NAME

RDKitCalculatePartialCharges.py - Calculate partial atomic charges

SYNOPSIS

RDKitCalculatePartialCharges.py [--allowParamFailure <yes or no>] [--chargesSDFormat <AtomAliases or DataField>] [--dataFieldLabel <text>] [--infileParams <Name,Value,...>] [--mode <Gasteiger or MMFF>] [--numIters <number>] [--outfileParams <Name,Value,...>] [--precision <number>] [--overwrite] [-w <dir>] -i <infile> -o <outfile>

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

DESCRIPTION

Calculate partial charges for atoms in molecules and write them out to a SD file. The hydrogens are automatically added to molecules before calculating partial charges.

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

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

OPTIONS

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

Allow calculation of Gasteiger partial charges to proceed for molecules containing atoms with unknown parameters. The atoms with unknown parameters are removed from the calculations by setting their values to zero.

-c, --chargesSDFormat <AtomAliases or DataField> [default: AtomAliases]

Format for writing out partial atomic charges to SD file. Possible values: AtomAliases or DataField.

The charges are stored as atom property named 'molFileAlias' for 'AtomAliases' format and may be retrieved using the RDKit function 'GetProp' for atoms: Aotm.GetProp('molFileAliases').

The charges are stored under a data field label speciefied using '-d, --dataFieldLabel' for 'DataField' format and may be retrieved using the RDKit function 'GetProp' for molecues.

-d, --dataFieldLabel <text> [default: PartialCharges]

Data field label to use for storing charged in SD file during 'DataField' value of '-c, --chargesSDFormat'.

-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.

-m, --mode <Gasteiger or MMFF> [default: Gasteiger]

Type of partial atomic charges to calculate. Possible values: Gasteiger [Ref 138] or Merk Molecular Mechanics Fore Field (MMFF) [Ref 83-87].

-n, --numIters <number> [default: 12]

Number of iterations to perform during calculation of Gasteiger charges.

-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
```

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

-p, --precision <number> [default: 3]

Floating point precision for writing the calculated partial atomic charges.

--overwrite

Overwrite existing files.

`-w, --workingdir <dir>`

Location of working directory which defaults to the current directory.

EXAMPLES

To calculate Gasteiger partial atomic charges for molecules in a SMILES file and write them out to a SD file as atom aliases, type:

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

To calculate MMFF forcefield partial atomic charges for molecules in a SD file and write them out to a SD file under 'PartialCharges' data field, type:

```
% RDKitCalculatePartialCharges.py  -m MMFF -c DataField -i Sample.sdf
  -o SampleOut.sdf
```

To calculate Gasteiger partial atomic charges for molecules in a SMILES file and write them out to a SD file under a data field named 'GasteigerCharges', type:

```
% RDKitCalculatePartialCharges.py  -m Gasteiger -c DataField
  -d GasteigerCharges -p 4 -i Sample.smi -o SampleOut.sdf
```

To calculate Gasteiger partial atomic charges for molecules in a CSV SMILES file, SMILES strings in column 1, name in column 2, and write out a SD file containing charges as atom aliases, type:

```
% RDKitCalculatePartialCharges.py --infileParams
  "smilesDelimiter,comma,smilesTitleLine,yes,smilesColumn,1,
  smilesNameColumn,2" --outfileParams "compute2DCoords,yes"
  -i SampleSMILES.csv -o SampleOut.sdf
```

AUTHOR

Manish Sud(msud@san.rr.com)

SEE ALSO

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

COPYRIGHT

The functionality available in this script is implemented using RDKit, an open source toolkit for cheminformatics developed by Greg Landrum.