For a 9 million library, it will take about 1 hour by using just one single CPU.

1.2 Analysis of a compounds library

$LeFrag –anal [Library of Tripos Mol2 molecules]

This will automatically analyze all small molecules in the library and the physicochemical properties of each molecule are written in a file named "anal.dat".

1.3 Scaffold hopping

$LeFrag –simi [Library of Tripos Mol2 molecules] [Small molecule in query]

This will automatically calculate the similarity index between the small molecule in query and each small molecule in the library. The results are written in a file named "simi.dat".

1.4 Split a compounds library into each individual small molecule

$LeFrag –spli [Library of Tripos Mol2 molecules]

This will automatically split the library into each individual small molecule. The multiple conformers (or protonation states) of one molecule, if any, will be named sequentially such as Molename_01.mol2.

1.5 Retrieval of parent molecules by substructure search

$LeFrag –subs [Library of Tripos Mol2 molecules]

To retrieve all parent molecules containing at least one of the fragments in query, please put all fragments in query into a file named "query.mol2". This command will then automatically search in the library and the retrieved molecules are written in a file named "subs.mol2".

1.6 Filter of a library

$LeFrag –phar [Library of Tripos Mol2 molecules]

This command will filter the library according to the input in the file named "phar.inp". The passed small molecules are written into a file named "f.[original library name], and a file named "filter.dat" saves the names of failed molecules. The input file "phar.inp" contains the following information:

mw	200	400
acceptor	1	99
donor	1	99
rotb	1	99
cring	1	6
sring	1	6
hb		

mw: molecular weight; acceptor: number of acceptors; donor: number of donors; rotb: number of rotatable bonds; cring: number of chemical rings including fused rings; sring: number of non-fused rings; hb: capability of forming two bi-dentate hydrogen bonds (one as acceptor and the other as donor). The first number is the lower limit and the second number is the upper limit. If one dose not want to use e.g., molecular weight as a filter, please simply remove the whole line of "mw …".

Reference

1. **Hongtao Zhao**, *et. al.*; Discovery of a novel chemotype of tyrosine kinase inhibitors by fragment-based docking and molecular dynamics simulations. *Acs Med Chem Lett* **2012**, 3, 834.