NAME

RDKitPerformMinimization.py - Perform structure minimization

SYNOPSIS

RDKitPerformMinimization.py [--addHydrogens <yes or no>] [--conformerGenerator <SDG, ETDG, KDG, ETKDG, None>] [--forceField <UFF or MMFF>] [--enforceChirality <yes or no>] [--infileParams <Name,Value,...>] [--maxConfs <number>] [--maxIters <number>] [--outfileParams <Name,Value,...>] [--overwrite] [--removeHydrogens <yes or no>] [--randomSeed <number>] [-w <dir>] -i <infile> -o <outfile>

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

DESCRIPTION

Generate 3D structures for molecules using combination of distance geometry and forcefield minimization or minimize existing 3D structures using a specified forcefield.

The supported input file formats are: Mol (.mol), SD (.sdf, .sd), SMILES (.smi) .csv, .tcsv .txt)

The supported output file formats are: SD (.sdf, .sd)

OPTIONS

-a, --addHydrogens <yes or no> [default: yes]

Add hydrogens before minimization.

-c, --conformerGenerator <SDG, ETDG, KDG, ETKDG, None> [default: ETKDG]

Conformation generation methodology for generating initial 3D coordinates. Possible values: Standard Distance Geometry, (SDG), Experimental Torsion-angle preference with Distance Geometry (ETDG), basic Knowledge-terms with Distance Geometry (KDG), and Experimental Torsion-angle preference along with basic Knowledge-terms with Distance Geometry (ETKDG) [Ref 129] .

The conformation generation step may be skipped by specifying 'None' value to perform only forcefield minimization of molecules with 3D structures in input file. This doesn't work for molecules in SMILES file or molecules in SD/MOL files containing 2D structures.

-f, --forceField <UFF or MMFF> [default: MMFF]

Forcefield method to use for energy minimization. Possible values: Universal Force Field (UFF) [Ref 81] or Merck Molecular Mechanics Force Field (MMFF) [Ref 83-87] .

--enforceChirality <yes or no> [default: Yes]

Enforce chirality for defined chiral centers.

-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
SMILES: smilesColumn,1,smilesNameColumn,2,smilesDelimiter,space,
    smilesTitleLine,auto,sanitize,yes
```

Possible values for smilesDelimiter: space, comma or tab.

--maxConfs <number> [default: 250]

Maximum number of conformations to generate for each molecule by conformation generation methodology for initial 3D coordinates. The conformations are minimized using the specified forcefield and the lowest energy conformation is written to the output file. This option is ignored during 'None' value of '-c --conformerGenerator' option.

--maxIters <number> [default: 500]

Maximum number of iterations to perform for each molecule during forcefield minimization.

-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: kekulize,no
```

--overwrite

Overwrite existing files.

-r, --removeHydrogens <yes or no> [default: Yes]

Remove hydrogens after minimization.

--randomSeed <number> [default: auto]

Seed for the random number generator for reproducing 3D coordinates. Default is to use a random seed.

-w, --workingdir <dir>

Location of working directory which defaults to the current directory.

EXAMPLES

To generate up to 250 conformations using ETKDG methodology followed by MMFF forcefield minimization for a maximum of 500 iterations for molecules in a SMILES file and write out a SD file containing minimum energy structurse corresponding to each molecule, type

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

To generate up to 150 conformations using ETKDG methodology followed by MMFF forcefield minimization for a maximum of 250 iterations along with a specified random seed for molecules in a SMILES file and write out a SD file containing minimum energy structures corresponding to each molecule, type

```
% RDKitPerformMinimization.py  --maxConfs 150  --randomSeed 201780117
  --maxIters 250  -i Sample.smi -o SampleOut.sdf
```

To minimize structures in a 3D SD file using UFF forcefield for a maximum of 150 iterations without generating any conformations and write out a SD file containing minimum energy structures corresponding to each molecule, type

```
% RDKitPerformMinimization.py  -c None -f UFF --maxIters 150
  -i Sample3D.sdf -o SampleOut.sdf
```

To generate up to 50 conformations using SDG methodology followed by UFF forcefield minimization for a maximum of 50 iterations for molecules in a CSV SMILES file, SMILES strings in column 1, name in column 2, and write out a SD file, type:

```
% RDKitPerformMinimization.py  --maxConfs 50  --maxIters 50 -c SDG
  -f UFF --infileParams "smilesDelimiter,comma,smilesTitleLine,yes,
  smilesColumn,1,smilesNameColumn,2"  -i SampleSMILES.csv
  -o SampleOut.sdf
```

AUTHOR

Manish Sud(msud@san.rr.com)

SEE ALSO

RDKitCalculateRMSD.py, RDKitCalculateMolecularDescriptors.py, RDKitCompareMoleculeShapes.py, RDKitConvertFileFormat.py, RDKitGenerateConformers.py

COPYRIGHT

Copyright (C) 2018 Manish Sud. All rights reserved.

The functionality available in this script is implemented using RDKit, an open source toolkit for cheminformatics

developed by Greg Landrum.

This file is part of MayaChemTools.

MayaChemTools is free software; you can redistribute it and/or modify it under the terms of the GNU Lesser General Public License as published by the Free Software Foundation; either version 3 of the License, or (at your option) any later version.