NAME

RDKitConvertFileFormat.py - Convert between molecular file formats

SYNOPSIS

RDKitConvertFileFormat.py [--infileParams <Name,Value,...>] [--outfileParams <Name,Value,...>] [--overwrite] [-w <dir>] -i <infile> -o <outfile>

RDKitConvertFileFormat.py -h | --help | -e | --examples

DESCRIPTION

Convert between molecular file formats.

The supported input file formats are: Mol (.mol), SD (.sdf, .sd), SMILES (.smi, .txt, .csv, .tsv), MOL2 (.mol2), PDB (.pdb)

The supported output file formats are: SD (.sdf, .sd), SMILES (.smi), PDB (.pdb)

OPTIONS

-e, --examples

Print examples.

-h, --help

Print this help message.

-i, --infile <infile>

Input file name.

--infileParams <Name,Value,...> [default: auto]

A comma delimited list of parameter name and value pairs for reading molecules from files. The supported parameter names for different file formats, along with their default values, are shown below:

```
SD, MOL: removeHydrogens,yes,sanitize,yes,strictParsing,yes
MOL2: removeHydrogens,yes,sanitize,yes
SMILES: smilesColumn,1,smilesNameColumn,2,smilesDelimiter,space,
    smilesTitleLine,auto,sanitize,yes
PDB: removeHydrogens,yes,sanitize,yes
```

Possible values for smilesDelimiter: space, comma or tab.

-o, --outfile <outfile>

Output file name.

--outfileParams <Name,Value,...> [default: auto]

A comma delimited list of parameter name and value pairs for writing molecules to files. The supported parameter names for different file formats, along with their default values, are shown below:

```
SD: compute2DCoords,auto,kekulize,no
SMILES: kekulize,no,smilesDelimiter,space, smilesIsomeric,yes,
    smilesTitleLine,yes
```

Default value for compute2DCoords: yes for SMILES input file; no for all other file types.

--overwrite

Overwrite existing files.

-w, --workingdir <dir>

Location of working directory which defaults to the current directory.

EXAMPLES

To convert a SD file into a isomeric SMILES file, type:

```
% RDKitConvertFileFormat.py -i Sample.sdf -o SampleOut.smi
```

To convert a SD file into a non isomeric SMILES file, type

```
% RDKitConvertFileFormat.py --outfileParams "smilesIsomeric,no"
  -i Sample.sdf -o SampleOut.smi
```

To convert a SMILES file into a SD file along with calculation of 2D coordinates, type:

```
% RDKitConvertFileFormat.py -i Sample.smi -o SampleOut.sdf
```

To convert a MDL MOL file into a PDB file, type:

```
% RDKitConvertFileFormat.py -i Sample.mol -o SampleOut.pdb
```

To convert a CSV SMILES file with column headers, SMILES strings in column 1, and name in column 2 into a SD file containing 2D coordinates, type:

```
% RDKitConvertFileFormat.py --infileParams "smilesDelimiter,comma,
  smilesTitleLine,yes,smilesColumn,1,smilesNameColumn,2" -i Sample.csv
  -o SampleOut.sdf
```

AUTHOR

Manish Sud(msud@san.rr.com)

SEE ALSO

RDKitDrawMolecules.py, RDKitRemoveDuplicateMolecules.py, RDKitSearchFunctionalGroups.py, RDKitSearchSMARTS.py

COPYRIGHT