

Manual of the vSDC program

The following syntax must be used to launch the vSDC program:

vSDC *input_table.dat* *number*

Input_table

The first argument *input_table.dat* is the file containing the data stored in a table. This table should be formatted as follows:

- The lines beginning by the character "#" are comment lines (useful to name the columns).
- The columns must be separated by space characters.
- The first column must contain the compound names (limited to 50 characters).
- The next columns contain the scores from different virtual screening programs.
- vSDC accepts up to 12 scores and 1 million compounds.
- If a score is missing for a given compound, the data must be replaced by a string where the first character is 'N' (case sensitive).

Many scoring functions provide a negative score approximating the binding affinity of the molecule. Thus, the lowest scores are attributed to the compounds with the most promising affinity. For this reason, the vSDC program considers that compounds with the lowest scores are the top-ranked molecules. But there are some programs in which the scores are positive, and the highest scores correspond to the top-ranked compounds. In this case, the minus sign should be added to the score in order to reverse them.

Example of table with positive Gold scores:

#1-LIGNAME	2-Glide	3-Surflex	4-Flexx	5-Gold
CHEMBL190	-8.41	-4.54	-22.74	33.72
CHEMBL8865	-7.25	NA	-26.46	61.90
CHEMBL16509	-9.08	-8.33	-30.59	56.67
...				

Example of a correctly formatted table:

#1-LIGNAME	2-Glide	3-Surflex	4-Flexx	5-Gold
CHEMBL190	-8.41	-4.54	-22.74	-33.72
CHEMBL8865	-7.25	NA	-26.46	-61.90
CHEMBL16509	-9.08	-8.33	-30.59	-56.67
...				

Number

Number is the number of consensus molecules required by the user.

Compilation

It is possible to re-compile the source code (written in C language):

```
gcc -lm vsdc.c -o vSDC
```