

## NAME

ElementalAnalysisSDFFiles.pl - Perform elemental analysis using formula data field in SDFFile(s)

## SYNOPSIS

ElementalAnalysisSDFFiles.pl SDFFile(s)...

ElementalAnalysisSDFFiles.pl [-d, --detail *infolevel*] [--fast] [--formulafield *SD data field name*] [-f, --formulamode *DataField | StructureData*] [--formulaout *yes or no*] [-m, --mode *All | "ElementalAnalysis, [MolecularWeight, ExactMass]"*] [-h, --help] [-o, --overwrite] [-r, --root *rootname*] [-v --valuefieldnames *Name, Label, [Name, Label,...]*] [-w, --workingdir *dirname*] SDFFile(s)...

## DESCRIPTION

Perform elemental analysis using molecular formula specified by a data field name or generated from structure data in *SDFFile(s)*.

In addition to straightforward molecular formulas - H<sub>2</sub>O, HCl, C<sub>3</sub>H<sub>7</sub>O<sub>2</sub>N - other supported variations are: Ca<sub>3</sub>(PO<sub>4</sub>)<sub>2</sub>, [PCl<sub>4</sub>]<sup>+</sup>, [Fe(CN)<sub>6</sub>]<sup>4-</sup>, C<sub>37</sub>H<sub>42</sub>N<sub>2</sub>O<sub>6</sub>+<sub>2</sub>, Na<sub>2</sub>CO<sub>3</sub>.10H<sub>2</sub>O, 8H<sub>2</sub>S.46H<sub>2</sub>O, and so on. Charges are simply ignored. Isotope symbols in formulas specification, including D and T, are not supported.

The file names are separated by space. 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

-d, --detail *infolevel*

Level of information to print about compound records being ignored. Default: 1. Possible values: 1, 2 or 3.

--fast

In this mode, the formula data field specified using -f, --formulafield option is assumed to contain valid molecular formula data and initial formula validation check is skipped.

--formulafield *SD data field name*

*SDFFile(s)* data field name containing molecular formulas used for performing elemental analysis during *DataField* value of -f, --formulamode option. Default value: *SD data field containing the word formula in its name*.

This option is ignore during *StructureData* value of -f, --formulamode option.

-f, --formulamode *DataField | StructureData*

Specify source of molecular formula used for performing elemental analysis: retrieve formula using *SDFFile(s)* data field name or generate formula from structure. Possible values: *DataField or StructureData*. Default value: *DataField*.

--formulaout *yes or no*

Specify whether to write out formula to SD file during *StructureData* value of -f, --formulamode option. Possible values: *Yes or No*. Default: *No*.

-m, --mode *All | "ElementalAnalysis,[MolecularWeight,ExactMass]"*

Specify what values to calculate using molecular formula data field or structure data from *SDFFile(s)*: calculate all supported values or specify a comma delimited list of values. Possible values: *All | "ElementalAnalysis, [MolecularWeight, ExactMass]"*. Default: *All*

-h, --help

Print this help message.

-o, --overwrite

Overwrite existing files.

-p, --precision *number*

Precision of calculated values in the output file. Default: up to 2 decimal places. Valid values: positive integers.

-r, --root *rootname*

New SD file name is generated using the root: <Root>.<Ext>. Default new file name: <InitialSDFFileName>ElementalAnalysis.<Ext>. This option is ignored for multiple input files.

---

`-v --valuefieldnames Name,Label,[Name,Label,...]`

Specify SD data field names to use for calculated values. In general, it's a comma delimited list of value name and SD field name pairs. Supported value names: *ElementalAnalysis*, *MolecularWeight*, *ExactMass*, and *MolecularFormula*. Default labels: *ElementalAnalysis*, *MolecularWeight*, *ExactMass*, and *MolecularFormula*.

*MolecularFormula* label is only used during *StructureData* value of `-f`, `--formulamode` option.

`-w, --workingdir dirname`

Location of working directory. Default: current directory.

## EXAMPLES

To perform elemental analysis, calculate molecular weight and exact mass using SD field name value with the word Formula in its name and generate a new SD file NewSample1.sdf, type:

```
% ElementalAnalysisSDFiles.pl -o -r NewSample1 Sample1.sdf
```

To perform elemental analysis, calculate molecular weight and exact mass using structure data in SD file and generate a new SD file NewSample1.sdf, type:

```
% ElementalAnalysisSDFiles.pl --formulamode StructureData -o  
-r NewSample1 Sample1.sdf
```

To perform elemental analysis using formulas in SD field name Formula, use field name Analysis for calculated data, and generate a new SD file NewSample1.sdf, type:

```
% ElementalAnalysisSDFiles.pl --m ElementalAnalysis --formulafield  
Formula --valuefieldnames "ElementalAnalysis,Analysis" -o  
-r NewSample1 Sample1.sdf
```

To calculate molecular weight, using formulas in SD field name Formula, with four decimal precision and generate a new SD file NewSample1.sdf, type

```
% ElementalAnalysisSDFiles.pl --m MolecularWeight --formulafield  
Formula --precision 4 -o -r NewSample1 Sample1.sdf
```

## AUTHOR

Manish Sud <msud@san.rr.com>

## SEE ALSO

AnalyzeSDFilesData.pl, InfoSDFiles.pl, ExtractFromSDFiles.pl

## COPYRIGHT

Copyright (C) 2024 Manish Sud. All rights reserved.

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.