Table B – Drug Discovery Predictive Toxicology Application

An example of a predictive toxicology application in drug discovery is given using the data on antimalarial compounds made available at the ChEMBL Neglected Tropical Disease (NTD) archive (http://www.ebi.ac.uk/chemblntd/).

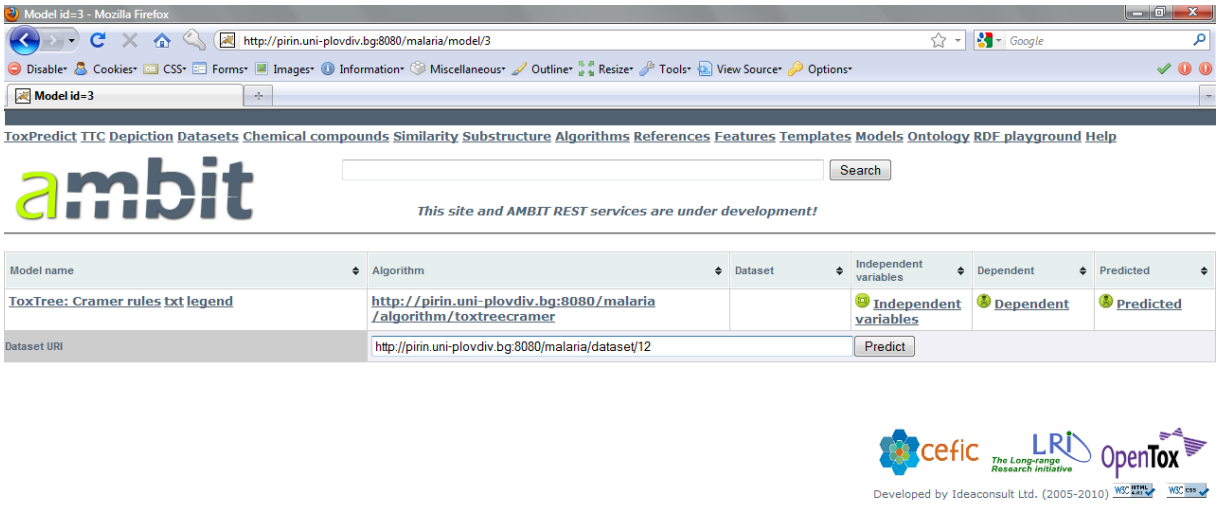
Activity A: Finding low toxicity drug candidates

In this exercise, the antimalarial compounds are prioritized based on a very conservative model for predicting oral toxicity. Experimentally determined cytotoxicities against human cells of the compounds predicted to be safe are further examined, and their mutagenicities predicted. Sites of cytochromeP450 metabolism are predicted for selected compounds with no mutagenicity alerts low human cytotoxicity, but high anti-malarial activity.

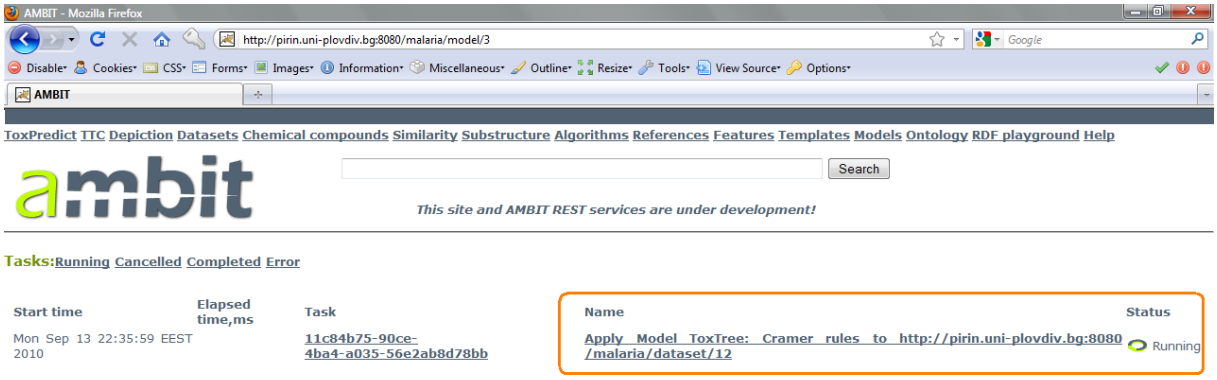
Step 1: Predicting Oral Toxicity

Go to the list of antimalarial datasets at http://pirin.uni-plovdiv.bg:8080/malaria/dataset. We'll first predict oral toxicity for the TCAMS dataset. Start by clicking on the TCAMS dataset link. The URL in the browser should read http://pirin.uni-plovdiv.bg:8080/malaria/dataset/12. You can browse the compounds.

In a new tab of your browser, go to the list of OpenTox models at http://pirin.uni-plovdiv.bg:8080/malaria/model (or follow the "Models" link at the top of the page listing the datasets). To predict oral toxicity we will use "Toxtree Cramer rules" model. Clicking on the Cramer rules link will open its page. OpenTox models accept dataset URLs as input (instead of file names). Enter (or paste) the TCAMS URL ("http://pirin.uni-plovdiv.bg:8080/malaria/dataset/12") into the text box . Click Predict.



Predict will launch calculations. You might click on the links to find out if the calculations are completed.



When completed, clicking on the link will lead to a dataset with the results.

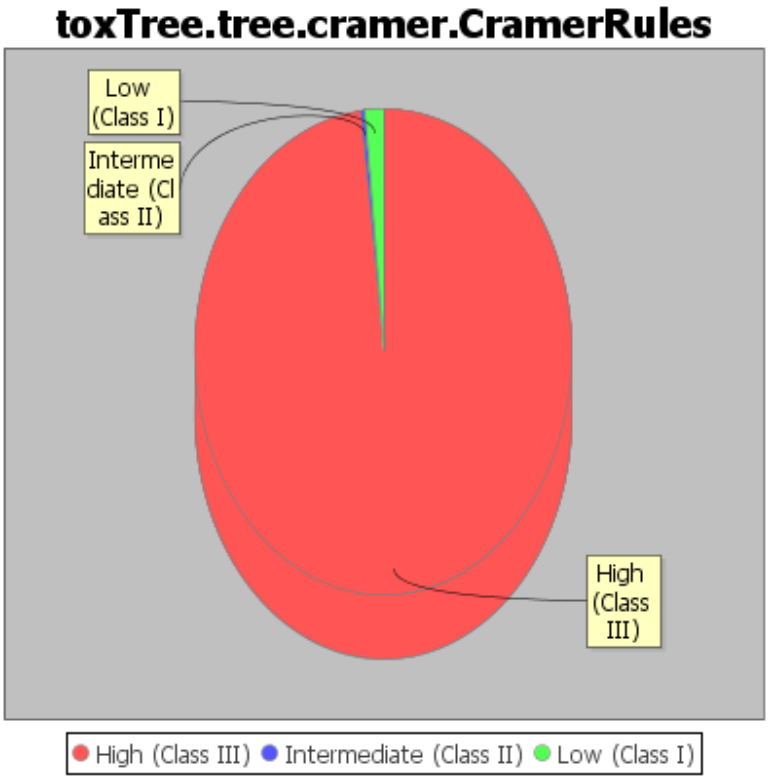
The Cramer rules model is an implementation of Cramer et al., *Estimation of Toxic Hazard – A Decision Tree Approach*, J Cosmet Toxicol, Vol. 16, pp. 255-276, Pergamon Press, 1978. It comprises 33 structural rules and places evaluated compounds into one of three classes:

- Class I substances are simple chemical structures with efficient modes of metabolism suggesting a low order of oral toxicity;
- Class III substances are those that permit no strong initial presumption of safety, or may even suggest significant toxicity or have reactive functional groups; and finally,
- Class II are intermediate. This model is very conservative and places most of the compounds in Class III.

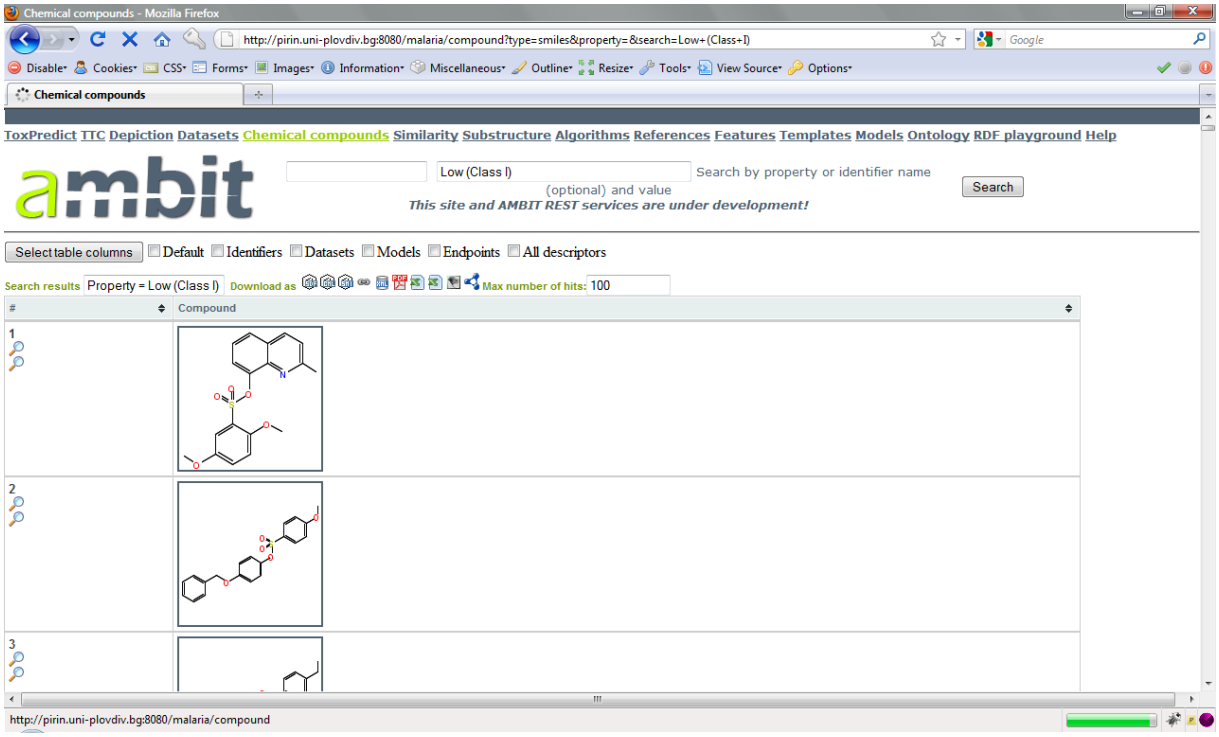
During this exercise, we'll look for compounds of low toxicity (Class I) and high antimalarial activity.

There are small number of Class I compounds, the distribution can be seen via the OpenTox chart generation service

http://pirin.uni-plovdiv.bg:8080/malaria/chart/pie?dataset_uri=http://pirin.uni-plovdiv.bg:8080/malaria/dataset/12&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/212



To filter for Class I compounds, click the "Chemical compounds" links on the top, and enter "Low (Class I)" in the search box. This results in the following web address http://pirin.uni-plovdiv.bg:8080/malaria/compound?type=smiles&property=&search=Low+%28Class+I%29 (which could be also used directly, instead of typing the search query in the text box). The results can be browsed as below.

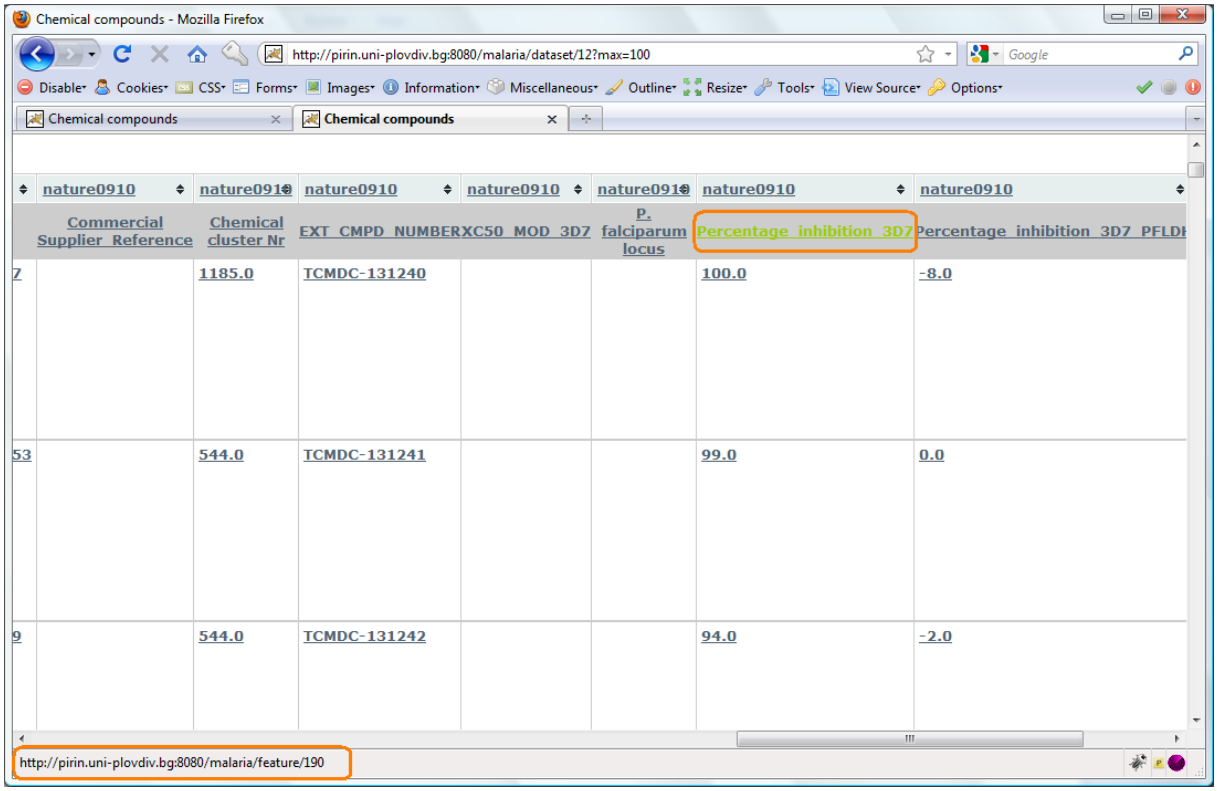


Step 2: Analyze Cytotoxicities of the Cramer Class I compounds

From the previous step we ended up with a list of compounds considered safe according to Cramer rules. However, we would like to have some more information than just the chemical structures. For example, we would like to know the antimalarial activity of these compounds.

To add such a column, we need to edit the URL by adding an entry denoting the antimalarial activity given in the TCAMS Dataset. All data columns in OpenTox have their unique URL, and in this example, the URL of the data indicating the percentage inhibition of the growth of the *P. falciparum* strain 3D7 (column "Percentage_Inhibition_3D7" in the TCAMS dataset) is

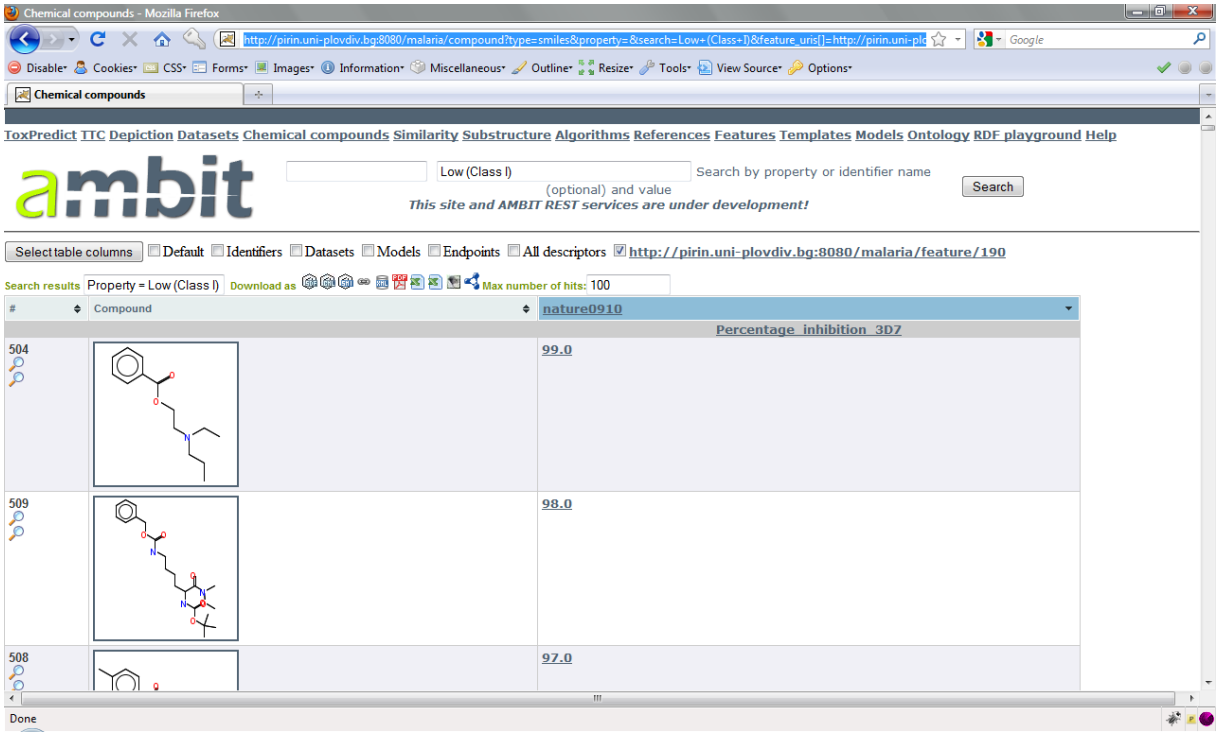
http://pirin.uni-plovdiv.bg:8080/malaria/feature/190



To add this column to our filtered list of compounds considered safe according to Cramer rules (Cramer class I), we simply add a feature_uris[] parameter to the URL of our filtered list:

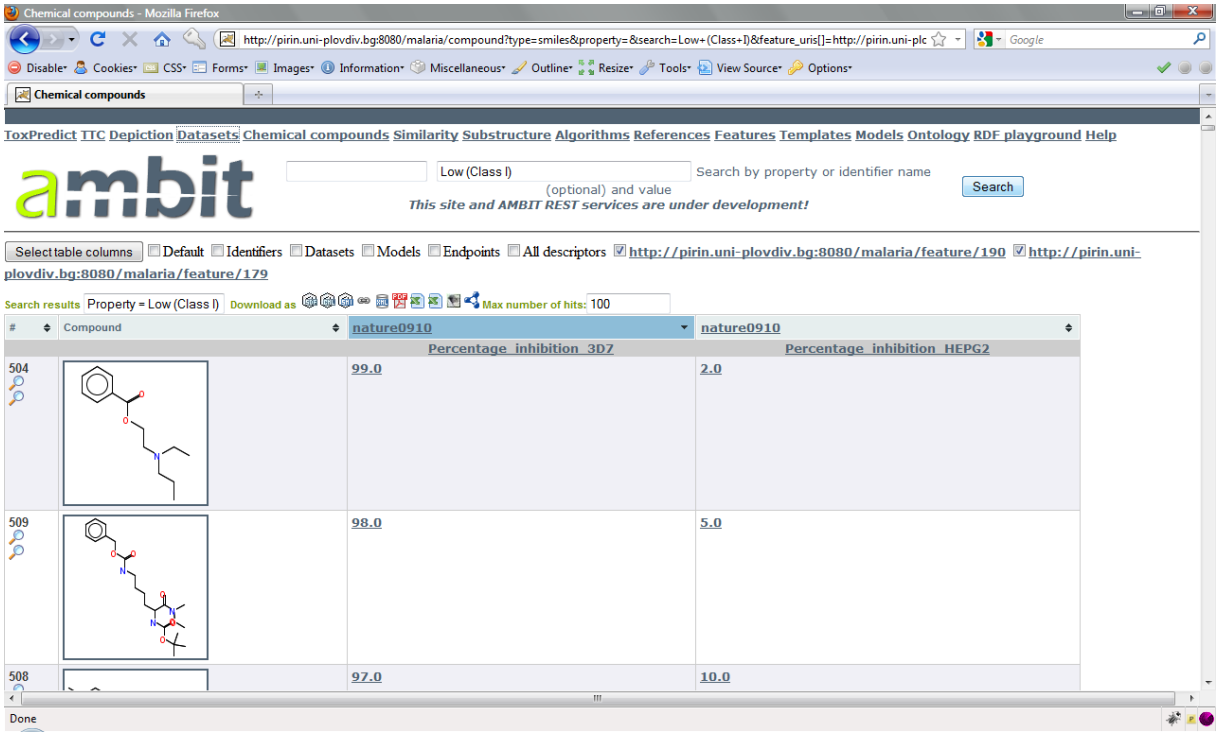
http://pirin.uni-plovdiv.bg:8080/malaria/compound?type=smiles&property=&search=Low+%28Class+I%29**&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/190**

Copy this address into the web browser. There will be a small number of nonempty entries in the second column.



We're not only interested in the antimalarial activity, but would also like to take into account the experimentally determined human cytotoxicity. To do so, weadd a second data column to our filtered list, now with human cytotoxicity data from TCAMS dataset (Percentage_inhibition_HEPG2 , http://pirin.uni-plovdiv.bg:8080/malaria/feature/179). The combination of the two features – antimalarial activity and human cytotoxicity – will result in the following URL:

http://pirin.uni-plovdiv.bg:8080/malaria/compound?type=smiles&property=&search=Low+%28Class+I%29&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/190**&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/179**



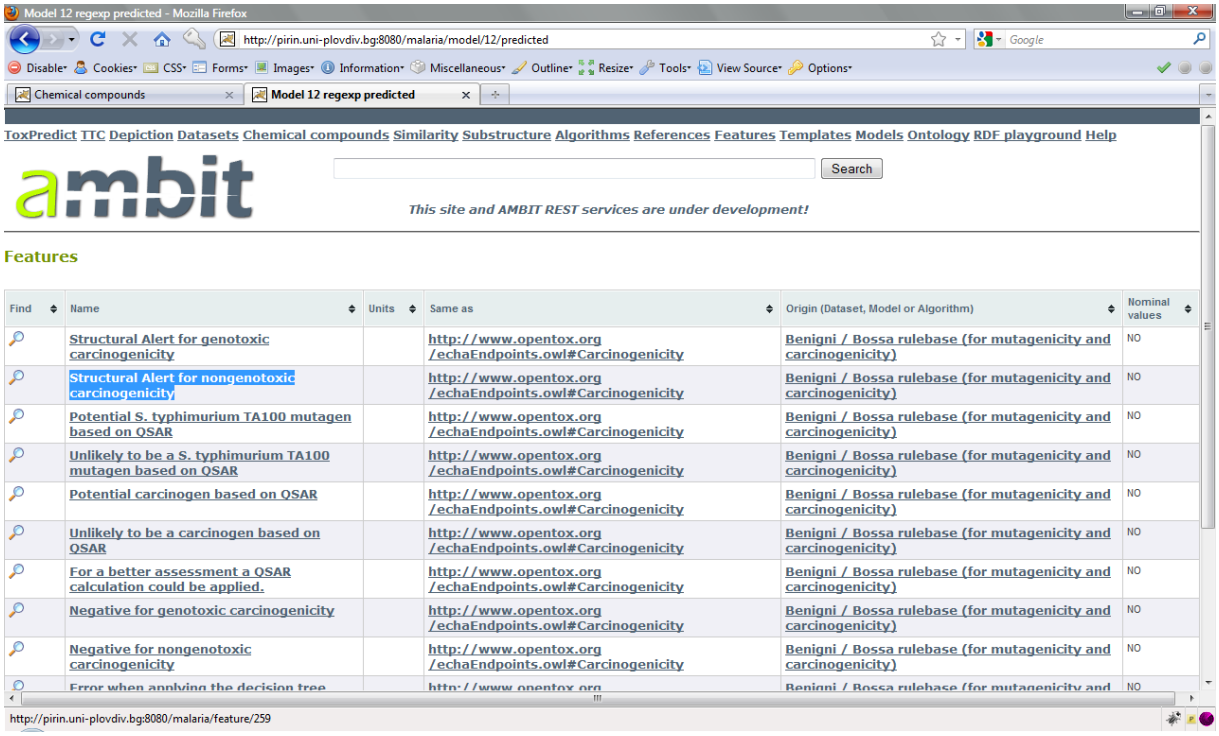
Step 3: Predicting the Mutagenicity of the Selected Compounds

To add a further criterion to be used when selecting our drug candidate, we predict the compounds' mutagenicities. To do so, we'll use the Toxtree Benigni/Bossa rules for mutagenicity and carcinogenicity (Benigni et al. 2007, *Mechanistic QSAR of aromatic amines: new models for discriminating between mutagens and nonmutagens, and validation of models for carcinogens*, Environ Mol Mutag **48**:754-771.). The URL of this model is http://pirin.uni-plovdiv.bg:8080/malaria/model/12.

Analogously as you have done for the Cramer rules, follow the URL of the Benigni/Bossa model (http://pirin.uni-plovdiv.bg:8080/malaria/model/12), type or paste the URL or the TCAMS dataset (http://pirin.uni-plovdiv.bg:8080/malaria/dataset/12) in the text box and click "Predict". Alternatively, the URL of the filtered list could be entered here, as well.

OpenTox models store the prediction results again under data columns with unique URL. These are available via http://host/model/{id}/predicted , which in our example corresponds to

http://pirin.uni-plovdiv.bg:8080/malaria/model/12/predicted



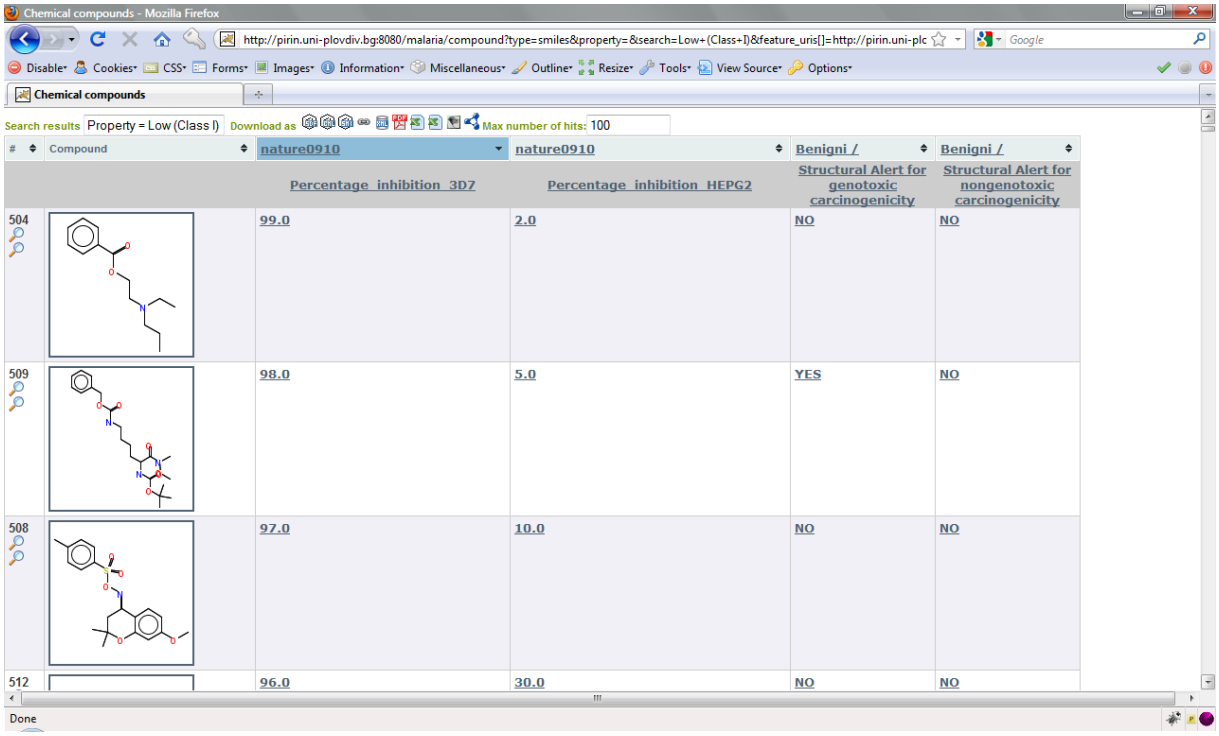
The Toxtree mutagenicity and carcinogenicity model predicts whether there are structural alerts for genotoxic or nongenotoxic carcinogenicity, and also uses linear discriminant model for specific classes of compounds.

For our purpose, we select the columns "Structural Alert for genotoxic carcinogenicity" (http://pirin.uni-plovdiv.bg:8080/malaria/feature/258) and "Structural Alert for nongenotoxic carcinogenicity" (http://pirin.uni-plovdiv.bg:8080/malaria/feature/259). As before, we add data columns for these structural alerts to our Cramer-class filtered list of compounds, again using the feature_uris[] method. The resulting URL is:

http://pirin.uni-plovdiv.bg:8080/malaria/compound?type=smiles&property=&search=Low+(Class+I)&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/190&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/179&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/258&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/259

The resulting table (as well as any other) can be sorted according to the values in any column by clicking on the column header.

In the following examples, we'll consider the first compound in the image below as our antimalarial drug candidate. It is a Cramer class I compound that inhibits growth of *P. falciparum* 3D7 by 99% at the concentration tested (2μM), has a very low human cytotoxicity and no structural alerts for carcinogenicity. (You may choose a different compound).



Similar to datasets and models, also each compound inOpentox services has its unique URL. You find the URL of a compound by clicking on its 2D structure, and stripping off the "?media=text/html" part at the end of the URL this brings you to.

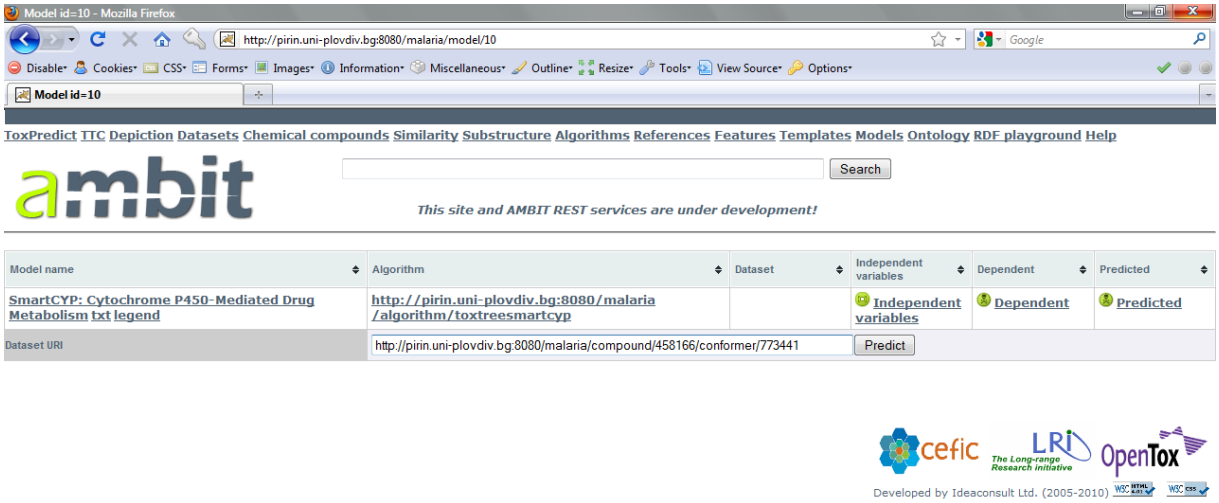
The URL of the compound selected above is
http://pirin.uni-plovdiv.bg:8080/malaria/compound/458166/conformer/773441.

Step 4: Predicting Sites of Cytochrome P450 Metabolism

The URL of our drug candidate will be used to submit this compound to two models predicting cytochrome P450 sites of metabolism, namely
SmartCYP [1] (http://pirin.uni-plovdiv.bg:8080/malaria/model/10) and
SOME [2] (http://pirin.uni-plovdiv.bg:8080/malaria/model/21)

Model prediction is done analogously to the two models used in this exercise. Go to
http://pirin.uni-plovdiv.bg:8080/malaria/model/10 and copy the compound URL into the text box.

[1] RydbergP et al. SMARTCyp: A 2D Method for Prediction of CytochromeP450-Mediated Drug Metabolism. ACS Medicinal Chemistry Letters20101(3), 96-100
[2] 7ZhengM et al. Site of metabolism prediction for six biotransformationsmediated by cytochromesP450. Bioinformatics (2009) 25(10): 1251-1258

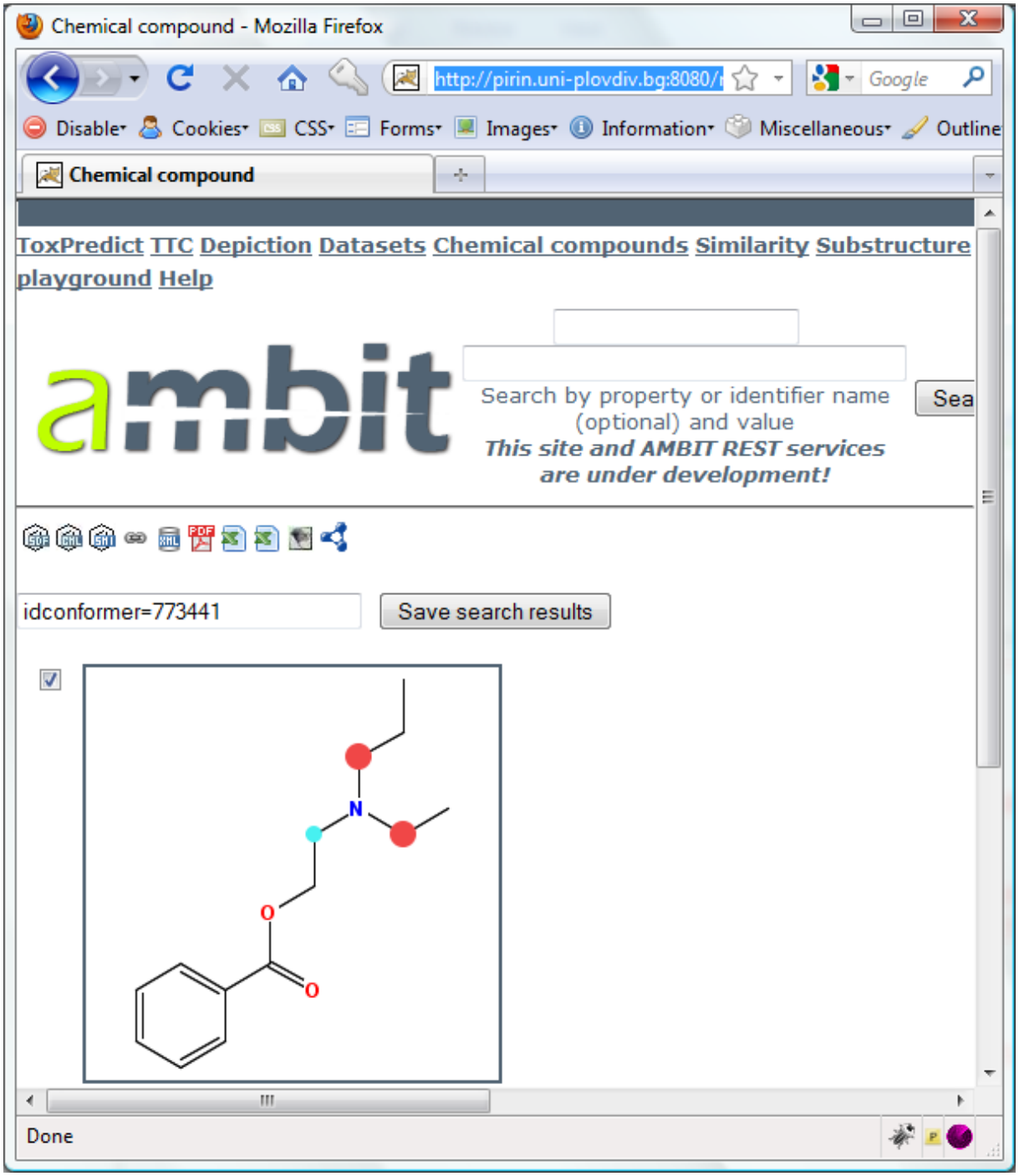


When completed, the results will be available at

http://pirin.uni-plovdiv.bg:8080/malaria/compound/458166/conformer/773441?feature_uris[]=http%3A%2F%2Fpirin.uni-plovdiv.bg%3A8080%2Fmalaria%2Fmodel%2F10%2Fpredicted

and will consists of information which atoms are of rank 1, 2, 3 or lower. Higher rank means more labile site. This information will be best viewed graphically, which could be achieved by the following URL

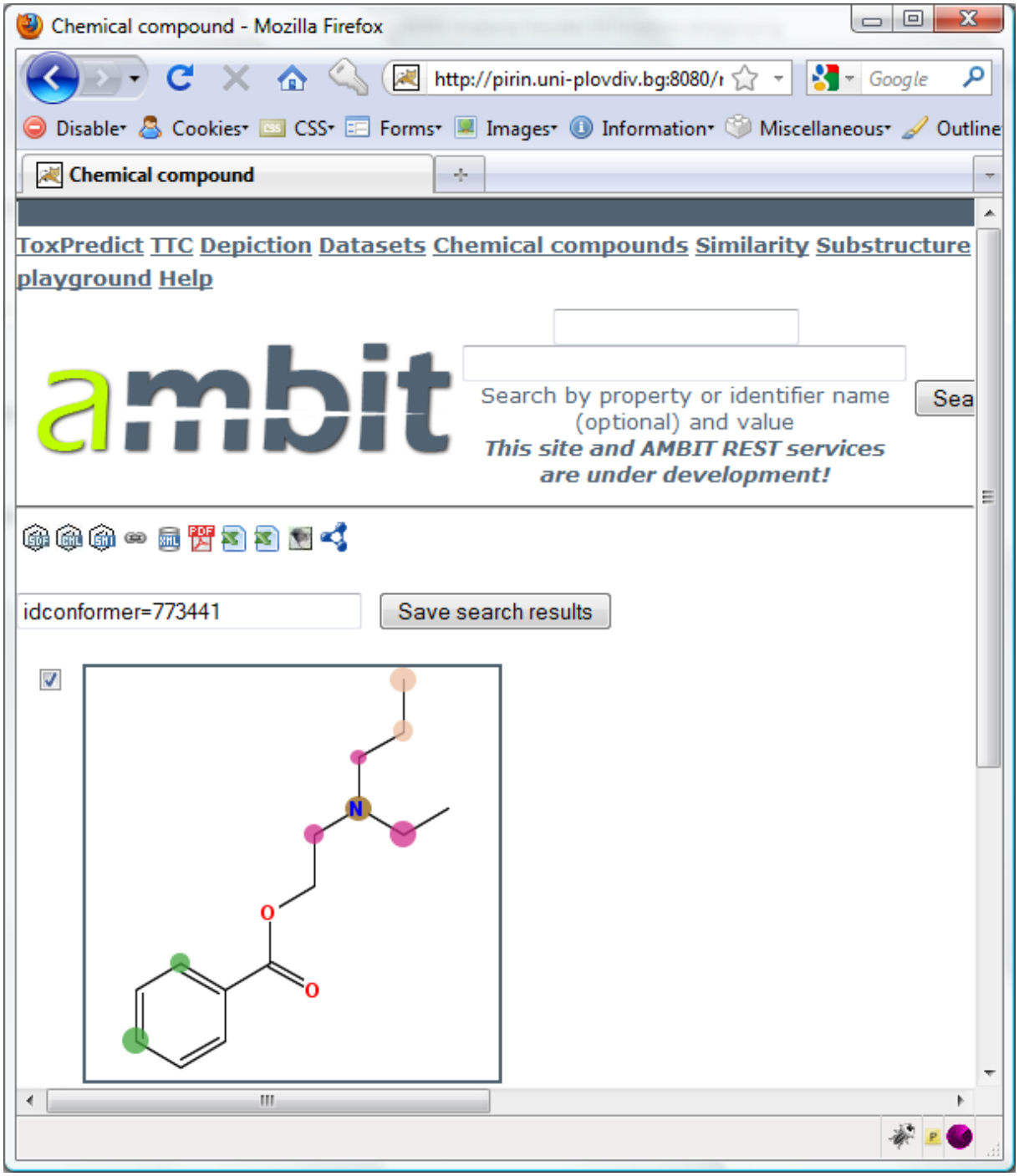
http://pirin.uni-plovdiv.bg:8080/malaria/compound/458166/conformer/773441?model_uri=http://pirin.uni-plovdiv.bg:8080/malaria/model/10



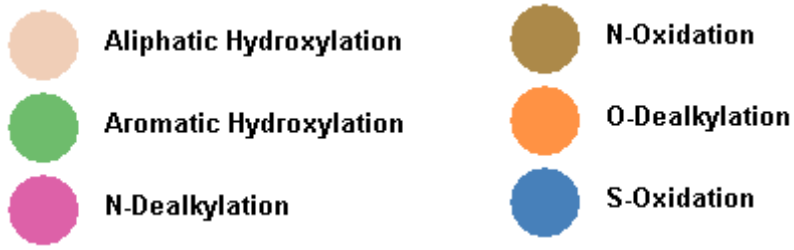
The color code can be found by clicking on the "legend" link on model page.

Similarly, the SOME model predictions are visualized via

http://pirin.uni-plovdiv.bg:8080/malaria/compound/458166/conformer/773441?model_uri=http://pirin.uni-plovdiv.bg:8080/malaria/model/21

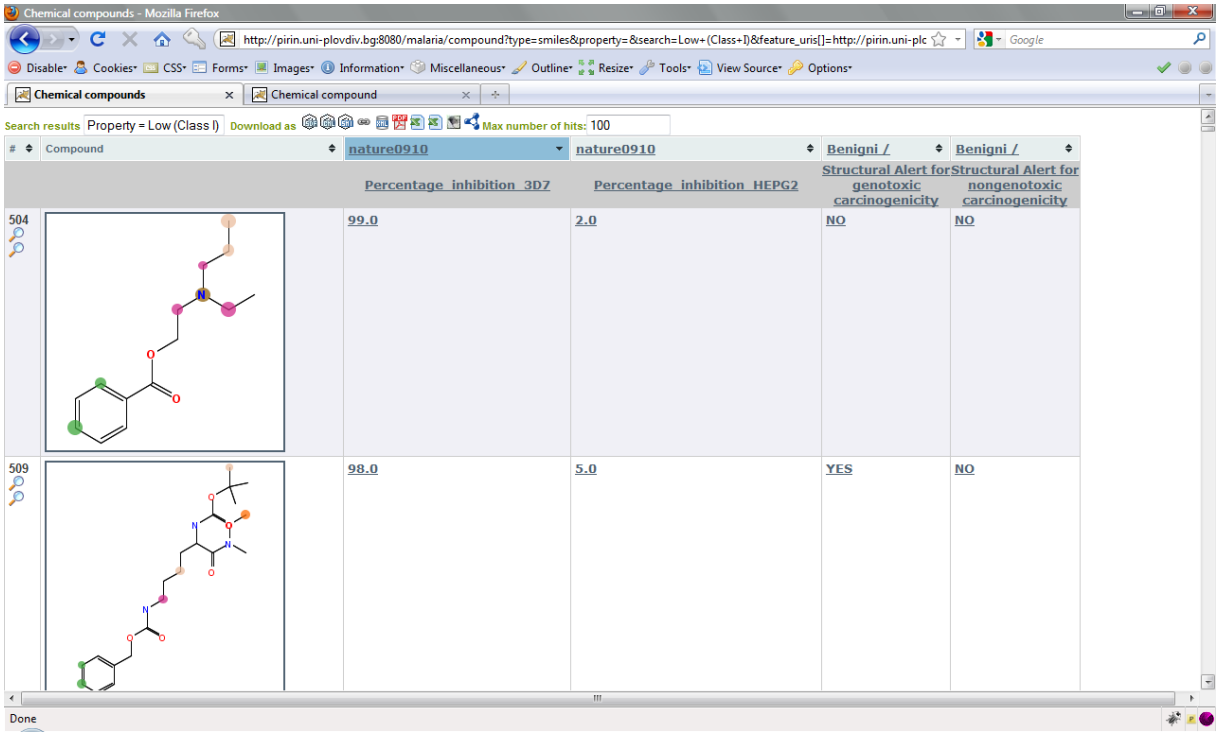


And the color code is



The color-coding of the metabolic sites according to the type of metabolic reaction taking place allows – with a little knowledge in organic chemistry – the work out the metabolites of the compound.

To obtain the predictions of the sites of metabolism for the entire dataset, use the following URL: http://pirin.uni-plovdiv.bg:8080/malaria/compound?type=smiles&property=&search=Low+%28Class+I%29&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/190&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/179&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/258&feature_uris[]=http://pirin.uni-plovdiv.bg:8080/malaria/feature/259&model_uri=http://pirin.uni-plovdiv.bg:8080/malaria/model/21&w=250&h=250



Disussion topics:

- Did we find a promising drug candidate?
- What kind of screening/toxicity assessment would be useful to apply as a next step?
- Comments/suggestions?