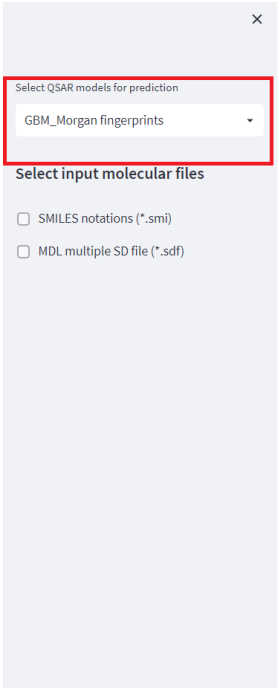
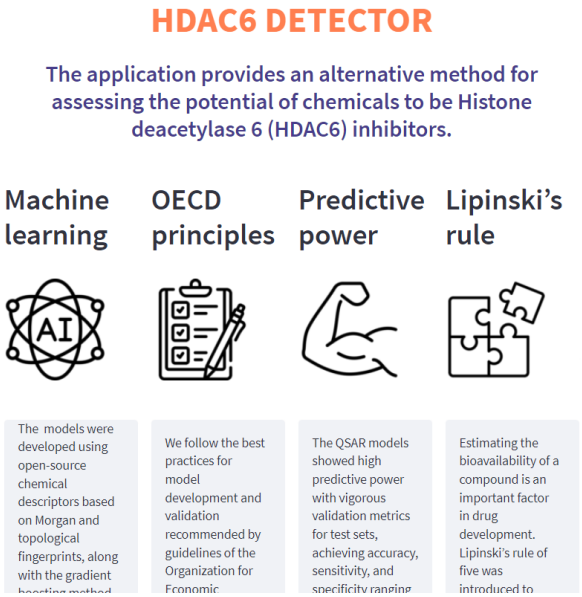
Manual of HDAC6 DETECTOR

open-source software

Step 1. *Select QSAR models for prediction*.

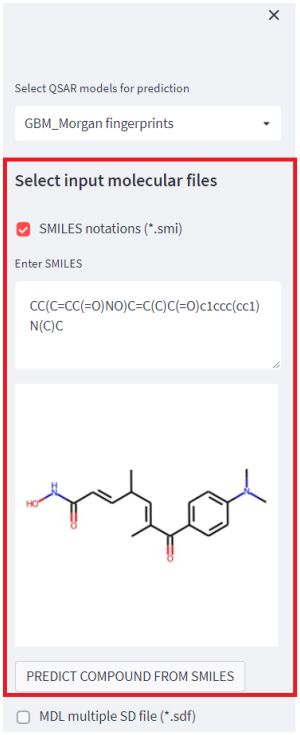
You can choose a model developed using the Morgan fingerprints and gradient boosting method (GBM) or a model developed using topological fingerprints, and multilayer perceptron (MLP) classifier.

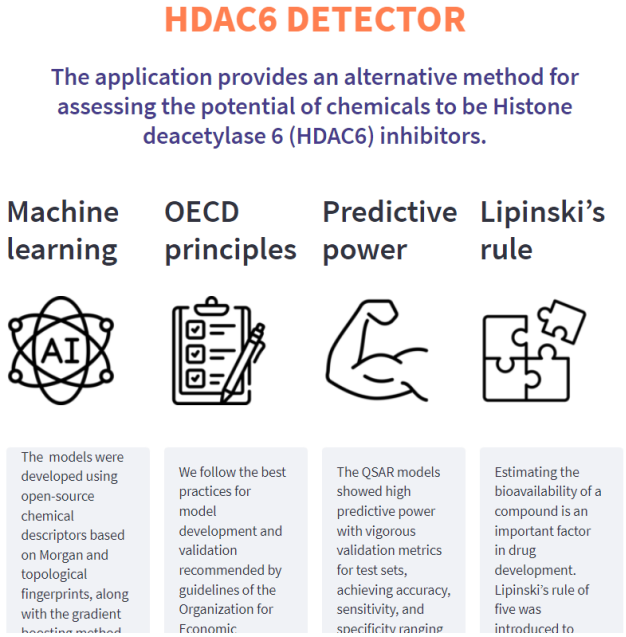




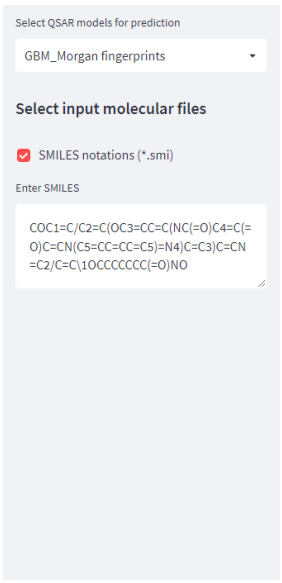
Step 2. *Select input molecular files.*

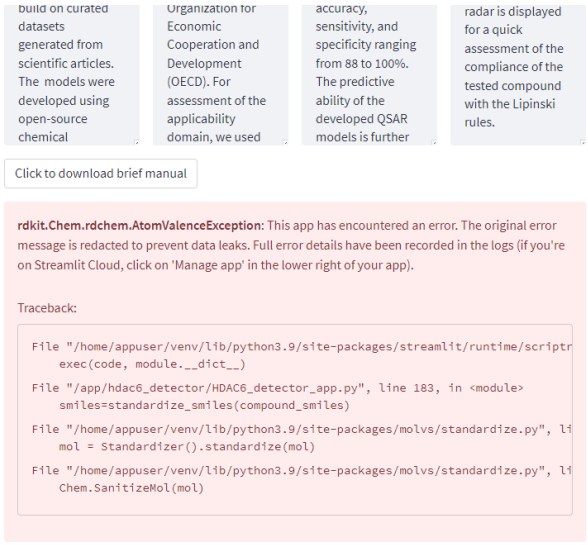
If you choose smiles, please, directly paste the SMILES representation of the desired chemical structure and press Ctrl+Enter. If the entered chemical structure is correct, the application will generate a 2D image of the studied compound. Click on the "PREDICT COMPOUND FROM SMILES" button.



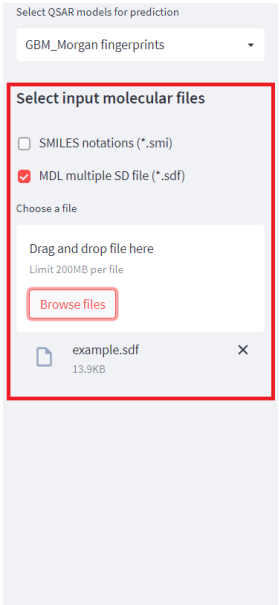


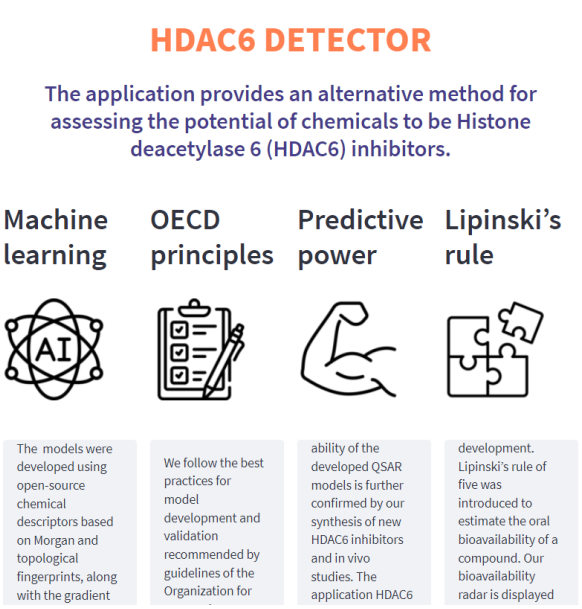
If the entered structure is incorrect, the application reports an error that has occurred.



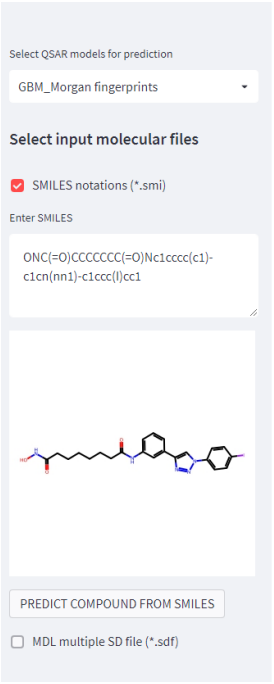


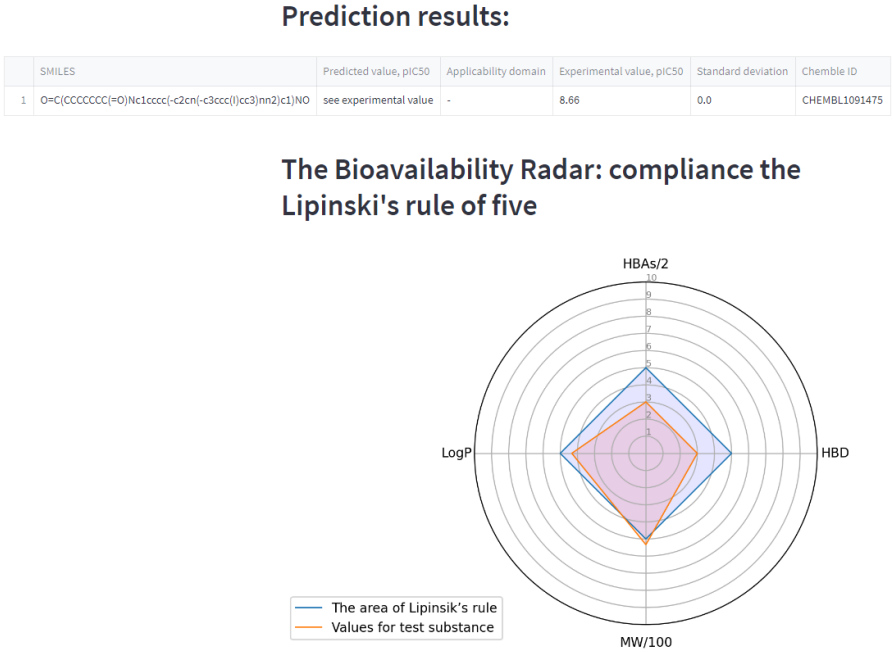
If you choose a file *sdf, that may contain a different number of chemical structures, please specify the path to this file on your computer's hard drive. In this case, you need to click the " Browse files" button.





Step 3. *Prediction results*.





The final table contains the following columns:

1) ***SMILES*** - the chemical structure is displayed in the SMILES notation

2) ***Predicted value, pIC50*** - the predicted value of the activity to inhibit the HDAC6 enzyme, expressed in pIC_{50}, where pIC_{50} is the negative logarithm of IC_{50} in molar concentration. If experimental data is available in the ChEMBL database, the label "see experimental value" is displayed in this cell.

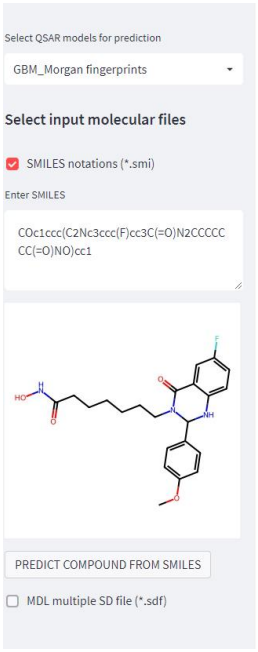
3) ***Applicability domain*** - compliance of the chemical compound with Applicability domain. If experimental data is available in the ChEMBL database, the label "-" is displayed in this cell.

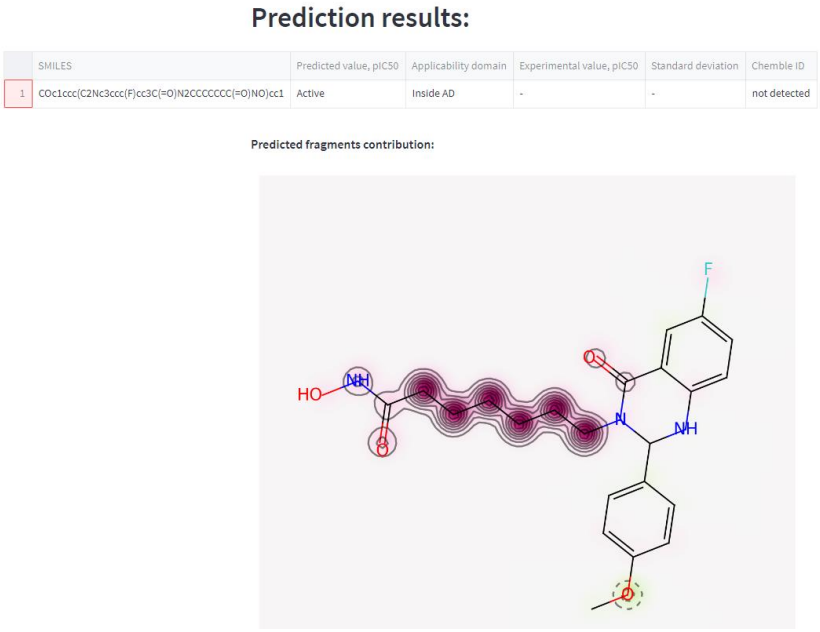
4) ***Experimental value, pIC50***- experimental data presented in the ChEMBL database. Where pIC_{50} is the negative logarithm of IC_{50} in molar concentration. If there is more than one value in the database, then the average value is given in the cell.

5) ***Standard deviation*** - ehe cell indicates the standard deviation of the experimental activity values presented in the ChEMBL database.

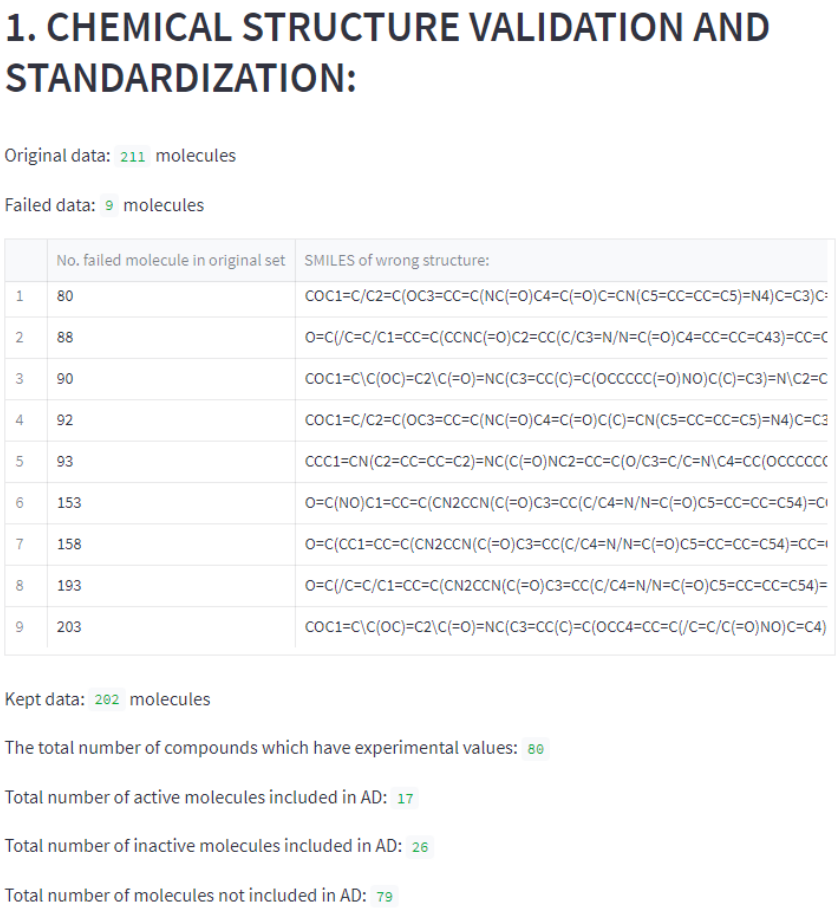
6) ***Chemble ID*** - identifier from the ChEMBL chemical database of molecule

The form of presentation of the results depends on the type of descriptors selected, as well as the format of the input chemical data. For example, when selecting SMILES, the results will be displayed for a single molecule. The chemical fragments are colored in green (predicted to reduce inhibitory activity) or magenta (predicted to increase activity HDAC6 inhibitors). The gray isolines separate positive and negative contributions.





If you select a file *sdf, the results will be presented in tabular form. If incorrect structures are detected in the file *sdf, the corresponding information will appear in the section "CHEMICAL STRUCTURE VALIDATION AND STANDARDIZATION"



1. CHEMICAL STRUCTURE VALIDATION AND STANDARDIZATION:

Original data: `211` molecules

Failed data: `9` molecules

	No. failed molecule in original set	SMILES of wrong structure:
1	80	COC1=C/C2=C(OC3=CC=C(NC(=O)C4=C(=O)C=CN(C5=CC=CC=C5)=N4)C=C3)C=
2	88	O=C(/C=C/C1=CC=C(CCNC(=O)C2=CC(C/C3=N/N=C(=O)C4=CC=CC=C43)=CC=C
3	90	COC1=C\C(OC)=C2\C(=O)=NC(C3=CC(C)=C(OCCCCC(=O)NO)C(C)=C3)=N\C2=C
4	92	COC1=C/C2=C(OC3=CC=C(NC(=O)C4=C(=O)C(C)=CN(C5=CC=CC=C5)=N4)C=C3
5	93	CCC1=CN(C2=CC=CC=C2)=NC(C(=O)NC2=CC=C(O/C3=C/C=N\C4=CC(OCCCCC(
6	153	O=C(NO)C1=CC=C(CN2CCN(C(=O)C3=CC(C/C4=N/N=C(=O)C5=CC=CC=C54)=C(
7	158	O=C(CC1=CC=C(CN2CCN(C(=O)C3=CC(C/C4=N/N=C(=O)C5=CC=CC=C54)=CC=(
8	193	O=C(/C=C/C1=CC=C(CN2CCN(C(=O)C3=CC(C/C4=N/N=C(=O)C5=CC=CC=C54)=
9	203	COC1=C\C(OC)=C2\C(=O)=NC(C3=CC(C)=C(OCC4=CC=C(/C=C/C(=O)NO)C=C4)

Kept data: `202` molecules

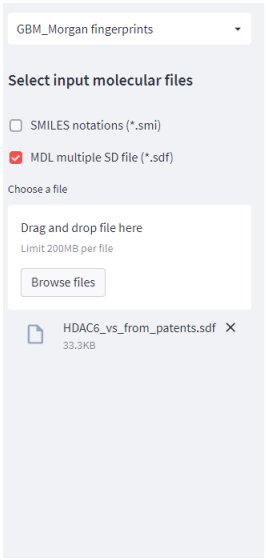
The total number of compounds which have experimental values: `80`

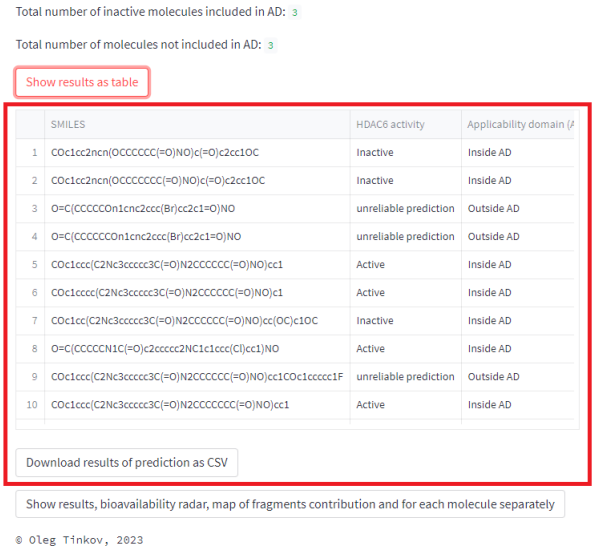
Total number of active molecules included in AD: `17`

Total number of inactive molecules included in AD: `26`

Total number of molecules not included in AD: `79`

The prediction results for correct chemical structures are displayed in a table that can be downloaded, or separately for each molecule if you click the " Show results and map of fragments contribution for each molecule separately" button.





	SMILES	HDAC6 activity	Applicability domain (A
1	COc1cc2ncn(OCCCCCC(=O)NO)c(=O)c2cc1OC	Inactive	Inside AD
2	COc1cc2ncn(OCCCCCCC(=O)NO)c(=O)c2cc1OC	Inactive	Inside AD
3	O=C(CCCCCOn1cnc2ccc(Br)cc2c1=O)NO	unreliable prediction	Outside AD
4	O=C(CCCCCCOn1cnc2ccc(Br)cc2c1=O)NO	unreliable prediction	Outside AD
5	COc1ccc(C2Nc3ccccc3C(=O)N2CCCCCC(=O)NO)cc1	Active	Inside AD
6	COc1cccc(C2Nc3ccccc3C(=O)N2CCCCCC(=O)NO)c1	Active	Inside AD
7	COc1cc(C2Nc3ccccc3C(=O)N2CCCCCC(=O)NO)cc(OC)c1OC	Inactive	Inside AD
8	O=C(CCCCCN1C(=O)c2ccccc2NC1c1ccc(Cl)cc1)NO	Active	Inside AD
9	COc1ccc(C2Nc3ccccc3C(=O)N2CCCCCC(=O)NO)cc1COc1ccccc1F	unreliable prediction	Outside AD
10	COc1ccc(C2Nc3ccccc3C(=O)N2CCCCCCC(=O)NO)cc1	Active	Inside AD