

QSAR plugin for Bioclipse getting in shape



Published June 27, 2007

Citation

Willighagen, E. (2007, June 27). QSAR plugin for Bioclipse getting in shape. *Chem-bla-ics*. <https://doi.org/10.59350/ht8k2-aed07>

Keywords

Bioclipse, Qsar, Cdk, Ambit

Abstract

Over the last few weeks I continued the work on getting (descriptor-based) QSAR/QSPR implemented in Bioclipse. JOELib (GPL) and the CDK (LGPL) being two prominent opensource engines that can calculate molecular descriptors, and AMBIT a front-end.

Copyright

Copyright © None 2007. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

chem-bla-ics

Over the last few weeks I continued the work on getting (descriptor-based) [QSAR](#)/[QSPR](#) implemented in [Bioclipse](#). [JOELib](#) (GPL) and the [CDK](#) (LGPL) being two prominent opensource engines that can calculate molecular descriptors, and [AMBIT](#) a front-end.

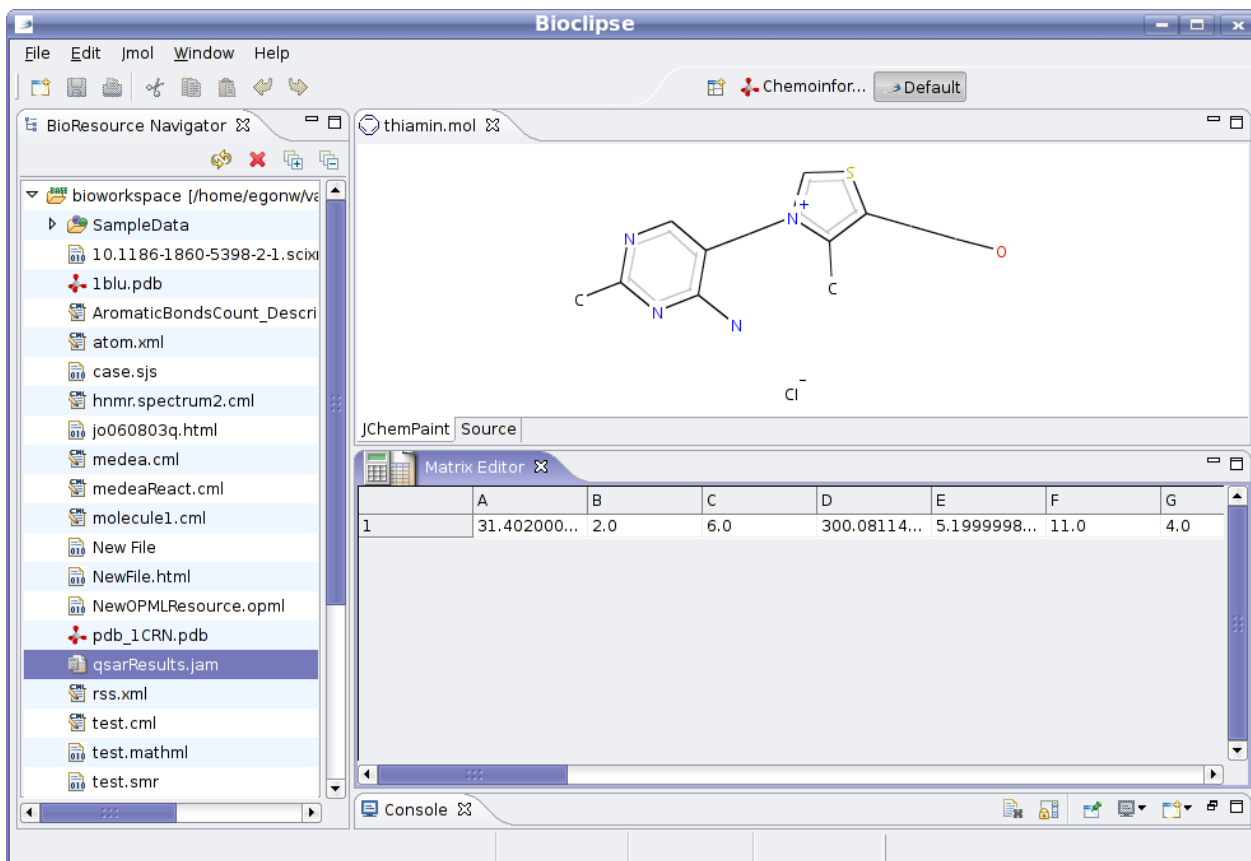
To be able to do QSAR/QSPR model building from start to end in Bioclipse, I worked in April [on an architecture for selecting descriptors](#). Being busy with so many things, it took me some time to get around to completing that, but here are the screenshots:

The screenshot displays the Bioclipse application window. The top menu bar includes File, Edit, Jmol, Window, and Help. Below the menu is a toolbar with various icons. The main workspace is divided into several panes. On the left is the 'BioResource Navigator' showing a file tree under 'bioworkspace [/home/egonw/...]'. The tree includes 'SampleData' and a list of files such as '10.1186-1860-5398-2-1.scix', '1blu.pdb', 'AromaticBondsCount_Descri', 'atom.xml', 'case.sjs', 'hnmr.spectrum2.cml', 'jo060803q.html', 'medea.cml', 'medeaReact.cml', 'molecule1.cml', 'New File', 'NewFile.html', 'NewOPMLResource.opml', 'pdb_1CRN.pdb', 'qsarResults.jam' (highlighted), 'rss.xml', 'test.cml', 'test.mathml', and 'test.smr'. The central pane shows the chemical structure of 'thiamin.mol', which is a thiazolium ring substituted with a pyrimidin-2-yl group and a methyl group, with a chloride ion (Cl⁻) nearby. Below the structure is a 'JChemPaint Source' pane. At the bottom is a 'Matrix Editor' window displaying a table with one row and seven columns labeled A through G.

	A	B	C	D	E	F	G
1	31.402000...	2.0	6.0	300.08114...	5.1999998...	11.0	4.0

The funny characters and the whitespace is gone. Right now, it still only lists one provider, but I plan to add JOELib plugin soon. The list of actual descriptors is provided by the extension.

What Bioclipse then does, is have the extension calculate the descriptor values for the selected **CDKResource** in the BioNavigator using the selected descriptors. This will then create a new **MatrixResource** in the Bioclipse workspace (currently called qsarResult.jam), and which is opened in the Matrix editor:



There is still enough work left to do. For example, the columns are not yet labeled according to the descriptor name, and selecting more than one `CDKResource` in the navigator does not give a multirow matrix yet.