Manual of HDAC2 SCAN

open-source software

Step 1. *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.

HDAC2 SCAN v.1.0

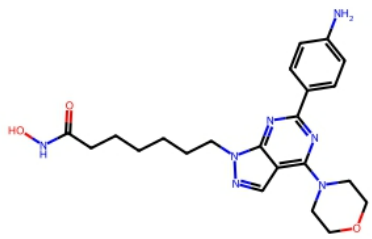
A machine learning Web application to assess the potential of histone deacetylase 2 (HDAC2) inhibitors.

Select input molecular files

SMILES

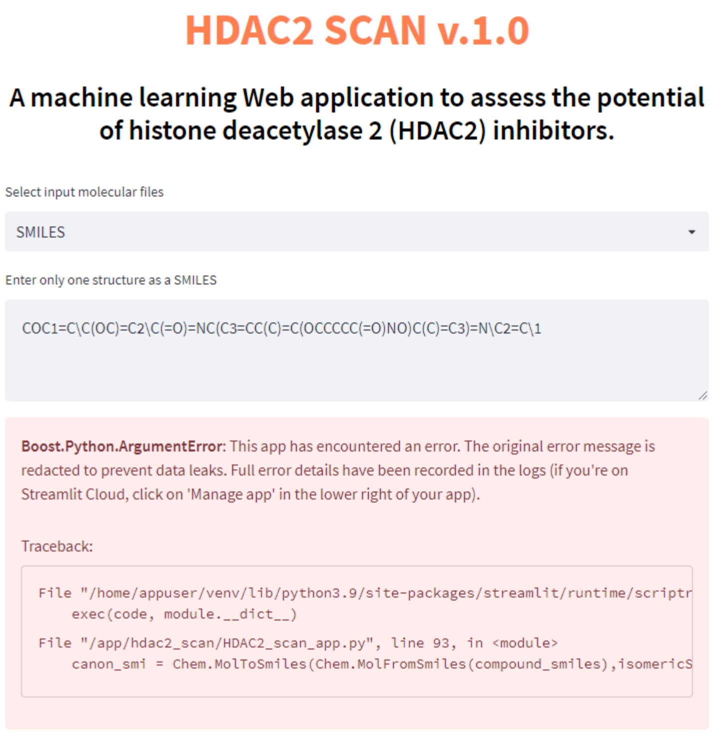
Enter only one structure as a SMILES

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

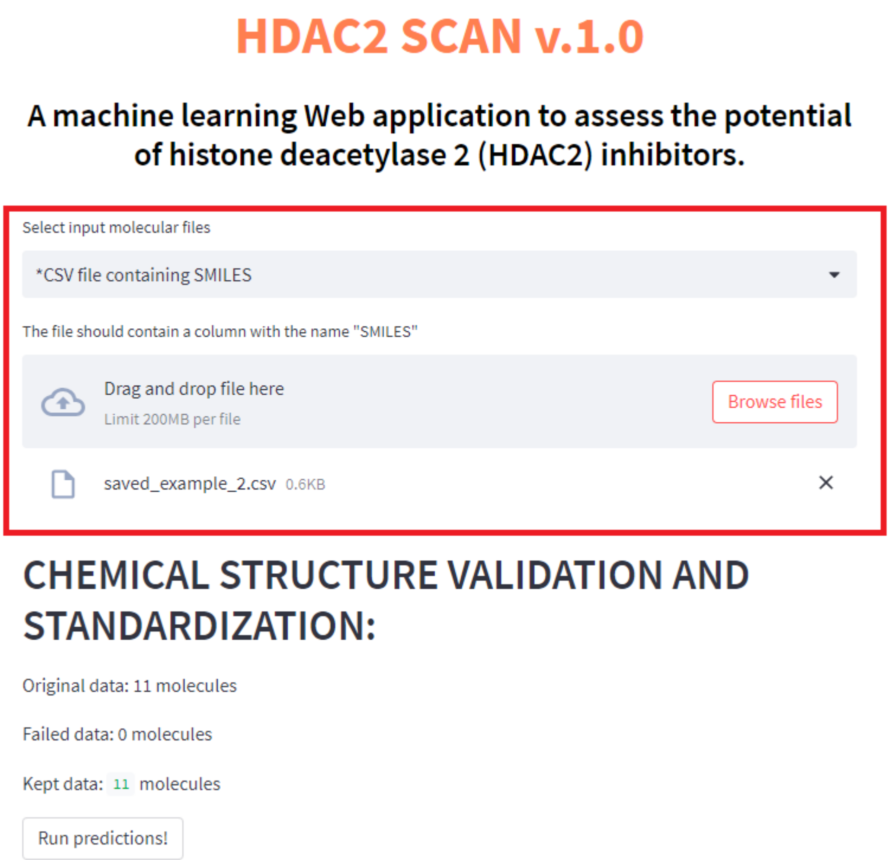


Run predictions!

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



If you choose a file *sdf or *csv, 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. It is important to note that if you choose a file with the *csv extension, the file should contain a column with the name "SMILES"



If incorrect structures are detected in the file *sdf or *csv, the corresponding information will appear in the section "CHEMICAL STRUCTURE VALIDATION AND STANDARDIZATION"

Select input molecular files

MDL multiple SD file (*.sdf)

Choose a SDF file

Drag and drop file here
Limit 200MB per file

Browse files

211_prop.sdf 0.7MB

CHEMICAL STRUCTURE VALIDATION AND STANDARDIZATION:

Original data: 211 molecules

Failed data: 9 molecules

No.	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
2	88	O=C(/C=C/C1=CC=C(CCNC(=O)C2=CC(C/C3=N/N=C(=O)C4=CC=CC=C43)=CC
3	90	COC1=C\C(OC)=C2\C(=O)=NC(C3=CC(C)=C(OCCCCC(=O)NO)C(C)=C3)=N\C2
4	92	COC1=C/C2=C(OC3=CC=C(NC(=O)C4=C(=O)C(C)=CN(C5=CC=CC=C5)=N4)C=
5	93	CCC1=CN(C2=CC=CC=C2)=NC(C(=O)NC2=CC=C(O/C3=C/C=N\C4=CC(OCCCC
6	153	O=C(NO)C1=CC=C(CN2CCN(C(=O)C3=CC(C/C4=N/N=C(=O)C5=CC=CC=C54)
7	158	O=C(CC1=CC=C(CN2CCN(C(=O)C3=CC(C/C4=N/N=C(=O)C5=CC=CC=C54)=C
8	193	O=C(/C=C/C1=CC=C(CN2CCN(C(=O)C3=CC(C/C4=N/N=C(=O)C5=CC=CC=C5
9	203	COC1=C\C(OC)=C2\C(=O)=NC(C3=CC(C)=C(OCC4=CC=C(/C=C/C(=O)NO)C=C

Kept data: 202 molecules

Run predictions!

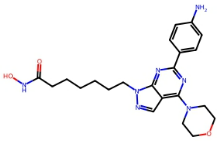
Step 2. Click on the "Run predictions!" button for prediction.
Step 3. *Prediction results*.

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. When displaying the results on the screen, it is taken into account whether there are experimental values of activity and toxicity for the studied compound.

SMILES

Enter only one structure as a SMILES

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



Run predictions!

Prediction results:

	SMILES	Predicted value, pIC50	Applicability domain	Experimental value, pIC50	Standard deviation	Chemble ID	Predicted value toxicity, Ld50, mg/kg	Applicability domain_tox	Experimental value toxicity, Ld50	CAS number
1	Nc1ccc(-c2nc(N3CCOCC3)	see experimental value	-	8.8	nan	CHEMBL4552057	1452.8057	Outside AD	-	not detected

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 HDAC2 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

7) ***Predicted value toxicity, Ld50, mg/kg*** - predicted value of acute toxicity when administered orally to rats. If experimental data is available in the PubChem database, the label "see experimental value" is displayed in this cell.

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

9) ***Experimental value toxicity, Ld50*** - experimental data presented in the PubChem database. Toxicity was measured by a dose of LD_{50} when administered orally to rats

10) ***CAS number*** - a unique identification number assigned by the Chemical Abstracts Service (CAS)

If you choose a file *sdf or *csv, the prediction results for correct chemical structures are displayed in a table that can be downloaded.

Original data: 11 molecules

Failed data: 0 molecules

Kept data: `11` molecules

Run predictions!

	SMILES	Predicted value, pIC50	Ap
2	CCOc1cc(NC(=O)CCCCCCC(=O)Nc2cc(-c3ccccc3)ccc2N)cc2c1OCC([N+](=O)[O-])=C2	see experimental value	-
3	[N-]=[N+]=NCCC(=O)Nc1ccc(C(=O)Nc2cc(N=[N+]=[N-])ccc2N)cc1	see experimental value	-
4	[N-]=[N+]=Nc1cccc(COC(=O)NCc2ccc(C(=O)Nc3cc(-c4ccccc4)ccc3N)cc2)c1	see experimental value	-
5	[N-]=[N+]=NCc1cc(N=[N+]=[N-])cc(C(=O)Nc2cc(-c3cccs3)ccc2N)c1	see experimental value	-
6	C#CCOc1ccc(C=NNC(=O)c2cccc(C(=O)NO)c2)cc1	see experimental value	-
7	O=C(NO)c1cnc(NC2(c3ccccc3)CCC(F)(F)CC2)nc1	see experimental value	-
8	CN1CCc2c(c3ccccc3n2Cc2ccc(C(=O)NO)cc2)C1	see experimental value	-
9	CCCNNC(=O)c1ccc(CNC(=O)C(Cc2c[nH]c3ccccc23)NC(=O)c2ccc(OC)cc2)cc1	see experimental value	-
10	COc1ccc(C(=O)c2ccc3c(ccn3Cc3ccc(C=CC(=O)NO)cc3)c2)cc1	see experimental value	-
11	CC(C=CC(=O)NO)=CC(C)C(=O)c1ccc(N(C)C)cc1	see experimental value	-

Download results of prediction as CSV