

Book1 - Microsoft Excel

Home Insert Page Layout Formulas Data Review View Developer Helium JChem

Clipboard Font Alignment Number Styles Cells Editing

AutoSum Fill Clear Sort & Filter Find & Select

chembl100

	A	B	C	D	E	F	G	H	I	J	K
1											
2											
3	chembl100										
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											

RDKit in Excel with Helium

Paul Wallace

Ceiba Solutions Inc.

paul.wallace@ceibasolutions.com

www.ceibasolutions.com

Helium

Data type: ChEMBL ID

- Get ChEMBL Compound Bioactivities
- Get ChEMBL Compound Record
- Get ChEMBL Target Bioactivities
- Get SMILES

Sheet1 Sheet2 Sheet3

Ready 100%

Book1 - Microsoft Excel

Home Insert Page Layout Formulas Data Review View Developer Helium JChem Format

Insert Shapes Edit Shape Text Box Shape Fill Shape Outline Shape Effects Shape Styles WordArt Styles Text Fill Text Outline Text Effects Arrange Size

TextBox 1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1																
2	chembl100															
3	chembl101															
4	chembl102															
5	chembl103															
6	chembl104															
7	chembl105															
8																
9																
10																
11																
12																
13																
14																
15																
16																
17																
18																
19																
20																
21																
22																
23																
24																
25																
26																
27																
28																
29																
30																
31																
32																

1. <= Select some data
Identified as Chembl IDs
Relevant functions are displayed =>

2. Click on a function.....

Helium

Data type: Null

- Get All ChEMBL Targets
- Get All KEGG Compounds
- Get All KEGG Drugs
- Get All KEGG Glycans
- Get Basic Data Types
- Get Biological Data
- test2
- Get Compounds for Project ID
- Get EnhancedCritres Data Test
- Get Similar Compounds
- LoadCsvFile
- OnlyVisibleInExcel
- Protein Data Bank Search
- Random Normal
- Search for ChEBI Entities
- Search for PubMed Articles
- Search Sub-structure in databases
- TEST HTML dialog
- VisibleInExcelAndSpotfire

Sheet1 Sheet2 Sheet3

Ready 100%

Book1 - Microsoft Excel

Home Insert Page Layout Formulas Data Review View Developer Helium JChem

Clipboard Font Alignment Number Styles Cells Editing

B2 CC1(C)Oc2ccc(cc2[C@H]([C@@H]1O)N3CCCC3=O)C#N

	A	B	C	D	E	F	G	H	I
1		SMILES							
2	chembl100	CC1(C)Oc2ccc(cc2[C@H]([C@@H]1O)N3CCCC3=O)C#N							
3	chembl101	CCCCC1C(=O)N(N(C1=O)c2ccccc2)c3ccccc3							
4	chembl102	CO\N=C(\C(=O)N[C@H]1[C@H]2SCC(=C(N2C1=O)C(=O)O)COC(=O)C)/c3csc(N)n3							
5	chembl103	CC(=O)[C@H]1CC[C@H]2[C@H]3CCC4=CC(=O)CC[C@]4(C)[C@H]3CC[C@]12C							
6	chembl104	Clc1ccccc1C(c2ccccc2)(c3ccccc3)n4ccnc4							
7	chembl105	CO[C@]12[C@H]3N[C@H]3CN1C4=C([C@H]2COC(=O)N)C(=O)C(=C(C)C4=O)N							
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
26									
27									
28									
29									
30									
31									
32									

3. The Smiles are returned to Excel

and repeat:

<= Select some data
Identified as SMILES
Relevant functions are displayed =>

Click on a function
Notice the functions using RDKit, others are easily added:
Calculate Properties
Smiles To Molfile
Smiles to 3D Molfile
Smiles to Image

Helium

- Calculate Derived Properties
- Calculate Properties (RDKit)
- Calculate Simple Properties
- mw
- Canonicalise
- ChemAxon CalculateNmr
- ChemAxon Cluster
- ChemAxon Properties
- Generate R-groups
- Get ChEMBL Compound Record
- Get Compound Number
- Get InChI
- Get PubChem Radar Plot
- Get PubChem Record
- Get PubChem Record (CONFIG)
- Get PubChem Simple Properties
- GetRadarPlot
- Protein Data Bank Search
- PubChem Substructure Search
- PubChem Superstructure Search
- RemoveIsotopes
- Remove Salts / Solvates
- Remove Stereochemistry
- Run PipelinePilot Workflow
- Smiles To 3D Molfile (RDKit)
- Smiles To Image (RDKit)
- Smiles To Molfile (RDKit)
- Standardise Valences
- Sub-structure Search within table

Ready Sheet1 Sheet2 Sheet3 Count: 6 100%

Book1 - Microsoft Excel

Home Insert Page Layout Formulas Data Review View Developer Helium JChem

Clipboard Font Alignment Number Styles Cells Editing

Calibri 11 A A

B I U Merge & Center

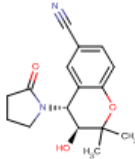
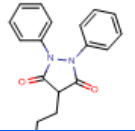
General

Conditional Formatting Format as Table Cell Styles

Insert Delete Format

AutoSum Fill Clear Sort & Filter Find & Select

C2

	A	B	C	D	E	F	G	H	I	J	K
1		SMILES	Conformation	Molfile	Mol Weight	logP	MR	TPSA			
2	chembl100	CC1(C)Oc2ccc(cc2[C@H]	<div>RDKit 3D 23 25 0 0 0 0 0 0 0999 V2000 -0.9697 3.5330</div>		314.469	4.7235	90.881	34.14			
			<div>RDKit 3D 23 25 0 0 0 0 0 0 0999 V2000</div>		344.845	5.3767	101.843	17.82			

Properties, Molfiles and Conformers returned to Excel

Molfiles can be shown with JChemForExcel or as embedded images

Conformations can be shown with PyMol

PyMOL Viewer

all A S H L C
Molfile1 A S H L C

Mouse Mode 3-Button Viewing
Buttons L M R Wheel
& Keys Rota Move MovZ Slab
ShFt +Box -Box Clip MovS
Ctrl +/- PkAt Pk1 MvSZ
CtSh Sele Orig Clip MovZ
SnglClk +/- Cent Menu
DblClk Menu - PkAt
Selecting Residues
State 1/ 1

PyMOL>_

Ready

100%