

Manual of HDAC1 Predictor

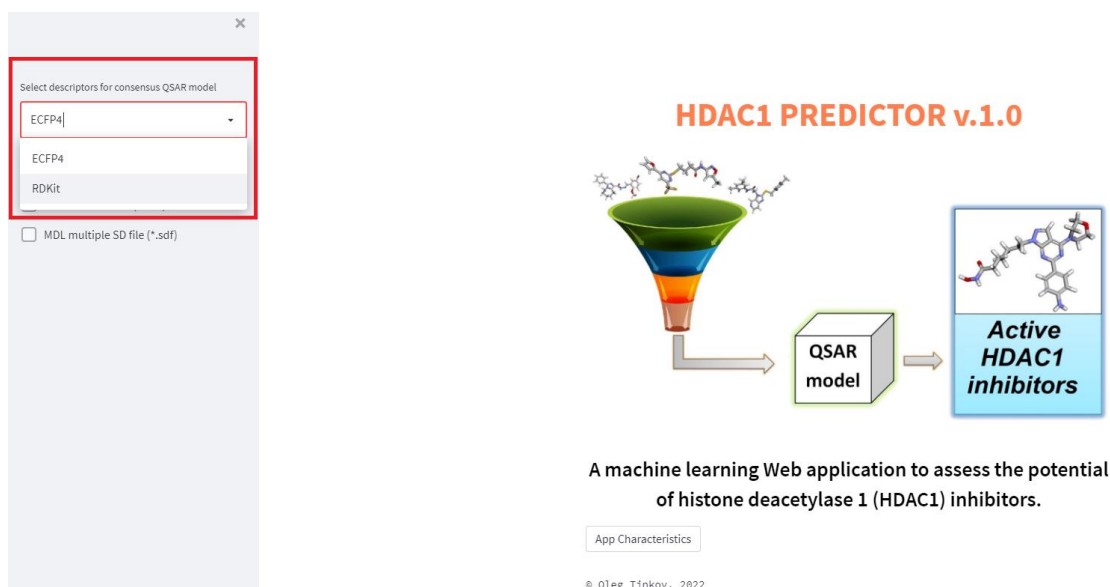
open-source software

version 1.0.

Overview

The HDAC1 Predictor application provides an alternative method for assessing the potential of chemicals to be Histone deacetylase 1 (HDAC1) inhibitors. Compound is classified as active if the predicted IC_{50} value is lower than mean IC_{50} value of the reference drug Vorinostat (11.08 nM) otherwise compound is labeled as inactive. This application makes predictions based on Quantitative Structure-Activity Relationship (QSAR) models build on curated datasets generated from ChEMBL database (ID: ChEMBL325). The consensus models were developed using open-source chemical descriptors based on ECFP4-like Morgan fingerprints and 2D RDKit descriptors, along with the random forest (RF), gradient boosting (GBM), support vector machines (SVM) algorithms, using Python 3.7. The models were generated applying the best practices for QSAR model development and validation widely accepted by the community. Batch processing is available through <https://github.com/ovttiras/HDAC1-inhibitors>. For more information, please refer to our paper:

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.



The image displays the HDAC1 Predictor v.1.0 web application interface on the left and a flow diagram on the right. The interface shows a dropdown menu for selecting descriptors, with 'ECFP4' selected. Below the dropdown, there are checkboxes for 'ECFP4' and 'RDKit'. A checkbox for 'MDL multiple SD file (*.sdf)' is also present. The flow diagram illustrates the process: chemical structures enter a funnel, which leads to a 'QSAR model' box, and finally to a box labeled 'Active HDAC1 inhibitors'.

HDAC1 PREDICTOR v.1.0

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.

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

App Characteristics

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.

HDAC1 PREDICTOR v.1.0

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

App Characteristics

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.

Select descriptors for consensus QSAR model

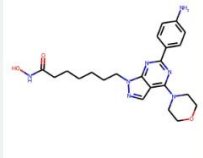
ECFP4

Select input: **molecular files**

☒ SMILES notations (*.smi)

Enter SMILES

```
Nc1ccc(cc1)-c1nc(N2CCCCC2)c2cnn(CCCCCC(=O)NO)c2n1
```



PREDICT COMPOUND FROM SMILES

☐ MDL multiple SD file (*.sdf)

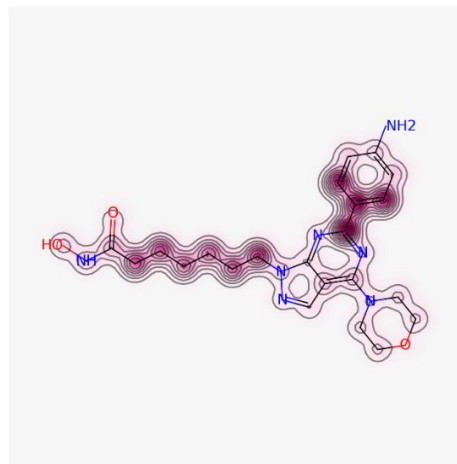
App Characteristics

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/C(OC3=CC=C...
2	88	O=C(/C=C/C1=CC=C(CCN(C...
3	90	COC1=C(C(OC)=C2(C(=O)...
4	92	COC1=C/C2=C/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(CN2CCN(...
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

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.

×

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

0.6MB

×

		SMILES
5	93	CCCC1=CN(C2=CC=CC=C2)...
6	153	O=C(NO)C1=CC=C(CN2CC...
7	158	O=C(C(C1=CC=C(CN2CCN...
8	193	O=C(C)C=C(C1=CC=C(CN2C...
9	203	COC1=C(C)(OC)=C2(C(=O)...

Kept data: 262 molecules

2. RESULTS OF PREDICTION:

Show results as table

	SMILES	HDAC1 activity	Applicability domain (AD)
1	C[C@H]1CN(C(=O)Nc2ccc(-c...	Active	Inside AD
2	Nc1ccc(-c2nc(N3CCOCC3)c3cn...	Active	Inside AD
3	COc1ccc(-c2nc(N3CCOCC3)c3...	Active	Inside AD
4	CN(Cc1nc2c(N3CCOCC3)nc(-c...	Active	Inside AD
5	CN(Cc1nc2c(N3CCOCC3)nc(-c...	Active	Inside AD
6	COc1ccc(Br)c(C=N/NC(=O)CC...	Active	Inside AD
7	CN(Cc1nc2c(N3CCOCC3)nc(-c...	Active	Inside AD
8	CCC(=O)CCCCC[C@H]1NC(=...	Active	Inside AD
9	O=C(CCCCCCn1ccc(-c2ccc3ncn...	Active	Inside AD
10	CN(C)C1=CC=C(C2=CC=CC=C2)...	Active	Inside AD

Download results of prediction as CSV

Show results and map of fragments contribution for each molecule separately

© Oleg Tinkov, 2022

Share

Manage app