

AVCpred: An integrated web server for prediction and design of antiviral compounds

Abid Qureshi , Gazaldeep Kaur and Manoj Kumar*

Bioinformatics Centre, Institute of Microbial Technology, Council of Scientific and Industrial Research, Sector 39-A, Chandigarh-160036, India

*Correspondence to: Manoj Kumar (E-mail: manojk@imtech.res.in)

Supplementary information

Table S1 Details of the AVCpred dataset used for the development of general prediction model.

Table S1. Details of statistical parameters used for the development of the QSAR models

Table S3 Details of chemical descriptors used in the development of QSAR models.

Supplementary Table S1. Details of the AVCpred dataset used for the development of general prediction model.

S. No.	Virus	Overall data	Data filter		
			Percent inhibition	Reference	Non redundant
1	Dengue virus 1	3	1	1	1
2	Dengue virus 2	22	16	16	16
3	Enterovirus	72	46	46	30
4	Human adenovirus 5	83	74	74	41
5	Human cox B1	13	4	4	4
6	Human cox B5	46	21	21	21
7	Human echovirus 13	3	3	3	3
8	Human echovirus 9	9	3	3	2
9	Human enterovirus 71	28	20	20	19
10	Human enterovirus C	2	2	2	1
11	Human polio virus 1	12	4	4	4
12	Human rhinovirus	139	13	2	1
13	Human rhinovirus 14	35	33	33	29
14	Human rhinovirus 1B	52	40	40	18
15	Human rhinovirus 2	2	2	2	2
16	Human T lymphotropic virus	47	45	45	42
17	Influenza A	624	90	52	36
18	Influenza A (H1N1)	37	16	16	16
19	Influenza B	148	1	1	1
20	Monkeypox virus	5	5	5	1
21	Respiratory syncytial virus	156	16	16	4
22	Rift valley fever virus (cercopithecidae)	1	1	1	1
23	Sandfly fever sicilian virus	2	2	2	2
24	SARS Corornavirus	77	61	61	23
25	Simian virus 40	45	45	45	45
26	Sindbis virus	168	4	4	4
27	Vaccinia virus	170	12	12	12
28	Vaccinia virus WR	23	22	22	22
29	Variola virus	8	6	6	1
30	Vescicular stomatitis virus	451	126	126	63
31	West Nile Virus	53	26	26	17
32	Yellow fever virus	52	51	51	51
	Total	2588	811	762	533

Supplementary Table S2. Details of statistical parameters used for the development of the QSAR models

S. No.	Virus	Training				Testing			
		R	R ²	MAE	RMSE	R	R ²	MAE	RMSE
1	Human immunodeficiency virus (HIV)	0.72	0.52	11.24	16.41	0.63	0.40	19.23	27.32
2	Hepatitis C virus (HCV)	0.74	0.55	10.32	15.64	0.65	0.42	17.53	25.14
3	Hepatitis B virus (HBV)	0.66	0.44	16.45	24.42	0.61	0.37	20.14	29.13
4	Human herpesvirus (HHV)	0.68	0.46	15.44	22.53	0.64	0.41	18.11	26.21
5	General (26 viruses)	0.71	0.50	12.01	19.22	0.67	0.45	16.34	23.24

R :Pearson correlation coefficient. It is a measure of correlation between two variables. A value of 1 denotes total positive correlation, 0 is no correlation and -1 is total negative correlation.
R² :Coefficient of determination. It indicates how well data fit a statistical model. An R² of 1 indicates that the model perfectly fits the data, while an R² of 0 means that the model does not fit the data at all.
MAE :Mean absolute error. This measure indicates how close the predictions are to the eventual outcomes. MAEs are negatively-oriented scores, that is, lower values are better.
RMSE :Root-mean-square error measures the average magnitude of the error. RMSEs are also negatively-oriented scores, that is, lower values are better.

Supplementary Table S3. Details of chemical descriptors used in the development of QSAR models.

Chemical Descriptor	Type	Details
ATSc4	Autocorrelation	Centered Broto-Moreau autocorrelation
ATSm3	Autocorrelation	Broto-Moreau autocorrelation
BCUTc-1h	BCUT	nlow highest partial charge weighted BCUTS
BCUTp-1h	BCUT	nlow highest polarizability weighted BCUTS
EStateFP7	sCH3	[CD1H3]-*
ETA_EtaP_B	Extended topochemical atom	Branching index EtaB relative to molecular size
ExtFP112, ExtFP128, ExtFP148, ExtFP165, ExtFP188, ExtFP223, ExtFP225, ExtFP266, ExtFP266, ExtFP283, ExtFP336, ExtFP346, ExtFP346, ExtFP347, ExtFP440, ExtFP461, ExtFP505, ExtFP506, ExtFP534, ExtFP575, ExtFP581, ExtFP610, ExtFP634, ExtFP637, ExtFP65, ExtFP69, ExtFP763, ExtFP766, ExtFP8, ExtFP805, ExtFP819, ExtFP833, ExtFP840, ExtFP848, ExtFP866, ExtFP888, ExtFP919, ExtFP960, ExtFP962, ExtFP978	Extended Fingerprinter	Extends the Fingerprinter with additional bits describing ring features
FP127, FP129, FP129, FP144, FP146, FP150, FP156, FP164, FP170, FP170, FP173, FP186, FP218, FP230, FP234, FP239, FP256, FP268, FP28, FP301, FP344, FP358, FP368, FP393, FP410, FP437, FP46, FP460, FP462, FP510, FP550, FP587, FP627, FP640, FP654, FP663, FP704, FP714, FP732, FP733, FP748, FP773, FP782, FP785, FP793, FP812, FP827, FP831, FP836, FP852, FP874, FP893, FP901, FP903, FP910, FP919, FP977, FP984	Fingerprinter	Path based fingerprints
GraphFP121, GraphFP144, GraphFP253, GraphFP278, GraphFP292, GraphFP396, GraphFP455, GraphFP506, GraphFP648, GraphFP770, GraphFP778, GraphFP796, GraphFP840, GraphFP948, GraphFP955	Graph Only Fingerprinter	Specialized version of the Fingerprinter which does not take bond orders into account
KRFP299	Klekota-Roth fingerprint	[!#1][CH2][CH](!#1)[CH3]
KRFP3349		CC(C)(C)c1cccc1
KRFP3435		CC(O)CN(C)C
KRFP3520		Cc1cc(O)cc(O)c1
KRFP366		[!#1][CH2][CH2][CH2][NH2]
KRFP3697		CCCC(=O)O
KRFP3744		CCCO
KRFP3764		CCN(CC)CCNC=O

	fingerprint	#8X2][#6;!\$(C=[O,N,S])]
SubFP90	Substructure fingerprint count	[CX3;\$([R0][#6]),\$([H1R0])](=[OX1])[SX2][#6;\$!(C=[O,N,S])]
SubFPC295		[#6]~[#7,#8,#16]
SubFPC300	Chi chain	[\$([#7X2,OX1,SX1]=*!H0;!\$([a;!n])),\$([#7X3,OX2,SX2;!H0]*=*),\$([#7X3,OX2,SX2;!H0]*:n)]
VCH-6		Valence chain, order 6
WPOL	Wiener numbers	Weiner polarity number
WTPT-4	Weighted path	Sum of path lengths starting from oxygens