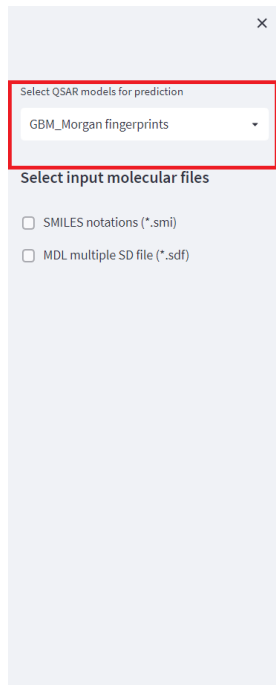


Manual of HDAC6 DETECTOR

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.



HDAC6 DETECTOR

The application provides an alternative method for assessing the potential of chemicals to be Histone deacetylase 6 (HDAC6) inhibitors.

Machine learning



The models were developed using open-source chemical descriptors based on Morgan and topological fingerprints, along with the gradient boosting method.

OECD principles



We follow the best practices for model development and validation recommended by guidelines of the Organization for Economic

Predictive power



The QSAR models showed high predictive power with vigorous validation metrics for test sets, achieving accuracy, sensitivity, and specificity ranging

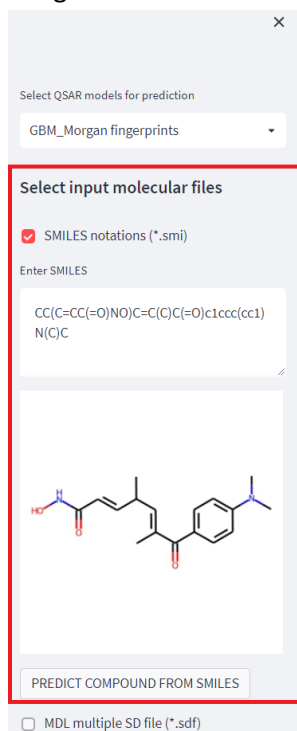
Lipinski's rule



Estimating the bioavailability of a compound is an important factor in drug development. Lipinski's rule of five was introduced to

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.



HDAC6 DETECTOR

The application provides an alternative method for assessing the potential of chemicals to be Histone deacetylase 6 (HDAC6) inhibitors.

Machine learning



The models were developed using open-source chemical descriptors based on Morgan and topological fingerprints, along with the gradient boosting method.

OECD principles



We follow the best practices for model development and validation recommended by guidelines of the Organization for Economic

Predictive power



The QSAR models showed high predictive power with vigorous validation metrics for test sets, achieving accuracy, sensitivity, and specificity ranging

Lipinski's rule



Estimating the bioavailability of a compound is an important factor in drug development. Lipinski's rule of five was introduced to

If the entered structure is incorrect, the application reports an error that has occurred, the details of which can be found through the application manager.

rdkit.Chem.rdchem.AtomValenceException: This app has encountered an error. The original error message is redacted to prevent data leaks. Full error details have been recorded in the logs (if you're on Streamlit Cloud, click on 'Manage app' in the lower right of your app).

Traceback:

```
File "/home/appuser/venv/lib/python3.9/site-packages/streamlit/runtime/script_runner.py", line 385, in exec(code, module.__dict__)
File "/app/hdac6_detector/HDAC6_detector_app.py", line 182, in <module>
  smiles=standardize_smiles(compound_smiles)
File "/home/appuser/venv/lib/python3.9/site-packages/molvs/standardize.py", line 385, in standardize_smiles
  mol = Standardizer().standardize(mol)
File "/home/appuser/venv/lib/python3.9/site-packages/molvs/standardize.py", line 385, in standardize
  Chem.SanitizeMol(mol)
```

To clarify the details of the error, please click here

< Manage app

rdkit.Chem.rdchem.AtomValenceException: This app has encountered an error. The original error message is redacted to prevent data leaks. Full error details have been recorded in the logs (if you're on Streamlit Cloud, click on 'Manage app' in the lower right of your app).

Traceback:

```
File "/home/appuser/venv/lib/python3.9/site-packages/streamlit/runtime/script_runner.py", line 385, in exec(code, module.__dict__)
File "/app/hdac6_detector/HDAC6_detector_app.py", line 182, in <module>
  smiles=standardize_smiles(compound_smiles)
File "/home/appuser/venv/lib/python3.9/site-packages/molvs/standardize.py", line 385, in standardize_smiles
  mol = Standardizer().standardize(mol)
File "/home/appuser/venv/lib/python3.9/site-packages/molvs/standardize.py", line 385, in standardize
  Chem.SanitizeMol(mol)
```

```
[10:20:19] ■ Processed dependencies!
Collecting usage statistics. To deactivate, set browser.gatherUsageStats to False.
[17:00:36] Explicit valence for atom # 16 N, 4, is greater than permitted
[17:00:36] Explicit valence for atom # 20 C, 5, is greater than permitted
[17:00:36] Explicit valence for atom # 8 C, 5, is greater than permitted
[17:00:36] Explicit valence for atom # 16 N, 4, is greater than permitted
[17:00:36] Explicit valence for atom # 4 N, 4, is greater than permitted
[17:00:36] Explicit valence for atom # 22 C, 5, is greater than permitted
[17:00:36] Explicit valence for atom # 21 C, 5, is greater than permitted
[17:00:36] Explicit valence for atom # 22 C, 5, is greater than permitted
[17:00:36] Explicit valence for atom # 8 C, 5, is greater than permitted
[17:02:11] Explicit valence for atom # 15 C, 5, is greater than permitted
2023-01-15 17:02:11.513 Uncaught app exception
Traceback (most recent call last):
  File "/home/appuser/venv/lib/python3.9/site-packages/streamlit/runtime/script_runner.py", line 385, in exec(code, module.__dict__)
  File "/app/hdac6_detector/HDAC6_detector_app.py", line 182, in <module>
    smiles=standardize_smiles(compound_smiles)
  File "/home/appuser/venv/lib/python3.9/site-packages/molvs/standardize.py", line 385, in standardize_smiles
    mol = Standardizer().standardize(mol)
  File "/home/appuser/venv/lib/python3.9/site-packages/molvs/standardize.py", line 385, in standardize
    Chem.SanitizeMol(mol)
rdkit.Chem.rdchem.AtomValenceException: Explicit valence for atom # 15 C, 5, is greater than permitted
```

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.

Select QSAR models for prediction

GBM_Morgan fingerprints

Select input molecular files

☐ SMILES notations (*.smi)

☒ MDL multiple SD file (*.sdf)

Choose a file

Drag and drop file here

Limit 200MB per file

Browse files

example.sdf

13.9KB

HDAC6 DETECTOR

The application provides an alternative method for assessing the potential of chemicals to be Histone deacetylase 6 (HDAC6) inhibitors.

Machine learning



The models were developed using open-source chemical descriptors based on Morgan and topological fingerprints, along with the gradient

OECD principles



We follow the best practices for model development and validation recommended by guidelines of the Organization for

Predictive power



ability of the developed QSAR models is further confirmed by our synthesis of new HDAC6 inhibitors and in vivo studies. The application HDAC6

Lipinski's rule



development. Lipinski's rule of five was introduced to estimate the oral bioavailability of a compound. Our bioavailability radar is displayed

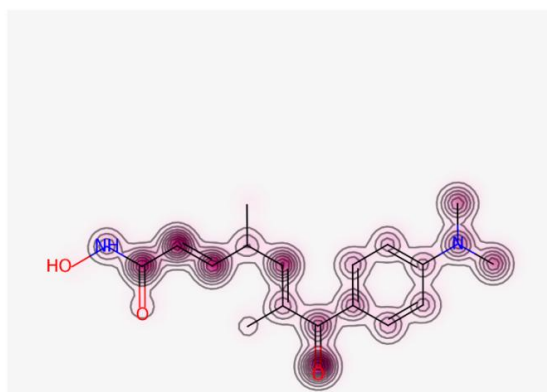
Step 3. *Prediction results*. The form of presentation of the results depends on the type of descriptors selected, as well as the format of the input chemical data. For example, when selecting SMILES, the results will be displayed for a single molecule. The chemical fragments are colored in green (predicted to reduce inhibitory activity) or magenta (predicted to increase activity HDAC6 inhibitors). The gray isolines separate positive and negative contributions.

Prediction results:

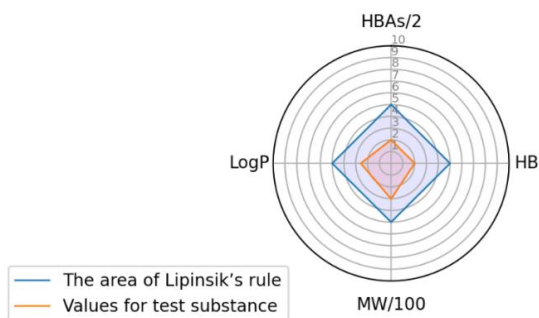
HDAC1: Active

Applicability domain (AD): Inside AD

Predicted fragments contribution:

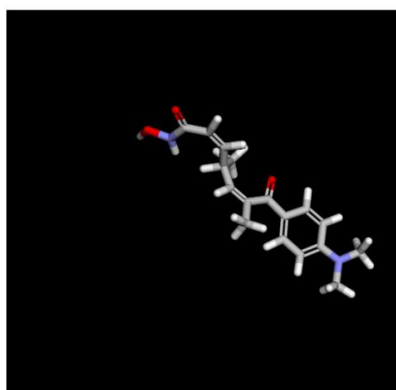


The Bioavailability Radar: compliance the Lipinski's rule of five



	Values for the test substance	Reference value of Lipinski's rule
Molecular weight(MW), Da	302.3740	500.0000
Octanol-water coefficient(LogP)	2.5793	5.0000
Number of hydrogen bond donors (HBD)	2.0000	5.0000
Number of hydrogen bond acceptors(HBAs)	4.0000	10.0000

3D structure of the studied compound:



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 ori...	SMILES of wrong structure:
1	80	<chem>COC1=C/C2=C(OC3=CC=C...</chem>
2	88	<chem>O=C(/C=C/C1=CC=C(CCN...</chem>
3	90	<chem>COC1=C/C(C(OC)=C2\C(=O)...</chem>
4	92	<chem>COC1=C/C2=C(OC3=CC=C...</chem>
5	93	<chem>CCC1=CN(C2=CC=CC=C2)...</chem>
6	153	<chem>O=C(NO)C1=CC=C(CN2CC...</chem>
7	158	<chem>O=C(CC1=CC=C(CN2CCN(...</chem>
8	193	<chem>O=C(/C=C/C1=CC=C(CN2C...</chem>
9	203	<chem>COC1=C/C(C(OC)=C2\C(=O)...</chem>

Kept data: 202 molecules

2. RESULTS OF PREDICTION:

Show results as table

Show results and map of fragments contribution for each molecule separately

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.

5	93	<chem>CCC1=CN(C2=CC=CC=C2)...</chem>
6	153	<chem>O=C(NO)C1=CC=C(CN2CC...</chem>
7	158	<chem>O=C(CC1=CC=C(CN2CCN(...</chem>
8	193	<chem>O=C(/C=C/C1=CC=C(CN2C...</chem>
9	203	<chem>COC1=C/C(C(OC)=C2\C(=O)...</chem>

Kept data: 202 molecules

2. RESULTS OF PREDICTION:

Show results as table

	SMILES	HDAC1 activity	Applicability domain (AD)
1	<chem>C[C@H]1CN(C(=O)N2ccc(-c...</chem>	Active	Inside AD
2	<chem>Nc1ccc(-c2nc(N3CCOCC3)c3cn...</chem>	Active	Inside AD
3	<chem>COc1ccc(-c2nc(N3CCOCC3)c3...</chem>	Active	Inside AD
4	<chem>CN(Cc1nc2c(N3CCOCC3)nc(-c...</chem>	Active	Inside AD
5	<chem>CN(Cc1nc2c(N3CCOCC3)nc(-c...</chem>	Active	Inside AD
6	<chem>COc1ccc(Br)c(/C=N/NC(=O)CC...</chem>	Active	Inside AD
7	<chem>CN(Cc1nc2c(N3CCOCC3)nc(-c...</chem>	Active	Inside AD
8	<chem>CCC(=O)CCCCC(C@H)1NC(=...</chem>	Active	Inside AD
9	<chem>O=C(CCCCCCn1cc(-c2ccc3ncn...</chem>	Active	Inside AD
10	<chem>CN(C1c1ccc(-c2nc(N3CCOCC3)...</chem>	Active	Inside AD

Download results of prediction as CSV

Show results and map of fragments contribution for each molecule separately