NAME

FilterSDFiles.pl - Filter compounds from SDFile(s)

SYNOPSIS

FilterSDFiles.pl SDFile(s)...

FilterSDFiles.pl [-a, --all] [-e, --empty] [-c, --cleansalts] [-h, --help] [-k, --keep] [-m, --mismatch] [-o, --overwrite] [-r, --root *rootname*] [-s, --salts] [-u, --unknownatoms] [-w, --workingdir *dirname*] SDFile(s)...

DESCRIPTION

Filter specific compounds from *SDFile(s)*. Available choices are: wash or remove compounds with salts; take out compounds with no structural data; remove compounds with mismatched atom/bond blocks data; remove compounds which contain uknown atoms and so on. Multiple SDFile names are separated by spaces. The valid file extensions are *.sdf* and *.sd*. All other file names are ignored. All the SD files in a current directory can be specified either by **.sdf* or the current directory name.

OPTIONS

-a, --all

Use all options to filter compounds.

-e, --empty

Filter compounds with empty atom/bond blocks. This is default behavior.

-c, --cleansalts

Wash compounds which contain salts identified as disconnected structural units. The largest fragment is kept.

-h, --help

Print this help message.

-k, --keep

Keep the compounds which were filtered in a separate file. Default: Just ignore these compounds. Option -r --root is used to generate the new file name: <Root>Ignored.sdf. Default file name: <SDFileName>Ignored.sdf.

-m, --mismatch

Remove compounds with mismatched atom/bond blocks and counts line information specified by header block.

-o, --overwrite

Overwrite existing files.

-r, --root *rootname*

New SD file name is generated using the root: <Root>.sdf. Default file name:<SDFileName>Filtered.sdf. This option is ignored for multiple input files.

-s, --salts

Remove compounds which contain salts identified as disconnected structural units.

-u, --unknownatoms

Remove compounds with atom blocks containing special atom symbols such as L, Q, * ,LP, X, R#, or any other non periodic table symbols.

-w, --workingdir *dirname*

Location of working directory. Default: current directory.

EXAMPLES

To remove compounds from SD files which contain salts, unknown atoms, or mismatched atom/bonds block data or no structural data, type:

```
% FilterSDFiles.pl -a -o Sample.sdf
% FilterSDFiles.pl -a -o *.sdf
```

And to generate a new NewSampleIgnored.sdf file for filtered compounds, type:

```
% FilterSDFiles.pl -a -k -r NewSample -o Sample.sdf
```

To wash compounds in order to get rid of all disconnected fragments except for the largest one, type:

```
% FilterSDFiles.pl -c -o Sample.sdf
```

AUTHOR

Manish Sud <msud@san.rr.com>

SEE ALSO

ExtractFromSDFiles.pl, InfoSDFiles.pl, MergeTextFilesWithSD.pl

COPYRIGHT