

Tutorial: ADME driven compound design

Ismael Zamora
Molecular Discovery, Ltd
www.moldiscovery.com
www.leadmolecular.com
ismael.zamora@leadmolecular.com

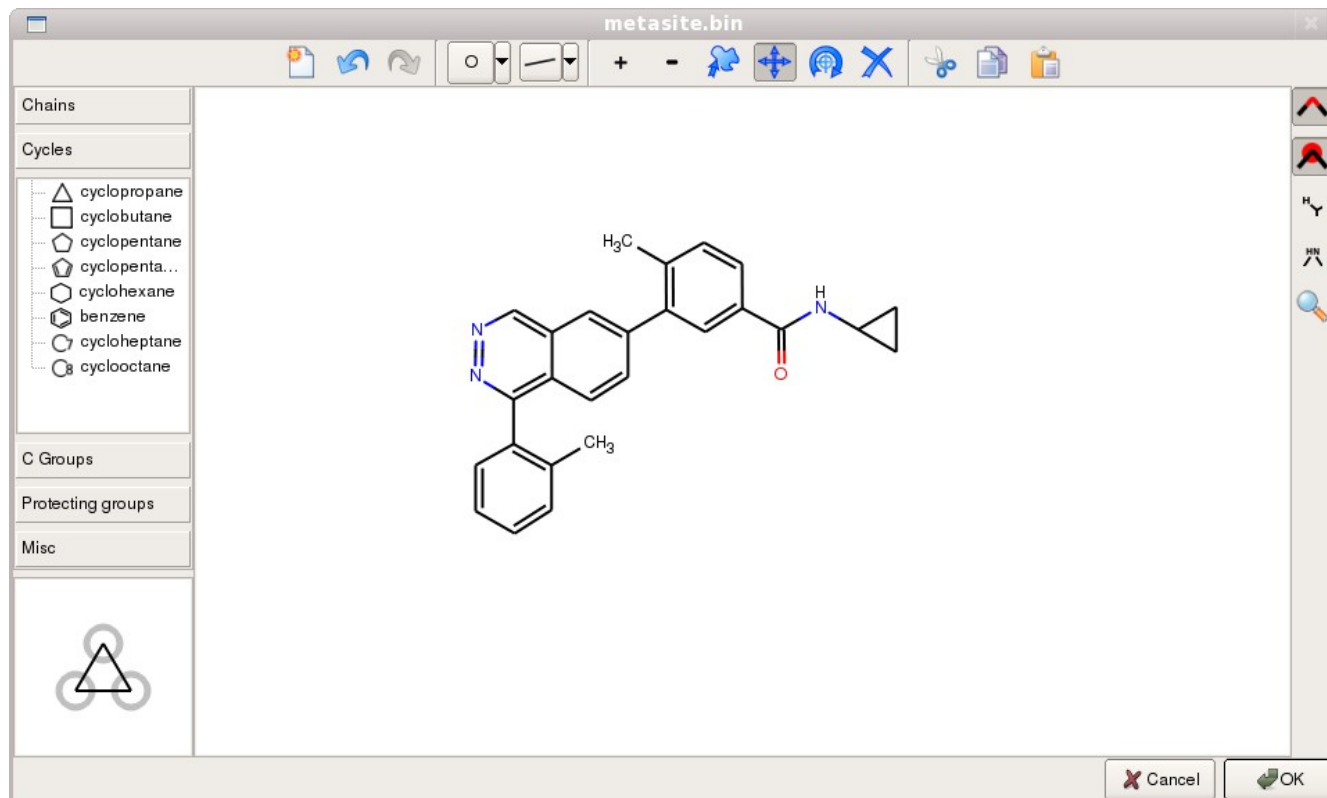
The aim of this tutorial is first to perform an analysis of the chemical structure of a good inhibitor of the p38 kinase from the point of view of ADME predictive tool. Then, to modify the chemical structures to see how these structural modifications affect the ADME properties. The example is based on the results reported on L.H.Pettus, et al *J.Med.Chem.***2008**, 51:6280-6292.

Analysis of the Site of Metabolism

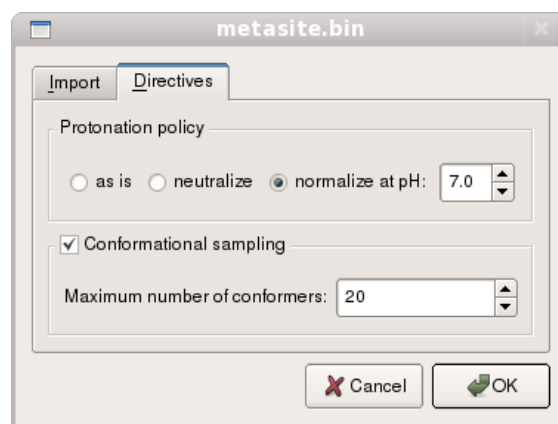
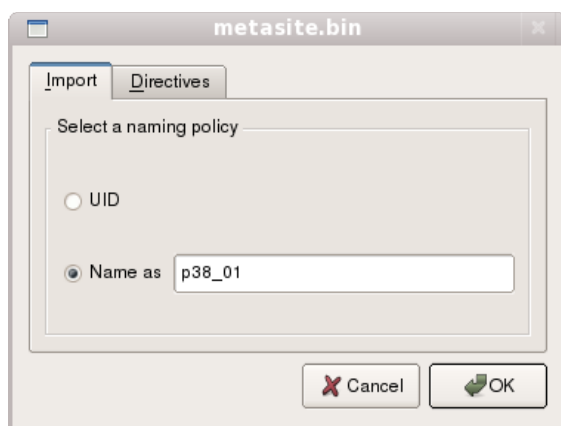
The first analysis will be done on the predicted site of metabolism using the program MetaSite 3.0 (www.moldiscovery.com).

Start the MetaSite graphical user interface by selecting the MetaSite shortcut from the Windows start menu.

Then, to open the molecular editor select: **File**→**import**→**Sketch** molecule and draw the substrate structure in the editor window, as shown in the graphic below (for an overview of the editor look at the tutorial Appendix A):



When drawing is complete, click on the OK button and the Objects Import dialog box (reported below) opens:



- Import tab:

Inside the frame labeled Select a naming policy , select the Name as option and assign the name *p38_01* to the structure.

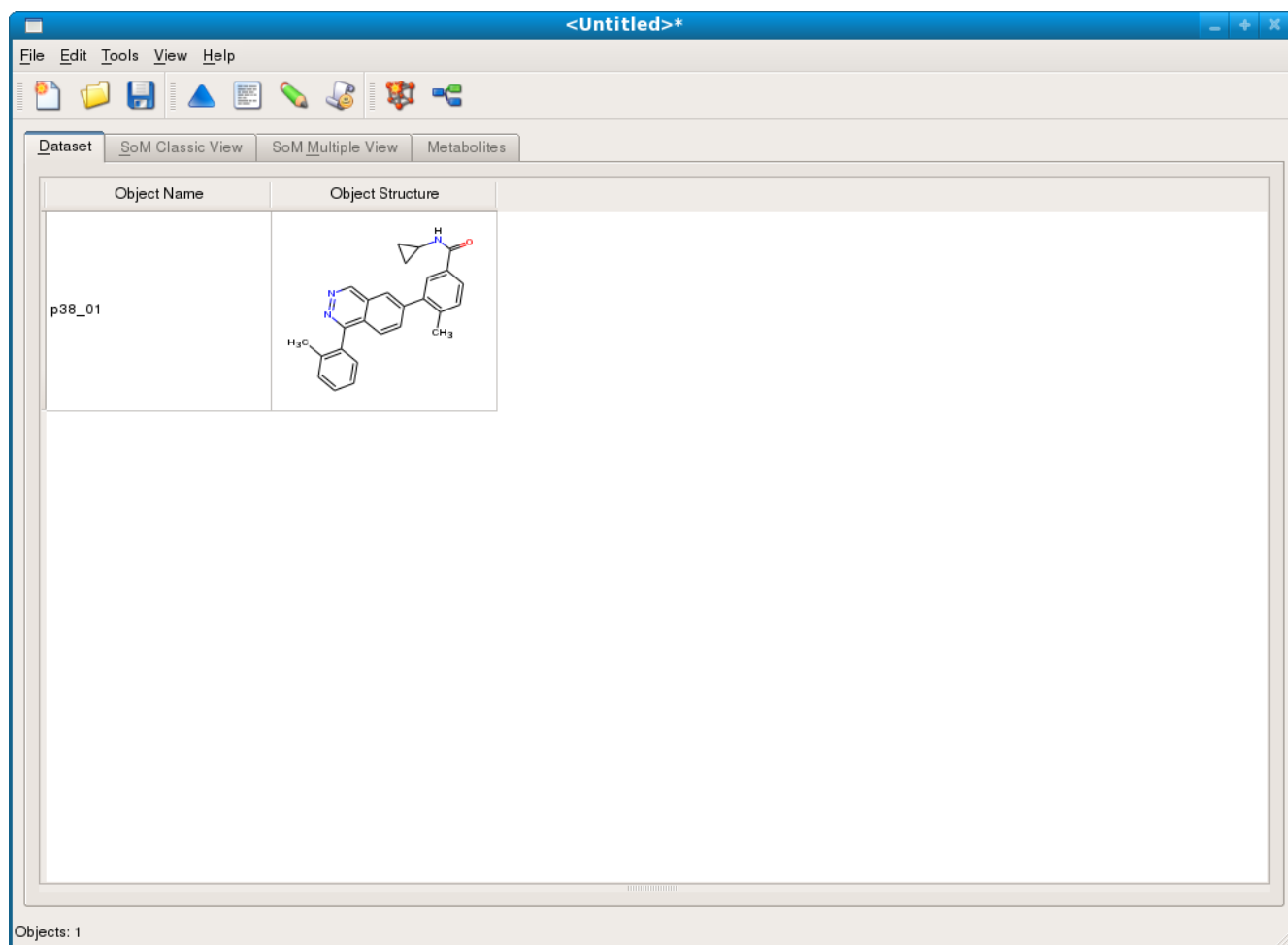
- Directives tab:

Change the parameters as shown in the figure and click on the OK button to start the import operation. During processing a progress dialog box is shown:

Tip: Check the option to Close this dialog upon completion if you want the progress dialog box to close automatically when processing has finished.

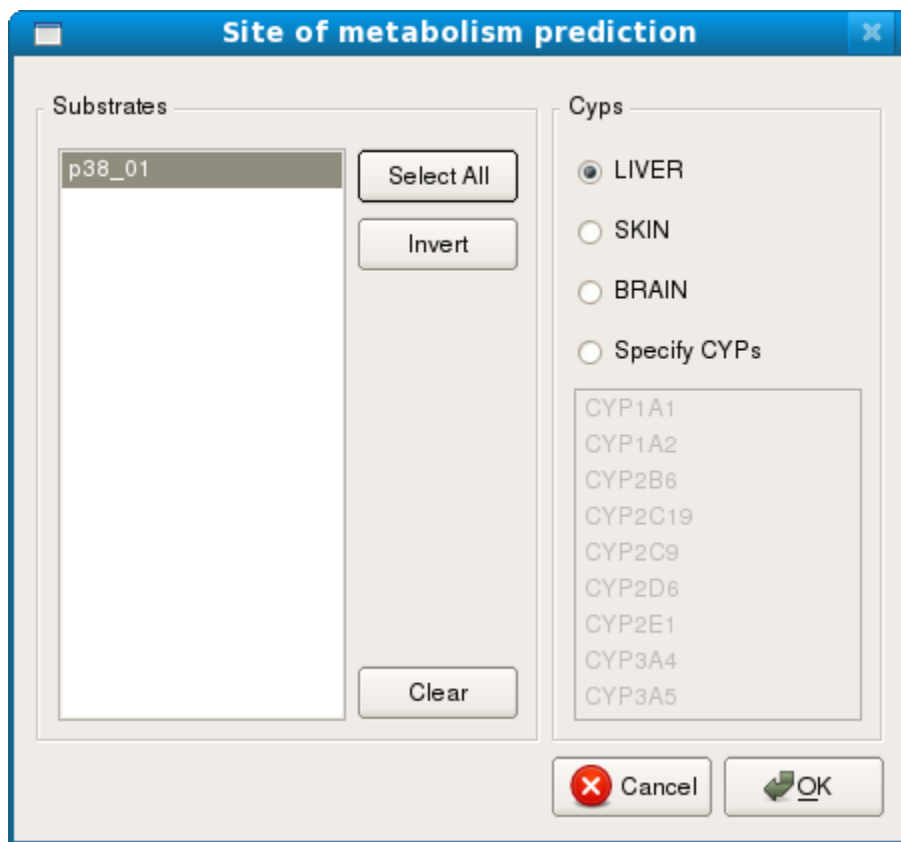
When the import process has been completed, click on the Close button in the progress dialog box.

The newly imported compound will be displayed in the Dataset tab page:



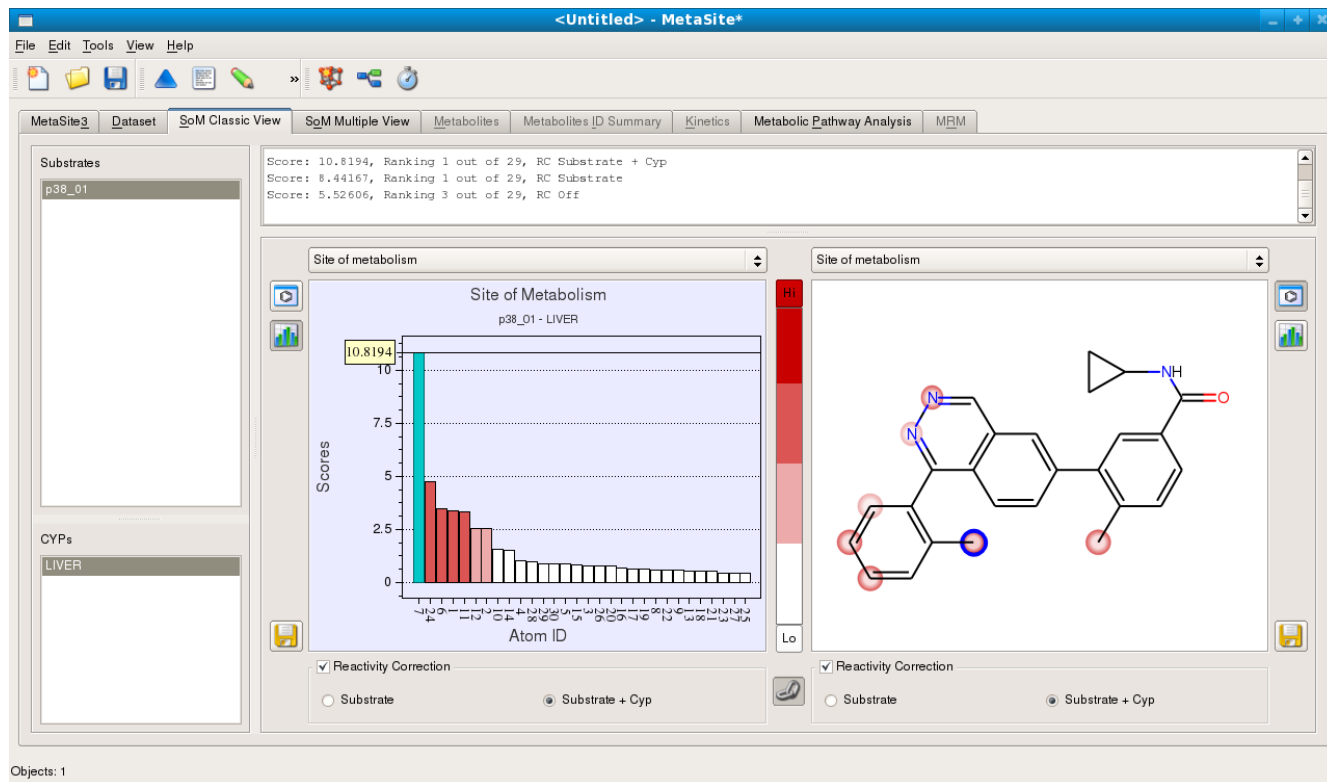
To start the Site of Metabolism (SoM) prediction, select **Tools**→**Site of metabolism prediction**.

A dialog box will be shown:



Select *p38_01* from the Substrates list box, and then activate the Specify CYPs option in the Cyps frame and select LIVER .

Finally, click on the OK button to start the computation. Once processing has been completed, the SoM prediction results are shown in the SoM Classic View tab page:



Reactivity Correction

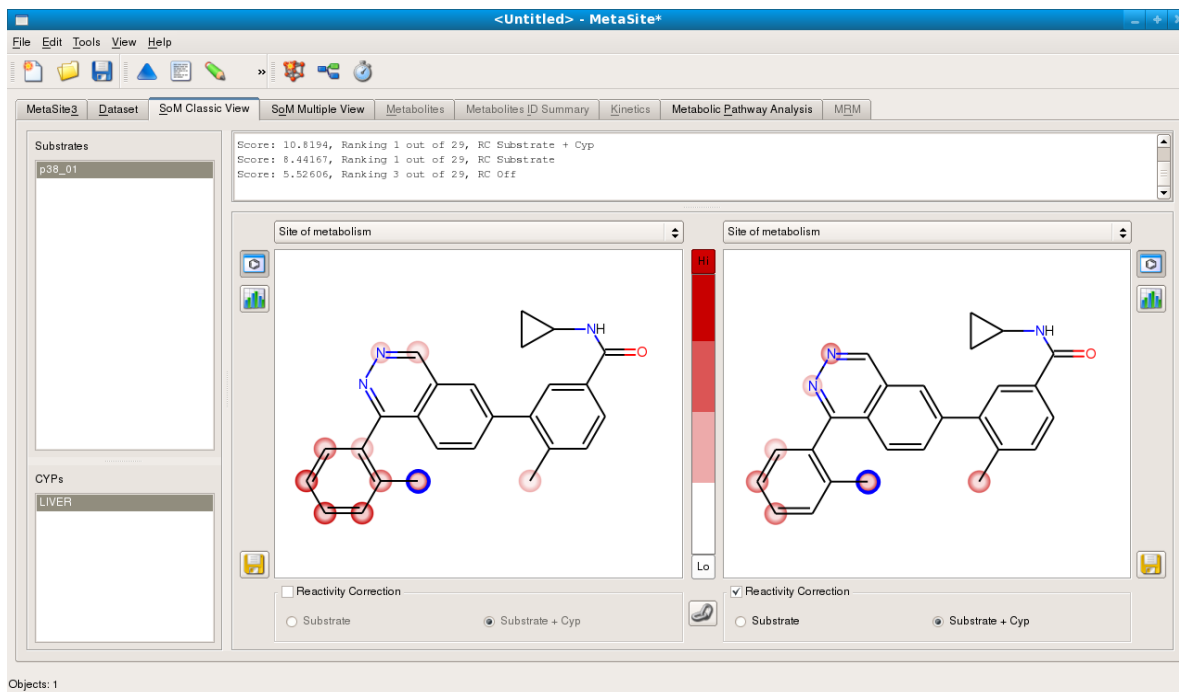
Click on the 2D button near the upper left hand corner of the left panel to switch from the bar graph view to the graphic of the substrate 2D structure. The atomic positions in the 2D graphic depiction are marked by colored spots, which highlight the most likely site of metabolism.



Click on the link button located between the two Reactivity Correction frames in order to unlink the reactivity correction settings for the two graphic panels.

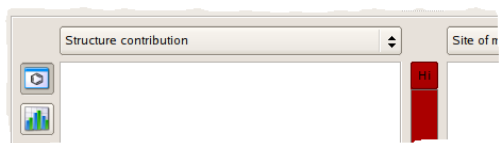
If the Reactivity Correction below the left hand panel is unchecked, the left hand panel shows only the site of metabolism likelihood based on the orientation of the compound inside the cytochrome cavity, while the right hand panel shows the full site of metabolism prediction taking both the protein and the reactivity factors into account.

Click again on the Link button to link both views.



Conclusion: The protein orientation factor and the reactivity factor point towards the same region of the molecule. Therefore, by chemically protecting the benzylic moiety may only solve the metabolic stability issue, but the compound may still be a cytochrome inhibitor.

Activity Contributions



Switch the pulldown on top of the left hand panel (set by default to Site of metabolism) to Structure contribution and click on the adjacent 2D button to switch the panel from the bar graph to the 2D graphic depiction of the

substrate currently selected.

In the right hand panel, click on the most likely site of metabolism (the more reddish atom) of the p38_01. This position is marked by MetaSite as the most likely (red) and this position. When this position is selected, the colored spots in the left hand panel refer to the contribution to orienting the site of metabolism towards the heme group made by the other atoms in the substrate structure.

Score: 10.8194, Ranking 1 out of 29, RC Substrate + Cyp
 Score: 8.44167, Ranking 1 out of 29, RC Substrate
 Score: 5.52606, Ranking 3 out of 29, RC Off

Structure contribution

Site of metabolism

Substrates
p38_01

CYPs
LIVER

Objects: 1

Conclusion: The group that contributes for the benzyl position to be the site of metabolism is the amide in the other site of the molecule. If one replace this group it may reduce the interaction of the compound with the cytochrome. This is observed in the mentioned article when a small difference in this part of the molecule change the IC50 for CYP3A4.

Export the molecule

In order to export the molecule click on the SoM Multiple View tab

Objects: 1

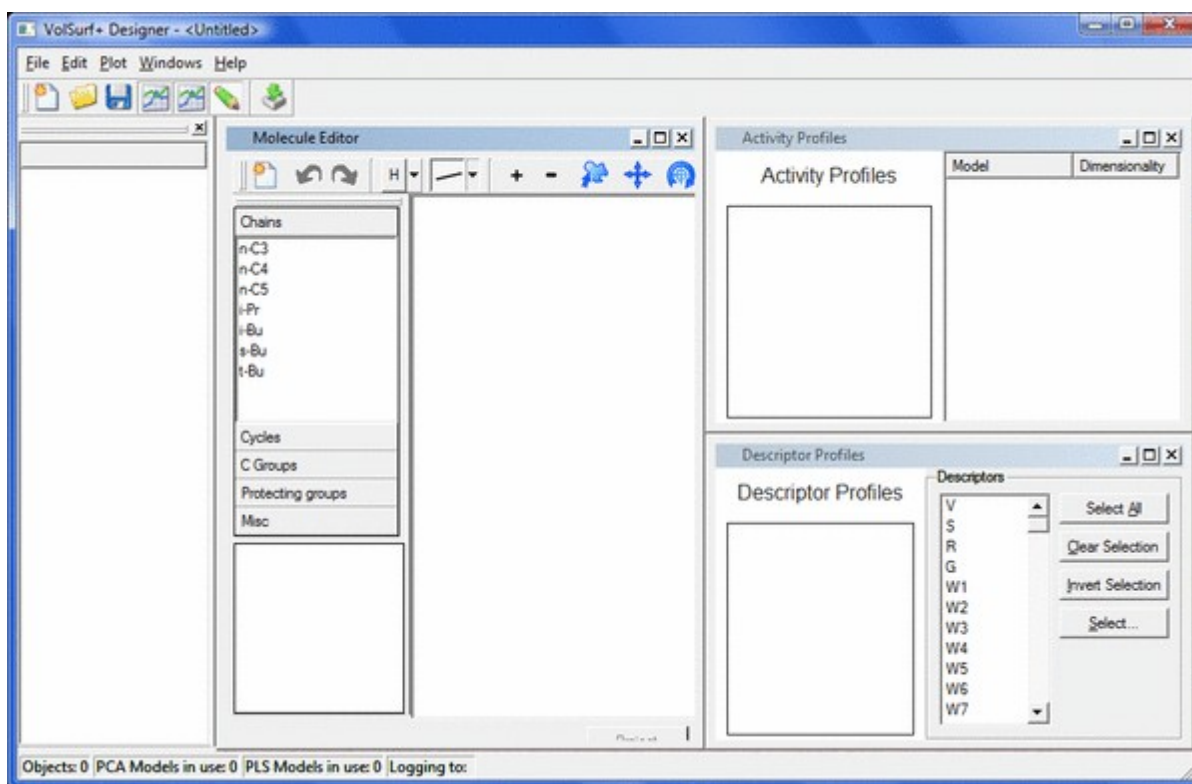
and click on export. Select the sdf file format and type as the name 2d6s_MSite.sdf. Then save you computation from **File**→**Save**. Choose the name of the file you prefer.

Analysis of ADME properties

Start the VoISurf+ program:

- for LINUX users typing: `$ vsdesigner`
- for Windows users double clicking the VoISurf+ Designer icon.

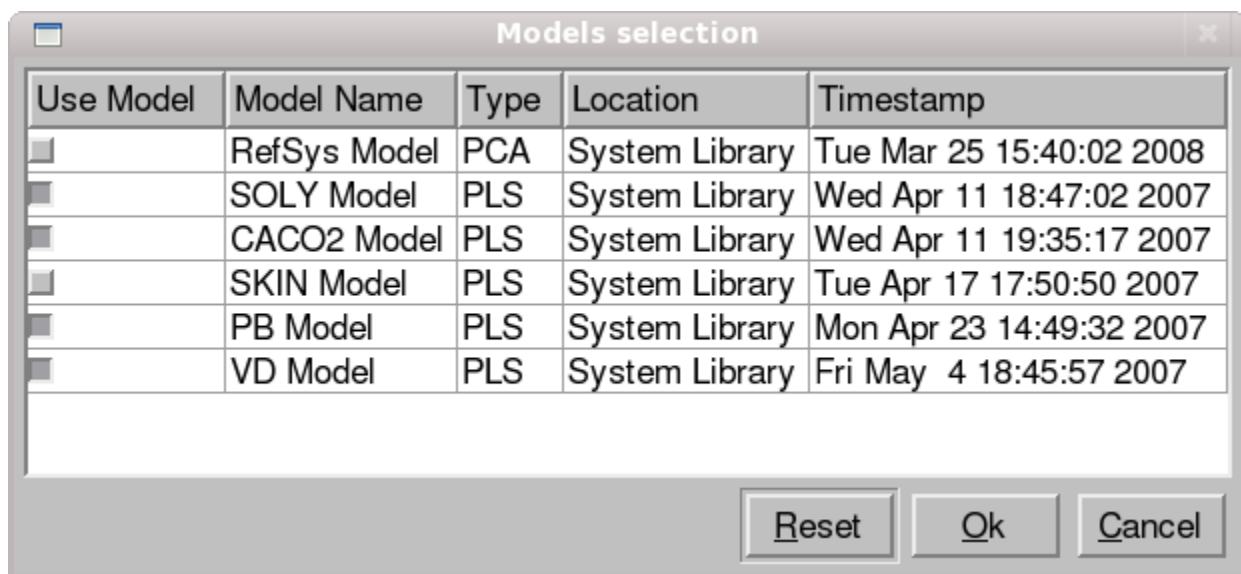
The VoISurf+ Designer interface will appear



Select the VoISurf+ library models by pressing:

[MENU] Edit→Select models

a new dialog appears: by clicking on the Use Model buttons, select the intrinsic aqueous solubility model (SOLY), passive permeability (CACO2), protein binding (PB) and the volume of distribution model (VD). Press Ok to accept the selection.



The screenshot shows a dialog box titled "Models selection" with a table containing the following data:

Use Model	Model Name	Type	Location	Timestamp
<input type="checkbox"/>	RefSys Model	PCA	System Library	Tue Mar 25 15:40:02 2008
<input type="checkbox"/>	SOLY Model	PLS	System Library	Wed Apr 11 18:47:02 2007
<input type="checkbox"/>	CACO2 Model	PLS	System Library	Wed Apr 11 19:35:17 2007
<input type="checkbox"/>	SKIN Model	PLS	System Library	Tue Apr 17 17:50:50 2007
<input type="checkbox"/>	PB Model	PLS	System Library	Mon Apr 23 14:49:32 2007
<input type="checkbox"/>	VD Model	PLS	System Library	Fri May 4 18:45:57 2007

At the bottom of the dialog box, there are three buttons: "Reset", "Ok", and "Cancel".

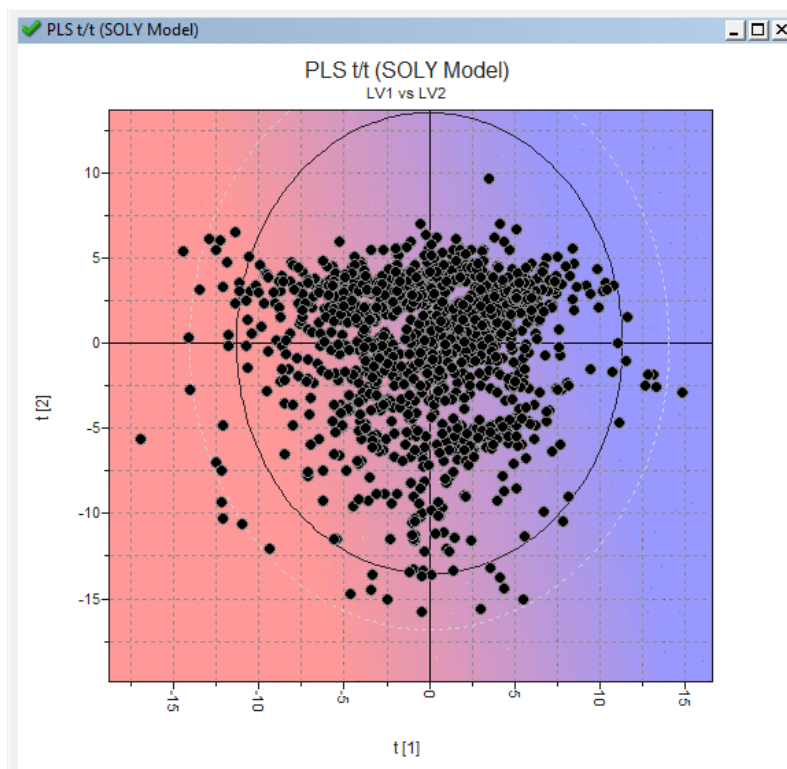
The models will be used by VolSurf+ Designer to make the external predictions.

Open the VD model plot, and the other models plot, using:

[MENU] Plot→PLS→t/t Scores

[DIALOG] select SOLY model and press Ok.

Repeat the procedure with the other models.



The background of the SOLY plot is color-coded according to the logarithm of the solubility (mole/L). The color blue identifies the plot region with higher solubility while red identifies the region with lower solubility.

The use of Caco2 cell monolayers as an in vivo human absorption surrogate has increased. However, due to the mechanisms involved, Caco2 cell permeability measurements exhibit certain limitations. Both passive and active pathways exist. Unstirred water can sensibly modify the penetration coefficient. Intersubject variability between laboratories are common problems. Quantitative comparison and modelling are almost impossible for all these problems. In order to avoid inconsistencies in the data, the Caco2 permeability values are transformed according the following schema:

- $P_{app} < 4 \cdot 10^{-6} \text{ cm/s} \implies \text{score } -1$
- $P_{app} > 8 \cdot 10^{-6} \text{ cm/s} \implies \text{score } +1$

However, different assumptions were made in special cases, when the experimental protocols were different or no internal standard compounds were used. A basic assumption used in the model is passive permeation.

The CACO2 model is a qualitative model containing a thousand related, but chemically diverse, compounds collected from the literature or experimentally measured in laboratories connected with our group. Data are either penetrating (score 1), or have little if any ability to penetrate the epithelial cells (score -1). PLS discriminant analysis was used to build the statistical model and two significant latent variables emerged from the cross validated PLS model.

The model can be used to project external compounds in the chemical space represented by the model in order to rank the Caco2 behaviour of external compounds

"In silico" quantitative models to predict binding affinity to Human Serum Albumin (HSA) are often useful in the pharmaceutical industry as they provide pharmacokinetic properties in an early phase of drug discovery. As HSA is the principal biological carrier of many drugs, it facilitates their transport through the circulatory system to the target tissues. Determining the probability of a molecule binding with a protein depends on the type of analysis used (dialysis, ultra-centrifugation, ultra-filtration, NMR, UV, HPLC and other chromatographic methods), the instruments used (type of dialysis membrane, type of spectrometer, type of chromatographic equipment) and the experimental conditions chosen in different laboratories (type of albumin, its concentration, temperature and the duration of the analysis).

The variation of these parameters not only dramatically affects the final results but also the experimental errors. Such huge variability of experimental conditions produces noise and makes interpretation of the data more difficult.

The Protein_Binding model is a qualitative model containing 500 related, but chemically different compounds partially collected from the literature or experimentally measured in

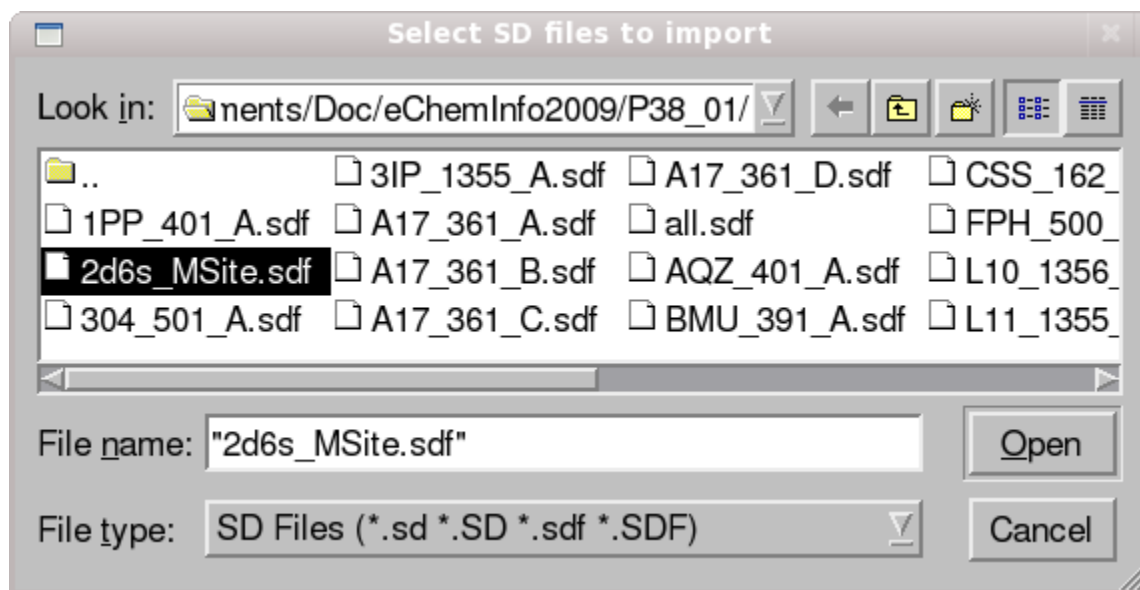
laboratories connected with our group. The data mainly report albumin protein binding values between 10% and 100% obtained using spectroscopic techniques. The average experimental error reported was 8%. Therefore, the model is not able to discriminate between protein binding values ranging from 95% to 100%.

The model can be used to project external compounds in the chemical space represented by the model in order to rank the protein binding profile of external compounds.

The volume of distribution (VD) for a drug is the volume that accounts for the total dose administration based on the observed plasma concentration. The plasma volume of the average adult is approximately 3 liters. Therefore, apparent volume of distribution larger than the plasma compartment (i.e. greater than 3 liters) indicates that the drug is also present in tissue or fluid outside the plasma compartment. Volume of distribution represents a complex combination of multiple chemical and biochemical phenomena. It also measures the relative partitioning of drug between plasma and the tissues. Although the volume of distribution cannot be used to determinate the actual site of distribution of a drug in the body, it is of extreme importance in estimating the loading dose necessary to rapidly achieve a desired plasma concentration.

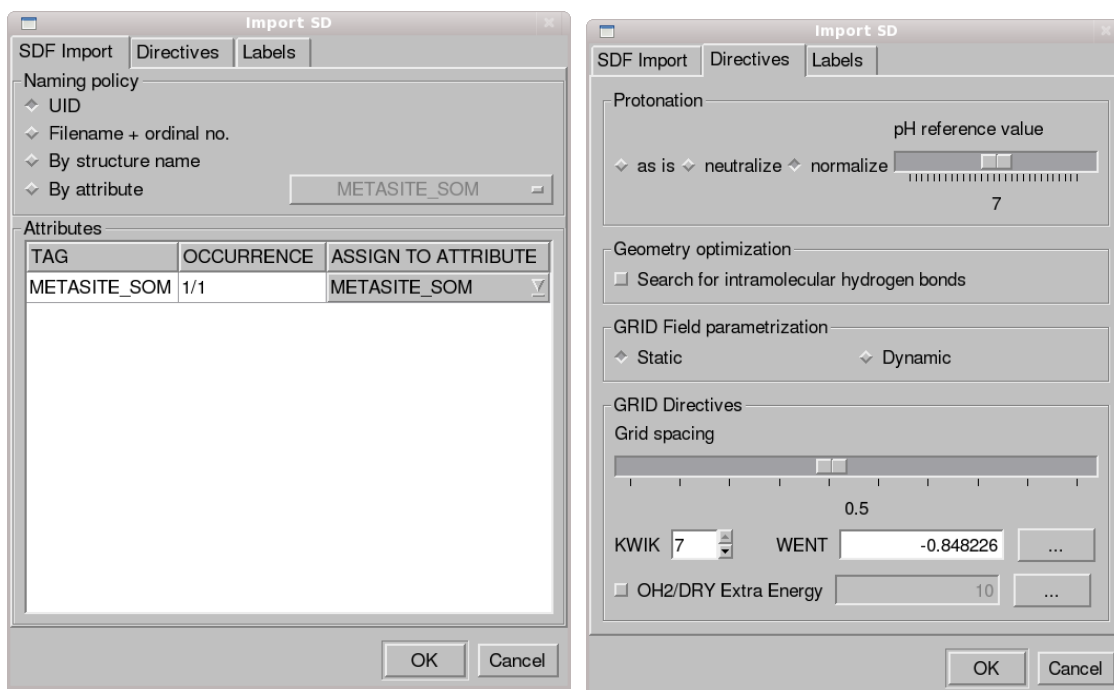
The Volume_Distribution model was obtained by collecting more than 600 compounds from the literature. The VD data (Litre/Kg) were converted into $-\text{Log}[\text{VD}]$ values. Low VD values mean low distribution into tissues while high VD values mean high distribution into tissues.

Import the molecule of interest
[MENU] File→Import→Sdf Files



Select the 2d6s_MSite.sdf file that we have just generated in MetaSite.

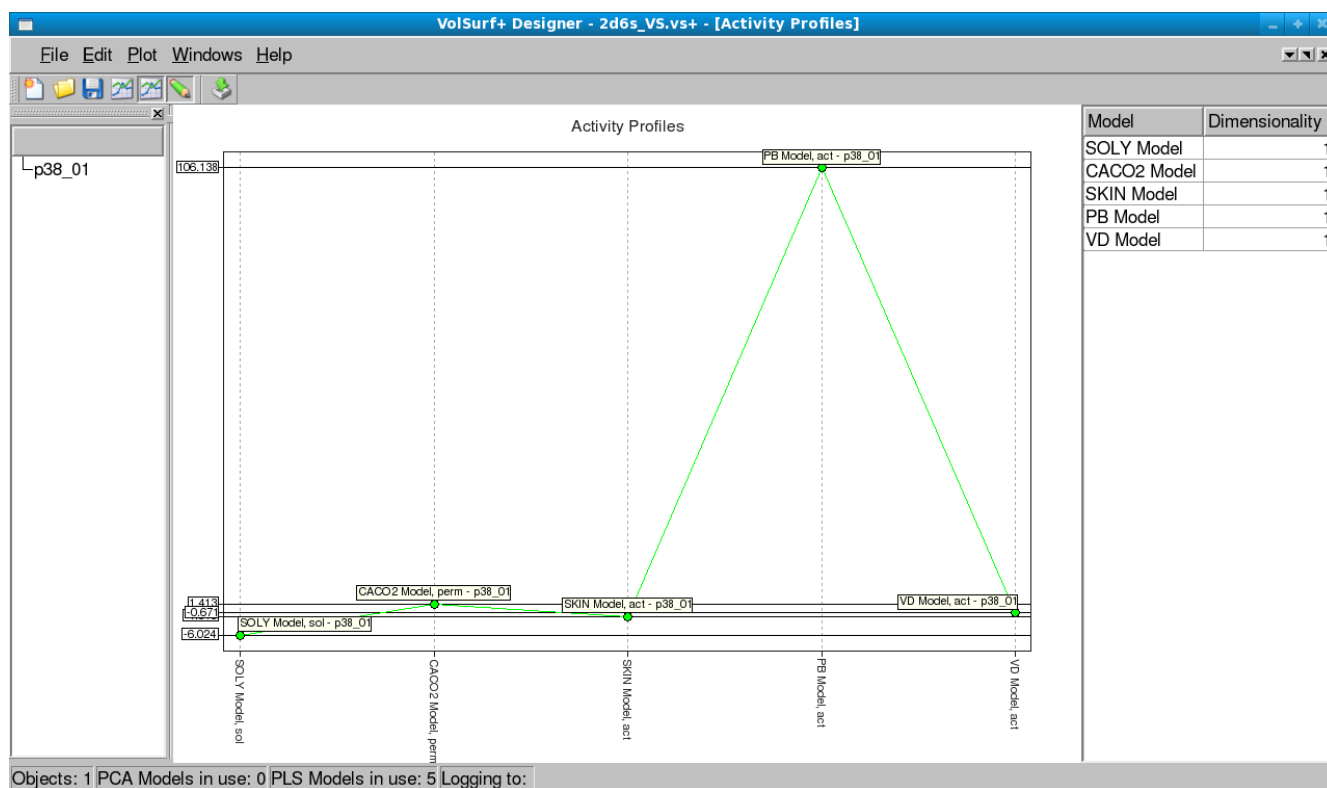
Select the options as indicated in the figures:



Click OK

The descriptors are produced automatically and the molecule projected (yellow points) in the statistical models as reported in the following figure (click on Close when the projection process has been completed):

The Activity Profiles plot will be upgraded with your profile profile. This profile can be used as a reference. Tune the dimensionality for the SOLY model to 3, Caco2, PB and VD to 2 by clicking in the Dimensionality box and then click on the yellow points in the Activity Profiles box to obtain the corresponding predicted values.



Save the results as Personal.vs+ file by selecting Save from the menu:

[MENU] File→Save As


Conclusions: The compound has low solubility, high protein binding and good Caco2 permeability.

Replace one fragment of the molecule – Ligand Based Approach


The aim of this tutorial is to find chemical substitutions for the molecules we have done ADME analysis, in order to solve some of the ADME problems or for whatever other reason.


Therefore the server has to be run:

In windows: execute the Shop server by selecting the *Start menu select the All Programs* option in Windows XP, then Shop menu, then the Server sub-menu and finally Server console and type `shopspd -e`

 In order to run the server it is necessary that you have the license key installed in the **ShopDirbin** directory (see manual).

In order to start the Shop client program:

In Windows: Double click on the Shop icon  or in the *Start menu select the All Programs* option and then *Shop sub-menu* and then the *Shop program*

Once the server is running the Shop application can be connected by selecting the *Connect* to server option in the *File* menu or clicking on the connect button in the Shop tool Bar . Since the Shop server is running locally the name of the Host is 127.0.0.1 or localhost, and the port will be 4949 (default communication ports). In the case you are running remotely you will have to contact the Shop administrator to know the Host name and port. These parameters are selected during the installation process. After clicking on *Connect* the list of databases available in the server will be shown to you, in this case including the pddbnd. Select the pddbnd database and click *OK*. The compounds in the database will be shown on the database tab in the Shop GUI.

This database was generated using fragments provided by cutting the single bonds from ligands obtained from the pddbnd dataset. In order to check the information for the pddbnd database select the menu *Database* in the Shop GUI and click on *Show Database details* (CTRL+T)

This will report the condition under the database was built. Check that there are:

7452 scaffolds for 1 anchor point


19286 scaffolds for 2 anchor points

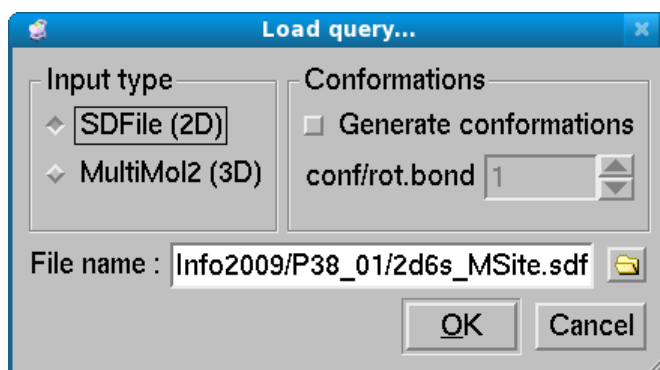
32442 scaffolds for 3 anchor points

0 scaffolds for 4 anchor points

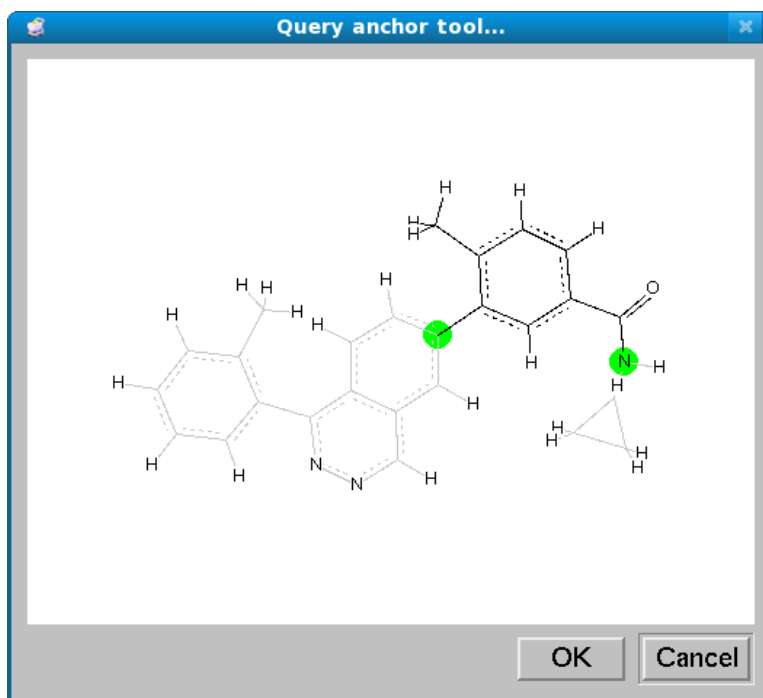
0 scaffolds for 5 anchor points

Click on the *OK* to close the window.

Load a SDFFile (2D) file without anchor points. In order to load the query molecules select the Load query from the Query menu or click on the Load query button on the Shop ToolBar (). A SDFFile (2D) without the anchor is imported. Do not select to Generate conformers. Select the 2d6s_MSite.sdf

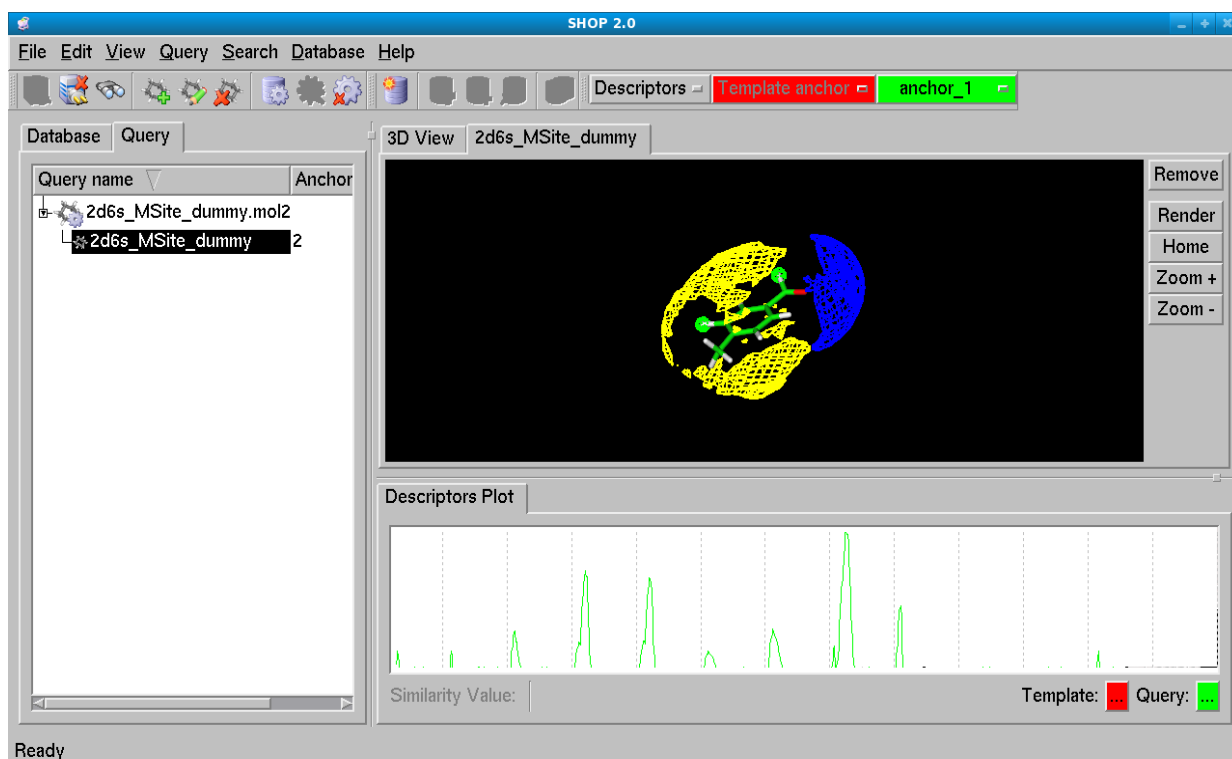



Since in this case the compound does not have an anchor point indicated, a window will pop-up where the user has to click the atom that will be the attachment point. The procedure to mark the atom is by clicking once on the atom the biggest fragment is selected as the fragment to change, clicking a second time on the same atom, the second biggest fragment is selected.

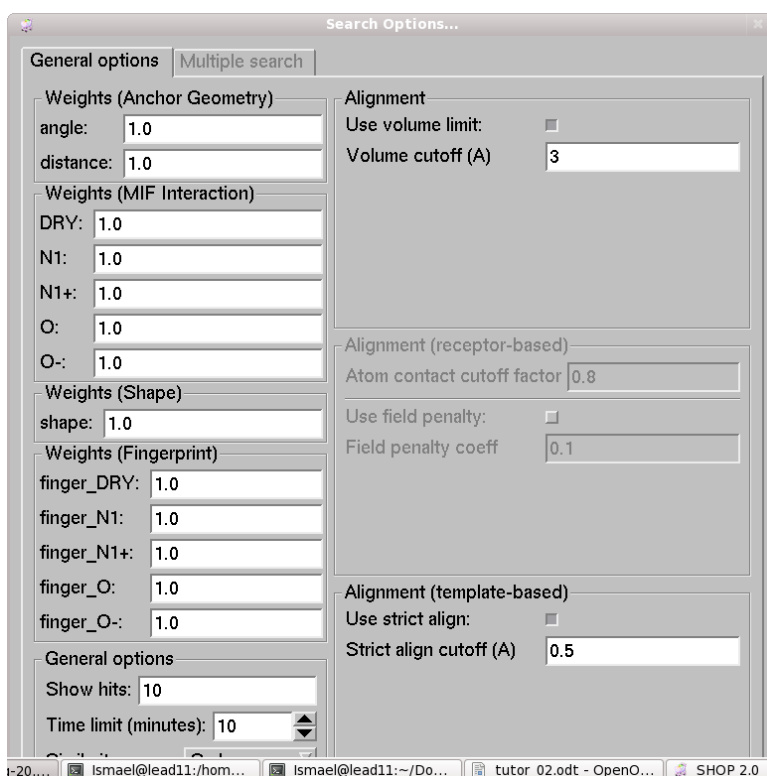


Visualize the query scaffold

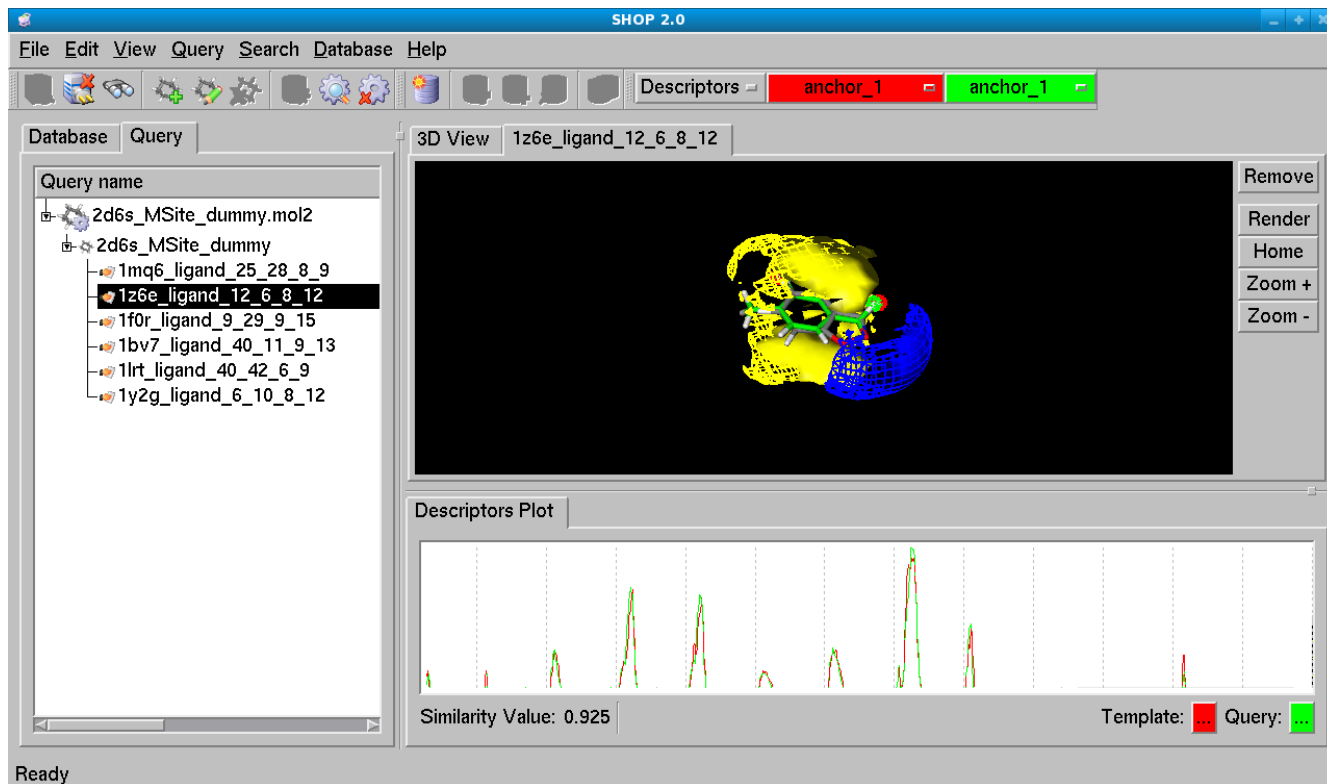
In the Query tab click on the 2d6s_MSite_dummy tree. Click on *2d6s_MSite_dummy* name from the query list (The anchor points are shown in green).



Perform a similarity search for the 2d6s_MSite_dummy. Select the compound name in the *Query Compound List* and select the *Search* option (CTRL+S) or click on the search option button ().



After 2-3 minutes the calculation is finished and the potential fragments to replace the one selected are shown in the query tab. Browse the solutions.



Conclusion: The second solution suggested by the methodology is the same structural modification proposed in the article as a bio-isosteric replacement for the ketone moiety. But, does this molecular modification improve the ADME properties?

In the case that the initial query did not have the attachment point, Shop is able to remember the fragment that is kept fixed from the tool used to define the fragment to change. Therefore, it is possible to rebuild an entire molecule with the new solutions + the fragment that is kept. Select the **Search/Rebuild ligand** option and click on the Rebuild All option.

After the process is finished the new molecules/suggestions have been rebuilt and saved a window will pop-up. Click OK and close the Rebuild window by clicking Close.

Exit the Shop Client by selecting.


[MENU] File→Exit

Replace one fragment of the molecule – Receptor Based Search


The aim of this tutorial is to find chemical substitutions for the molecules we have done ADME analysis, in order to solve some of the ADME problems or for whatever other reason.


Therefore the server has to be run:

In windows: execute the Shop server by selecting the *Start menu select the All Programs* option in Windows XP, then Shop menu, then the Server sub-menu and finally Server console and type `shopspd -e`

 In order to run the server it is necessary that you have the license key installed in the **ShopDirbin** directory (see manual).

In order to start the Shop client program:

In Windows: Double click on the Shop icon  or in the *Start menu select the All Programs* option and then *Shop sub-menu* and then the *Shop program*

Once the server is running the Shop application can be connected by selecting the *Connect* to server option in the *File* menu or clicking on the connect button in the Shop tool Bar . Since the Shop server is running locally the name of the Host is 127.0.0.1 or localhost, and the port will be 4949 (default communication ports). In the case you are running remotely you will have to contact the Shop administrator to know the Host name and port. These parameters are selected during the installation process. After clicking on *Connect* the list of databases available in the server will be shown to you, in this case including the `synthonix_2008`. Select the `synthonix_2008` database and click *OK*. The compounds in the database will be shown on the database tab in the Shop GUI.

This database was generated using fragments provided by Synthonix and a collection of chemical reactions. In order to check the information for the Synthonix database select the menu *Database* in the Shop GUI and click on *Show Database details* (CTRL+T)

This will report the condition under the database was built. Check that there are:

24235 scaffolds for 1 anchor point

43743 scaffolds for 2 anchor points

27027 scaffolds for 3 anchor points

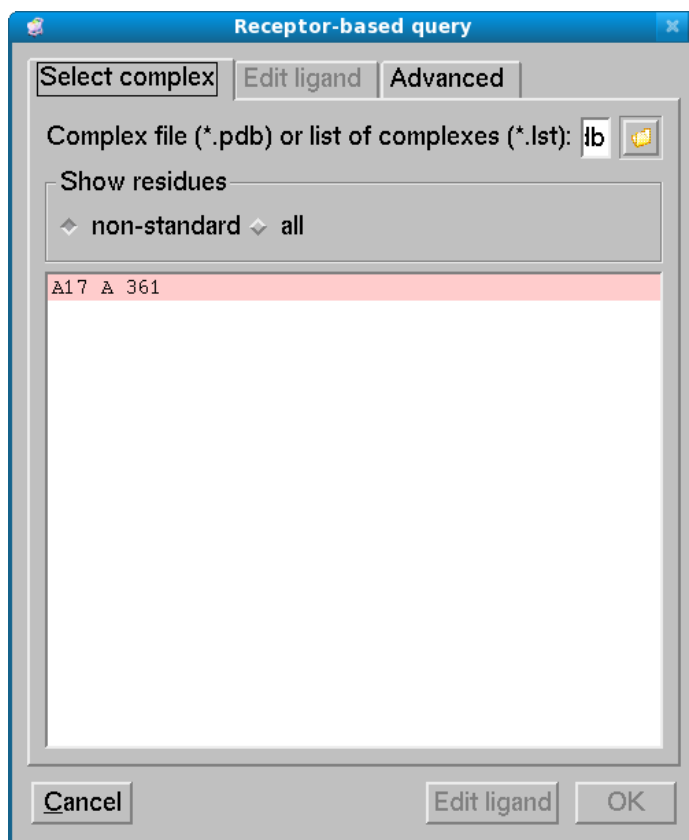
8150 scaffolds for 4 anchor points

6466 scaffolds for 5 anchor points

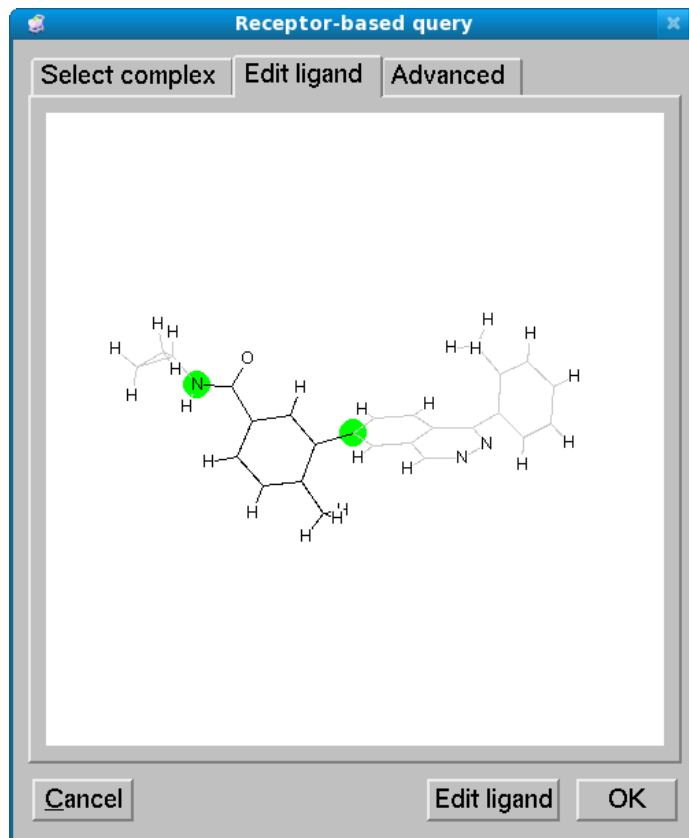
(Since the conformational search is done by a random methods the number of conformers could be slightly different on different systems)

Click on the *OK* to close the window.

Load a Receptor Based Query, which is a pdb file with the ligand structure without anchor points. In order to load a receptor based query select the Receptor Based Query from the Query menu (CTRL+O). Select the 3D6S_A.pdb

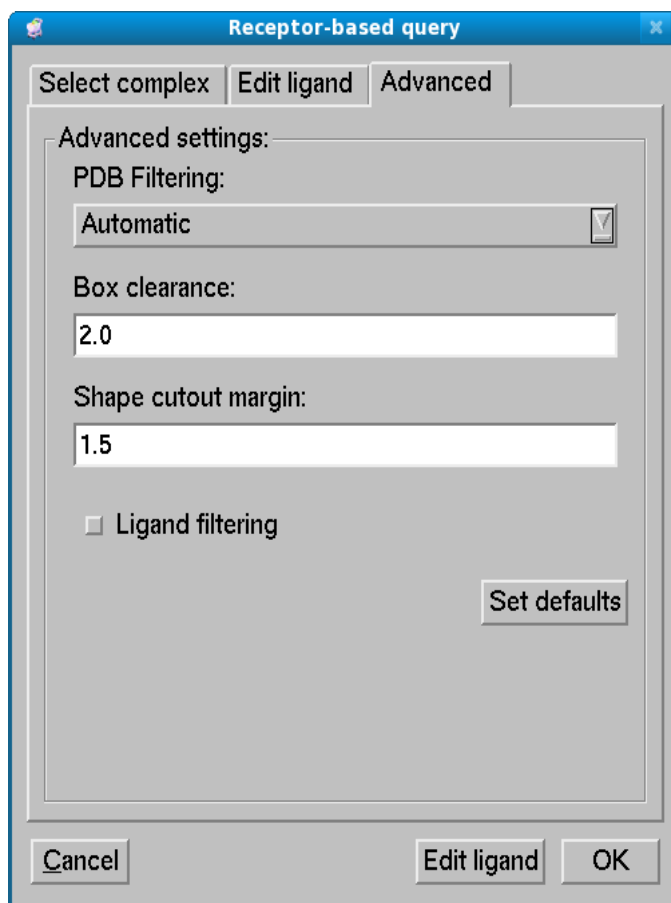


Select the A17 A 361 from the list of non-standard residues and click Edit Ligand.



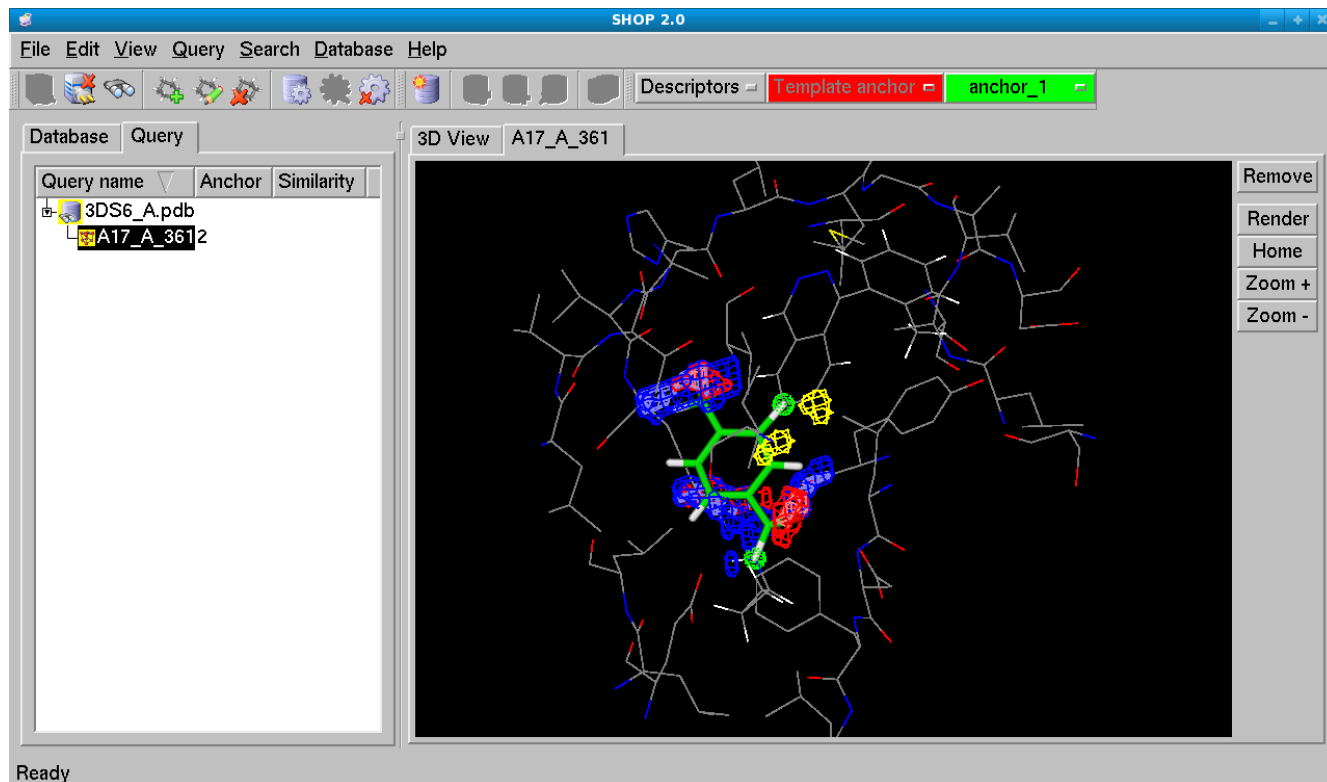
Since in this case the compound does not have an anchor point indicated, The Edit ligand tab is used to mark them, where the user has to click the atom that will be the attachment point. The procedure to mark the atom is by clicking once on the atom the biggest fragment is selected as the fragment to change, clicking a second time on the same atom, the second biggest fragment is selected. Select the fragment to replace as shown.

Click on the advanced tab and select the parameters as shown:




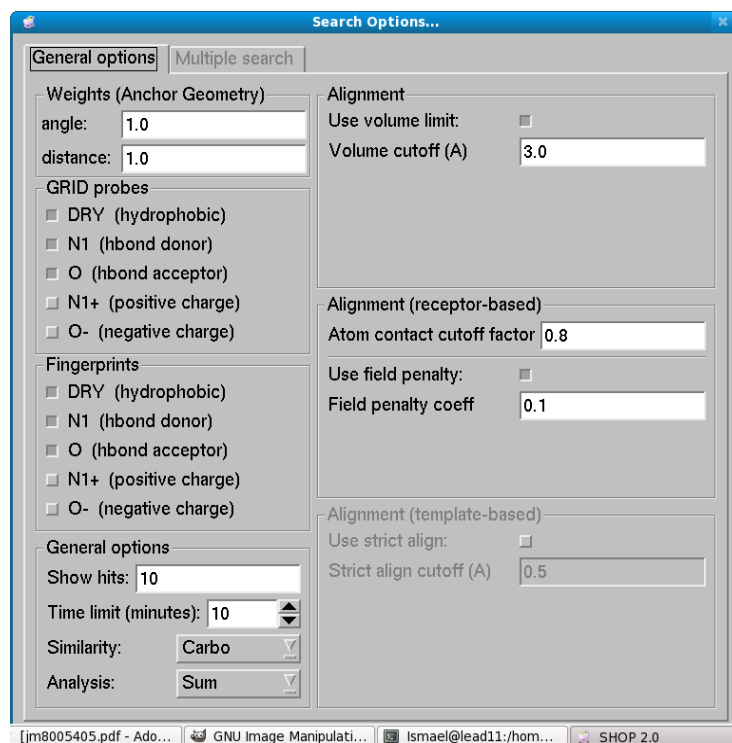
Visualize the query scaffold

In the Query tab click on the A17_A_361 tree. Click on A17_A_361 name from the query list (The anchor points are shown in green).

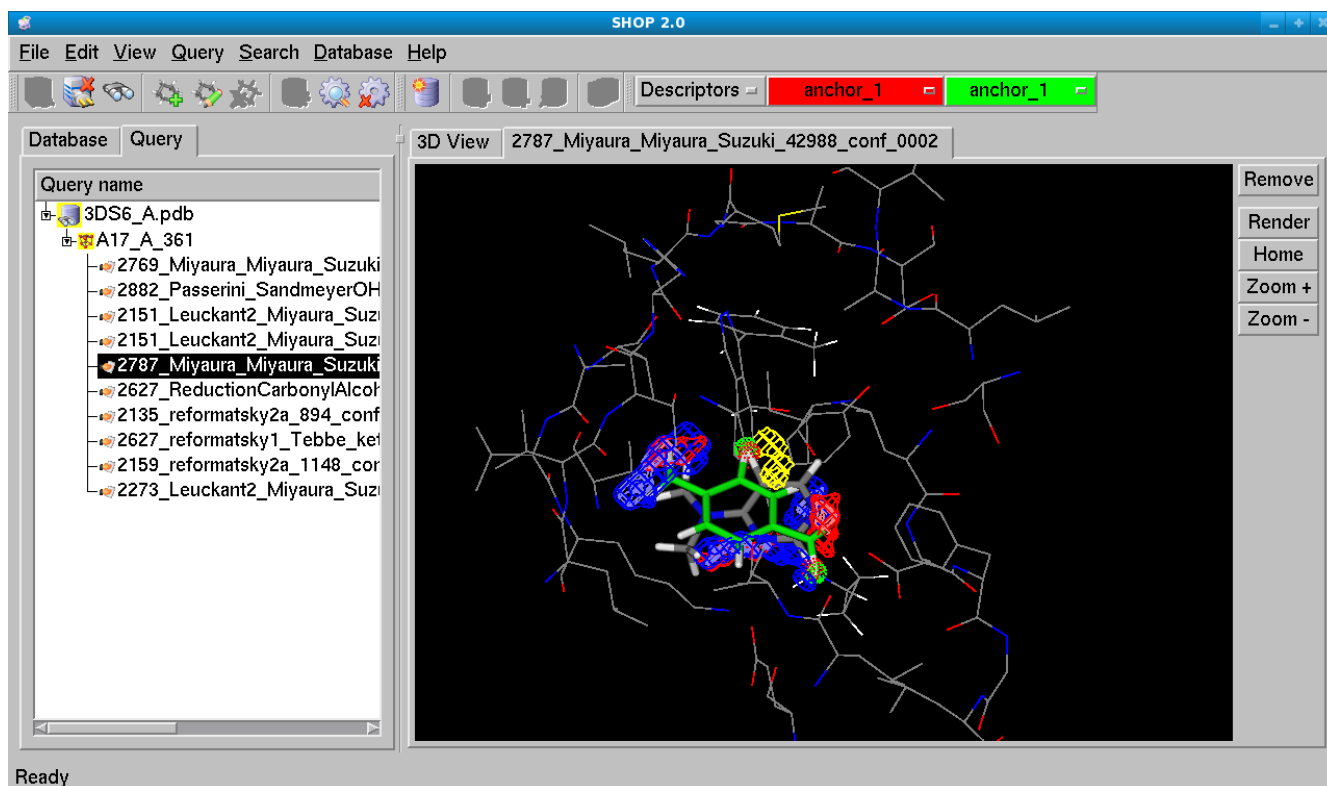


In this case the interactions are computed in the cavity without considering the ligand structure. It is observed that the protein offer more interactions that the ones fulfilled by the ligand in the crystal structure.

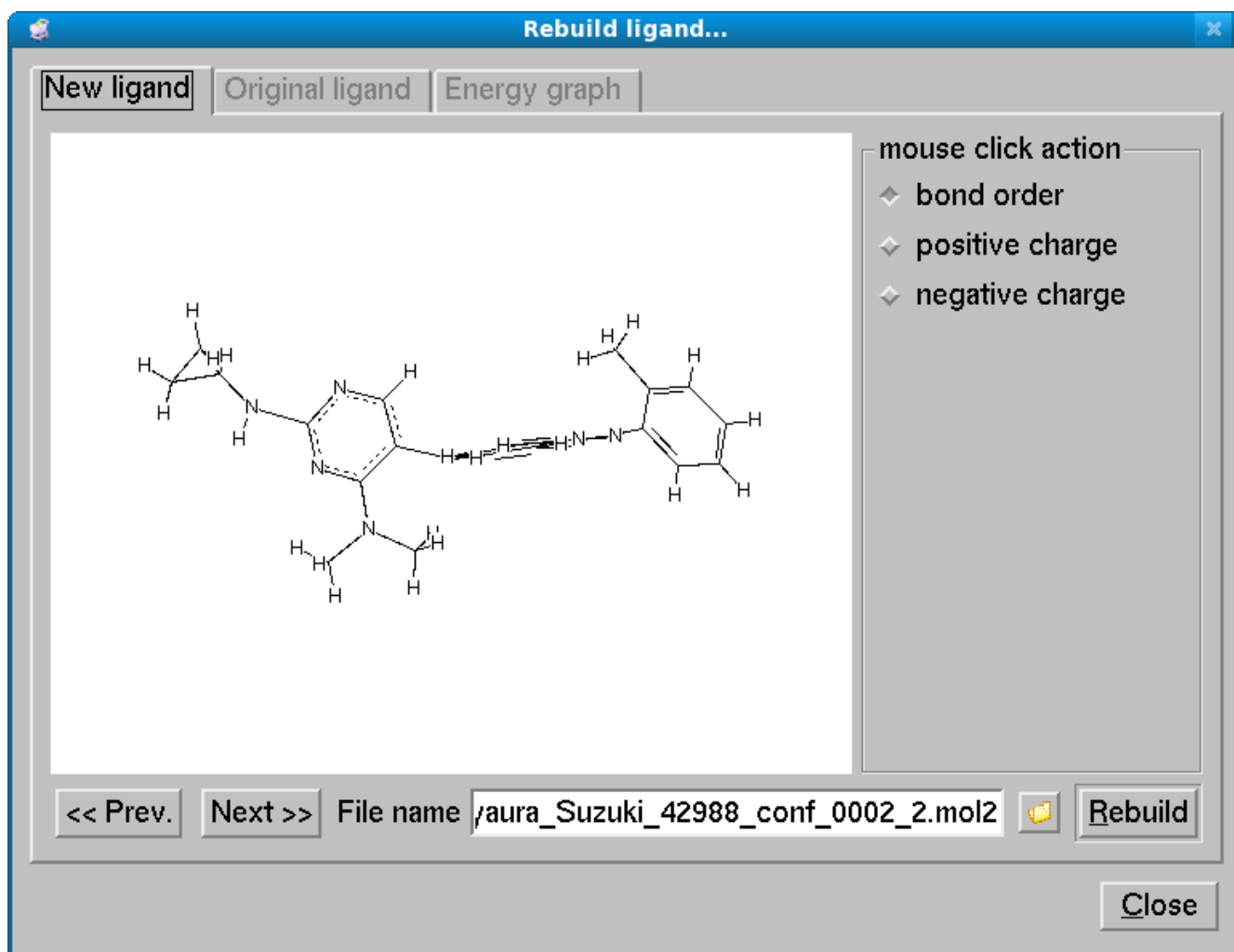
Perform a similarity search for the A17_A_361. Select the compound name in the *Query Compound List* and select the *Search* option (CTRL+S) or click on the search option button ().



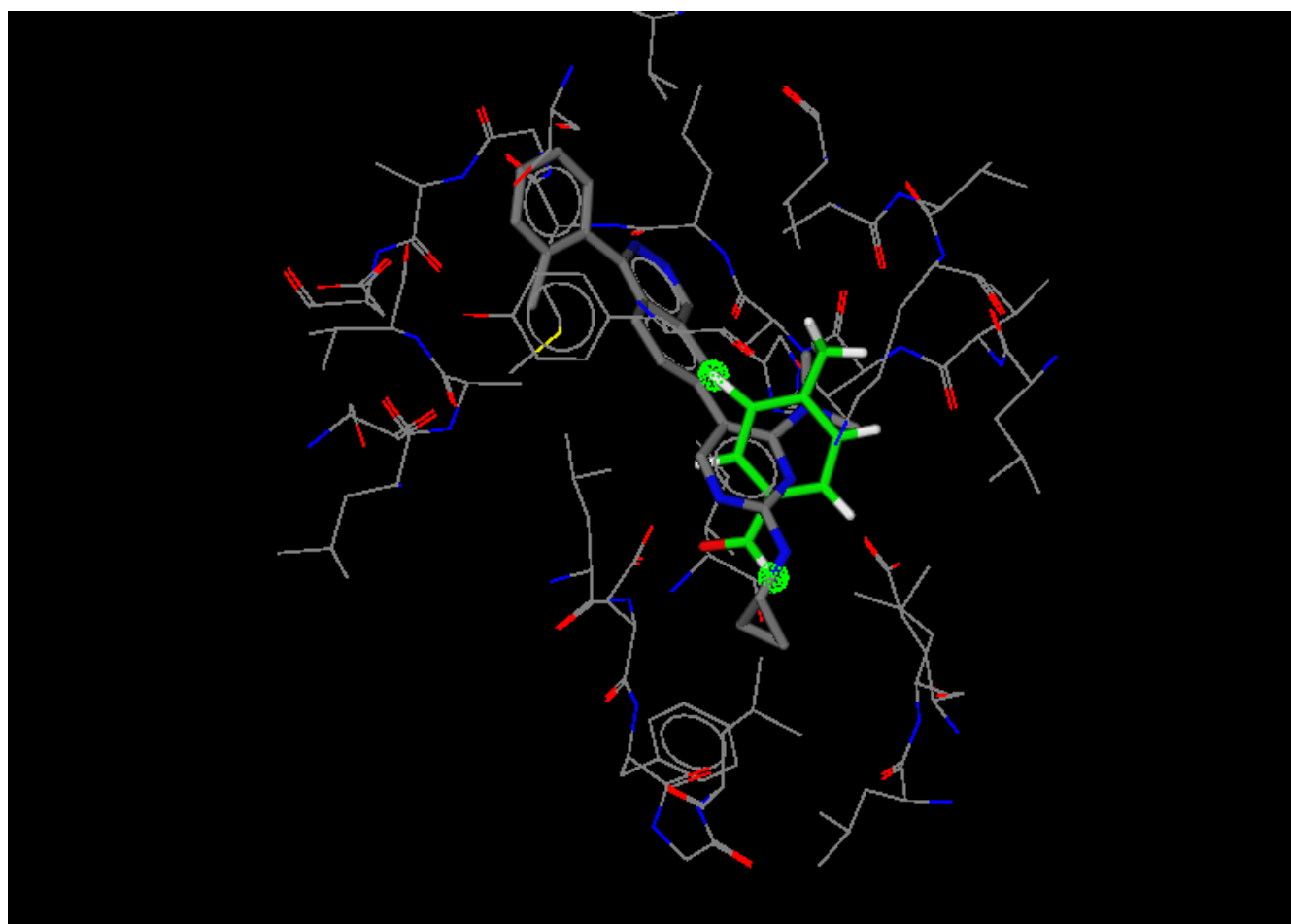
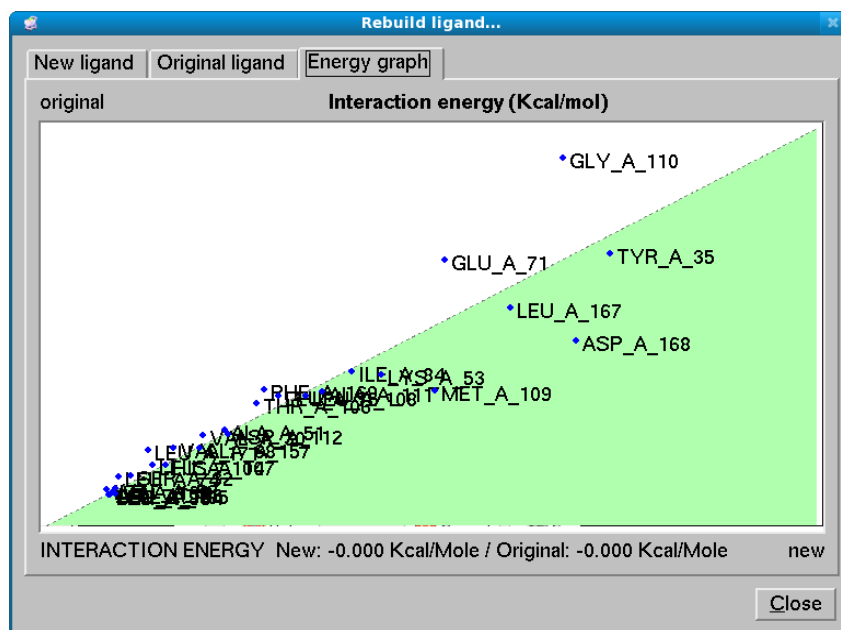
After 2-3 minutes the calculation is finished and the potential fragments to replace the one selected are shown in the query tab. Browse the solutions.



In the case that the initial query did not have the attachment point, Shop is able to remember the fragment that is kept fixed from the tool used to define the fragment to change. Therefore, it is possible to rebuild an entire molecule with the new solutions + the fragment that is kept, optimizing the new structure in the protein cavity. Select the **Search/Rebuild ligand** option and click on the Rebuild option.



After the rebuild process is finished the original ligand tab will be activated. Then, click on this tab and select analysis. When the computation is completed an analysis of the interaction of each molecule (the original one and the one under optimization) is shown.



After the process is finished the new optimized molecule in the protein cavity is saved in the 3D6S_A directory.

Exit the Shop Client by selecting.

[MENU] File→Exit

Calculate the ADME properties for the new suggestions

Import the molecules into MetaSite.

Original Molecule:

[MENU] File→Import and Select sdf files.

Import the 2d6s_MSite.sdf from the original folder.

Ligand based fragment replacement:

[MENU] File→Import and Select mol2 files.

The new file is located in Template folder and it is named as 2d6s_MSite_dummy.mol2. Use the same importing options as used in the first part of the tutorial.

Receptor based fragment replacement:

[MENU] File→Import and Select mol2 files.

The new file is located in Template folder and it is named as minimized_A17_A_361_2787_Miyaura_Miyaura_Suzuki_42988_conf_0002.mol2. If an error occur during the importation in MetaSite, edit the file with the NotePad and add the lines SMALL and NO_CHARGES as shown in the following example.

```
@<TRIPOS>MOLECULE
```

```
Ligand
```

```
54 58 1 0 0
```

```
SMALL
```

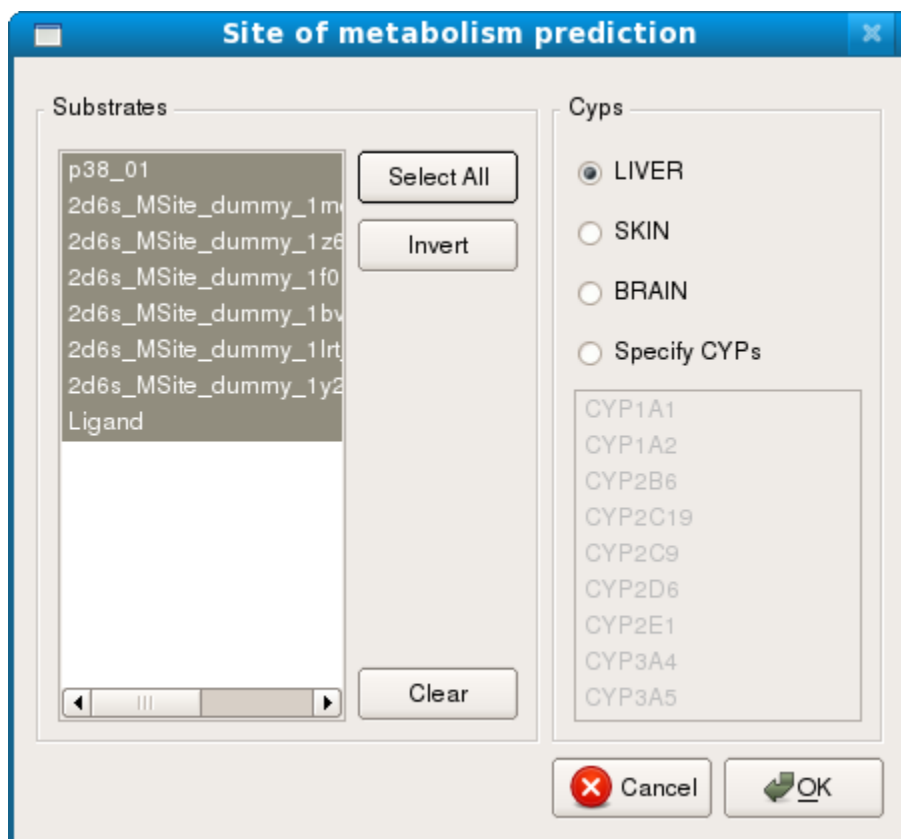
```
NO_CHARGES
```

Use the following importing options.

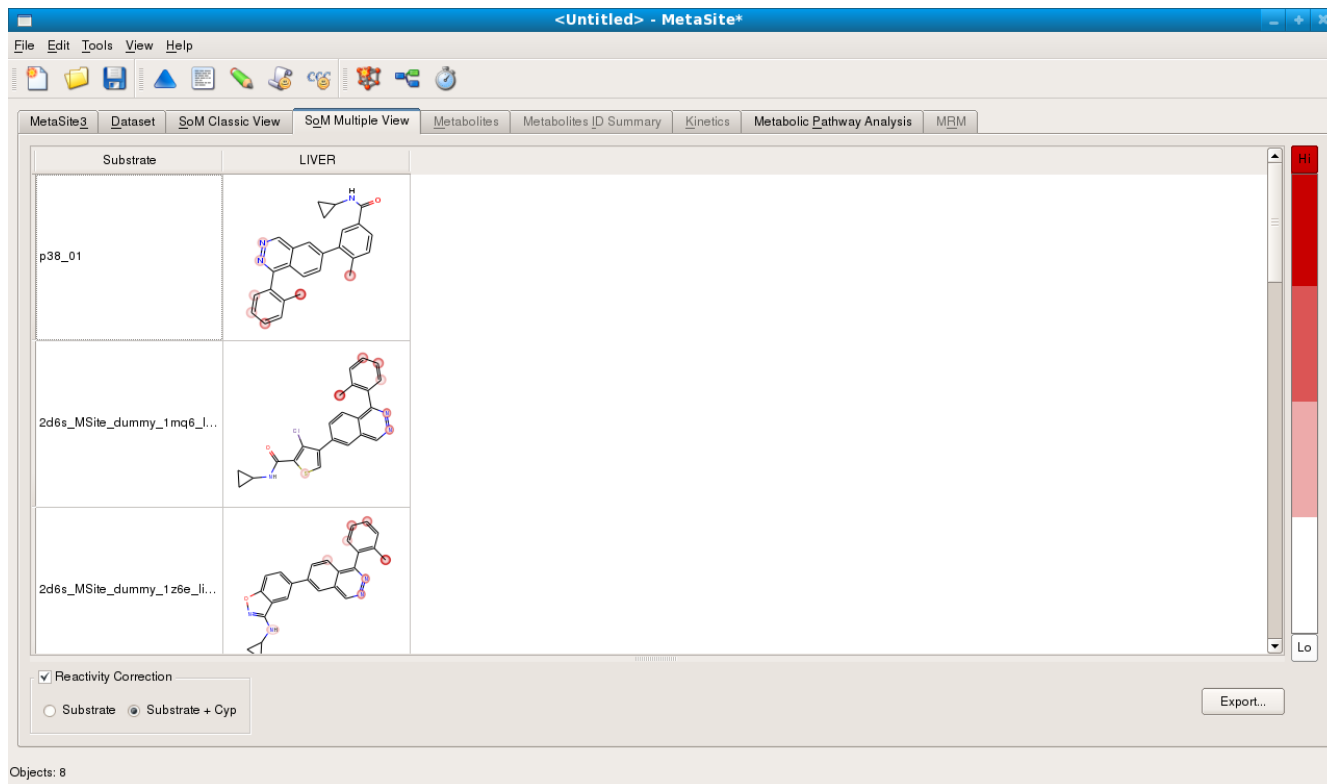
Use the same importing options as in the first step of this tutorial.

[MENU] Tools→Site of Metabolism Prediction.

Select All the compounds and the option LIVER and click OK.



Click on SoM Multiple View in order to compare if the site of oxidation has been modified compared to the initial compound. If that is the case, you have found a new chemical replacement for a fragment in your original molecule.



All the ligand based replacements of the selected fragment is not changing the region of the molecule where the predicted metabolism will take place (benzylic position), even with the bioisosteric replacement proposed by the authors, the same interaction with CYP3A4 is foreseen. In the case of the replacement suggested by the receptor based query, since it is not a bioisosteric replacement, but a complementarity interaction to the protein, the structural modification is more likely to disturb the interaction of the compound with the cytochrome.

Click in Export and select the name NewMolecules.sdf

Now you can close the MetaSite interface

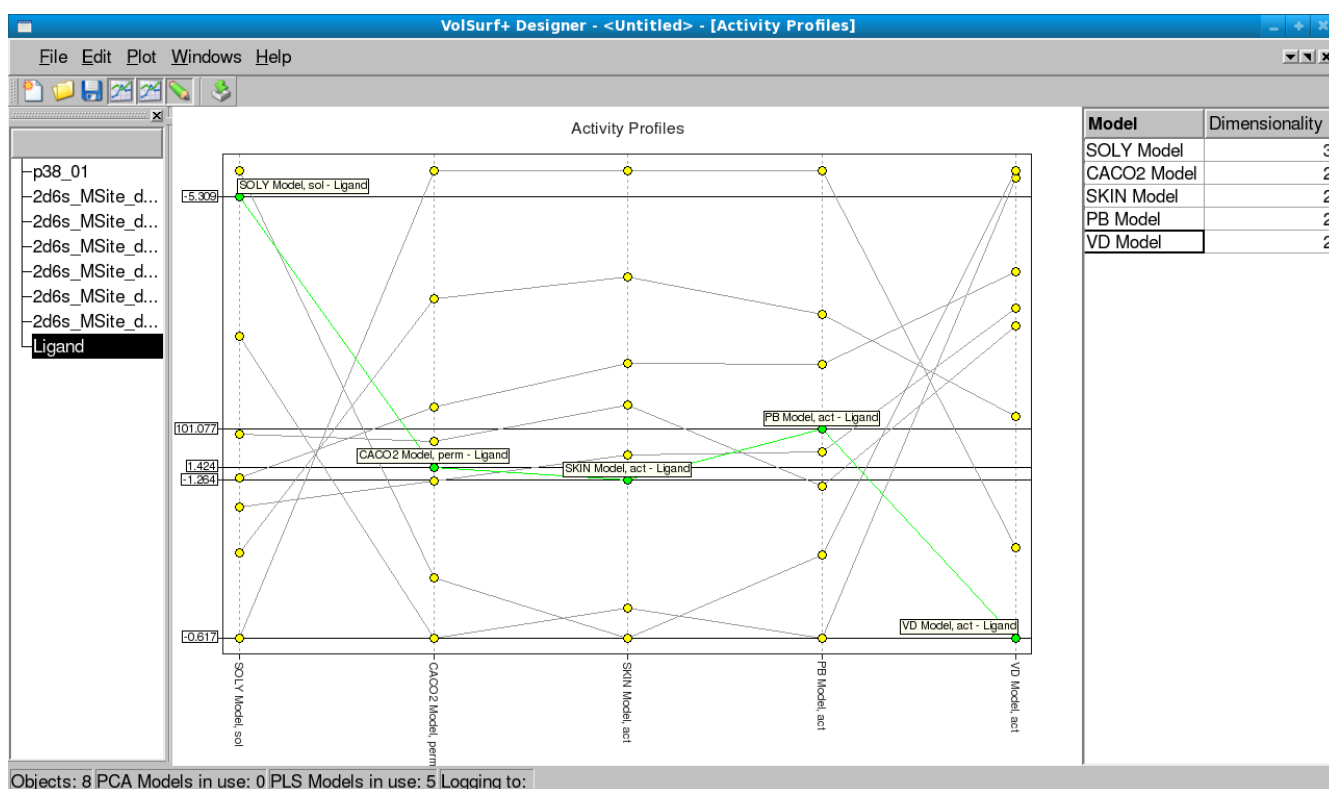
MENU] File→Exit.

Perform the calculation on the Volsurf+ models

[MENU] File→Import and Select mol2 files.

The new file is located in the folder you just exported all the compounds in sdf format from MetaSite and it is named as NewMolecules.mol2. Use the same importing options as used in the first part of the tutorial.

Right Click on the Activity profile window and select the scaling method Maximize, you will be able to see better the results



The new compounds are imported into the software and the predictions are done in the same models as have been selected. Are the new compound changing their ADME predicted properties? Are the changes going in the right direction?

Conclusion: There are solutions like the Ligand from the receptor based query that increase

solubility and decrease protein binding, while keeping permeability in the same range.

Now you can close the VolSurf+ interface

MENU] File→Exit.

In order to improve other properties one could re-do this analysis replacing the benzylic ring in the original molecule.