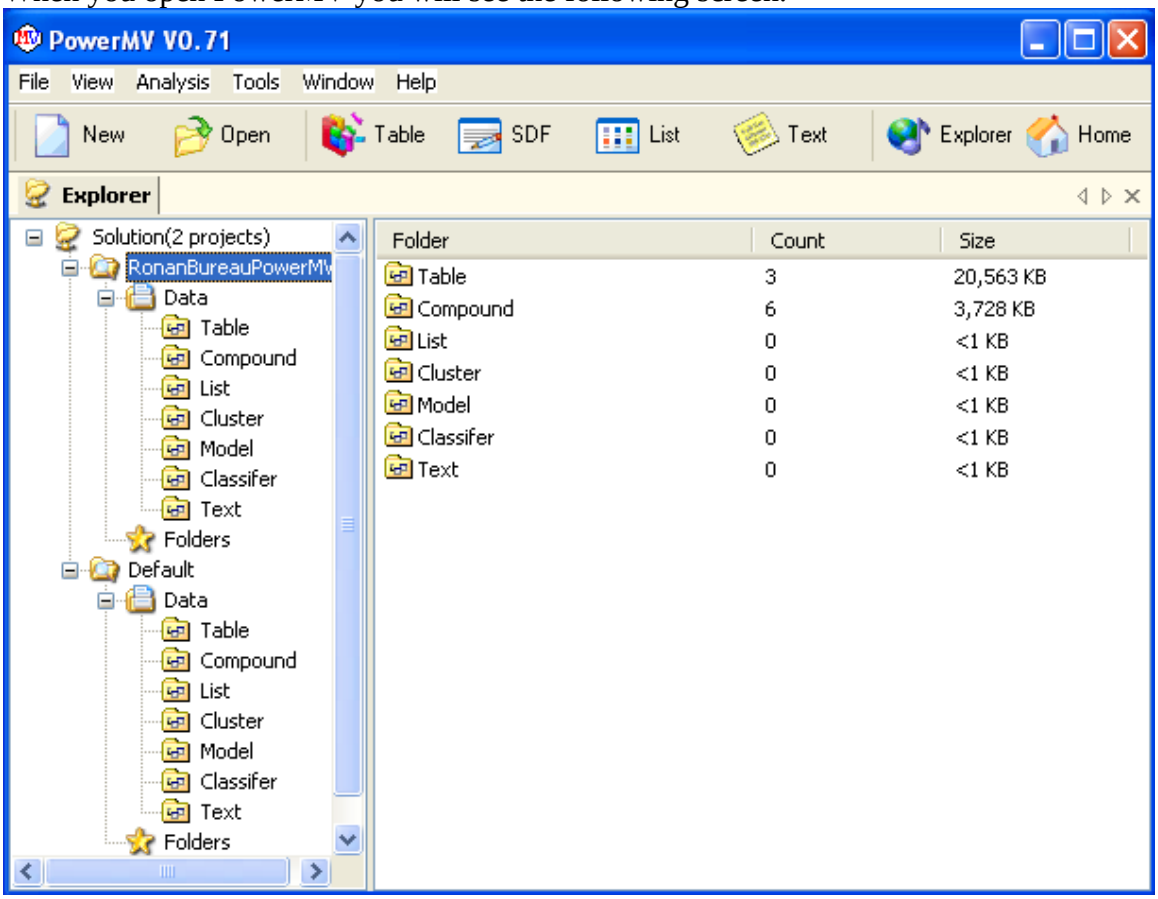
ReadMe PowerMV V02

A free version of PowerMV is available at www.niss.org/PowerMV.

(There is a more fully functional "**Affiliates' Version**" available to Affiliates of the National Institute of Statistical Science. Contact admin@niss.org for information on downloading that program.

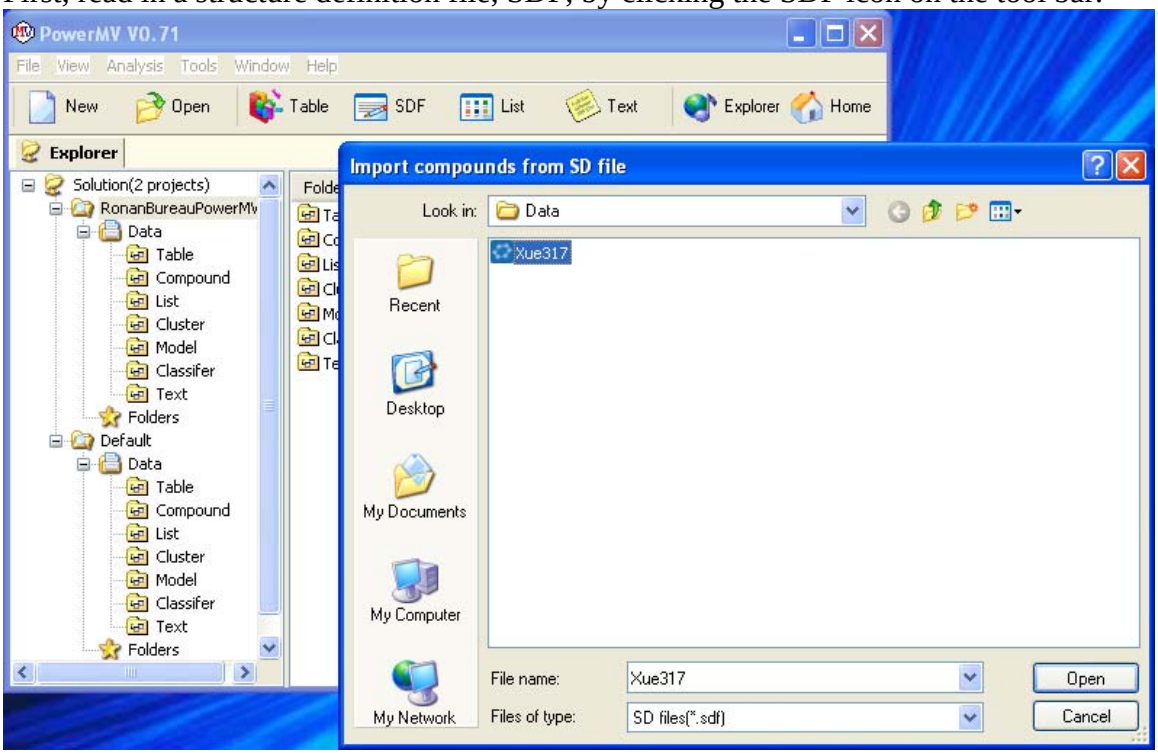
Note that you should install the current versions of Microsoft ".NET" and the statistics software "R" on your computer. Both of these can be downloaded from the NISS web site.

When you open PowerMV you will see the following screen.

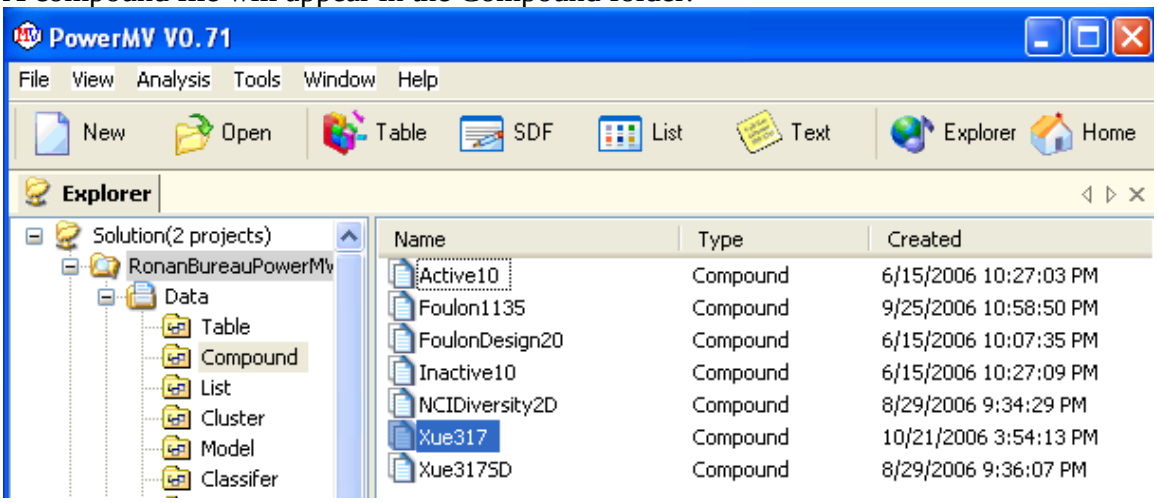


This ReadMe will walk you through several simple tasks that reflect usual chemometrics work.

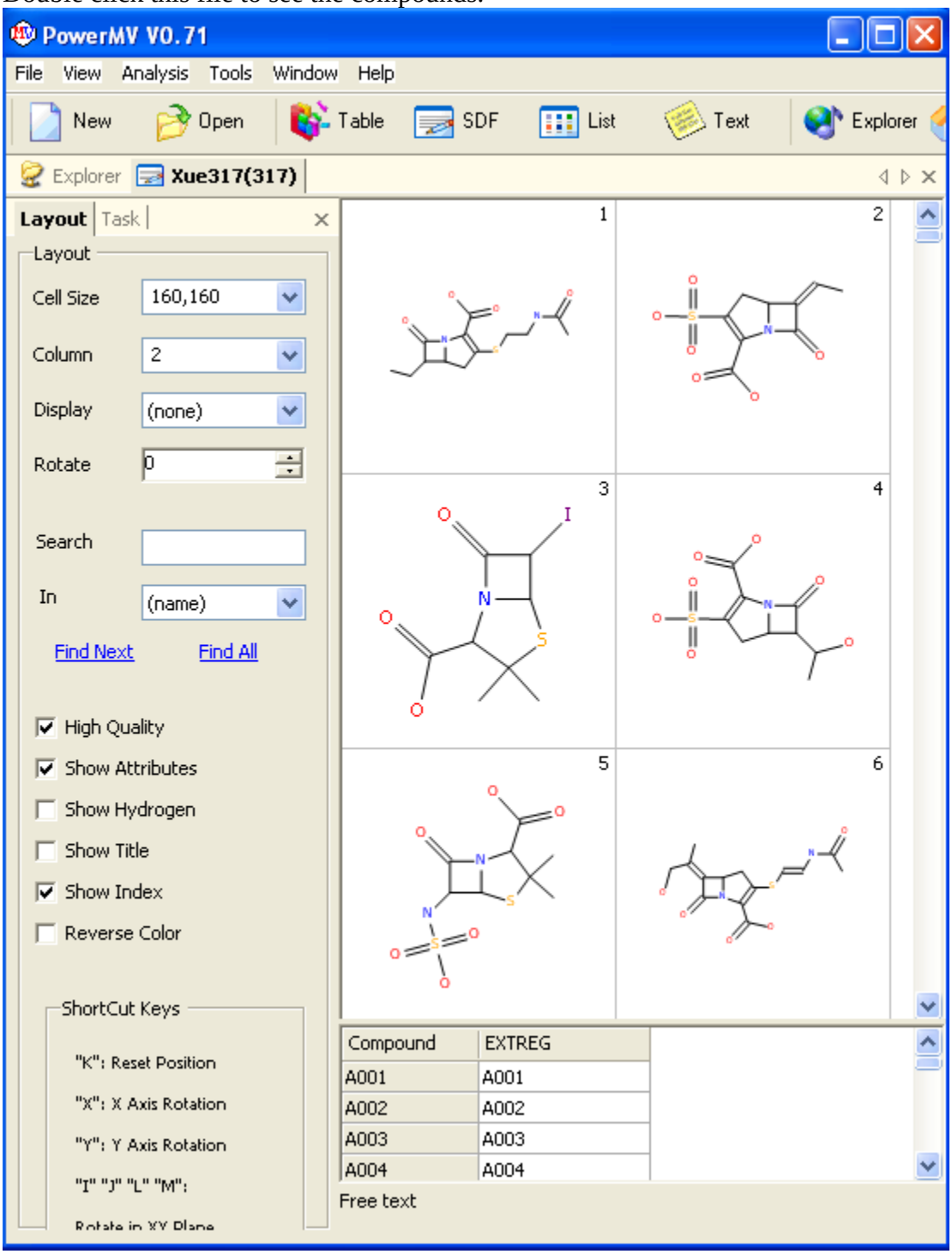
First, read in a structure definition file, SDF, by clicking the SDF icon on the tool bar.



A compound file will appear in the Compound folder.

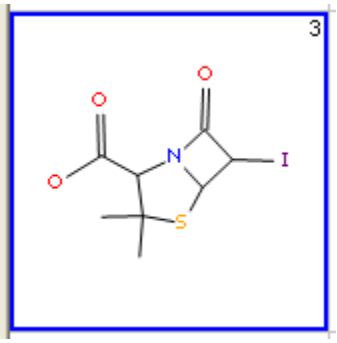


Double click this file to see the compounds.



Check the Show Attributes box to see any annotations to the SD file. Check the Show Index box to number the compounds. Note that you can resize the window and control the number of columns in the display. You can also control the size of the compounds.

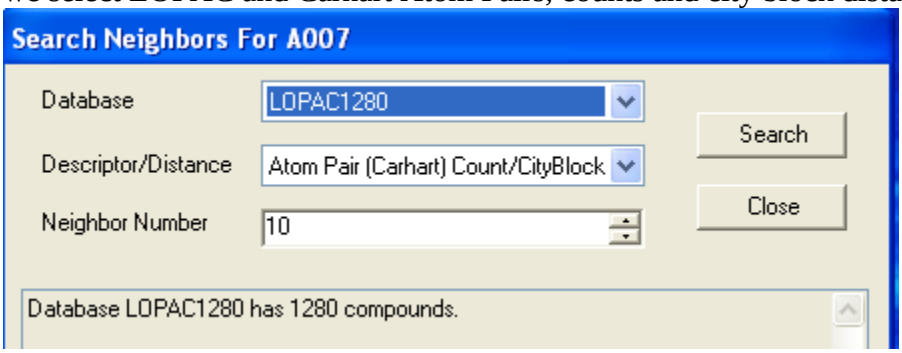
You can also rotate the compounds in a window.



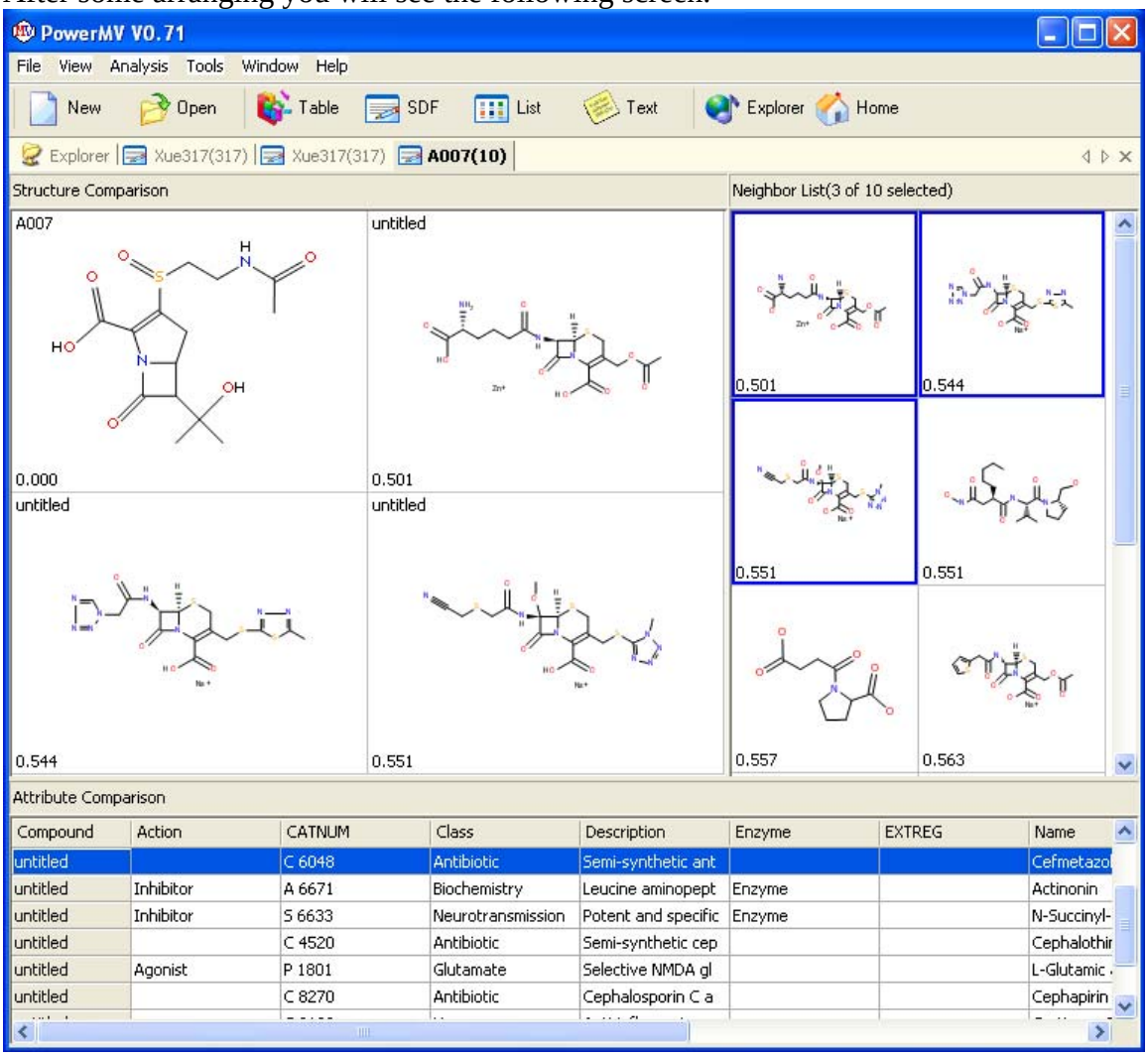
Click the Task tab to see various editing functions.

Note that if you click over a compound, you get a menu that includes the following items:
Search Database
View in 3D Mode
Substructure Search (current compound)
Substructure Search (through JME®)

Some background is useful here. In HTS active compounds are found. Often you will want to compare an active compound to compounds in a curated data set to get some idea how your compound is acting in this screen or potential side effects your compound might have. So if you have a curated data set, you can enter it into PowerMV to serve as a reference data set. Click the Tools toolbar to see a menu and select Generate Similarity Index for Database Searching. We have generated the search keys for the LOPAC data set of 1280 drugs. If you click on a compound to select it (blue border) and then right click, you will get a menu. Select search database. Select the curated data base you want and the molecular descriptors to be used. Here we select LOPAC and Carhart Atom Pairs, counts and city block distance.

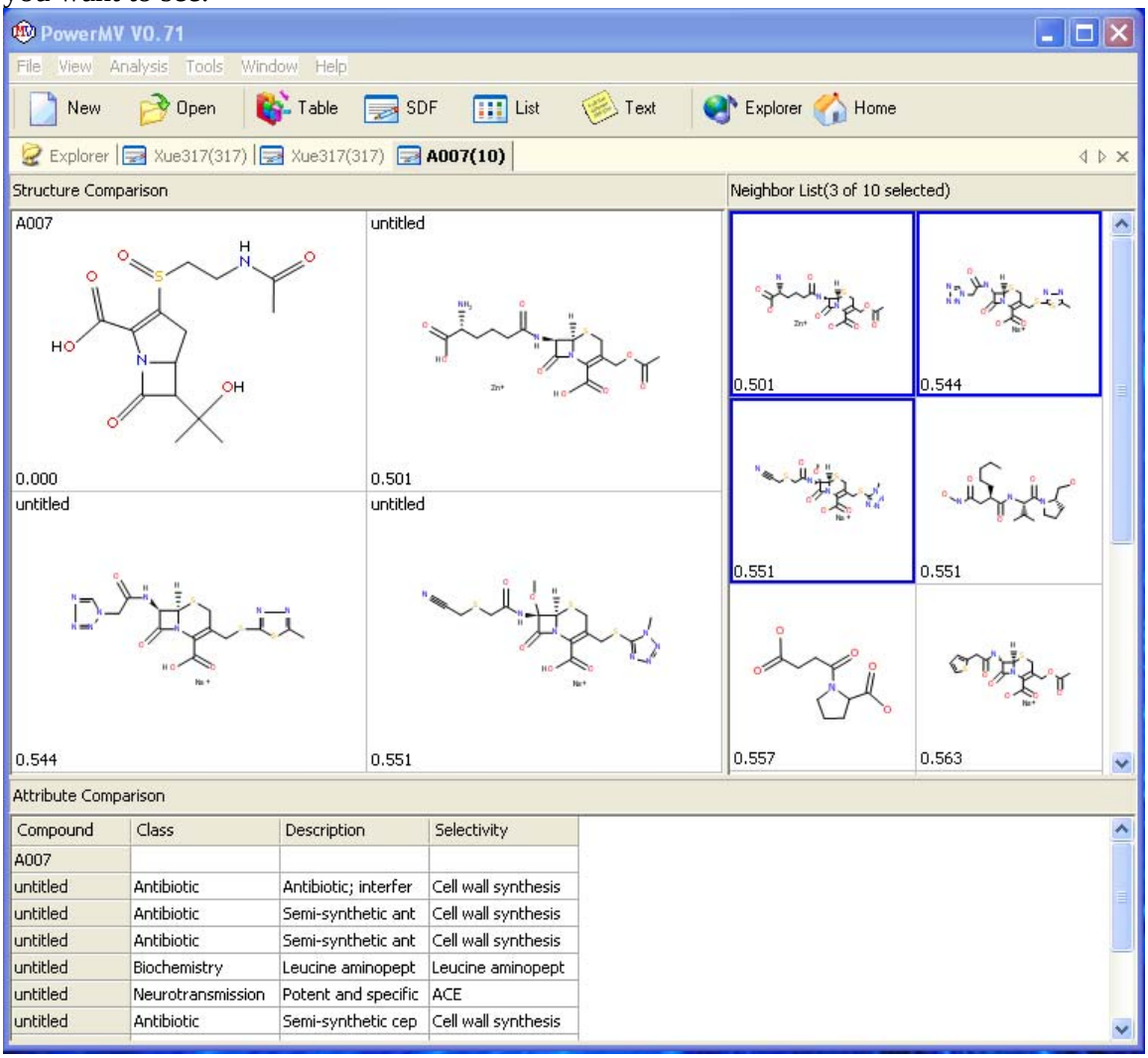


After some arranging you will see the following screen.



The upper left compound is the target compound that came from Xue317. Initially the 10 closest hit from LOPAC are given in two columns on the right side of the screen. If you hold down the Ctrl key and click a compound, it will move into the viewing screen. Here we see the three most similar compounds from the curated data set along with the target compound. The attributes of the near neighbor compounds are given along the bottom of the screen.

If you right click over the data table, you can select Show/Hide to select the compound attributes you want to see.



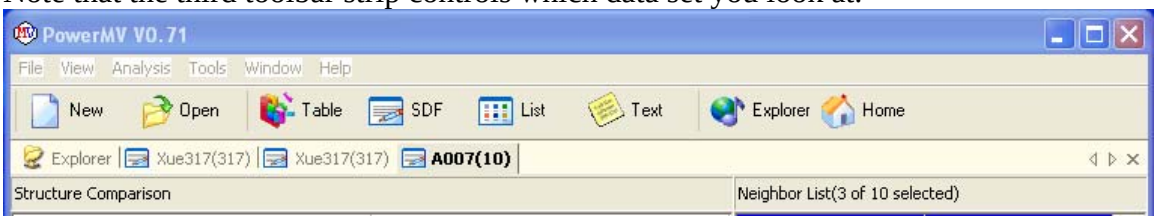
These compounds all have the distinct 4-member ring of antibiotics. Note that the columns can be sized.

Attribute Comparison

Compound	Class	Description	Selectivity
A007			
untitled	Antibiotic	Antibiotic; interferes with cell wall synthesis	Cell wall synthesis
untitled	Antibiotic	Semi-synthetic antibiotic; interferes with cell wall synthesis	Cell wall synthesis
untitled	Antibiotic	Semi-synthetic antibiotic derived from cephamycin C; interferes with cell	Cell wall synthesis
untitled	Biochemistry	Leucine aminopeptidase inhibitor	Leucine aminopept
untitled	Neurotransmission	Potent and specific angiotensin converting enzyme (ACE) inhibitor	ACE
untitled	Antibiotic	Semi-synthetic cephalosporin antibiotic; interferes with cell wall synthesis	Cell wall synthesis

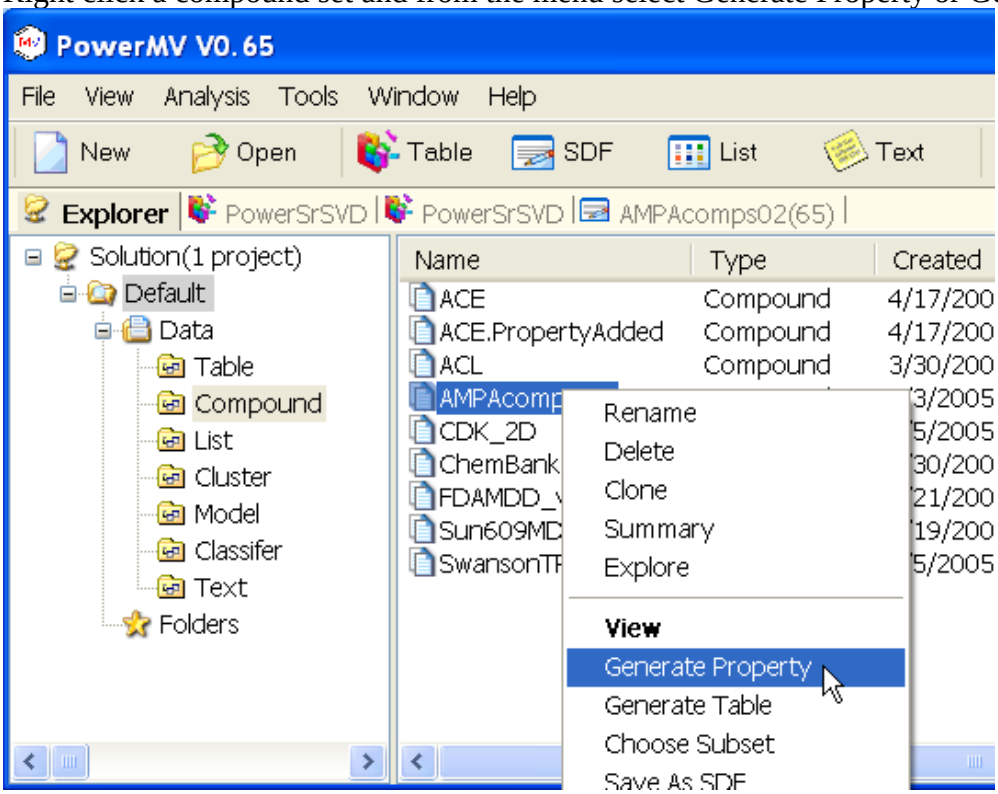
So if A007 were a hit from HTS, similar compounds in this curated data base would indicate the compound is likely to be an antibiotic that interferes with cell wall synthesis.

Note that the third toolbar strip controls which data set you look at.



Right now there are three data sets open, Xue317, Xue317 (I opened it twice! To remove one of the copies from the viewer, click the X on the right end of the tool strip.)

PowerMV can be used to compound mathematical properties of the molecules in a data set. Right click a compound set and from the menu select Generate Property or Generate Table.



If you select Generate Property, then the “Lipenski” properties will be computed and the resulting data set will appear in the Table folder. If you select Generate Table, then a menu will appear and you can select the properties you would like to compute. These properties along with any of the attributes from the SD file can then be exported to an Excel spreadsheet to feed into an external statistical analysis.

Note that useful data analysis functions are also included in PowerMV. It is possible to compute the singular value decomposition of a 2-way data table. Right click over the name of the data table and select the menu item, Analyze by SVD. If you select the Multiple Outlier Analysis, a robust outlier detection method is applied to the data set.

References:

Bradu, D. and Hawkins, D.M., (1982) Location of multiple outliers in two-way tables, using tetrads. Technometrics 24, 103-108.

Liu, K., Feng, J., and Young, S.S. (2005) PowerMV: A Software Environment for Molecular Viewing, Descriptor Generation, Data Analysis and Hit Evaluation *J. Chem. Inf. Model.* 2005, *45*, 515-522.

Liu, L., Hawkins, D.M., Ghosh, S., and Young, S.S. (2003). Robust singular value decomposition analysis of microarray data. *Proc. Natl. Acad. Sci. U. S. A.* 100, 13167-13172.

Xue, L. and Ju¨rgen Bajorath, J. (2002) Accurate partitioning of compounds belonging to diverse activity classes. *J. Chem. Inf. Comput. Sci.* 2002, *42*, 757-764