

AIDS Antiviral Screen Data

The DTP AIDS Antiviral Screen has checked tens of thousands of compounds for evidence of anti-HIV activity. Available are screening results and chemical structural data on compounds that are not covered by a confidentiality agreement.

Screening Results (May 2004 Release)

The results of the screening tests are evaluated and placed in one of three categories:

- **CA** - Confirmed active
- **CM** - Confirmed moderately active
- **CI** - Confirmed inactive

[aids_conc_may04.txt](#) is an ASCII comma delimited file containing the screening results for 43,850 compounds. The following data fields are provided for each compound:

1. NSC number - the NCI's internal ID number
2. Screening Result - one of the categories listed above

To download this file, click the right mouse button and choose "save link as"

Data is available for the concentrations necessary to see a protective effect on the infected cells (EC50). The [aids_ec50_may04.txt](#) file is a file containing seven data fields in comma delimited format for 39365 NSCs.

1. NSC - the NCI internal ID number.
2. $\text{Log}_{10}(\text{HiConc})$ - the Log_{10} of the highest concentration used in the dose response. Experiments with the same HiConc are averaged.
3. ConcUnit
 - M = molar
 - u = micrograms per ml.
 - V = Volumetric
4. Flag - when different experiments are averaged, if the EC50 value would be less than the lowest tested concentration, the the lowest tested concentration is used for the value of EC50
 - A > in this field indicates EC50 would be higher than the highest concentration tested in at least one of the experiments.
 - A < in this field indicates EC50 would be less than the lowest concentration tested in at least one of the experiments.
 - A = indicated that all the experiments reached EC50.
5. $\text{Log}_{10}\text{EC50}$ - Log_{10} of the concentration of compound that gives 50% protection of infected cells.
6. NumExp - number of experiments included in the average.
7. StdDev - Standard Deviation of the average $\text{Log}_{10}\text{EC50}$ across multiple experiments

Data is also available for the concentrations necessary to inhibit the growth of uninfected cells (IC50). The [aids_ic50_may04.txt](#) file is a file containing six data fields in comma delimited format for 39350 NSCs.

1. NSC - the NCI internal ID number.
2. HiConc - the highest concentration used in the dose response. Experiments with the same HiConc are averaged.
3. ConcUnit:
 - M = molar
 - u = micrograms per ml.
4. Flag - when different experiments are averaged, the highest tested concentration is used for the value of IC50 in experiments that don't actually reach an EC50.
 - A > in this field indicates that this happened in at least one of the experiments.
 - A = indicated that all the experiments reached IC50.
5. IC50 - concentration of compound that inhibits the growth of uninfected cells by 50%.
6. NumExp - number of experiments included in the average.
7. StdDev - Standard Deviation of the average $\text{Log}_{10}\text{EC50}$ across multiple experiments

NOTE: the conclusion data was determined by inspection of individual dose response curves and is an overall judgment by trained personnel. The EC50 and IC50 data are computer generated averages and don't necessarily capture everything that was considered when making the judgement.

Chemical Structural Data - AIDS Screened (October 99 release)

The 2D structure (connection table) for each of the 42,390 compounds was retrieved from the DTP's Drug Information System. Conversion to a 3D structure was accomplished using the program [Corina](#), created by Prof. Gasteiger's group.

Note: No stereochemical information was used in the build, nor was any geometry optimization performed. Please take these facts into account when deciding if these 3D coordinates are suitable for your particular use.

The data are presented in a compressed ASCII file in MDL's SDF file format. There are 2 data fields included in the SD file:

1. NSC - the NCI's internal ID number.
2. CAS_RN - the CAS Registry Number when known or 999-99-9 meaning the actual number is unknown.

Caution: The structural data file is almost 29MB compressed and about 140MB uncompressed. The file is compressed with the standard Unix compress program which is widely available for all platforms. The standard Netscape setting tries to uncompress the file on the fly and display it when the file has the usual .Z type, so I given it a file type .BIN to make sure Netscape asks you to write it directly to disk. You will probably have to rename the file from aids099sd.bin to aids099sd.Z to uncompress after which you should have an ASCII file in MDL SD file format.

