Drug-logS

The program to calculate aqueous solubility (logS)

Edited by Tingjun Hou and Wei Zhang

Drug-logS User Guide

What is Drug-logS

Drug-logS is a program developed by Tingjun Hou, Wei Zhang and coworkers to calculate aqueous solubility based on atomic addition.
In drug-logS, the aqueous solubility of an organic molecule can be obtained by summing the contribution of each atomic component and correction factors.

$$\log S = C_0 + \sum_i b_i s_i + \sum_i c_i B_i \qquad (1)$$

where C_0 is a constant; b_i and c_i are regression coefficients; s_i is the number of occurrence of the *i*th atom type in a molecule, and B_j is the number of occurrence of the *j*th correction factor.

Here we defined 64 basic atom types for the elements commonly found in organic molecules (H, C, O, N, P, S and halogens). The classification scheme differentiates atoms according to (i) element, (ii) hybridization state and (iii) nature of the neighboring atoms.
Here we used two correction factors: hydrophobic carbon (HYD) and the square of molecular weight (MW2). We defined sp3- or sp2- hybridized carbon without any attached heteroatom (any atom other than carbon) within five steps as 'hydrophobic carbon'. It should be noted that sp2-hybridized aromatic carbons were not considered as hydrophobic carbons. Moreover, the sp2- hybridized carbon in ring was also not considered as hydrophobic carbons, because the sp2- hybridized carbon in ring was relatively rigid and not easy to adjust conformation to form aggregation. For more detailed description of drug-logS can be found in the below reference:

[1]. Tingjun Hou, Ke Xia, Wei Zhang, Xiaojie Xu, ADME evaluation in drug discovery. 4. Prediction of aqueous solubility based on atom contribution approach. *J. Chem. Inf. Comput. Sci.* 2004, 44, 266-275.
[2]. Wei Zhang, Tingjun Hou, Xuebin Qiao, Xiaojie Xu, Some basic structures and algorithms for chemical generic programming. *J. Chem. Inf. Comput. Sci.* 2004, 44, 1571-1575.

Notes: The atom typing rules in drug-logS is different from those used in paper [1]. In the new version, the number of atom types is only 64. Moreover, the definition of the hydrophobic carbon is changed a little. In our paper, we defined the hydrophobic carbon as the sp3- or sp2- hybridized carbon without any attached heteroatom (any atom other than carbon) within four steps; while in the new version, we defined the hydrophobic carbon as the sp3- or sp2- hybridized carbon without any attached heteroatom (any atom other than carbon) within five steps.

Executive Platform

Linux Operation System

How to obtain the drug-logS program

Drug-logS is freely distributed to the public. The program can be downloaded from the supporting website (*http://modem.ucsd.edu/adme*) as a compressed file.

After uncompressing the program, a directory named as "drug-logS" will be created. Under the directory, you will find several subdirs: freelib, libltdl, mortdoc, mortsrc, and admepre.

mortsrc: There are MORT supporting library files developed by us. This library is used to recognize the SMARTS definition, define atom types, input and output files. Now the MORT library was not released. It will be formally released in AMBER10 next year. MORT is also the basis for gleap, a new module in AMBER10. When AMBER10 is release, the complete supporting documents will be released.
mortdoc: The simple supporting documents for mortsrc.
admepre: The solubility prediction model is located in this directory.

The Installation of drug-logS

To compile the drug-logS program, do the following:

1. Decompress the HlogS.tar.Z file using the unzip and tar command:
   unzip HLOGS.tar.gz
   tar xvf HLOGS.tar

   The program was protected by password. So when uncompressing the program using the "unzip" command, a password is necessary. Please fill the license file, and send it back by email. Then, you will get the password for uncompressing.

2. Compile MORT and the supporting libraries. In the HlogS directory, use the following commands:
   ./configure
   ./make

3. Compile the executive files for drug-logS.
   cd admepre/hlogs
   ./make

The usage of drug-logS

The drug-logS can input MDL/mol (sdf) file or Tripos/mol2 file. It usage is very simple, and you can use the following command:

hlogs ***.sdf or hlogs ***.mol2

If you want to generate the detailed information of solubility for all molecules, you can use the following command:

hlogs ***.sdf output or hlogs ***.mol2 output

Note: Before you use the drug-logS program, please add the hydrogen atoms for all molecules. If not, the calculated molecular weights will be not correct.

In the hlogs directory, there are two parameter files: typing.txt and logS.prm. The typing.txt file contains the SMARTS definitions and the logS.prm contains the paramters for all atom types. The explanation of SMARTS can be found in website: *http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html*

Our correspondents are:

Dr. Tingjun Hou
Dept. of Chemistry and Biochemistry, RM 4254
9500 Gilman Drive,
University of California, San Diego
La Jolla, CA 92093-0359, USA
E-mail: *tingjunhou@hotmail.com*

Dr. Wei Zhang
Department of Molecular Biology,
The Scripps Research Institute,
La Jolla, CA 92037, USA
E-mail: *zweig@scripps.edu*