JChem for Excel

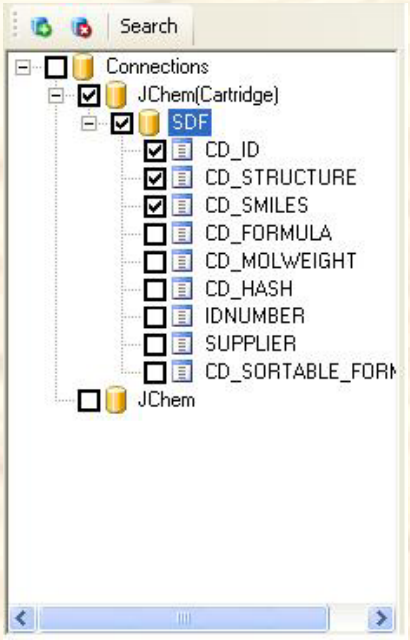
By Tamas Pelcz

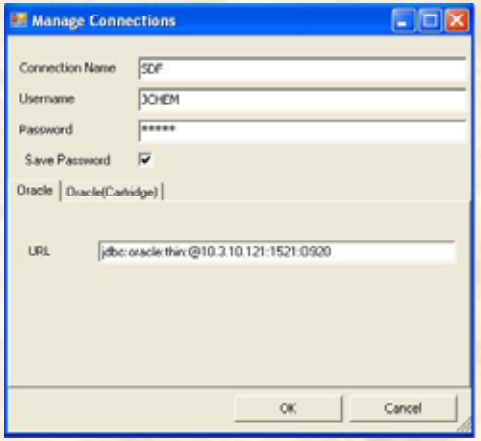
Introduction

- Microsoft Excel integrated solution for using Marvin and JChem functionality
- Use Excel's powerful features: Functions, Sorting, Filtering, Charts…
- Implemented in **C# .NET, and Visual Studio** proof that the ChemAxon APIs can be used in a .NET environment
- Easy to install and deploy

Import from Database –Manage Connections

- JChem Cartridge Databases
  - Structure search runs on the server
- JChem Databases
  - Structure search runs locally
  - Databases supported
    - Oracle
    - MSSQL
    - MySQL
    - …
- Connection settings are saved

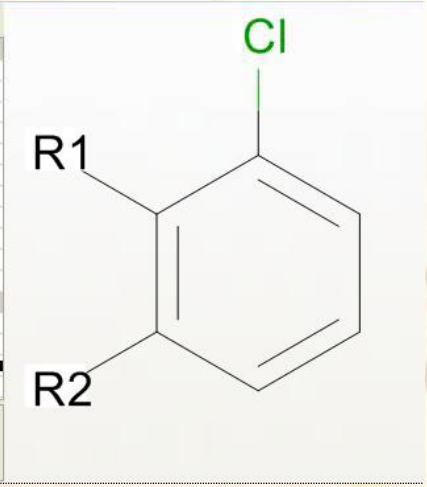




Import from Database - Query

Chemistry Search Options	
Absolute Stereo	Table Option
Charge Matching	Default
Double Bond Stereo M	Marked
Exact Query Atom Mat	False
Exact Stereo Matching	False
Hydrogen Count Matc	Auto
Isotope Matching	Default
Radical Matching	Default
Stereo Search	True
Tautomer Search	False
Vague Bond Level	Level 1
General Search Options	
Maximum Hits	
Maximum Time	
Query Structure	<?xml version="1.0" ?>
Search Type	Substructure

Query Structure

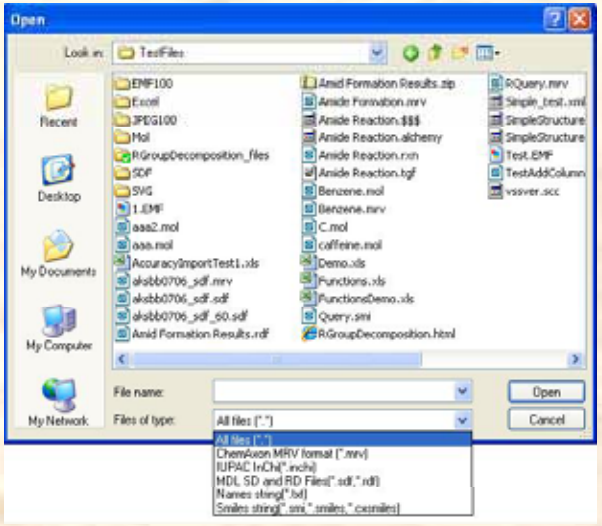


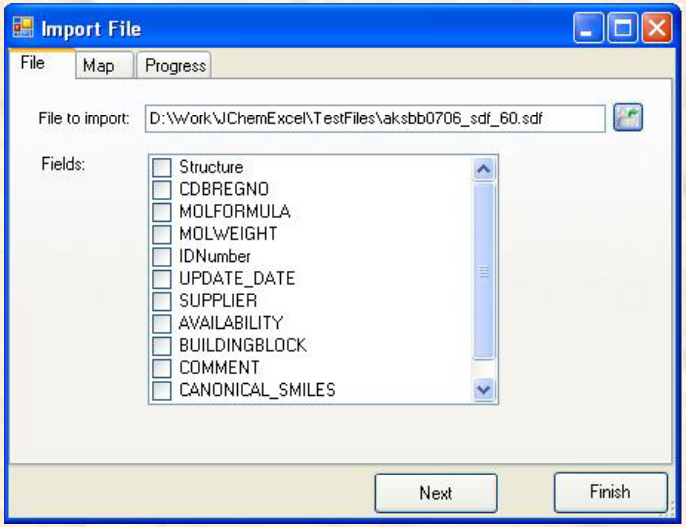
	AND/OR	Column	Operator	Value
	AND	CD_MOLWEIGHT	>=	300
▶	AND	CD_FORMULA	Like	'C10'
*				

- Common interface for specifying JChem or JChem Cartridge queries
- Easy to use control for specifying structure queries
- AND and OR type queries can be specified
- Specify Chemical Terms expressions for queries

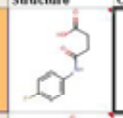
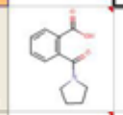
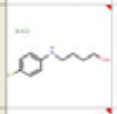
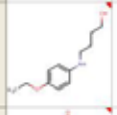
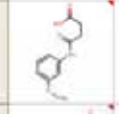
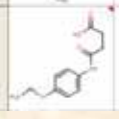
Import and Export files

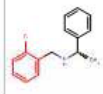
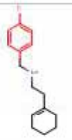
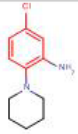
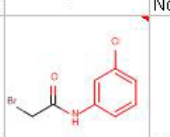
- All file formats supported by Marvin could be exported and imported
  - MRV,SDF,RDF,SMILES, Name,…
- Wizards guide you through the import and export process





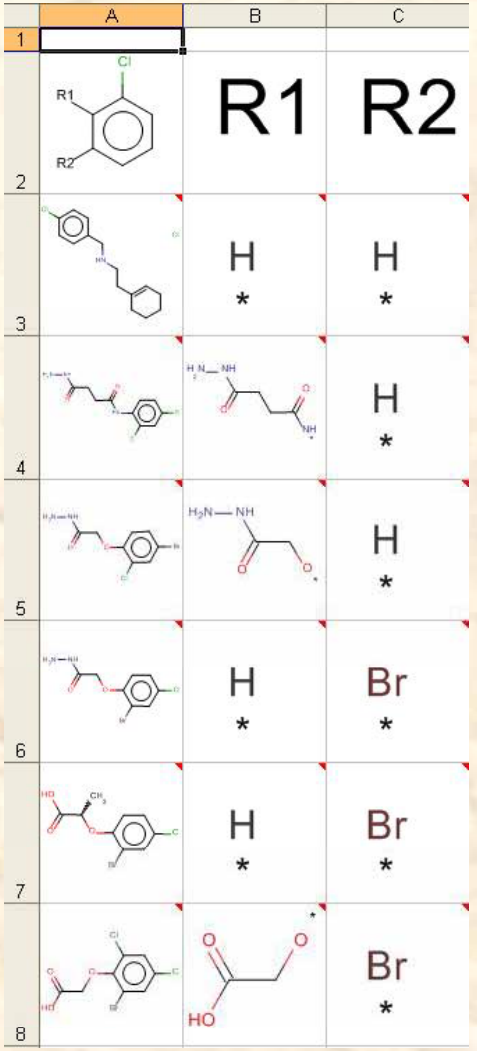
Import - Results

	A	B	C	D	E	F	G
1	Structure	CDBREGNO	MOLFORMULA	MOLWEIGHT	IDNumber	UPDATE_DATE	SUPPLIER
2		1	C10 H10 F N O3	211.19	AKS-BB/0001	12/7/2006	AKS
3		2	C12 H13 N O3	219.24	AKS-BB/0002	12/7/2006	AKS
4		3	C10 H14 F N O . Cl H	219.69	AKS-BB/0003	12/7/2006	AKS
5		4	C12 H19 N O2	209.29	AKS-BB/0005	12/7/2006	AKS
6		5	C11 H13 N O4	223.23	AKS-BB/0007	12/7/2006	AKS
7		6	C12 H15 N O4	237.25	AKS-BB/0008	12/7/2006	AKS

	A	B	C	D
1	CD_ID	CD_STRUCTURE	CD_SMILES	CD_FORMULA
2	46		[Cl-].CC(NCc1ccccc1Cl)c1ccccc1 |w:2.0|	C15H16Cl2N
3	47		[Cl-].Clc1ccc(CNCCC2=CCCCC2)cc1 |t:9|	C15H20Cl2N
4	68		Nc1cc(Cl)ccc1N1CCCCC1	C11H15ClN2
5	88		Clc1cccc(NC(=O)CBr)c1	C8H7BrClNO
6	118		[Cl-].CCNCc1ccc(Cl)cc1	C9H12Cl2N
7	227		Cc1cc(OCC(=O)NN)ccc1Cl	C9H11ClN2O2

- Up: SDF import results
- Right: Database import results with hit coloring

R-group Decomposition

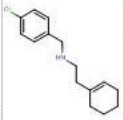
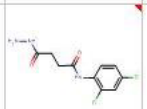


- Run Decomposition on
  - JChem Database
  - Files
  - Excel Worksheet

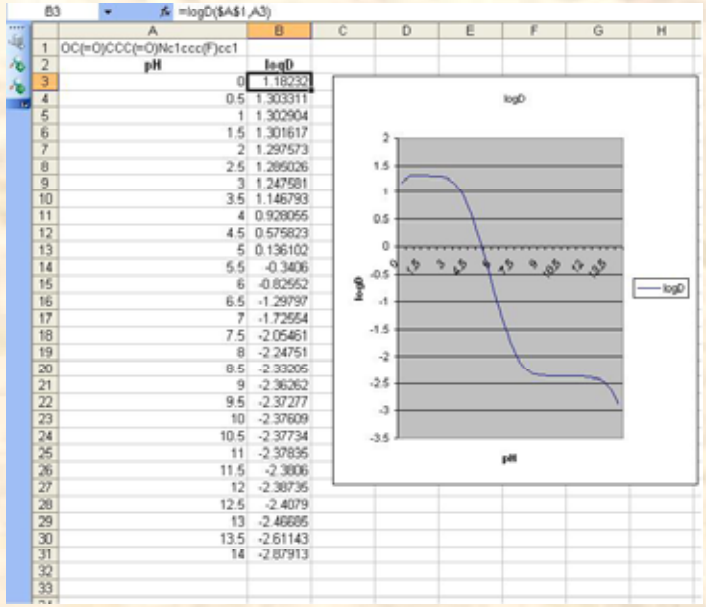
Excel Functions

- Most ChemTerms functions and Marvin calculations are implemented as Excel functions
  - Mass, MolFormula, Composition
  - logP / logD, pKa, H-bond donor / acceptors
  - Lipinski rule of 5, Bioavailability
  - Match, Dissimilarity
  - IUPAC Name
- Values are calculated from SMILES, Structure cell, Marvin OLE Object
- Values automatically updated when structures edited or added

Excel Functions

	A	B	C	D	E
1	CD_ID	CD_STRUCTURE	MolFormula	Composition	IUPAC Name
2	46		C15H16Cl2N	C (64.1%), H (5.7%), Cl (25.2%), N (5%)	[(2-chlorophenyl)methyl](1-phenylethyl)amine chloride
3	47		C15H20Cl2N	C (60.2%), H (7.1%), Cl (24.9%), N (4.9%)	[(4-chlorophenyl)methyl][2-(cyclohex-1-en-1-yl)ethyl]amine chloride
4	268		C10H11Cl2N3O2	C (43.5%), H (4%), Cl (25.7%), N (15.2%), O (11.6%)	N-(2,4-dichlorophenyl)-3-(hydrazinecarbonyl)propanamide

- Up: MolFormula, Composition, IUPAC Name calculated from CD_STRUCTURE
- Right: logD calculated from SMILES in A1 and pHs in A column



Deployment

- MSI Installer
  - No configuration needed
  - 8MB File with JChem and Marvin API included
  - Automatic deployment and updates through Active Directory
- 2. ClickOnce (planned)
  - Update from Excel at startup or when needed

.NET Interop

- Uses IKVM, a .NET implemented Java VM (deployed together, or separately if needed)
- Open Source solution, continuously improving
- JChem and Marvin functionalities are covered by unit tests
- Plan to release a .NET API based on IKVM

Technical Overview

- Implemented in C#, no VBA code.
- Interfaces with Excel as a Managed COM Add-In, and Automation Add-In for Excel functions
- Uses COM Shim
  - Security: no problems with High or Very High
  - Isolation: Separate application domain, does not crash with other add-ins
- Plan to use Visual Studio Tools for Office in the future
- These are the directions Microsoft is pushing for Office development

Future Plans

- Research demand for the product
- Feature ideas are welcome
  - Markush enumeration
  - Reactor
  - Standardizer
  - …?
- Release planned this summer

Thank you for your attention!

For more information please visit

www.chemaxon.com