B. Procedures of GEMDOCK

a. Overviews of GEMDOCK

GEMDOCK has three main folder:

- **Drug:** ligand files in the MDL mol and SYBYL mol2 format
- **WCavPDB:** files of binding cavities in the PDB format
- **Fit:** executing docking and output files
  - **PrePDB:** files of docked poses

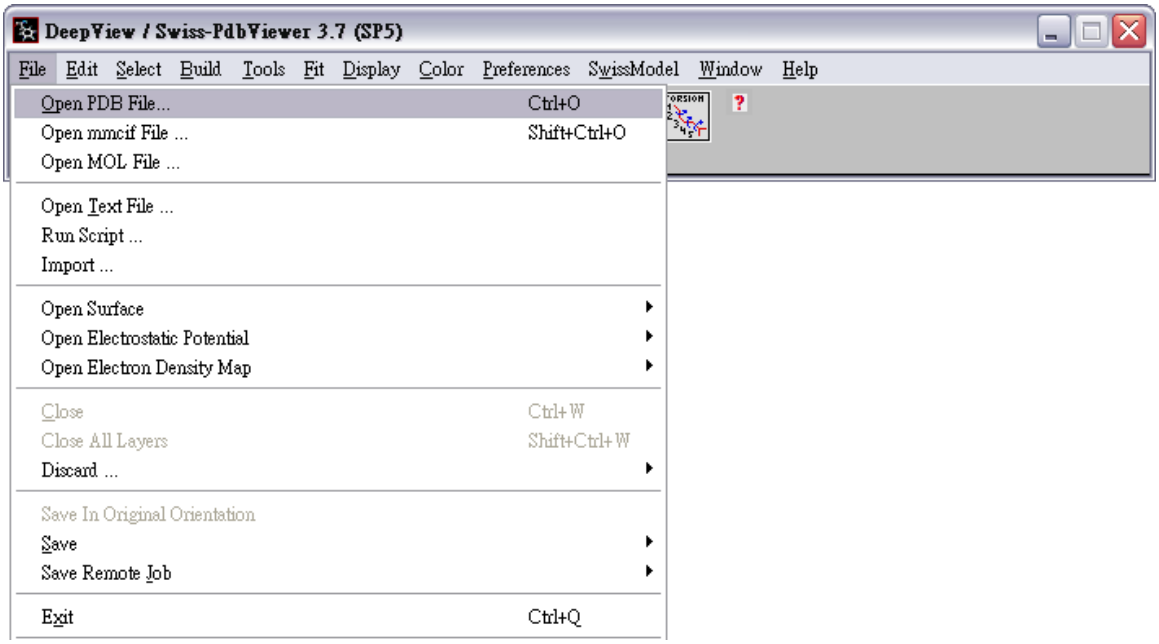
b. Setting up ligands

- Acceptable file formats: MDL mol and SYBYL mol2
- Don't need to add hydrogen atoms and atom charges on docked ligands. The program will automatically assign them when docking.
- After finishing the ligand files, please put them in the folder "Drug" in the mol and mol2 format.

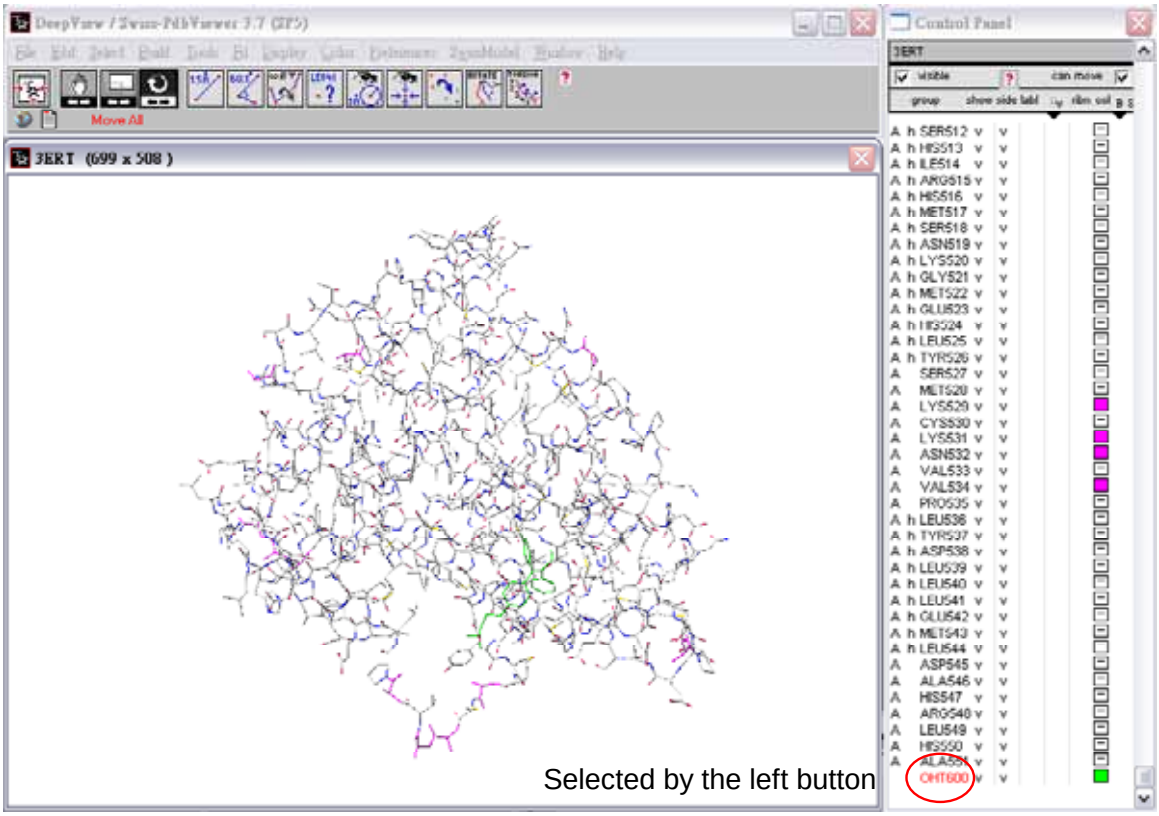
c. Setting up the protein

- Preparing the binding site of the protein.
- Acceptable protein file format: PDB
- After finishing the protein file, put it in the folder "WCavPDB".
- Tools for preparing the file of the binding site
  - Swiss PDB Viewer (http://au.expasy.org/spdbv/text/main.htm)
  - Any program for text edit, e.g. UltraEdit

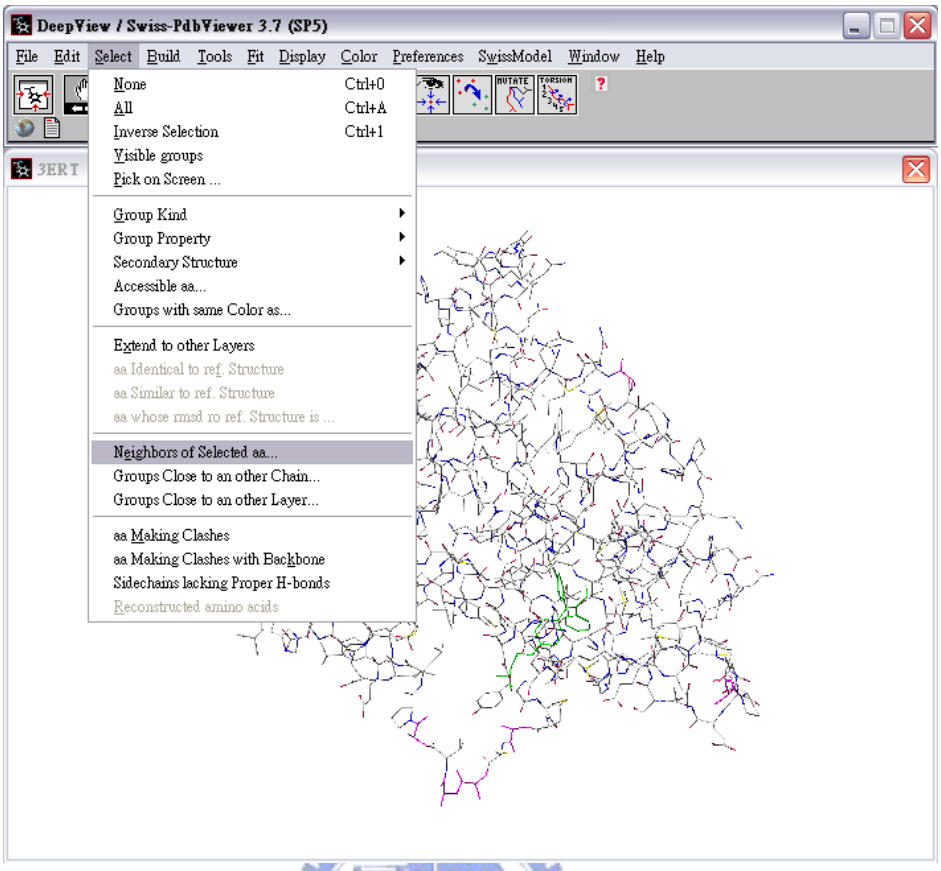
- Procedures:
  - The target protein complex with its ligand, e.g. 3ert complex with its ligand, OHT



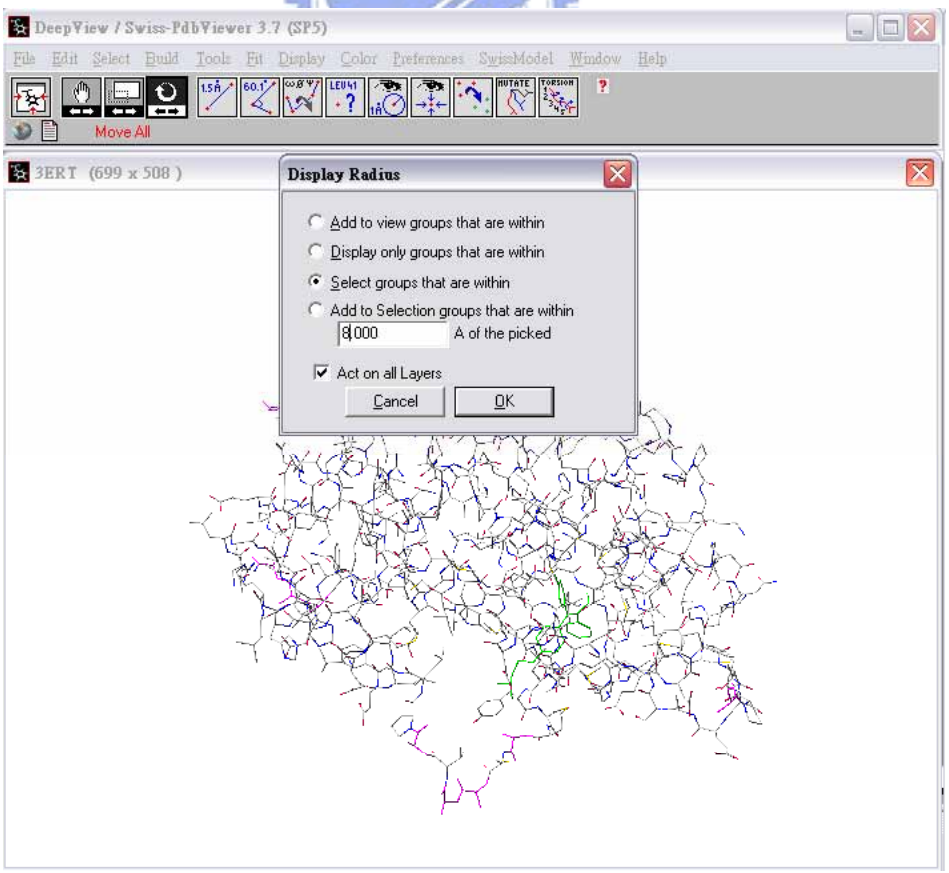
Step 1. Open the PDB file "3ert".



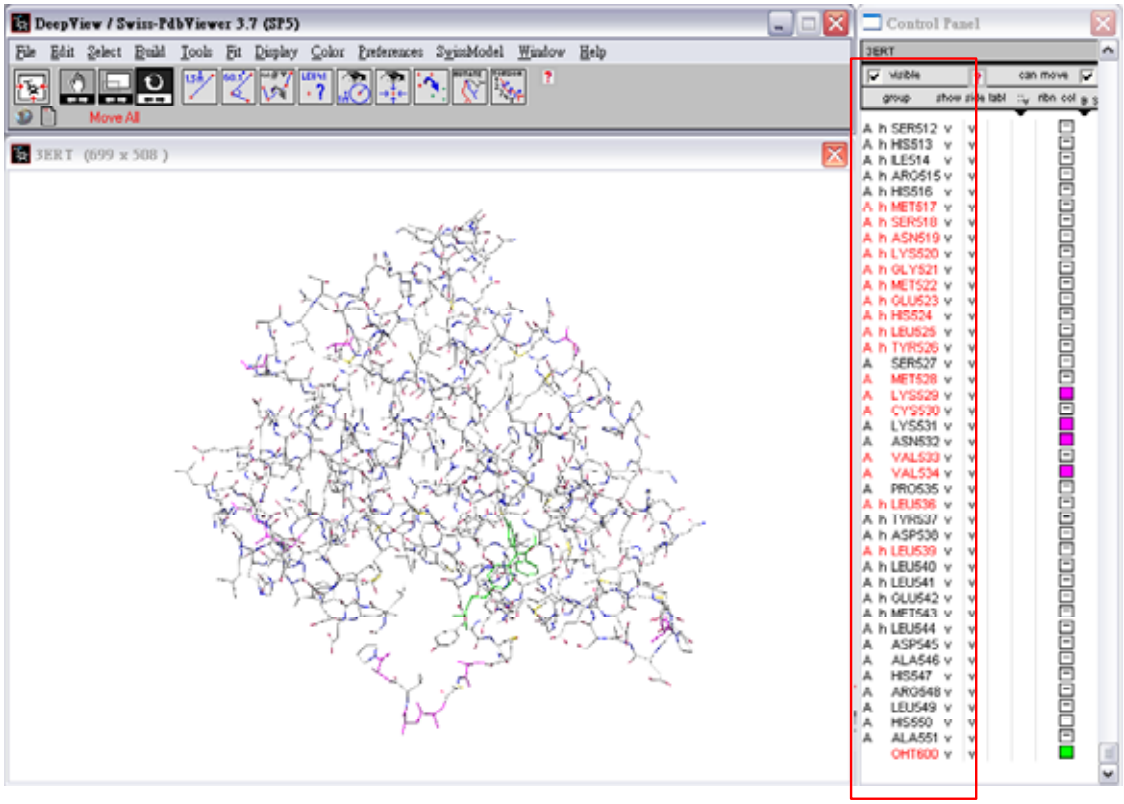
Step 2. Select the ligand "OHT" in the control panel.



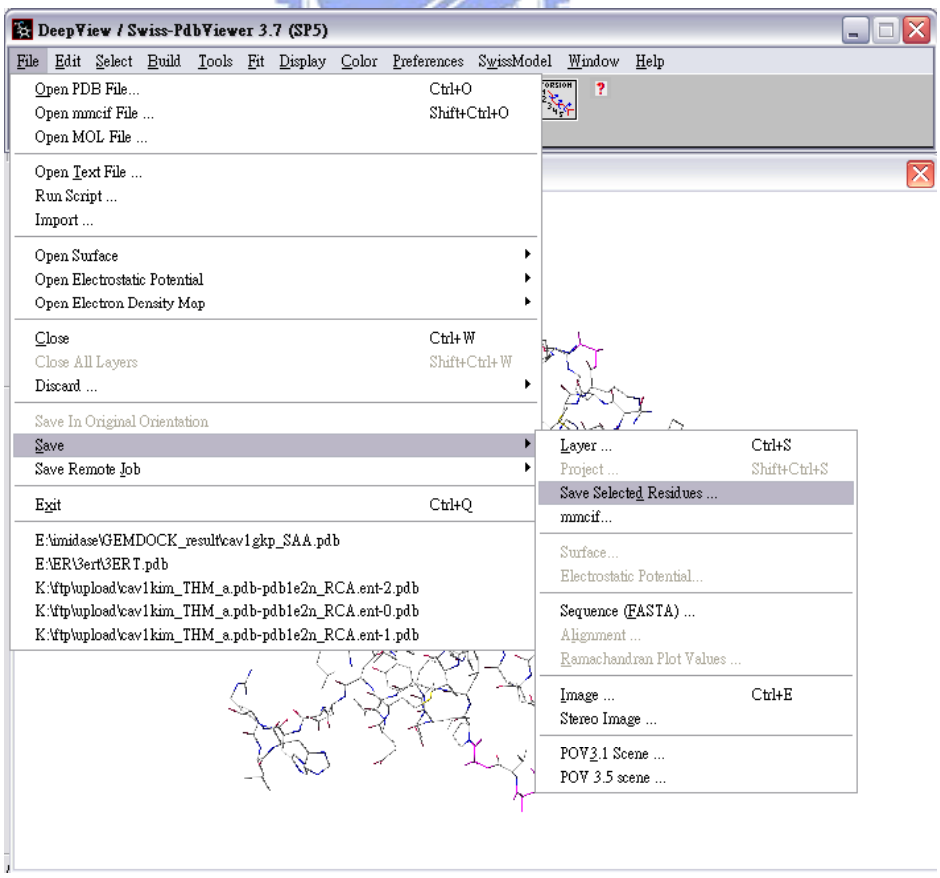
Step 3. Select neighbors of the selected ligand “OHT”.



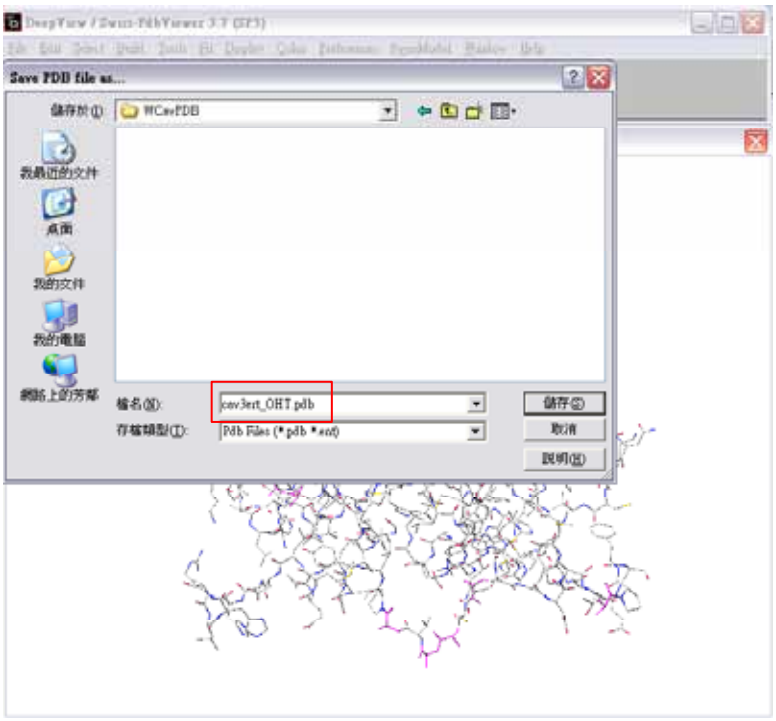
Step 4. Select neighbors of the selected ligand “OHT” within 8 Å.



Step 5. Neighbors of the selected ligand "OHT" within 8 Å were marked.



Step 6. Save selected residues in the cavity file.



Step 7. Input the name of the cavity file (cav3ert_OHT.pdb).

```
ATOM    435  N   PHE   179      15.837  16.915  21.325  1.00 18.62
ATOM    436  CA  PHE   179      16.922  17.352  22.203  1.00 19.17
ATOM    437  C   PHE   179      18.216  17.665  21.440  1.00 19.78
ATOM    438  O   PHE   179      18.545  16.796  20.646  1.00 20.51
ATOM    439  CB  PHE   179      17.208  16.345  23.286  1.00 17.15
ATOM    440  CG  PHE   179      16.130  15.913  24.181  1.00 16.44
ATOM    441  CD1 PHE   179      15.471  16.795  25.045  1.00 16.34
ATOM    442  CD2 PHE   179      15.667  14.577  24.082  1.00 16.52
ATOM    443  CE1 PHE   179      14.465  16.332  25.865  1.00 17.28
ATOM    444  CE2 PHE   179      14.660  14.152  24.894  1.00 16.72
ATOM    445  CZ  PHE   179      13.997  14.983  25.779  1.00 17.20
TER
HETATM  478 NC2* NAP   187       9.686  30.815  33.614  1.00 24.26
HETATM  479 NO2* NAP   187       8.616  29.934  33.924  1.00 25.50
HETATM  480 NC1* NAP   187      10.180  30.602  32.218  1.00 25.55
HETATM  481 NN1  NAP   187      11.465  29.747  32.009  1.00 22.49
HETATM  482 NC2  NAP   187      11.425  28.737  31.046  1.00 22.79
HETATM  483 NC3  NAP   187      12.612  28.091  30.815  1.00 21.00
HETATM  484 NC7  NAP   187      12.537  27.022  29.750  1.00 20.55
HETATM  485 NO7  NAP   187      13.548  26.448  29.502  1.00 17.75
HETATM  486 NN7  NAP   187      11.406  26.850  29.152  1.00 19.09
HETATM  487 NC4  NAP   187      13.909  28.338  31.585  1.00 21.30
HETATM  488 NC5  NAP   187      13.758  29.487  32.531  1.00 20.60
HETATM  489 NC6  NAP   187      12.588  30.085  32.749  1.00 22.66
```

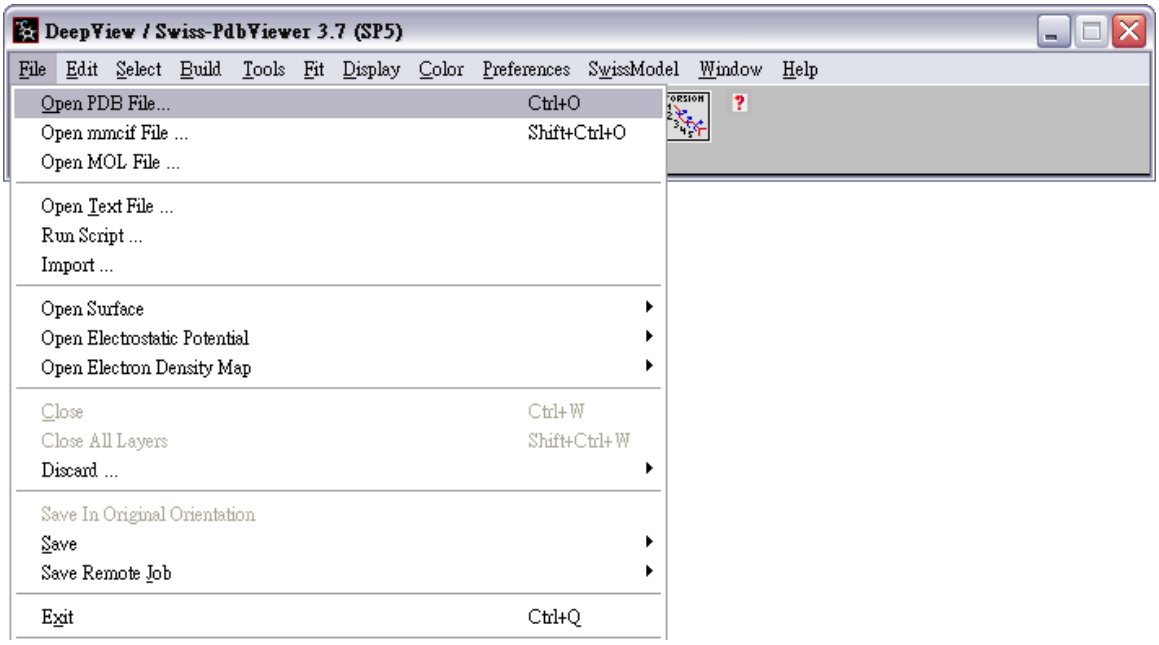
Before “TER”: protein

P
P
P
P
P
P
P
P
P
P
P
P

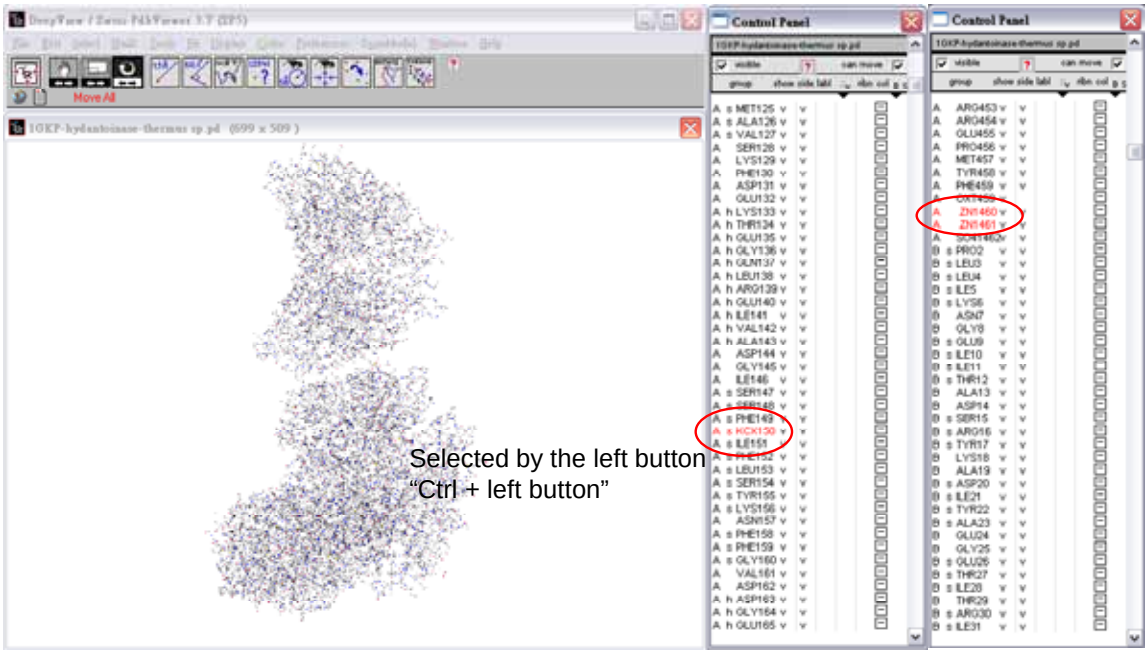
After “TER”: heterogen atom

Step 8. In the cavity file, remove other titles except “ATOM” and “HETATM”. Before the title “TER” are standard protein. After the title “TER”, only heterogen atoms with “P” in the 82th column are treated as the binding site.

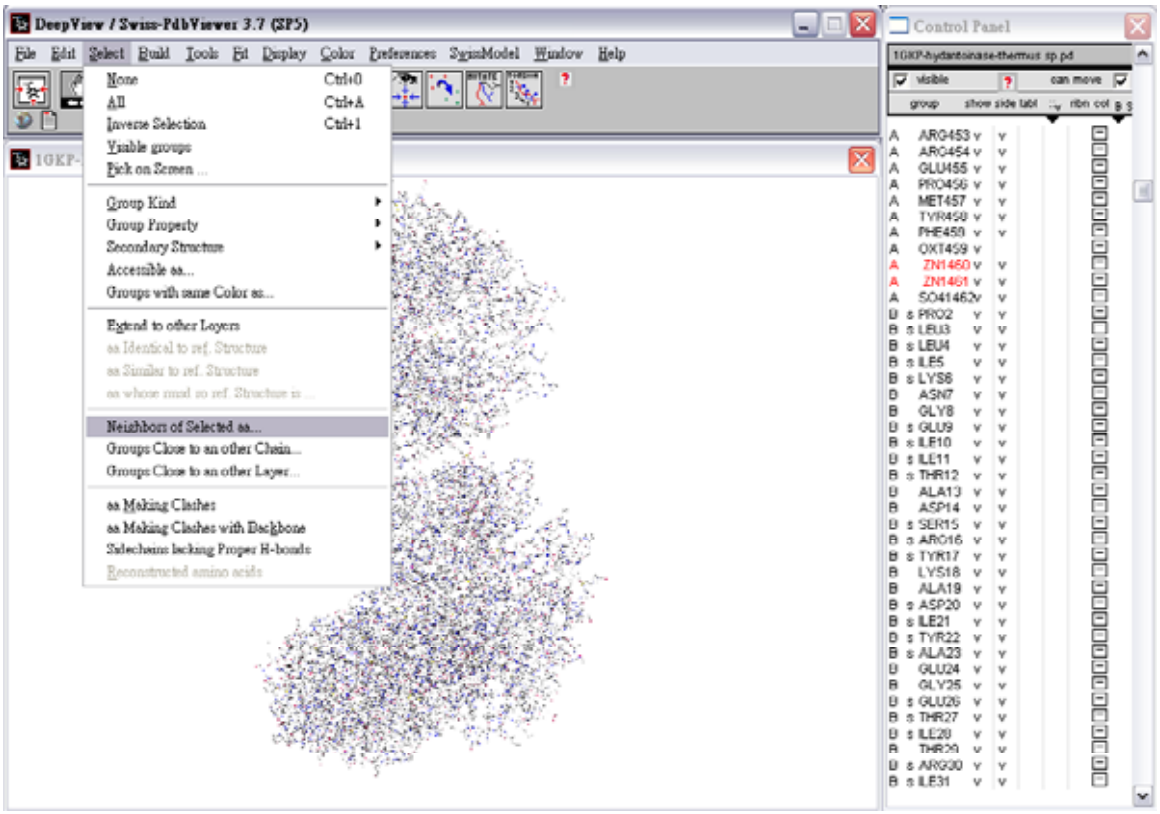
- The target protein complex with no ligand, e.g. 1gkp



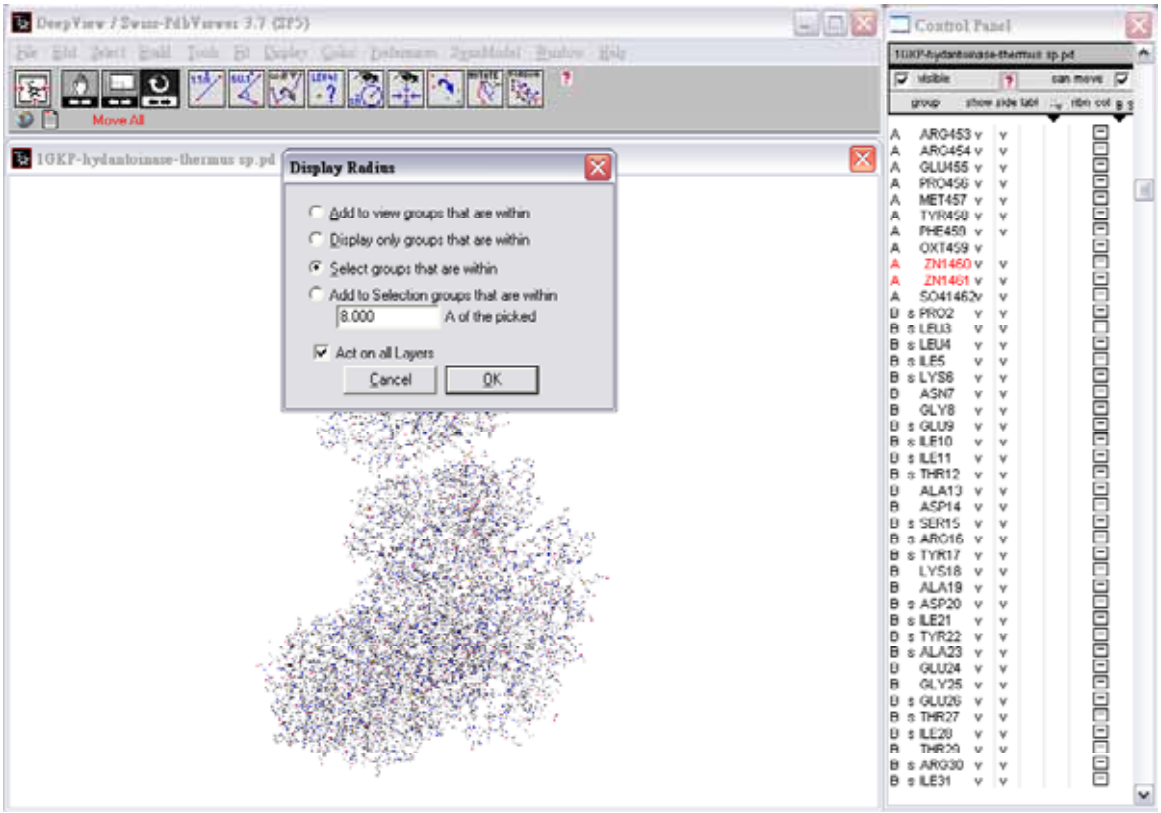
Step 1. Open the PDB file "1gkp"



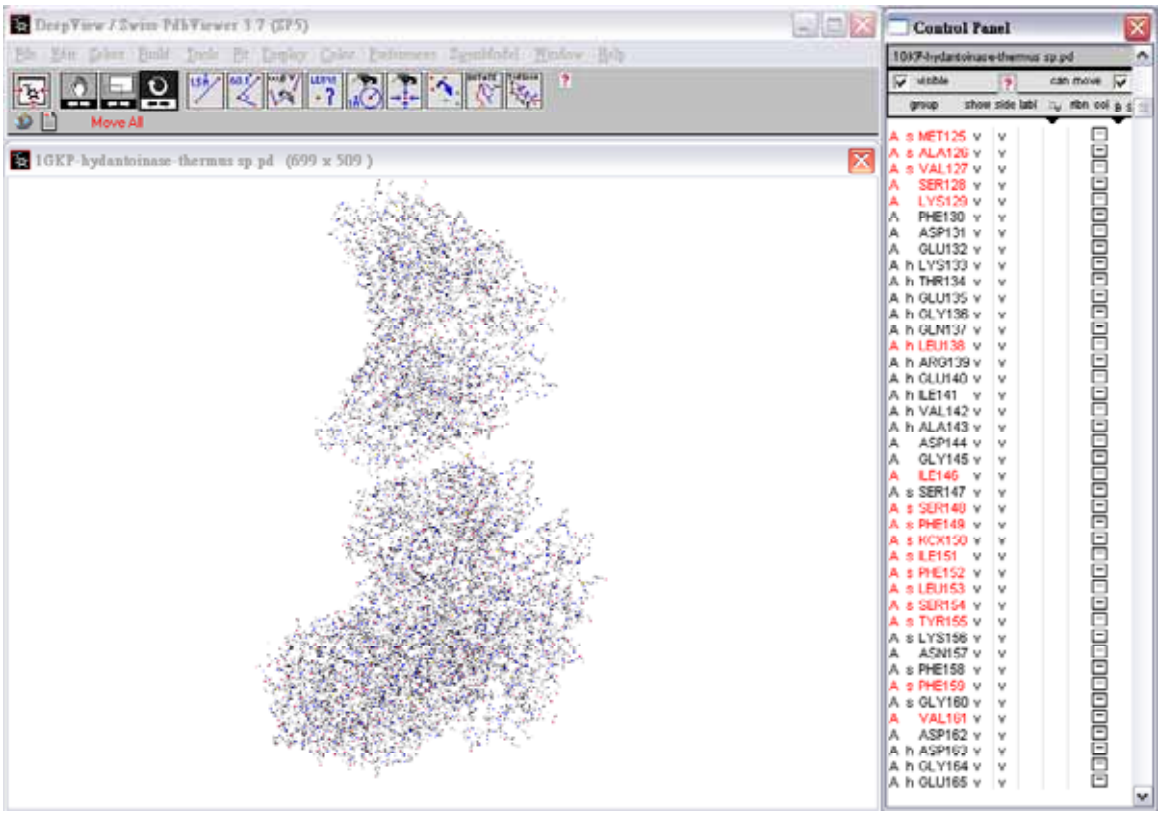
Step 2 . Select the residue KCX150 and two zinc ions in the control panel.



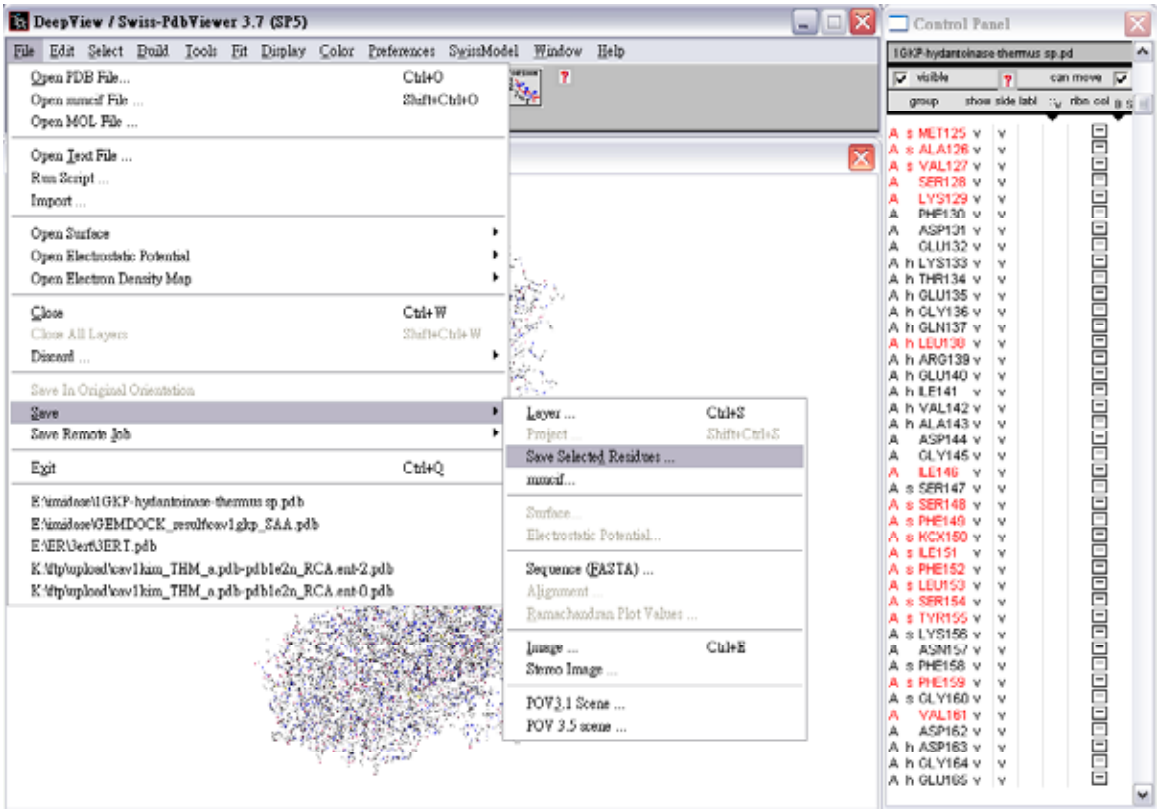
Step 3. Select neighbors of the selected group, “KCX” and two zinc ions.



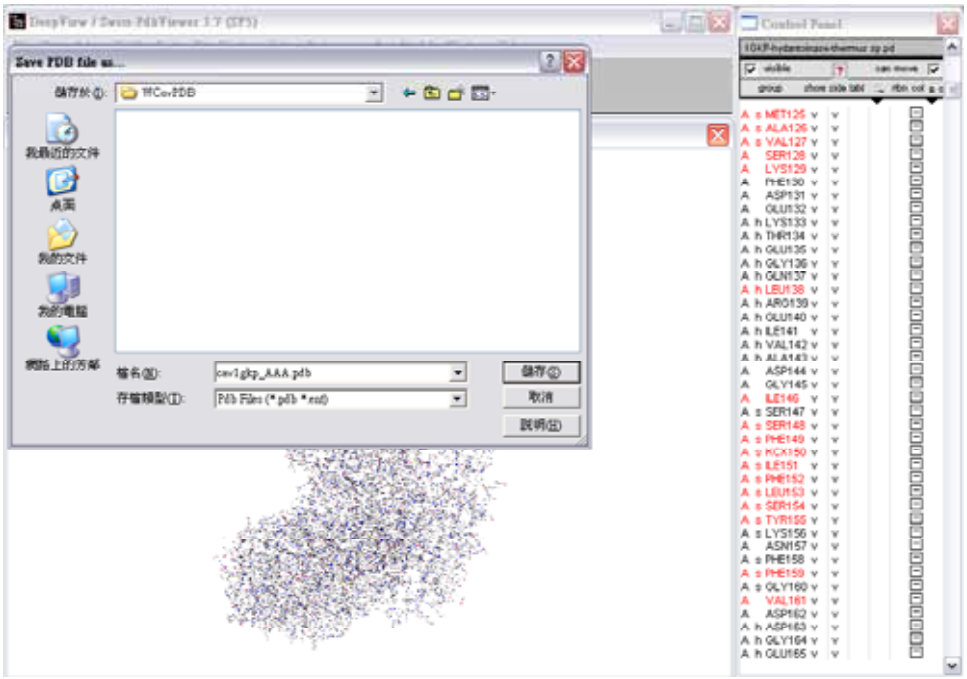
Step 4. Select neighbors of the selected group within 8 Å.



Step 5. Neighbors of the selected group within 8 Å were marked.



Step 6. Save selected residues in the cavity file.



Step 7. Input the name of the cavity file (cav1gkp_AAA.pdb).

```
ATOM    445  N   GLY A 337      16.224  87.656  71.362  1.00 16.38
ATOM    446  CA  GLY A 337      17.443  87.900  70.632  1.00 17.49
ATOM    447  C   GLY A 337      17.783  86.823  69.669  1.00 15.80
ATOM    448  O   GLY A 337      17.479  85.623  69.882  1.00 18.56
ATOM    449  N   ILE A 338      18.467  87.174  68.617  1.00 16.73
ATOM    450  CA  ILE A 338      19.009  86.274  67.628  1.00 17.53
ATOM    451  C   ILE A 338      20.454  86.702  67.417  1.00 16.46
ATOM    452  O   ILE A 338      20.781  87.904  67.390  1.00 18.14
ATOM    453  CB  ILE A 338      18.253  86.297  66.276  1.00 18.10
ATOM    454  CG1 ILE A 338      16.774  86.000  66.528  1.00 19.04
ATOM    455  CG2 ILE A 338      18.891  85.359  65.221  1.00 19.45
ATOM    456  CD1 ILE A 338      15.844  86.088  65.315  1.00 19.48
TER
HETATM  120  N   KCX A 150      20.204 105.981  73.899  1.00 13.49           P
HETATM  121  CA  KCX A 150      19.243 104.973  73.510  1.00 12.60           P
HETATM  122  C   KCX A 150      18.064 105.124  74.415  1.00 13.22           P
HETATM  123  O   KCX A 150      18.182 105.183  75.643  1.00 14.01           P
HETATM  124  CD  KCX A 150      18.219 102.256  72.098  1.00 14.67           P
HETATM  125  CE  KCX A 150      17.372 101.003  71.948  1.00 14.21           P
HETATM  126  NZ  KCX A 150      18.234  99.841  72.064  1.00 14.32           P
HETATM  127  CB  KCX A 150      19.796 103.563  73.691  1.00 14.32           P
HETATM  128  CG  KCX A 150      18.771 102.404  73.542  1.00 14.53           P
HETATM  129  CX  KCX A 150      17.788  98.618  71.769  1.00 17.60           P
HETATM  130  OX1 KCX A 150      16.582  98.423  71.334  1.00 16.17           P
HETATM  131  OX2 KCX A 150      18.605  97.635  71.975  1.00 17.99           P
HETATM  457 ZN    ZN A1460      15.412  96.822  70.941  1.00 19.62           P
HETATM  458 ZN    ZN A1461      18.896  95.941  70.768  1.00 19.49           P
```

Step 8. In the cavity file, remove other titles except “ATOM” and “HETATM”. Before the title “TER” are standard protein. “KCX” and two zinc ions should put after the title “TER” and add “P” at the 82th column to serve as a part of the binding site.

d. Running GEMDOCK for docking

- Edit GEMDOCK\Fit\t.bat
  - fcdock [modes] [population] [cavity file] [ligand file] [parameters]
- Standard parameters: (recommend)

```
fcdock -p   300  cav1ohj_COP.pdb pdb1ohj_COP.ent 6  1 1  1  2 0 80  1 1    0  0
       A   B     C                    D                  E F G H I  J
```

- Statement
  - A: -m (MOL); -p (PDB)
  - B: population size
  - C: protein file (binding site)
  - D: ligand file
  - E: option of the scoring function, default = 6
  - F: electrostatic preference of docked ligands
  - G: hydrophobic preference of docked ligands
  - H: intra-energy of ligands
  - I: family competition, default = 2
  - J: flexible or rigid docking, 0~9, e.g. J=1, fix 10% of rotable bonds in the docked ligand

e. Running GEMDOCK for screening

- Edit GEMDOCK\Fit\t.bat
  - fcdock [modes] [population] [cavity file] [ligand file] [parameters]

- Standard parameters: (recommend)

fcdock -p 300 cav1ohj_COP.pdb pdb1ohj_COP.ent 6 1 1 1 2 0 80 1 1 0 0

A B C D E F G H I J

- Statement
  - A: -l (list of the docked compounds)
  - B: population size
  - C: protein file (binding site)
  - D: ligand file
  - E: option of the scoring function, default = 6
  - F: electrostatic preference of docked ligands
  - G: hydrophobic preference of docked ligands
  - H: intra-energy of ligands
  - I: family competition, default = 2
  - J: flexible or rigid docking, 0~9, e.g. J=1, fix 10% of rotable bonds in the docked ligand
- The format of list files
  - The first line have to declare the file format of docked ligands. Then list the path and file names of docked ligands. e.g.

#!MOL

mddr/090235.mol

mddr/090918.mol

mddr/091050.mol

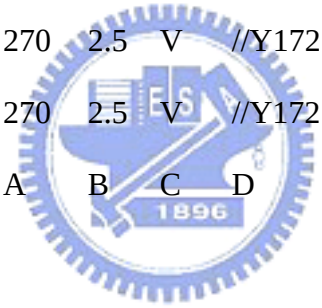
mddr/091114.mol

mddr/107859.mol

mddr/110597.mol

- Pharmacological consensus
  - Edit the file “imptatom.txt”. e.g.

```
127    2.0    H    // Y101 OH
148    4.0    H    // Q125 OE1
149    3.5    H    // Q125 NE2
206    1.5    H    // R163 NH1
270    2.5    V    // Y172 CG
270    2.5    V    //Y172 CD1
270    2.5    V    //Y172 CD2
270    2.5    V    //Y172 CE1
270    2.5    V    //Y172 CE2
270    2.5    V    //Y172 CZ
A      B      C    D
```



A: atom number

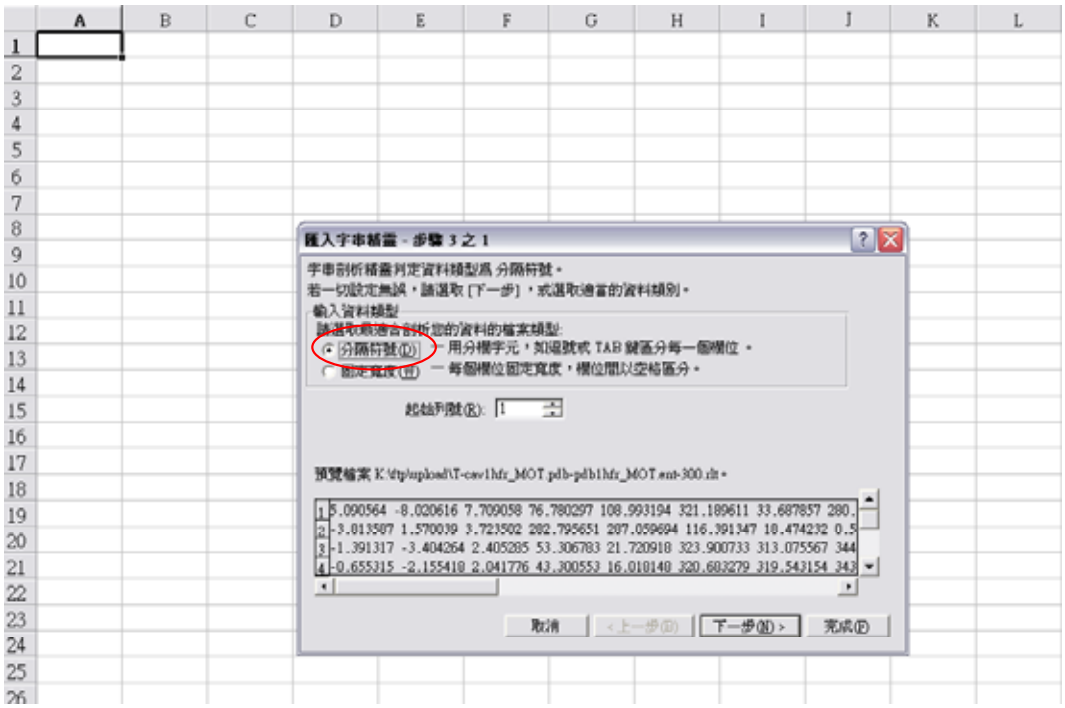
B: weights defined by users

C: “H” for hbonds, “V” for van der waal force, “A” for all

D: discription (optional)

f. Docking result analysis

When running docking, GEMDOCK will output docked ligands in the folder “PrePDB” and a result file (.rlt) including RMSD and fitness. We could observe the result file as follows,



Step 1. Edit the result file (.rlt) with the Microsoft Excel. Import the file and separate each column by the space character.

	N	O	P	Q	R	S	T	U	V	W
1	9.627064	335.3871	315.4299	0	1	-25.9411	0.080048	16.59386	1	300
2	354.1322	24.02281	336.6916	0	74	-77.5288	0.080822	3.588865	20	23100
3	62.82752	308.7834	356.4417	0	151	-127.243	0.216407	2.473188	40	47100
4	55.76602	310.2248	54.70344	0	233	-185.463	0.244611	2.075294	60	71100
5	55.76602	310.2248	54.70344	0	237	-185.463	0.244611	2.075294	61	72300
6	7.979064	88.28059	332.3328	0	1	-32.9694	0.177587	13.86293	1	300
7	276.3252	38.14104	87.99186	0	74	-79.3123	0.246277	5.960055	20	23100
8	274.4629	39.18238	88.16485	0	151	-85.277	0.273431	6.130755	40	47100
9	274.3271	41.137	74.43878	0	230	-107.634	0.213857	6.005676	60	71100
10	274.5836	41.11343	77.44951	0	234	-108.684	0.2267	5.986531	61	72300
11	296.4504	58.52321	62.87393	0	1	-15.6268	0.248825	19.0044	1	300
12	18.33781	357.7723	346.7146	0	74	-80.8373	0.081087	8.184809	20	23100
13	6.68761	9.516703	331.4798	0	151	-175.14	-0.12685	0.774657	40	47100
14	5.01626	8.823029	328.1116	0	232	-210.838	-0.13363	0.419872	60	71100
15	6.120086	8.34643	328.3246	0	237	-211.425	-0.13476	0.483985	61	72300
16										
17						fitness		RMSD	generation	
18										
19										

Step 2. The three columns records the fitness, RMSD and generation.

Sorting by this column (generation)

N	O	P	Q	R	S	T	U	V	W
55.76602	310.2248	54.70344	0	237	-185.463	0.244611	2.075294	61	run0 00
274.5836	41.11343	77.44951	0	234	-108.684	0.2267	5.986531	61	run1 00
6.120086	8.34643	328.3246	0	237	-211.425	-0.13476	0.483985	61	run2 00
55.76602	310.2248	54.70344	0	233	-185.463	0.244611	2.075294	60	71100
274.3271	41.137	74.43878	0	230	-107.634	0.213857	6.005676	60	71100
5.01626	8.823029	328.1116	0	232	-210.838	-0.13363	0.419872	60	71100
62.82752	308.7834	356.4417	0	151	-127.243	0.216407	2.473188	40	47100
274.4629	39.18238	88.16485	0	151	-85.277	0.273431	6.130755	40	47100
6.68761	9.516703	331.4798	0	151	-175.14	-0.12685	0.774657	40	47100
354.1322	24.02281	336.6916	0	74	-77.5288	0.080822	3.588865	20	23100
276.3252	38.14104	87.99186	0	74	-79.3123	0.246277	5.960055	20	23100
18.33781	357.7723	346.7146	0	74	-80.8373	0.081087	8.184809	20	23100
9.627064	335.3871	315.4299	0	1	-25.9411	0.080048	16.59386	1	300
7.979064	88.28059	332.3328	0	1	-32.9694	0.177587	13.86293	1	300
296.4504	58.52321	62.87393	0	1	-15.6268	0.248825	19.0044	1	300

fitness RMSD generation

Step 3. Sort by the column that records the generation. After sorting, the top three rows are results of the three docking solution.

g. Screening result analysis

When running screening, GEMDOCK will output docked ligands in the folder “PrePDB” and two result file (dock.log and run.log). The file “run.log” records the fitness value of each docked ligand and the file “dock.log” contains ranks of docked ligands.

```
#
# Top best docking result
#
# RANK FitnessValue   DrugName              RUN Atom Hbond Elect
     1    -195.632148  MFCD00010060.mol       0   24    11     0
     2    -178.100070  MFCD00005733.mol       2   20    11     0
     3    -174.206047  MFCD00006591.mol       1   31    10     0
     4    -172.416484  MFCD00006600.mol       0   23    13     0
     5    -171.069377  MFCD00005084.mol       2   24     5     0
     6    -170.655685  MFCD00006528.mol       2   18    10     0
     7    -170.613814  MFCD00010057.mol       1   22    10     9
     8    -168.818432  MFCD00006532.mol       2   18    10     0
     9    -167.632663  1ki3_penciclovir.mol   0   18     6     0
    10    -166.489039  MFCD00005048.mol       2   24     6     0
    11    -165.739889  1ki2_ganciclovir.mol   2   18    10     0
    12    -165.086313  MFCD00006628.mol       0   23    14     0
    13    -164.777292  MFCD00006606.mol       0   23    16     0
    14    -162.940455  MFCD00001238.mol       1   30     5     0
    15    -161.136848  MFCD00006529.mol       1   17     8     0
    16    -161.101503  1e2k_mct.mol           2   18     9     0
    17    -159.050144  MFCD00016886.mol       1   23     6     0
    18    -158.809893  1ki6_ahiu.mol          0   18     8     0
    19    -158.590124  MFCD00013264.mol       0   28     2     0
    20    -154.999935  MFCD00006626.mol       1   23    11     0
```

dock.log: record the best solution of each docked ligand and ranked by their fitness.

- Atom: the number of heavy atoms of each docked ligand
- Hbond: the number of hydrogen bonds formed between the protein and the docked ligand
- Elect: the number of ionic bonds formed between the protein and the docked ligand