MUDO

Introduction

MUDO (MolecUlar eDitOr) uses SMIRKS-encoded transforms to modify chemical structures in a controlled manner and is an update of the Leatherface molecular editor that makes use of the OEChem programming toolkit. The program is run as follows (optional commands in parentheses):

mudo --help (to get help)

mudo -i Structure.File -o Output.File -s SMIRKS.File -m Mode (-v Vector.Binding.File) (-e)

MUDO is based on Leatherface [1], which used SMARTS rather than SMIRKS to encode the structural transforms, and it provides all the capability of the ealier sofware.

Structure files

Molecular editing is most safely performed using 2D (i.e. lacking 3D coordinates) structure files. However, MUDO will accept structures in any format supported by OEChem and the primary function of the link function is to form covalent bonds in 3D structures (e.g. for covalent docking). Just try to think a bit about what you're trying to achieve and hopefully you can get the software to do something useful for you.

Mode of operation

The mode of operation is specified by -m Mode and three modes of operation are possible.

> normal: The SMIRKS transforms are applied sequentially to each structure and each transform is applied repeatedly to the structure until it fails to match. This mode is used primarily for standardisation (e.g. protonation, tautomers, de-salting). In this mode SMIRKS strings are used to create OEUniMolecularRxn objects.
>
> enum: Enumerates structures by applying SMIRKS. This mode is used for enumerating states and is also the usual starting point for Matched Molecular Pair Analysis [2,3]. In this mode SMIRKS strings are used to create OELibraryGen objects.
>
> link: Used for making bonds between disconnected atoms and is primarily for covalent docking. In this mode SMIRKS strings are used to create OELibraryGen objects.

Exhaustive

The exhaustive option is specified by the -e flag and it is only enabled for the normal and enum modes. The option applies the sequence of structural transforms repeatedly until either the structure (in normal mode) or the sorted list of canonical, isomeric SMILES (in enum mode) remains identical on successive edit cycles. This option can be used for stripping substituents off rings in a controlled manner [4] or for tautomer enumeration.

SMIRKS

The format of the SMIRKS file is as follows (with comments specified using # as first character on the line):

```
SMIRKS.String
```

Vector Bindings

An optional file of vector bindings can be used both in profiling and filtering and the format of this file is as follows (with comments specified using # as first character on the line):

```
Vector.Binding.Name  Vector.Binding.Definition
```

References

1. Kenny & Sadowski (2005) Structure modification in chemical databases. Methods and principles in medicinal chemistry. In: Oprea T (ed) Chemoinformatics in drug discovery, vol 23, pp 271–285. doi: http://dx.doi.org/10.1002/3527603743.ch11
2. Leach *et al* (2006): Matched Molecular Pairs as a Guide in the Optimization of Pharmaceutical Properties; a Study of Aqueous Solubility, Plasma Protein Binding and Oral Exposure. JMC 49:6672-6682. doi: http://dx.doi.org/10.1021/jm0605233
3. Birch *et al* (2009) Matched molecular pair analysis of activity and properties of glycogen phosphorylase inhibitors. BMCL 19:850-853. doi: http://dx.doi.org/10.1016/j.bmcl.2008.12.003
4. Morley *et al* (2009) 5-Aminopyrimidin-2-ylnitriles as Cathepsin K inhibitors. BMCL 19:1658-1661. doi: http://dx.doi.org/10.1016/j.bmcl.2009.01.110

Peter W. Kenny, São Carlos, 13-Mar-2013