

NAME

RDKitDrawMolecules.py - Draw molecules and generate an image or HTML file

SYNOPSIS

```
RDKitDrawMolecules.py [--alignmentSMARTS <SMARTS>] [--atomLabelFontSize <number>] [
--bondLineWidth <number>] [--compute2DCoords <yes | no>] [--counterCol <yes or no>] [--colVisibility
<yes or no>] [--fontBold <yes or no>] [--footer <text>] [--footerClass <text>] [--header <text>] [
--headerStyle <text>] [--highlightSMARTS <SMARTS>] [--infileParams <Name,Value,...>] [--kekulize <yes
or no>] [--keysNavigation <yes or no>] [--molImageSize <width,height>] [--molImageEncoded <yes or
no>] [--numOfMolsPerRow <number>] [--overwrite] [--paging <yes or no>] [--pagingType <numbers,
simple, ...>] [--pageLength <number>] [--popover <yes or no>] [--popoverDataCount <number>] [
--popoverTextWidth <number>] [--showMolName <yes or no>] [--scrollX <yes or no>] [--scrollY <yes or
no>] [--scrollYSize <number>] [--tableFooter <yes or no>] [--tableHeader <yes or no>] [
--tableHeaderStyle <thead-dark,thead-light,...>] [--tableStyle <table,table-striped,...>] [-w <dir>] -i
<infile> -o <outfile>
```

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

DESCRIPTION

Draw molecules in a grid and write them out as an image file or a HTML table file. The SVG image or HTML table file appears to be the best among all the available image file options, as rendered in a browser. The Python modules aggdraw/cairo are required to generate high quality PNG images.

The drawing of the molecules are embedded in HTML table columns as in line SVG images. The HTML table is an interactive table and requires internet access for viewing in a browser. It employs the following frameworks: JQuery, Bootstrap, and DataTable.

The options '-atomLabelFontSize' and '-bondLineWidth' don't appear to work during the generation of a SVG image.

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

The output image file can be saved in any format supported by the Python Image Library (PIL). The image format is automatically detected from the output file extension.

Some of the most common output image file formats are: GIF (.gif), JPEG (.jpg), PNG (.png), SVG (.svg), TIFF (.tif). In addition, a HTML (.html) file format containing a table is supported.

OPTIONS

-a, --alignmentSMARTS <SMARTS> [default: none]

SMARTS pattern for aligning molecules to a common template.

--atomLabelFontSize <number> [default: 12]

Font size for drawing atom labels. This option is ignored during the generation of a SVG and HTML output file.

-b, --bondLineWidth <number> [default: 1.2]

Line width for drawing bonds. This option is ignored during the generation of a SVG and HTML output file.

-c, --compute2DCoords <yes or no> [default: auto]

Compute 2D coordinates of molecules before drawing. Default: yes for all file formats.

--counterCol <yes or no> [default: yes]

Show a counter column as the first column in the table. It contains the position for each row in the HTML table. This option is only used during the generation of a HTML table file.

--colVisibility <yes or no> [default: yes]

Show a dropdown button to toggle visibility of columns in the table. This option is only used during the generation of a HTML table file.

-e, --examples

Print examples.

-f --fontBold <yes or no> [default: yes]
Make all text fonts bold during the generation of a SVG and HTML output file. This option is ignored for all other output files.

--footer <text> [default: none]
Footer text to insert at the bottom of the HTML page after the table. This option is only used during the generation of a HTML table file.

--footerClass <text> [default: small text-center text-muted]
Footer class style to use with <p> tag. This option is only used during the generation of a HTML table file.

-h, --help
Print this help message.

--header <text> [default: none]
Header text to insert at the top of the HTML page before the table. This option is only used during the generation of a HTML table file.

--headerStyle <text> [default: h5]
Header style to use. Possible values: h1 to h6. This option is only used during the generation of a HTML table file.

--highlightSMARTS <SMARTS> [default: none]
SMARTS pattern for highlighting atoms and bonds in molecules. All matched substructures are highlighted.

-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.

-k, --kekulize <yes or no> [default: yes]
Perform kekulization on molecules. This option is ignored during the generation of a SVG and HTML output file.

--keysNavigation <yes or no> [default: yes]
Provide Excel like keyboard cell navigation for the table. This option is only used during the generation of a HTML table file.

-m, --molImageSize <width,height> [default: 250,200]
Image size of a molecule in pixels.

--molImageEncoded <yes or no> [default: yes]
Base64 encode SVG image of a molecule for inline embedding in a HTML page. The inline SVG image may fail to display in browsers without encoding.

-n, --numOfMolsPerRow <number> [default: 4]
Number of molecules to draw in a row.

-o, --outfile <outfile>
Output file name.

--overwrite
Overwrite existing files.

-p, --paging <yes or no> [default: yes]

Provide page navigation for browsing data in the table. This option is only used during the generation of a HTML table file.

--pagingType <numbers, simple, ...> [default: full_numbers]

Type of page navigation. Possible values: numbers, simple, simple_numbers, full, full_numbers, or first_last_numbers.

- numbers - Page number buttons only
- simple - 'Previous' and 'Next' buttons only
- simple_numbers - 'Previous' and 'Next' buttons, plus page numbers
- full - 'First', 'Previous', 'Next' and 'Last' buttons
- full_numbers - 'First', 'Previous', 'Next' and 'Last' buttons, plus page numbers
- first_last_numbers - 'First' and 'Last' buttons, plus page numbers

This option is only used during the generation of a HTML table file.

--pageLength <number> [default: 5]

Number of rows to show per page. This option is only used during the generation of a HTML table file.

--popover <yes or no> [default: yes]

Display a popover window containing additional information about the molecule. The popover is opened after a click on the drawing of a molecule. A subsequent click on the same drawing closes the popover. This option is only used during the generation of a HTML table file.

--popoverDataCount <number> [default: 25]

Maximum number of data fields to show in a popover window. This option is only used during the generation of a HTML table file.

--popoverTextWidth <number> [default: 50]

Maximum width in characters for text display in a popover window before truncating the text. This option is only used during the generation of a HTML table file.

-s, --showMolName <yes or no> [default: yes]

Show molecule names under the images. This option is only used during the generation of a HTML table file.

--scrollX <yes or no> [default: yes]

Provide horizontal scroll bar in the table as needed. This option is only used during the generation of a HTML table file.

--scrollY <yes or no> [default: yes]

Provide vertical scroll bar in the table as needed. This option is only used during the generation of a HTML table file.

--scrollYSize <number> [default: 75vh]

Maximum height of table viewport either in pixels or percentage of the browser window height before providing a vertical scroll bar. Default: 75% of the height of browser window. This option is only used during the generation of a HTML table file.

-t, --tableStyle <table,table-striped,...> [default: table,table-hover,table-sm]

Style of table. Possible values: table, table-striped, table-bordered, table-hover, table-dark, table-sm, none, or All. Default: 'table,table-hover'. A comma delimited list of any valid Bootstrap table styles is also supported

This option is only used during the generation of a HTML table file.

--tableFooter <yes or no> [default: yes]

Show Excel style column headers at the end of the table. This option is only used during the generation of a HTML table file.

--tableHeader <yes or no> [default: yes]

Show Excel style column headers in the table. This option is only used during the generation of a HTML table file.

--tableHeaderStyle <thead-dark,thead-light,...> [default: thead-dark]

Style of table header. Possible values: thead-dark, thead-light, or none. The names of the following contextual color classes are also supported: table-primary (Blue), table-success (Green), table-danger (Red), table-info (Light blue), table-warning (Orange), table-active (Grey), table-light (Light grey), and table-dark (Dark grey).

This option is only used during the generation of a HTML table file.

-w, --workingdir <dir>

Location of working directory which defaults to the current directory.

EXAMPLES

To automatically compute 2D coordinates for molecules in a SMILES file and generate a SVG image file containing 4 molecules per row in a grid with cell size of 250 x 200 pixels, type:

```
% RDKitDrawMolecules.py -i Sample.smi -o SampleOut.svg
```

To automatically compute 2D coordinates for molecules in a SMILES file and generate a SVG image file containing 2 molecules per row in a grid with cell size of 400 x 300 pixels and without any keulization along with highlighting a specific set of atoms and bonds indicated by a SMARTS pattern, type:

```
% RDKitDrawMolecules.py -n 2 -m "400,300" -k no --fontBold no
--highlightSMARTS 'c1ccccc1' -i Sample.smi -o SampleOut.svg
```

To generate a PNG image file for molecules in a SD file using existing 2D coordinates, type

```
% RDKitDrawMolecules.py --compute2DCoords no -i Sample.sdf
-o SampleOut.png
```

To automatically compute 2D coordinates for molecules in a SD file and generate a HTML file containing 4 molecules per row in a table, along with all the bells and whistles to interact with the table, type:

```
% RDKitDrawMolecules.py -i Sample.sdf -o SampleOut.html
```

To automatically compute 2D coordinates for molecules in a SD file and generate a HTML file containing 4 molecules per row in a table without any bells and whistles to interact with the table, type:

```
% RDKitDrawMolecules.py --counterCol no --colVisibility no
--keysNavigation no --paging no --popover no --scrollX no
--scroll no --tableFooter no --tableHeader no -i Sample.sdf
-o SampleOut.html
```

To automatically compute 2D coordinates for molecules in a CSV SMILES file with column headers, SMILES strings in column 1, and name in column 2 and generate a PDF image file, type:

```
% RDKitDrawMolecules.py --infileParams "smilesDelimiter,comma,
smilesTitleLine,yes,smilesColumn,1,smilesNameColumn,2"
-i SampleSMILES.csv -o SampleOut.pdf
```

AUTHOR

Manish Sud(msud@san.rr.com)

SEE ALSO

RDKitConvertFileFormat.py, RDKitDrawMoleculesAndDataTable.py, RDKitRemoveDuplicateMolecules.py, RDKitSearchFunctionalGroups.py, RDKitSearchSMARTS.py

COPYRIGHT

Copyright (C) 2020 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.