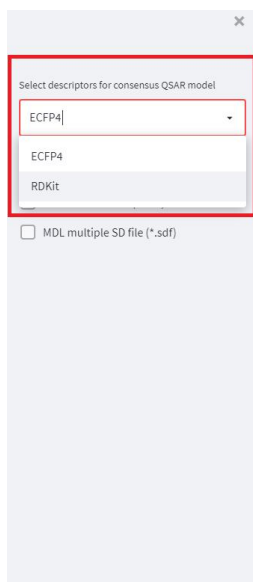


## Manual of HDAC1 Predictor

open-source software

version 1.0.

**Step 1.** Select descriptors for consensus QSAR model. You can choose one of two types of molecular descriptors that were used in the development of the QSAR model. If you choose ECFP4 descriptors, an additional analysis of the contributions of molecular fragments of the studied compound to the inhibition of HDAC1 will be carried out.



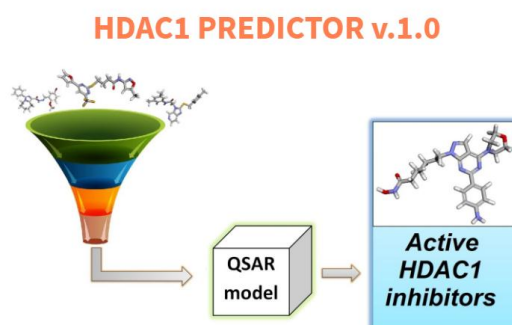
Select descriptors for consensus QSAR model

ECFP4

ECFP4

RDKit

☐ MDL multiple SD file (\*.sdf)



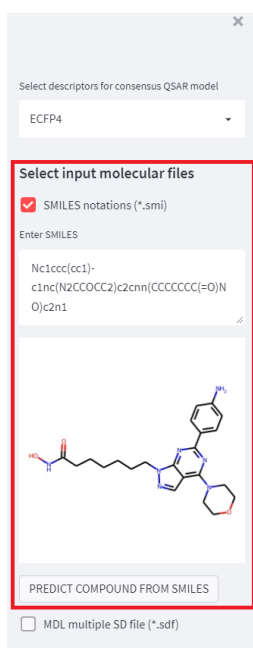
A machine learning Web application to assess the potential of histone deacetylase 1 (HDAC1) inhibitors.

App Characteristics

© Oleg Tinkov, 2022

**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.



Select descriptors for consensus QSAR model

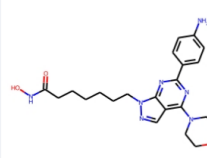
ECFP4

Select input molecular files

☒ SMILES notations (\*.smi)

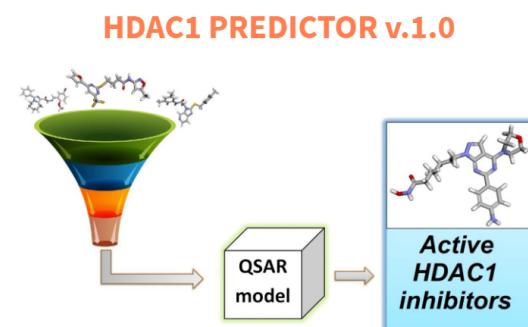
Enter SMILES

Nc1ccc(cc1)-c1nc(N2CCOCC2)c2cnn(CCCCCC(=O)NO)c2n1



PREDICT COMPOUND FROM SMILES

☐ MDL multiple SD file (\*.sdf)



A machine learning Web application to assess the potential of histone deacetylase 1 (HDAC1) inhibitors.

App Characteristics

© Oleg Tinkov, 2022

The screenshot displays the user interface of the HDAC1 Inhibitor Predictor. On the left, there are input sections for selecting descriptors (ECFP4) and molecular files (SMILES). The main area features a diagram illustrating the workflow: a funnel representing descriptor selection leads into a box labeled 'QSAR model', which then points to a box labeled 'Active HDAC1 inhibitors' containing chemical structures. Below this diagram, a text box states: "A machine learning Web application to assess the potential of histone deacetylase 1 (HDAC1) inhibitors." To the right of this text is a button labeled "App Characteristics".

An error message is displayed below the "App Characteristics" button:

```
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.
```

Below the error message, a "Traceback:" section shows the following code snippet:

```
File ~/home/appuser/.venv/lib/python3.7/site-packages/streamlit/scriptrunner/w
axes(codey, module,..._file_)
File ~/app/hdac1-inhibitors/hdac1_predictor_app.py, line 74, in <module>
    smarts_to_standardizer(smarts_to_standardizer.smarts)
File ~/home/appuser/.venv/lib/python3.7/site-packages/molvs/standardize.py, l
mol = Standardizer().standardize(mol)
File ~/home/appuser/.venv/lib/python3.7/site-packages/molvs/standardize.py, l
Chem.SanitizeMol(mol)
```

A red arrow points from the text "To clarify the details of the error, please click here" to the "Manage app" button in the bottom right corner of the interface.

[illegible]

The diagram illustrates the workflow of the HDAC1 Predictor v.1.0. On the left, a screenshot of the web application interface shows the 'Select input molecular files' section, which includes options for SMILES notations and MDL multiple SD files, a file upload area, and a 'Browse files' button. On the right, a flowchart shows chemical structures entering a funnel, leading to a 'QSAR model' box, which then points to a box labeled 'Active HDAC1 inhibitors' containing a chemical structure.

**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.

App Characteristics

**Prediction results:**

HDAC1: Active

Applicability domain (AD): Inside AD

Predicted fragments contribution:

The chemical fragments are colored in green (predicted to reduce inhibitory activity) or magenta (predicted to increase activity HDAC1 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"



A machine learning Web application to assess the potential of histone deacetylase 1 (HDAC1) inhibitors.

App Characteristics

## 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	COC1=C/C2=C(OC3=CC=C...
2	88	O=C(/C=C/C1=CC=C(CCN...
3	90	COC1=C(C(OC)=C2(C(=O)...
4	92	COC1=C/C2=C(OC3=CC=C...
5	93	CCC1=CN(C2=CC=CC=C2)...
6	153	O=C(NO)C1=CC=C(CN2CC...
7	158	O=C(CC1=CC=C(CN2CCN1...
8	193	O=C(/C=C/C1=CC=C(CN2C...
9	203	COC1=C(C(OC)=C2(C(=O)...

Kept data: 202 molecules

## 2. RESULTS OF PREDICTION:

Show results as table

Show results and map of fragments contribution for each molecule separately

Select descriptors for consensus QSAR model

ECFP4

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

211.sdf  
0.5MB

Kept data: 292 molecules

## 2. RESULTS OF PREDICTION:

Show results as table

	SMILES	HDAC1 activity	Applicability domain (AD)
1	<chem>C[C@@H]1CN(C)[C@H](O)Nc2cccc(c...)</chem>	Active	Inside AD
2	<chem>Nc1ccc(c2nc(N3CCOC(C3)c3cn...</chem>	Active	Inside AD
3	<chem>COc1ccc(c2nc(N3CCOC(C3)c3cn...</chem>	Active	Inside AD
4	<chem>CN(Cc1nc2c(N3CCOC(C3)nc(c...)</chem>	Active	Inside AD
5	<chem>CN(Cc1nc2c(N3CCOC(C3)nc(c...)</chem>	Active	Inside AD
6	<chem>COc1ccc(Br)cc1C(=O)N(C)[C@H](O)CC...</chem>	Active	Inside AD
7	<chem>CN(Cc1nc2c(N3CCOC(C3)nc(c...)</chem>	Active	Inside AD
8	<chem>CCC(=O)CCCC[C@H](C)[C@H]1NC(=O...</chem>	Active	Inside AD
9	<chem>O=C(C(CCCCCn1cc(c2ccc3ncn...</chem>	Active	Inside AD
10	<chem>CN(C1c1ccc(c2nc(N3CCOC(C3)...</chem>	Active	Inside AD

Download results of prediction as CSV

Show results and map of fragments contribution for each molecule separately

< Manage app