Small Molecule Antiviral Compound Collection (SMACC): A comprehensive, highly curated database to support the discovery of broad-spectrum antiviral drug molecules.

Holli-Joi Martin[1], Cleber C. Melo-Filho[1], Daniel Korn[1], Richard Eastman[2], Ganesha Rai[2], Anton Simeonov[2], Alexey V. Zakharov[2,*], Eugene Muratov[1,*], Alexander Tropsha[1,*]

Table S1. Overview of Database Entries After Data Curation

Property	Phenotypic	Target-based
Viruses	13	10
Entries	21,392	11,123
Assays	1,124	245
Cell types	53	-
Target types	-	15

Identification of compounds with multiple antiviral activities

A threshold of 10 μM, irrespective of the type of activity measurement, was applied to define the outcome (i.e., if a compound was active or inactive). We chose 10 μM as the activity threshold in accordance with the threshold set for an initial hit compound in standard medicinal chemistry campaigns. When the compound activity was reported with ambiguous operators (greater than, ">", or less than, "<", certain value), it was annotated as inconclusive. The final definition of the activity call for each compound was based on the concordance of all compound replicate entries tested against the same virus but in different assays (or the same viral target for the target-based dataset). Three outcomes were possible: (i) the compound was active when tested in all assays; (ii) it was active in some assays and inactive in others; and (iii) it was inactive in all assays. In case (i), the compound was considered active while in case (iii), inactive. Any compound in case (ii), with discordant activity calls resulting from different assays, i.e., with at least one activity call different from other ones for the same virus, was considered inconclusive and was not used for the overlap analysis to identify compounds with multiple antiviral activities. A compound was also annotated as inconclusive if the assay reported the compound's activity as "Not

Determined". Finally, all compounds tested in different viruses (or viral targets in the target-based dataset) were analyzed and those showing activity against two or more viruses were selected as potential BSAs. **Table S2** summarizes our protocols to decide on the final activity calls for compounds included in SMACC database.

Table S2. Rules for making the final activity calls for compounds in SMACC database.

Virus	Assay	Cell Type	Activity discordance*	Standard Value	Activity Call
same	same	same	no	<10 uM or >50%	active
same	same	same	no	>10 uM or <50%	inactive
any	any	any	no	"Not Determined"	inconclusive
same	same	same	yes	Any	inconclusive
same	different	same	yes	Any	inconclusive
same	different	different	yes	Any	inconclusive

* At least one discordant duplicate/replicate compound

Ontological examination and curation of phenotypic assays reported in ChEMBL

Table S3. Curation issues of phenotypic data

Virus	# of Entries	# of Assays	Changing BAO Assay Annotation	Assay Type Missing in Description	Cell Type Missing	Cell Type Available in Description	Cell Type Completely Missing
SARS-CoV-2	18,190	21	18,190	2	7	0	7
MERS-CoV	49	9	49	0	49	49	0
HCoV-229E	164	11	164	7	69	65	5
Dengue	2,685	581	2,685	191	1,682	1,495	187
Yellow Fever	930	66	930	41	169	72	97
Zika	357	91	357	19	334	308	26
West Nile	514	102	514	81	308	148	160
Powassan	24	1	0	0	0	0	0
RSV	2,906	239	2,862	586	608	141	467
HPIV-3	1,632	161	1,566	435	520	96	421

H1N2	61	7	43	0	18	18	0
H7N7	26	7	26	0	18	0	18
Sandfly Fever	24	3	24	0	2	0	2

Analysis of Cell Types

The 21,392 compounds integrated into our database were tested in 53 unique cell types. The most common cell types were Vero C1008 (37.6% of entries), Caco-2 (24.2%), Vero (9.1%), Huh-7 (4.6%), and Hep-2 (3.8%). The high propensity of testing in VeroC1008 cells is explained by a single assay screening against SARS-CoV-2 (8043 entries). Other cell types, such as Caco-2, were used for testing in multiple viruses and amongst various assays. Interestingly, Dengue virus had the greatest number of cell types tested (26), followed by RSV (17), and Zika virus (10) (**Figure 4**). Conversely, Vero cells were tested in the largest number of viruses (9) across 1,956 entries, followed by A549 (7 viruses), and Huh-7 cells (6 viruses).

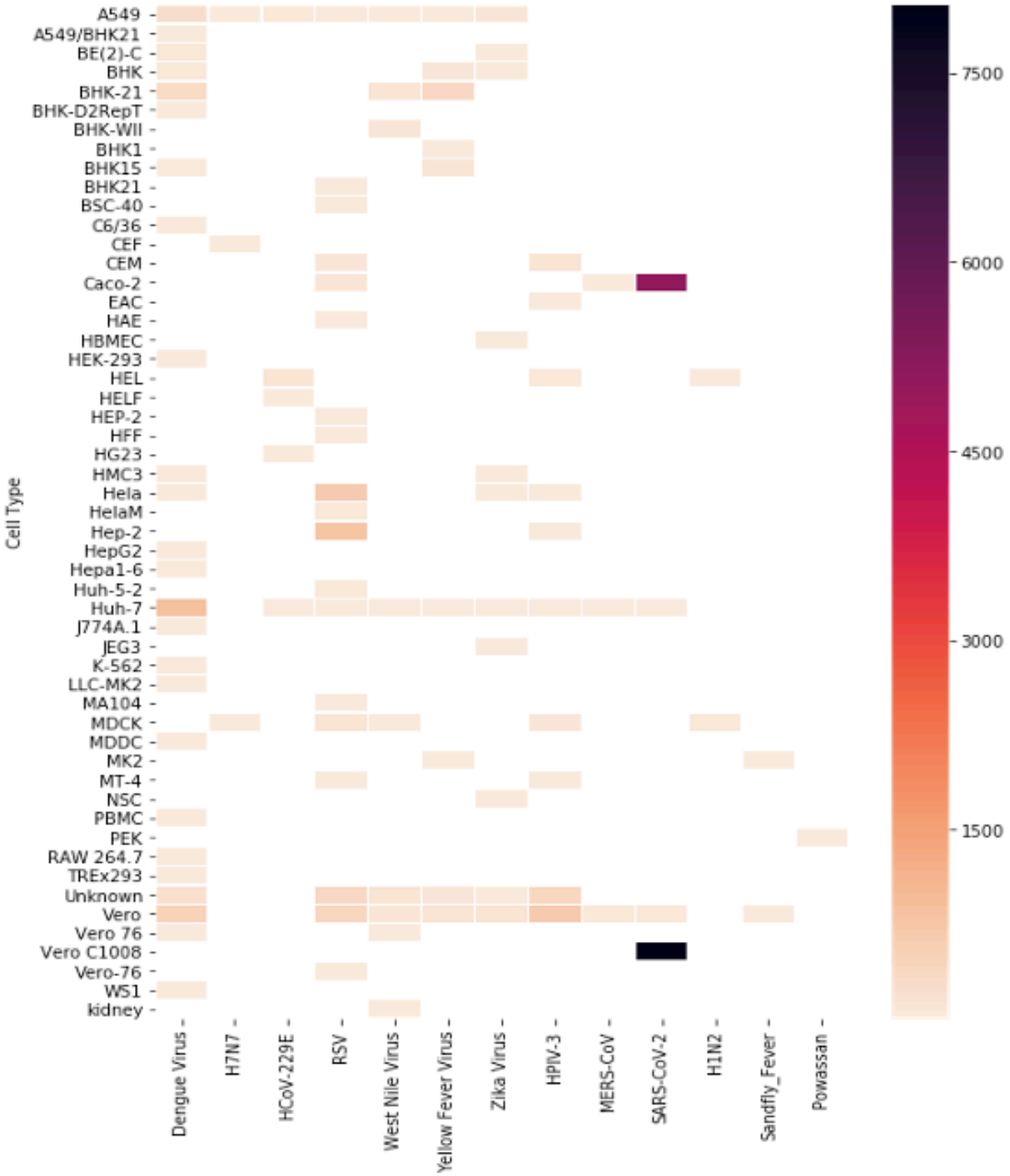


Figure S1. Heat map showing distribution of compounds tested in phenotypic screens for different cell types and different viruses.

We analyzed the distribution of compounds tested in phenotypic screens for different cell types and different viruses (Figure S1). We further examined the effect of cell type on the resulting

activity determination using 27 compounds tested in the largest number of phenotypic assays in our database (Table S4). As expected, there were some inconsistencies in the activities determined when stratifying compounds by the virus they were tested against and the cell type for that virus and then comparing their activities. Unknown cell types and inconclusive activity results were ignored. We identified 26 assay results when a compound was tested in the same virus and cell line but showed conflicting results. In another 19 cases, a compound was tested against the same virus but in different cell lines and had different results. In contrast, for 10 cases, we observed completely consistent activity testing results (in 2+ entries) for a compound assayed in the same virus in the same cell line. We also observed 22 cases of consistent activity when compounds were tested for the same virus but in multiple cell lines. There were also many cases reporting a compound tested for a single virus and multiple cell types. In this case, we only analyzed whether or not the activities reported for each cell line were consistent. we observe that the choice of cell type can influence the outcome of the assay. This observation was reported previously (He et al., 2021) and thus, the annotation of a compound as active or inactive against any virus should be always reported strictly in the context of the specific underlying assay. Consequently, integration of data across multiple cell lines, for instance, to increase the size of the data for QSAR model development, should be done with care, i.e., only when evidence exists that compounds show similar activities when tested in different cell lines.

Table S4. Analysis of cell type effect on activity of compounds with 10+ entries. Compounds tested in the same virus and cell line that showed conflicting results within that cell line have the cell line annotated as "#". Compounds tested against the same virus but different cell lines that had different results have that virus designated with "*". Compounds with consistent activity testing results (in 2+ entries) for a compound assayed in the same virus in the same cell line show that virus are annotated with "‡". Compounds with consistent activity when tested for the same virus but in multiple cell lines show that virus are annotated with "+".

ChEMBL ID	Virus	Cell Type	Activity (count)

CHEMBL1421	Dengue Virus+	C6/36	Active(3)
		Huh-7	Active(4), Inconclusive(1)
		Unknown	Active(2)
		Vero	Active(5), Inconclusive(4)
	SARS-CoV-2*	Caco-2	Active(1)
		Vero C1008	Inactive(1)
CHEMBL160	Dengue Virus*	Huh-7 #	Active(3), Inactive(1)
		Vero	Inconclusive(1), Active(1)
	SARS-CoV-2	Vero	Active(1)
	West Nile Virus*	Huh-7 #	Active(3), Inactive(6)
		Vero	Inconclusive(1), Active(1)
	Yellow Fever Virus	Huh-7 #	Active(2), Inactive(1)
CHEMBL1630219	Dengue Virus+	A549	Active(9)
		PBMC	Active(1)
CHEMBL1630220	Dengue Virus+	A549	Active(8)
		BHK-21	Active(4)
		PBMC	Active(1)
		Vero	Active(1)
	West Nile Virus	Vero	Active(1)
CHEMBL1630221	Dengue Virus*	A549 #	Inactive(1), Active(6)
		BHK-21 #	Active(11), Inactive(2), Inconclusive(1)
		HEK-293	Active(2)
		PBMC	Active(1)
		Vero	Inconclusive(1), Active(5)
CHEMBL1631361	Dengue Virus*	A549 #	Active(5), Inactive(1), Inconclusive(1)
		BHK-21	Active(5)
	West Nile Virus	BHK-21	Active(1)
CHEMBL1643	Dengue Virus*	BHK-21	Inconclusive(1), Active(1)
		Huh-7 #	Inactive(7), Active(1)
		Unknown	Inactive(3)
		Vero #	Active(1), Inconclusive(1), Inactive(5)
	H7N7‡	CEF	Inactive(2)
	HCoV-229E+	HEL	Inactive(1)
		HELF	Inactive(1)
	HPIV-3*	CEM	Inactive(1)
		Hela	Inactive(1)
		Unknown	Active(1), Inactive(15)
		Vero #	Active(2), Inactive(20)
	RSV*	Caco-2	Inactive(1)
		Hela #	Inactive(7), Active(11)
		Hep-2	Inconclusive(1), Inactive(12)
		Unknown #	Inactive(17), Active(5)
		Vero #	Active(3), Inactive(1)
	SARS-CoV-2+	Caco-2	Inactive(1)
		Vero C1008	Inactive(3)
	Sandfly Fever	MK2	Active(1)
	West Nile Virus+	A549	Inactive(1)

		BHK-21	Inconclusive(1), Inactive(1)
		Huh-7	Inactive(1)
		Unknown	Inactive(1)
		Vero	Inactive(2)
	Yellow Fever Virus*	BHK	Inactive(1)
		BHK-21	Inactive(3)
		MK2	Active(1)
		Vero	Inactive(2)
	Zika Virus	Unknown	Inactive(1)
		Vero	Inconclusive(1)
CHEMBL184	Dengue Virus	BHK-21	Inconclusive(1)
		Vero	Inconclusive(4)
	HPIV-3	Vero	Inactive(1), Inconclusive(2)
	RSV+	Hela	Inactive(1)
		Vero	Inconclusive(1), Inactive(3)
	SARS-CoV-2+	Caco-2	Inactive(1)
		Vero C1008	Inactive(1)
	West Nile Virus	BHK-21	Inconclusive(1)
	Yellow Fever Virus+	BHK	Inactive(1)
		BHK-21	Inactive(3)
CHEMBL202626	RSV	Hep-2 #	Active(8), Inactive(2)
CHEMBL203727	West Nile Virus*	BHK-21 #	Inactive(3), Active(4)
		Vero #	Inactive(3), Active(4)
CHEMBL204738	RSV	Hep-2 #	Inactive(6), Active(10)
	SARS-CoV-2+	Caco-2	Inactive(1)
		Vero C1008	Inactive(1)
CHEMBL217092	Dengue Virus*	C6/36	Active(3)
		Huh-7 #	Active(3), Inconclusive(1), Inactive(1)
		Unknown	Active(2)
		Vero	Active(4), Inconclusive(1)
	SARS-CoV-2+	Caco-2	Inactive(1)
		Vero C1008	Inactive(1)
CHEMBL219900	RSV	Hep-2 #	Inconclusive(2), Active(10), Inactive(3)
CHEMBL223402	HPIV-3	Hep-2	Inactive(1)
	RSV*	HAE	Active(7)
		Hep-2 #	Active(21), Inactive(5)
	SARS-CoV-2+	Caco-2	Inactive(1)
		Vero C1008	Inactive(1)
CHEMBL31634	HPIV-3‡	Unknown	Inactive(5)
		Vero	Inactive(4)
	RSV‡	Hela	Inactive(3)
		Unknown	Inactive(6)
	SARS-CoV-2+	Caco-2	Inactive(1)
		Vero C1008	Inactive(4)
CHEMBL387388	RSV	Hep-2 #	Inconclusive(2), Active(9), Inactive(4)
CHEMBL408963	RSV+	Hela	Active(8)
		HelaM	Active(1)
		Unknown	Active(8)

	SARS-CoV-2+	Caco-2	Inactive(1)
		Vero C1008	Inactive(1)
CHEMBL4446551	Dengue Virus‡	Huh-7	Inconclusive(2), Active(14)
CHEMBL4538323	Dengue Virus‡	Huh-7	Inconclusive(12), Active(10)
CHEMBL4556306	Dengue Virus*	BE(2)-C #	Active(15), Inactive(8), Inconclusive(5)
		HMC3	Active(2)
	Zika Virus*	BE(2)-C #	Active(14), Inactive(7), Inconclusive(2)
		HMC3	Active(2)
CHEMBL4571759	West Nile Virus*	A549	Active(1)
		Huh-7 #	Active(4), Inactive(4)
		Vero	Inactive(1)
CHEMBL494759	HPIV-3	Unknown	Inactive(11)
	HPIV-3‡	Vero	Inactive(12)
	RSV*	Hela #	Inactive(10), Active(1)
		Unknown	Inactive(10)
CHEMBL564201	Dengue Virus	BHK-21	Inconclusive(1)
		Unknown	Active(1)
	RSV‡	Vero	Active(8)
	West Nile Virus	BHK-21	Inconclusive(1)
		Unknown	Active(1)
	Yellow Fever Virus+	BHK	Inactive(3)
		BHK-21	Inactive(4)
CHEMBL682	Dengue Virus‡	BHK-21	Active(4)
	SARS-CoV-2+	Huh-7	Active(2)
		Vero	Active(1)
		Vero C1008	Active(2)
	Zika Virus+	NSC	Active(1)
		Vero	Active(1)
CHEMBL6948	HPIV-3	Vero	Inactive(1)
	RSV‡	Hela	Active(9)
	Yellow Fever Virus	Vero	Inactive(1)
CHEMBL76	Dengue Virus	BHK-21	Active(1)
	SARS-CoV-2+	Huh-7	Active(2)
		Vero	Active(1)
	Zika Virus*	HBMEC	Active(1)
		NSC	Active(1)
		Vero #	Active(1), Inactive(3)
CHEMBL866	Dengue Virus+	BHK-21	Inconclusive(1)
		BHK-D2RepT	Active(1)
		Huh-7	Active(1)
		Vero	Active(2)
	HPIV-3+	Hela	Active(1)
		Vero	Active(1)
	RSV‡	Vero	Inconclusive(1), Active(3)
	SARS-CoV-2*	Caco-2	Inactive(1)
		Vero C1008 #	Active(2), Inactive(5)
	West Nile Virus+	BHK-21	Inconclusive(1), Active(1)
		Vero	Active(1)

	Yellow Fever Virus*	BHK	Inactive(1)
		BHK-21	Inactive(3)
		Vero	Active(2)

Compounds with activity against 2 viruses in phenotypic assays

Table S5. Compounds active against two viruses in phenotypic assays.

Molecule ID	SMILES	Active	Inactive
CHEMBL393857	NC1C=CN(C2OC(CO)([N-][N+]#N)C(O)C2O)C(=O)N=1	RSV, Dengue Virus	SARS-CoV-2
CHEMBL3769592	CCN1CCN(CC1)C(c1ccccc1)c1ccc(Cl)cc1	RSV, Dengue Virus	
CHEMBL2016758	CC1(OC(CO)C(O)C1O)c1ccc2c(N)[n]c[n][n]12	RSV, HPIV-3	engue Virus, West Nile Virus
CHEMBL3417243	NC1C=CN(C2OC(CO)(CCl)C(O)C2F)C(=O)N=1	RSV, HPIV-3	SARS-CoV-2
CHEMBL2017637	Cc1ccc(NC(=O)C(C)Sc2[n]c3ccccc3[n]2-c2ccc(cc2)OC)c(Cl)c1	RSV, HPIV-3	
CHEMBL491005	Clc1cc(ccc1)N1CCN(CC1)C=CN=Nc1ccccc1	RSV, Yellow Fever Virus	
CHEMBL491006	Clc1ccc(cc1)N1CCN(CC1)C=CN=Nc1ccccc1	RSV, Yellow Fever Virus	
CHEMBL491007	FC(F)(F)c1cc(ccc1)N1CCN(CC1)C=CN=Nc1ccccc1	RSV, Yellow Fever Virus	
CHEMBL2304286	)c1ccc(CC2C(=O)OC(C)C(NC(=O)C(CC(C)C)N(C)C()C3CCCN3C(=O)C(C)O)C(=O)NC(C(O)CC(=O)OC(C)C)C(=O)C(C)C(=O)NC(CC(C)C)C(=O)N3CCCC3C(=O)N2C)C(C)CC)cc1	Sandfly Fever, Yellow Fever Virus	
CHEMBL2059155	Nc1[n]cc(Br)c2c1[n]c[n]2C1C=C(CO)C(O)C1O	HPIV-3, Dengue Virus	West Nile Virus
CHEMBL4441971	=C(C=Cc1ccccc1)NC(CC1CCCCC1)C(=O)NC(CC1CCNC1=O)C(=O)C(=O)NCc1ccccc1	MERS-CoV, HCoV-229E	
CHEMBL4590737	C(C)CC(NC(=O)C=Cc1ccccc1)C(=O)NC(CC1CCNC1=O)C(=O)C(=O)NCc1ccccc1	MERS-CoV, HCoV-229E	
CHEMBL1197690	c1cc(Nc2ccc(cc2)OCc2ccccc2Cl)c2cc(ccc2[n]1)OCC	Dengue Virus, West Nile Virus	
CHEMBL1630220	C(C)C(=O)OCC1OC([n]2cc(c3c2[n]c[n]c3N)C(N)=O)C(O)(C#C)C1OC(=O)C(C)C	Dengue Virus, West Nile Virus	
CHEMBL1652113	OCC1C(O)C(O)C(O)CN1CCCCCC1(O)CCCCC1	Dengue Virus, West Nile Virus	
CHEMBL3542231	CC(C)C1OC2(CC3CC(CC=C(C)C(OC4CC(OC5CC(OC(O)C(C)O5)C(OC)C(C)O4)C(C)C=CC=C4COC5C(O)C(C)=CC(C(=O)O3)C45O)O2)CCC1C	Dengue Virus, West Nile Virus	
CHEMBL3578024	C1=CN(C2CC(C(CO[Si](C)(C)C(C)(C)C)O2)[n]2[n][n]cc2-c2ccc3cc(ccc3c2)OC)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578025	C1=CN(C2CC(C(CO[Si](C(C)C)(C(C)C)C(C)C)O2)[n]2[n][n]cc2-c2ccc3cc(ccc3c2)OC)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578026	C1=CN(C2CC(C(CO[Si](c3ccccc3)(c3ccccc3)C(C)(C)O2)[n]2[n][n]cc2-c2ccc3cc(ccc3c2)OC)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578034	C1=CN(C2CC(C(CO[Si](C)(C)C(C)(C)C)O2)[n]2[n][n]cc2-c2cccc3ccccc32)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578035	C1=CN(C2CC(C(CO[Si](C)(C)C(C)(C)C)O2)[n]2[n][n]cc2-c2ccc(cc2)-c2ccccc2)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578036	C1=CN(C2CC(C(CO[Si](C)(C)C(C)(C)C)O2)[n]2[n][n]cc2-c2ccc(cc2)Oc2ccccc2)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578039	C1=CN(C2CC(C(CO[Si](C)(C)C(C)(C)C)O2)[n]2[n][n]cc2C2CCCCC2)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578040	C1=CN(C2CC(C(CO[Si](C)(C)C(C)(C)C)O2)[n]2cc([n][n]2)-c2ccc3cc(ccc3c2)OC)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578041	C1=CN(C2CC(C(CO[Si](C)(C)C(C)(C)C)O2)[n]2cc([n][n]2)-c2cc3ccccc3c3ccccc32)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578049	C1=CN(C2CC(NC(=O)c3cc4ccc(cc4cc3)OC)C(CO[Si](C)(C)C(C)(C)C)O2)C(=O)NC1=O	Dengue Virus, West Nile Virus	

CHEMBL3578057	1=CN(C2CC(NC(=O)c3cc4ccccc4cc3)C(CO[Si](C)(C)C)C(C)(C)O2)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578063	1=CN(C2CC(NS(=O)(=O)c3cc4ccccc4cc3)C(CO[Si](C)(C)C)C(C)(C)O2)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578066	1=CN(C2CC(NC(=O)c3ccc(cc3)N3CCOCC3)C(CO[Si](C)(C)C)C(C)(C)O2)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578067	1=CN(C2CC(NC(=O)Nc3ccccc3)C(CO[Si](C)(C)C(C)(C)C)O2)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL3578068	1=CN(C2CC(NC(=O)NS(=O)(=O)c3ccc(C)cc3)C(CO[Si](C)(C)C)C(C)(C)O2)C(=O)NC1=O	Dengue Virus, West Nile Virus	
CHEMBL408500	CCCCCCCCCN1CC(O)C(O)C(O)C1CO	Dengue Virus, West Nile Virus	
CHEMBL4164410	Cc1c[n]([n][n]1)-c1ccc(cc1)NC(=O)Nc1ccccc1C	Dengue Virus, West Nile Virus	
CHEMBL73809	CC1(O)C(OC(CO)C1O)[n]1c[n]c2c(N)[n]c[n]c12	Dengue Virus, West Nile Virus	
CHEMBL251056	OC1CC2C(NC(=O)c3cc4OCOc4cc32)C(O)C1O	engue Virus, Yellow Fever Virus	Sandfly_Fever
CHEMBL3736249	Oc1ccc(cc1)-[n]1[n]c2cc(Cl)c(Cl)cc2[n]1	engue Virus, Yellow Fever Virus	West Nile Virus
CHEMBL400092	OC1C=C2CCN3Cc4cc5OCOc5cc4C(C32)C1O	engue Virus, Yellow Fever Virus	Sandfly_Fever
CHEMBL419335	c1c2OCOc2cc2C3C(NC(=O)c12)C(O)C(O)C(O)C3O	engue Virus, Yellow Fever Virus	Sandfly_Fever
CHEMBL476486	)c1c2C(=O)NC3C(O)C(O)C(O)CC=3c2cc2OCOc12	engue Virus, Yellow Fever Virus	Sandfly_Fever
CHEMBL586091	COc1cc2CN3CCC4=CC(O)C(O)C(C34)c2cc1O	engue Virus, Yellow Fever Virus	Sandfly_Fever
CHEMBL98745	)c1c2OCOc2cc2c1C(=O)NC1C(O)C(O)C(O)C=C21	engue Virus, Yellow Fever Virus	Sandfly_Fever
CHEMBL3339396	O=C(NCc1ccc(cc1)-c1cc[n]cc1)c1[n]cc[n]c1C(=O)Nc1ccc(F)cc1	engue Virus, Yellow Fever Virus	
CHEMBL3339398	O=C(NCc1ccc(cc1)-c1c[n]ccc1)c1[n]cc[n]c1C(=O)Nc1ccc(F)cc1	engue Virus, Yellow Fever Virus	
CHEMBL463048	OC1C(O)C2NC(=O)c3cc4OCOc4cc3C2=CC1O	engue Virus, Yellow Fever Virus	
CHEMBL487798	OC1C2C(NC(=O)c3cc4OCOc4cc32)C(O)C(O)C1O	engue Virus, Yellow Fever Virus	
CHEMBL98481	COc1c(cc(cc1OC)C=Cc1cc(O)c(cc1)OC)OC	engue Virus, Yellow Fever Virus	
CHEMBL67	Oc1cc2c(NC(C)CCCN(CC)CC)c3ccc(Cl)cc3[n]c2cc1	Dengue Virus, Zika Virus	SARS-CoV-2
CHEMBL7568	Oc1cc2c(NC(C)CCCN(CC)CC)c3ccc(Cl)cc3[n]c2cc1	Dengue Virus, Zika Virus	SARS-CoV-2
CHEMBL172	C(C1CCCCN1)c1cc([n]c2c(cccc21)C(F)(F)F)C(F)(F)F	Dengue Virus, Zika Virus	
CHEMBL330407	COc1c(cc(cc1OC)C=Cc1cc(N)c(cc1)OC)OC	Dengue Virus, Zika Virus	
CHEMBL4454990	CCCCOCc1c([n]c([n]c1Nc1ccc(cc1)S(N)(=O)=O)N1CCOCC1)C(=O)NCCOc1ccccc1	Dengue Virus, Zika Virus	
CHEMBL4518162	c1c(cc(cc1OC)C=Cc1cc(NC(=O)C(CO)NC(=O)C(CCC(N)=N)NC(=O)C(CCCNC(N)=N)NC(=O)c2ccccc2)c(cc1)OC)OC	Dengue Virus, Zika Virus	
CHEMBL464432	Oc1c2C(=O)NC3C(CC(O)C(O)C3O)c2cc2OCOc12	ellow Fever Virus, Dengue Virus	Sandfly_Fever
CHEMBL448322	COC(=O)c1[s]c([n]c1C(Br)Br)-c1ccc(Cl)cc1	ellow Fever Virus, Dengue Virus	
CHEMBL236207	CC1(O)C(OC(CO)C1O)[n]1ccc2c(N)[n]c[n]c12	West Nile Virus, Zika Virus	Dengue Virus

Cluster Analysis Methodology

The curated compound structures were submitted to hierarchical cluster analysis in KNIME v.4.1.4 (Berthold et al., 2007) integrated with python v.3.7.3 (SciPy and Matplotlib libraries) and using RDKit descriptors (RDKit v.4.2.0). The optimal number of clusters was determined by the software default Euclidean distance cut-off.

Cluster analysis of active compounds in phenotypic assays

The structural clustering of all compounds tested in phenotypic cell-based assays revealed ten clusters (**Figure S2, Figure S3, Table S6**). The top BSA compounds, CHEMBL4437334 and CHEMBL4454780, active against four different viruses, are in clusters #7 and #5, respectively. The subcluster containing CHEMBL4437334 (cluster #7) has 847 compounds. Among them, some nearest neighbors of CHEMBL4437334 (**Figure S3**) were active against SARS-CoV-2 and could be further tested against a panel of flaviviruses (Dengue, West Nile, Yellow Fever, and Zika). The subcluster of CHEMBL4454780 (cluster #5) contains 1,406 compounds; the nearest neighbors of CHEMBL4454780 are presented in **Figure S3**. Compounds CHEMBL1197690, CHEMBL3581155, and CHEMBL7568 were active against one or two flaviviruses and could be further tested against additional flaviviruses and viruses from other families like RSV and MERS-CoV. Likewise, CHEMBL4303559 was only tested and active against SARS-CoV-2 and could be tested against members of *Flaviviridae* and other coronaviruses such as MERS-CoV.

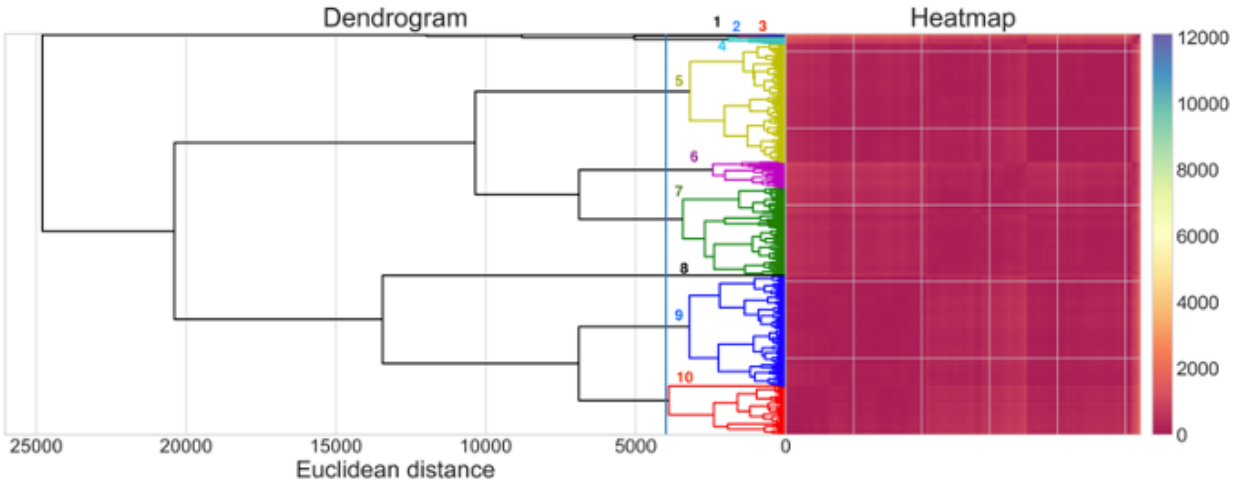


Figure S2. Clustering of compounds from the phenotypic cell-based assays by chemical structure. The colors from the heatmap are based on the Euclidean distances between compounds. Colors nearer to dark red indicate a shorter distance between molecules.

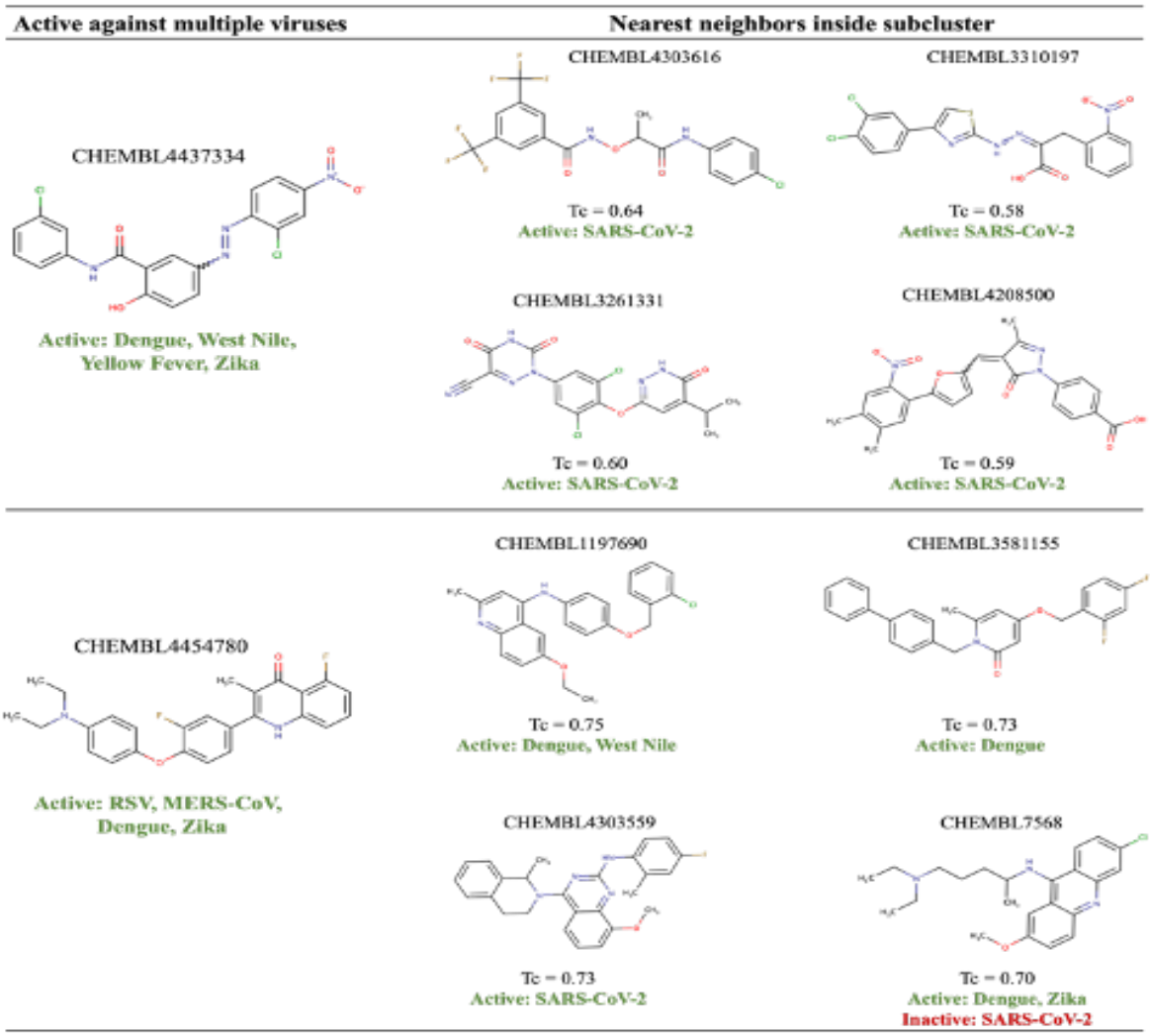
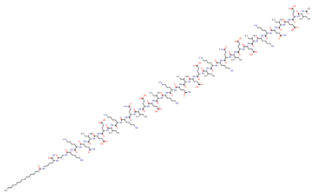
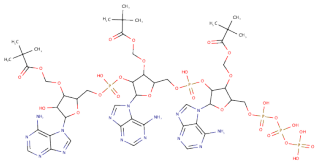
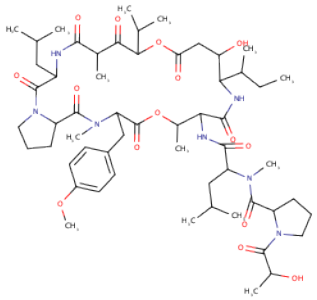
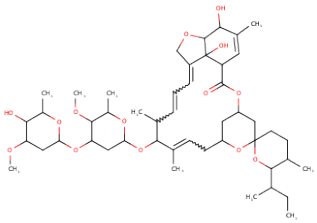
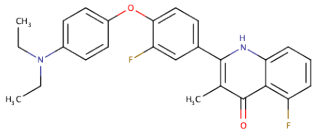
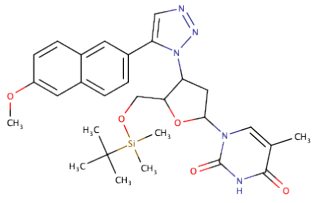
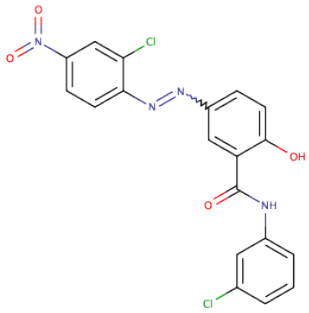
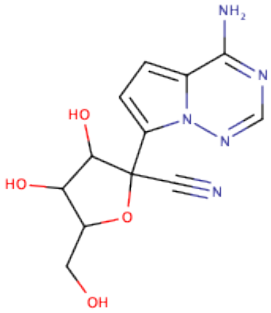
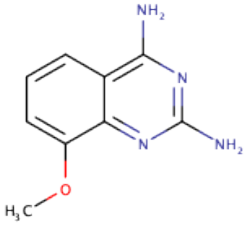
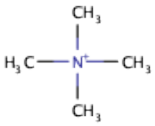


Figure S3. Examples of compounds similar to experimental broad-spectrum hits that could be further tested against multiple viruses of interest.

Table S6. Examples of compounds of each cluster from phenotypic data.

Cluster	Representative compound	Structure	Active against	Inactive against

1	CHEMBL4206744		MERS-CoV
2	CHEMBL2414866		RSV
3	CHEMBL2304286		Sandfly Fever, Yellow Fever Virus
4	CHEMBL3542231		Dengue Virus, West Nile Virus
5	CHEMBL4454780		RSV, MERS-CoV, Dengue Virus, Zika Virus
6	CHEMBL3578024		Dengue Virus, West Nile Virus

7	CHEMBL4437334		Dengue Virus, West Nile Virus, Yellow Fever Virus, Zika Virus	
8	CHEMBL2016757		RSV, HPIV-3, Dengue Virus	West Nile Virus
9	CHEMBL2042315		Dengue Virus	
10	CHEMBL46486			SARS-CoV-2

Target-based Clustering Analysis

Structural clustering of all compounds revealed 11 clusters (**Figure S4, Figure S5, Table S7**). CHEMBL4544781 was active against targets from different viral families (NS5 of Zika Virus and RDRP of MERS-CoV) and is in cluster #7 along with 867 other compounds; nearest neighbors of CHEMBL4544781 are presented in **Figure S5**. CHEMBL1630221 was only tested (and active) against the NS5 Polymerase of Zika Virus and could be further tested against other polymerases from other flaviviruses and the RNA-Dependent RNA Polymerase (RdRP) of MERS-CoV. Three other nearest neighbors of CHEMBL4544781 were only tested (and active) against SARS-CoV-2 Main Protease (M^{pro}). These compounds could be further tested against polymerases of Zika and MERS-CoV.

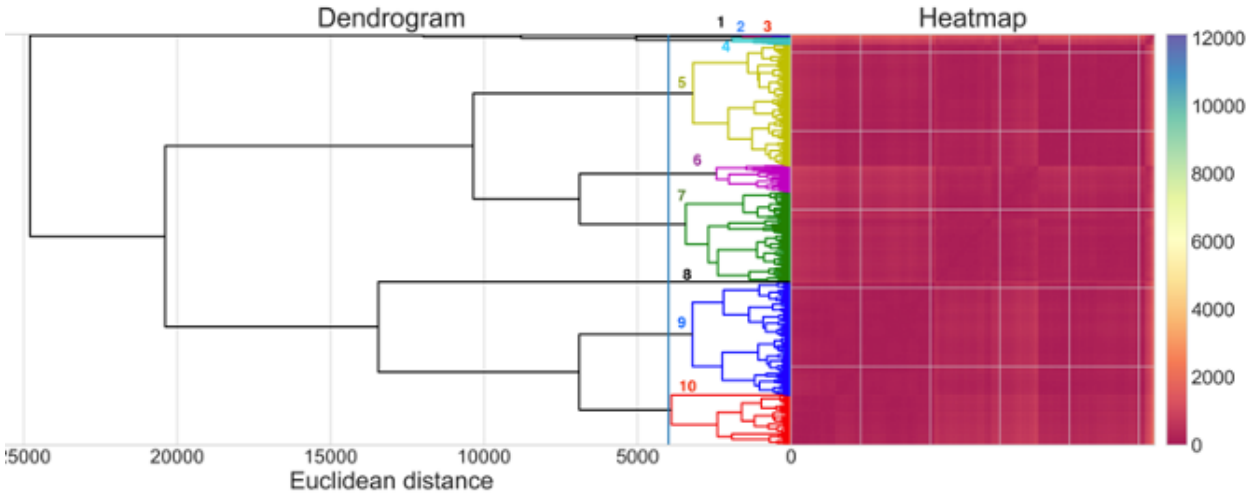


Figure S4. Clustering analysis of compounds from the target-based assays. The colors from the heatmap are based on the Euclidean distances between compounds. Colors nearer to dark red indicate a shorter distance between molecules.

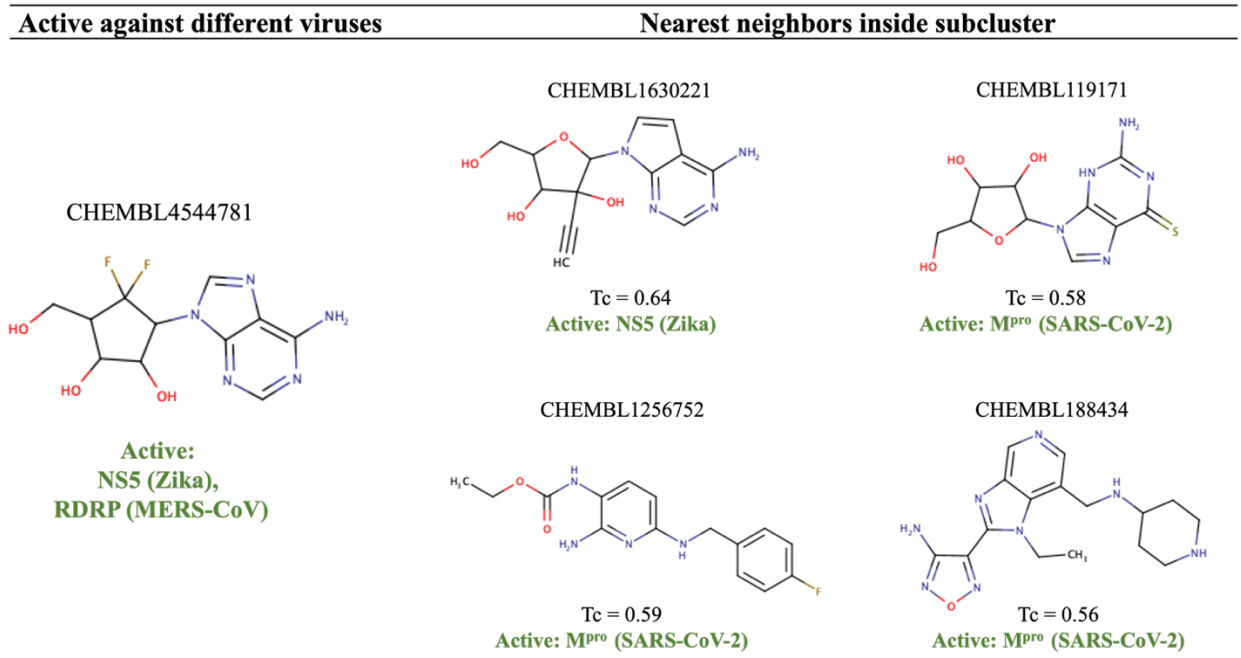
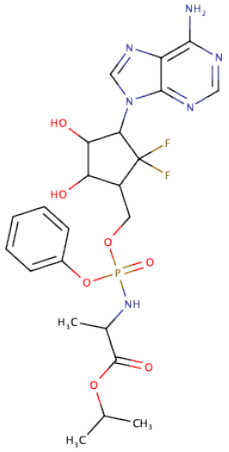
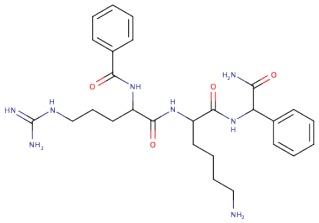
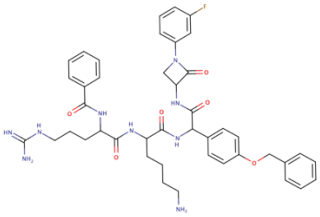
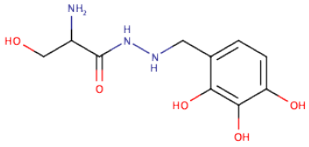
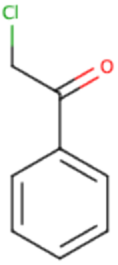
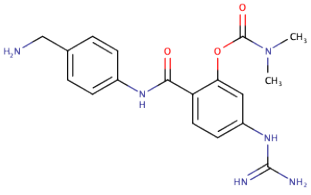
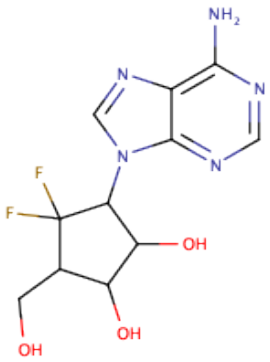
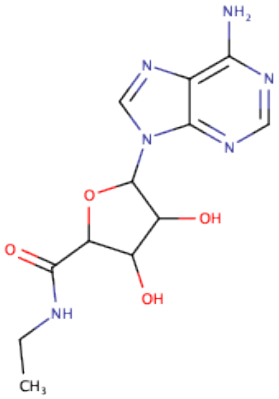
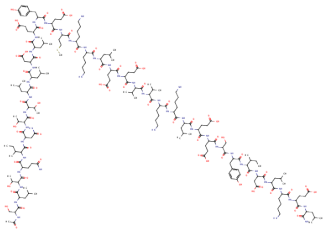
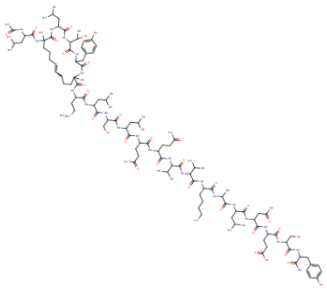


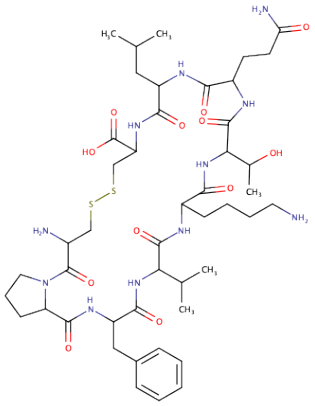
Figure S5. Examples of compounds that could be further tested against different viral targets of interest due to their chemical similarity to an active molecule with multiple antiviral activities.

Table S7. Examples of compounds of each cluster from target-based data.

Cluster	Representative compound	Structure	Active against	Inactive against
1	CHEMBL4522602		NS5 (Zika), RDRP (MERS-CoV)	

2	CHEMBL3740277		NS2B-NS3 Protease (Dengue), NS2B-NS3 (West Nile)
3	CHEMBL4537775		NS2B-NS3 Protease (Dengue), NS2B-NS3 (West Nile)
4	CHEMBL1096979		M^{pro} (SARS-CoV-2)
5	CHEMBL105712		M^{pro} (SARS-CoV-2)
6	CHEMBL4536920		NS2B-NS3 Protease (Dengue), NS2B-NS3 (West Nile)

7	CHEMBL4544781		NS5 (Zika), RDRP (MERS-CoV)
8	CHEMBL464859		M^{pro} (SARS-CoV-2)
9	CHEMBL4127413		Spike protein (MERS-CoV)
10	CHEMBL4125813		Spike protein (MERS-CoV)

11	CHEMBL1630963		Integrin alpha-V/beta-3 (SNV)

Concordance between the phenotypic and target-based data

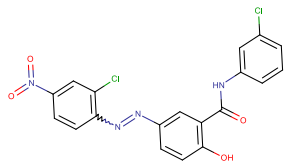
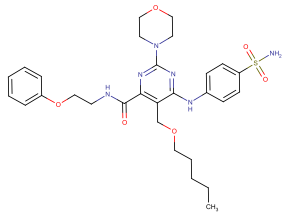
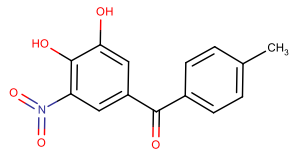
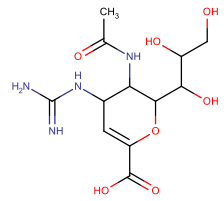
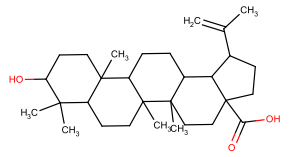
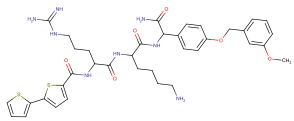
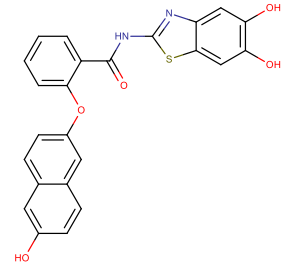
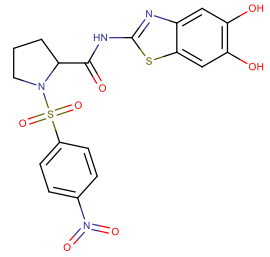
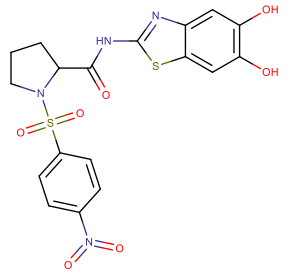
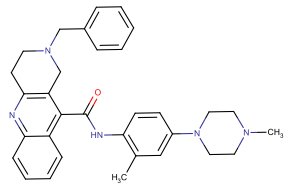
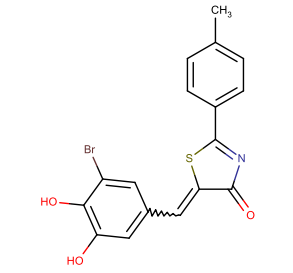
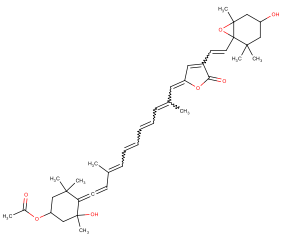
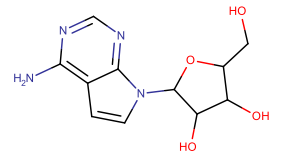
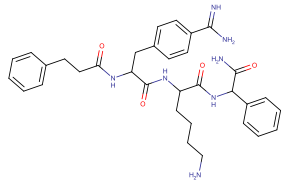
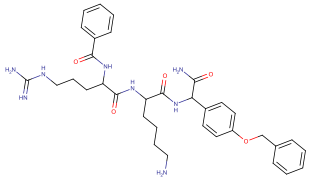
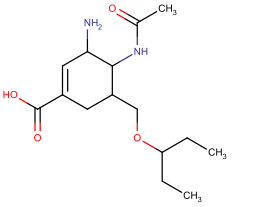
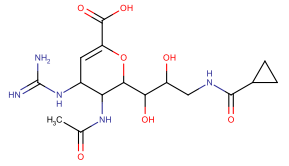
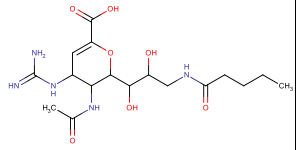
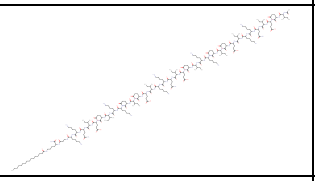
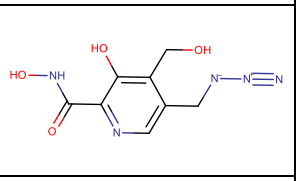
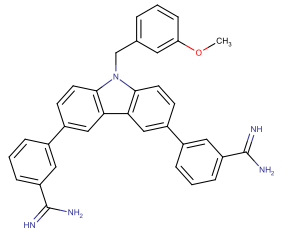
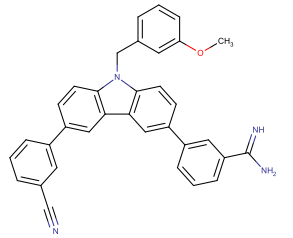
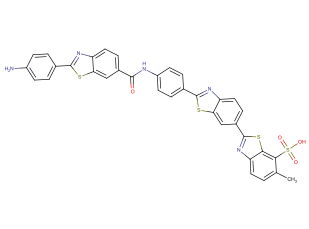
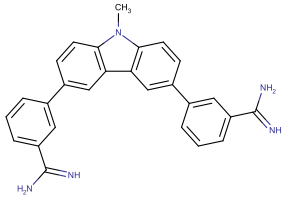
Table S8. Compounds active in at least one phenotypic and one target-based assay

	Phenotypic Assays		Target-Based Assays			
Molecule ID	**Viruses (active)**	**Viruses (inactive)**	**Viruses (active)**	**Targets**	**Viruses (inactive)**	**Targets**
CHEMBL4437334	Dengue Virus, West Nile Virus, Yellow Fever Virus, Zika Virus		Dengue, Zika	NS2B-NS3 Protease		
CHEMBL4454990	Dengue Virus, Zika Virus		Zika	NS5		
CHEMBL1324	Dengue Virus	SARS-CoV-2	Dengue	NS2B-NS3 Protease	SARS-CoV-2	Main Protease (3CLpro, Mpro)
CHEMBL222813	H7N7	SARS-CoV-2	H7N7	Neuraminidase	SARS-CoV-2	Main Protease (3CLpro, Mpro)
CHEMBL269277	Dengue Virus	SARS-CoV-2	Dengue	NS5	SARS-CoV-2	Main Protease (3CLpro, Mpro)
CHEMBL3741713	Dengue Virus	West Nile Virus	Dengue	NS2B-NS3 Protease		
CHEMBL4440832	Dengue Virus	Zika Virus	Dengue, Zika	NS2B-NS3 Protease		
CHEMBL4462325	Dengue Virus	Zika Virus	Zika	NS2B-NS3 Protease		
CHEMBL4583315	Dengue Virus	Zika Virus	Zika	NS2B-NS3 Protease		

CHEMBL1370977	Dengue Virus	Dengue	NS2B-NS3 Protease
CHEMBL1458891	Dengue Virus	Dengue	NS2B-NS3 Protease
CHEMBL1980535	Dengue Virus	Dengue	NS2B-NS3 Protease
CHEMBL267099	HPIV-3	Zika	NS5
CHEMBL3628278	Dengue Virus	West_Nile	NS2B-NS3 Protease
CHEMBL3741422	Dengue Virus	Dengue, West_Nile	NS2B-NS3 Protease
CHEMBL3818159	H7N7	H7N7	Neuraminidase
CHEMBL3818654	H7N7	H7N7	Neuraminidase
CHEMBL3819510	H7N7	H7N7	Neuraminidase
CHEMBL4206744	MERS-CoV	MERS-CoV	Spike protein
CHEMBL4439416	Zika Virus	Zika	NS5
CHEMBL4446364	Zika Virus	Zika	NS2B-NS3 Protease
CHEMBL4447165	Zika Virus	Zika	NS2B-NS3 Protease
CHEMBL4447800	Dengue Virus	Dengue	NS2B-NS3 Protease
CHEMBL4448497	Zika Virus	Zika	NS2B-NS3 Protease
CHEMBL4449109	Dengue Virus	Dengue	NS5
CHEMBL4457232	Zika Virus	Zika	NS2B-NS3 Protease
CHEMBL4522006	Dengue Virus	Dengue	NS2B-NS3 Protease
CHEMBL4526128	Dengue Virus	Dengue	NS5
CHEMBL4532866	Zika Virus	Zika	NS2B-NS3 Protease
CHEMBL4576745	Dengue Virus	Dengue, West_Nile	NS2B-NS3 Protease
CHEMBL4587069	Dengue Virus	Dengue	NS5
CHEMBL522355	West Nile Virus	West_Nile	NS2B-NS3 Protease
CHEMBL610669	West Nile Virus	West_Nile	not defined
CHEMBL611003	West Nile Virus	West_Nile	not defined
CHEMBL82242	Dengue Virus	Dengue	NS5

The chemical structures are presented below

CHEMBL4437334	**CHEMBL4454990**	**CHEMBL1324**	**CHEMBL222813**

			
CHEMBL269277	CHEMBL3741713	CHEMBL4440832	CHEMBL4462325
			
CHEMBL4583315	CHEMBL1370977	CHEMBL1458891	CHEMBL1980535
			
CHEMBL267099	CHEMBL3628278	CHEMBL3741422	CHEMBL3818159
			
CHEMBL3818654	CHEMBL3819510	CHEMBL4206744	CHEMBL4439416
			
CHEMBL4446364	CHEMBL4447165	CHEMBL4447800	CHEMBL4448497
			
CHEMBL4449109	CHEMBL4457232	CHEMBL4522006	CHEMBL4526128

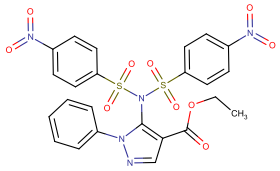
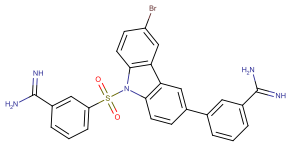
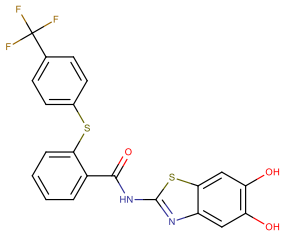
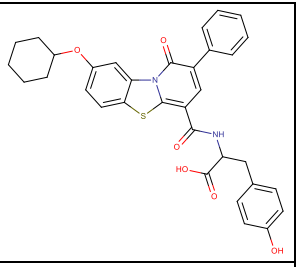
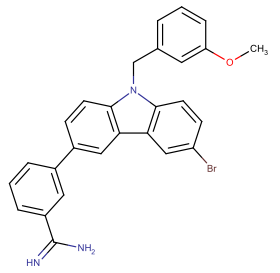
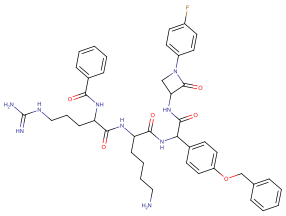
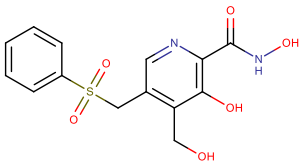
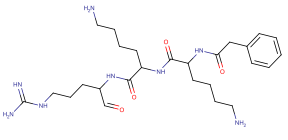
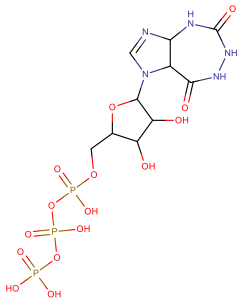
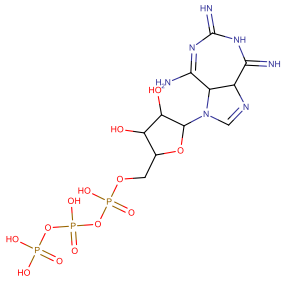
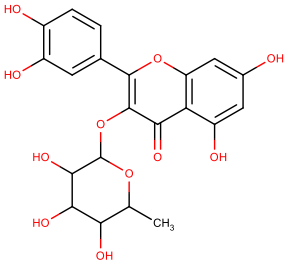
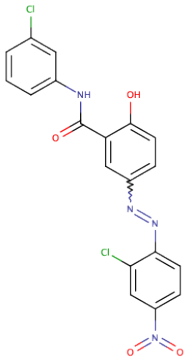
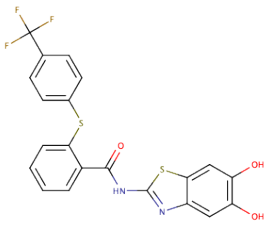
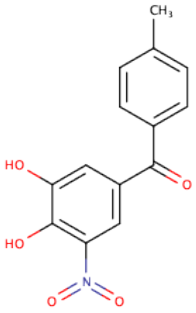
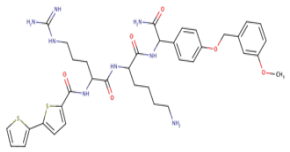
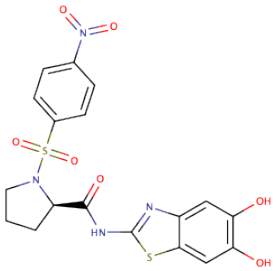
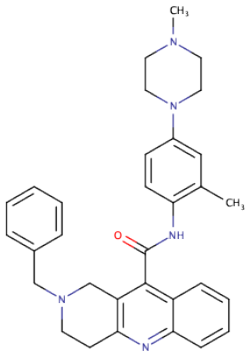
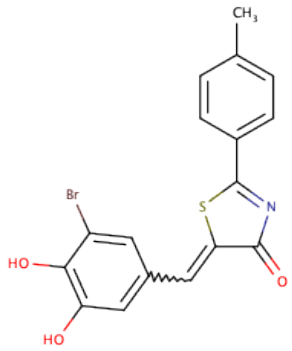
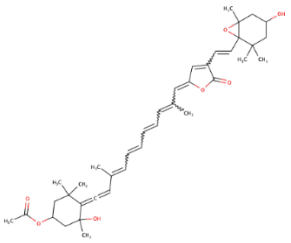
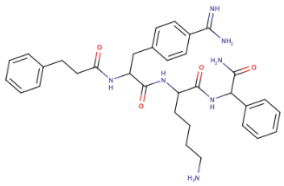
			
CHEMBL4532866	**CHEMBL4576745**	**CHEMBL4587069**	**CHEMBL522355**
			
CHEMBL610669	**CHEMBL611003**	**CHEMBL82242**	
			

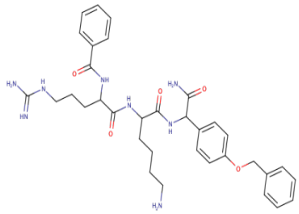
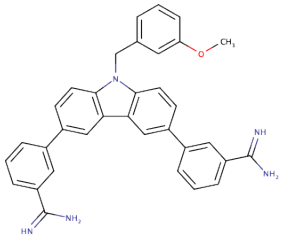
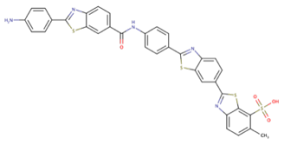
Table S9. Selection of compounds nominated for experimental testing as potential BSA agents

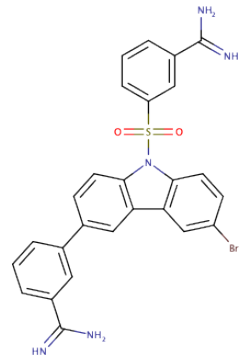
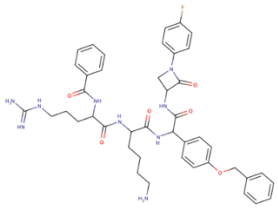
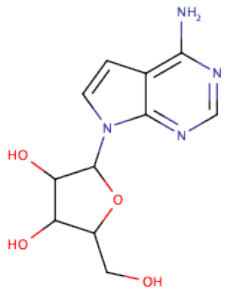
CHEMBL ID	Structure	Target active against	Virus(es) active against in cell-based assays	Suggested target-virus combination for testing	Reasoning for testing

CHEMBL4437334		NS2B-NS3 Protease (Dengue, Zika)	Dengue Virus, West Nile Virus, Yellow Fever Virus, Zika Virus	NS2B-NS3 homologs from other flaviviruses (West Nile, Yellow Fever)	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3
CHEMBL4522006		NS2B-NS3 Protease (Dengue)	Dengue Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3
CHEMBL1324		NS2B-NS3 Protease (Dengue)	Dengue Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3

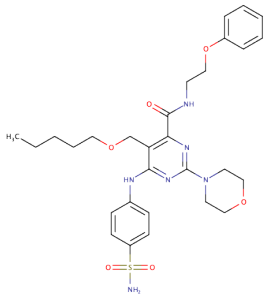
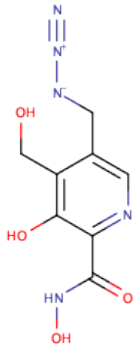
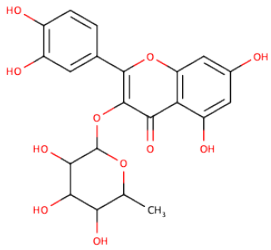
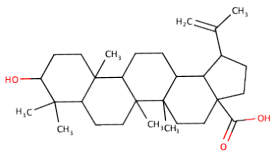
CHEMBL3741713		NS2B-NS3 Protease (Dengue)	Dengue Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3
CHEMBL4462325		NS2B-NS3 Protease (Zika)	Dengue Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3
CHEMBL1370977		NS2B-NS3 Protease (Dengue)	Dengue Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3

CHEMBL1458891		NS2B-NS3 Protease (Dengue)	Dengue Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3
CHEMBL1980535		NS2B-NS3 Protease (Dengue)	Dengue Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3
CHEMBL3628278		NS2B-NS3 Protease (West Nile)	Dengue Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3

CHEMBL3741422		NS2B-NS3 Protease (Dengue, West Nile)	Dengue Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3
CHEMBL4446364		NS2B-NS3 Protease (Zika)	Zika Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3
CHEMBL4447800		NS2B-NS3 Protease (Dengue)	Dengue Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3

CHEMBL4457232		NS2B-NS3 Protease (Zika)	Zika Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3
CHEMBL4576745		NS2B-NS3 Protease (Dengue, West Nile)	Dengue Virus	NS2B-NS3 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS2B-NS3
CHEMBL267099		NS5 (Zika)	HPIV-3	Zika and other flaviviruses (cell-based assays) NS5 homologs in other flaviviruses (target-based) A panel of HPIV-3 targets	Cell-based assays against Zika and other flaviviruses to confirm NS5 inhibition as a possible mechanism of action

					Active against NS5 of Zika that could be tested against NS5 homologs from other flaviviruses Explore possible mechanism of action for HPIV-3 inhibition observed in a cell-based assay

CHEMBL4454990		NS5 (Zika)	Dengue Virus, Zika Virus	NS5 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS5
CHEMBL4439416		NS5 (Zika)	Zika Virus	NS5 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS5
CHEMBL82242		NS5 (Dengue)	Dengue Virus	NS5 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS5
CHEMBL269277		NS5 (Dengue)	Dengue Virus	NS5 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS5

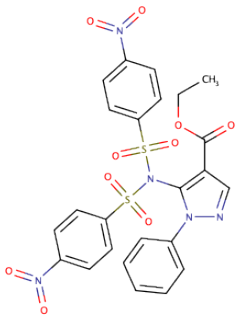
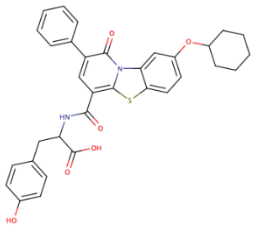
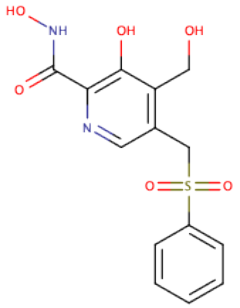
CHEMBL4449109		NS5 (Dengue)	Dengue Virus	NS5 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS5
CHEMBL4526128		NS5 (Dengue)	Dengue Virus	NS5 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS5
CHEMBL4587069		NS5 (Dengue)	Dengue Virus	NS5 homologs from other flaviviruses Other flaviviruses in cell-based assays	Cell-based and target-based assays support the hypothesis of activity against flaviviruses by targeting NS5

Table S10. Compounds active in at least one phenotypic assay and inactive in a target-based assay.

ChEMBL ID	SMILES
CHEMBL941	CN1CCN(Cc2ccc(cc2)C(=O)Nc2cc(Nc3[n]c(cc[n]3)-c3c[n]ccc3)c(C)cc2)CC1
CHEMBL29197	COc1cc2c(Nc3cc(Br)ccc3)[n]c[n]c2cc1OC
CHEMBL97771	COc1cc2[n]c[n]c(Nc3ccc(cc3)Oc3ccccc3)c2cc1OC
CHEMBL483847	NC(=O)c1cc(ccc1)-c1cc(Nc2ccc(cc2)OC(F)(F)F)[n]c[n]1
CHEMBL296407	NC(=O)C(=Cc1cc(O)c(O)cc1)C#N
CHEMBL249089	O=S(=O)(c1cccc2c(I)cccc12)N1CCCNCC1
CHEMBL1797882	CC(C)(C)NC(=O)C(=O)C=Cc1cc(OC)c(cc1)OC
CHEMBL304087	COc1cc(ccc1-c1c[n]co1)NC(=O)Nc1cc(CNC(=O)OC2COCC2)ccc1

CHEMBL6741	CC1(C)N=C(N)N=C(N)N1c1ccc(Br)cc1
CHEMBL7130	CC1(C)N=C(N)N=C(N)N1c1cccc(Cl)c1
CHEMBL7301	CC1CCCC(C)(C)C=1C=CC(C)=CC=CC(C)=CC(=O)Nc1ccc(O)cc1
CHEMBL7568	COc1cc2c(NC(C)CCCN(CC)CC)c3ccc(Cl)cc3[n]c2cc1
CHEMBL3410450	Cc1c[n]2[n]c(cc2[n]c1N1CC(N)CC1)C1CCCCN1C(=O)c1cc(Cl)ccc1NS(C)(=O)=O
CHEMBL408963	Cc1ccc(CCCO)c(c1)NCc1cc2c(cc1)[n]c(NCCCN1CCOCC1)[n]2Cc1[n]c(C)ccc1O
CHEMBL513181	NC1C=CN(C2OC(CO)([N-][N+]#N)C(O)C2O)C(=O)N=1
CHEMBL3417243	NC1C=CN(C2OC(CO)(CCl)C(O)C2F)C(=O)N=1
CHEMBL393857	NC1C=CN(C2OC(CO)([N-][N+]#N)C(O)C2O)C(=O)N=1
CHEMBL682	CCN(Cc1cc(ccc1O)Nc1cc[n]c2cc(Cl)ccc21)CC
CHEMBL3781415	Nc1[n]c(CC2CC2)c(c[n]1)-c1cc[n]c(Nc2cc[n]cc2)[n]1
CHEMBL1205	Cc1c(cccc1O)C(=O)NC(CSc1ccccc1)C(O)CN1CC2CCCCC2CC1C(=O)NC(C)(C)C
CHEMBL3407547	Cc1c2c([s]c1-c1[n]cco1)N(CC(OC1CCOCC1)c1ccccc1OC)C(=O)N(C2=O)C(C)(C)C(O)=O
CHEMBL137648	O=C1N(Cc2cc(F)ccc2)CN(c2ccccc2)C21CCN(CCCC(=O)c1ccc(F)cc1)CC2
CHEMBL119	Cc1c(CNc2cc(OC)c(OC)c(c2)OC)ccc2[n]c(N)[n]c(N)c12
CHEMBL38434	Cc1c([n]c2ccc(F)cc2c1C(O)=O)-c1ccc(cc1)-c1ccccc1F
CHEMBL1078	CC(C)[n]1c2ccccc2c(c1C=CC(O)CC(O)CC(O)=O)-c1ccc(F)cc1
CHEMBL54440	CCC(C)C(=O)OC1CCC=C2C=CC(C)C(CCC3CC(O)CC(=O)O3)C21
CHEMBL503	CCC(C)C(=O)OC1CC(C)C=C2C=CC(C)C(CCC3CC(O)CC(=O)O3)C21
CHEMBL1064	CCC(C)(C)C(=O)OC1CC(C)C=C2C=CC(C)C(CCC3CC(O)CC(=O)O3)C21
CHEMBL4533598	O=C(Nc1ccc[s]1)c1ccccc1Sc1ccc(cc1)N(=O)=O
CHEMBL4457482	CCCCCCCCCCCCCCCCC[N+](CO)(CO)CC
CHEMBL553	COCCOc1cc2[n]c[n]c(Nc3cc(ccc3)C#C)c2cc1OCCOC
CHEMBL67	COc1c(cc(cc1OC)C=Cc1cc(O)c(cc1)OC)OC
CHEMBL41	CNCCC(Oc1ccc(cc1)C(F)(F)F)c1ccccc1
CHEMBL1263	Oc1ccc(cc1CO)C(O)CNCCCCCCOCCCCc1ccccc1
CHEMBL251254	OCC1OC(OC2C(=O)c3c(cc(O)cc3O)OC=2c2cc(O)c(O)cc2)C(O)C(O)C1O
CHEMBL226335	CC1OC(OCC2OC(OC3C(=O)c4c(cc(O)cc4O)OC=3c3cc(O)c(O)cc3)C(O)C(O)C2O)C(O)C(O)C1O
CHEMBL289277	Cc1cc(O)c2c(c1)C(=O)c1cc(O)cc(O)c1C2=O
CHEMBL4470866	C[n]1c2ccc(Br)cc2c2cc(ccc12)-c1cc(ccc1)C(N)=N
CHEMBL4444928	NC(=N)c1cc(C[n]2c3ccc(Br)cc3c3cc(Br)ccc23)ccc1
CHEMBL4437316	NC(=N)c1cc(ccc1)S(=O)(=O)[n]1c2ccc(Br)cc2c2cc(Br)ccc12
CHEMBL4544053	NC(=N)c1cc(ccc1)-c1cc2cc(ccc2cc1)-c1cc(ccc1)C(N)=N
CHEMBL430893	COc1ccc(cc1OC1CCCC1)C1CC(=O)NC1
CHEMBL325795	COc1ccc(cc1OC1CCCC1)C1CC(=O)NC1
CHEMBL113142	Nc1[n]c(NCCc2ccc(O)cc2)[n]c2[n]c([n][n]21)-c1ccco1
CHEMBL63	COc1ccc(cc1OC1CCCC1)C1CC(=O)NC1
CHEMBL240624	COCCOc1ccc(cc1)N1CCN(CC[n]2[n]cc3c4[n]c([n][n]4c(N)[n]c23)-c2ccco2)CC1
CHEMBL16901	C=CCc1cc(-c2cc(CC=C)c(O)cc2)c(O)cc1
CHEMBL2146722	CC1CCN(CC1)C(c1ccc(F)cc1)c1ccc(cc1-c1ccc(cc1)C(F)(F)F)C(C)C(O)=O
CHEMBL1491	CC1NC(=COCCN)C(C(C=1C(=O)OC)c1ccccc1Cl)C(=O)OCC
CHEMBL3039369	COC1CC(OC(C)C1O)OC1C(C)C=CC=C2COC3C(=NO)C(C)=CC(C(=O)OC4CC(CC=C1C)OC1(C4)CCC(C)C(O1)C1CCCCC1)C23O
CHEMBL430226	CC(=O)ON=C1C(Nc2ccccc12)=C1c2ccc(Br)cc2NC1=O
CHEMBL4518162	COc1c(cc(cc1OC)C=Cc1cc(NC(=O)C(CO)NC(=O)C(CCCNC(N)=N)NC(=O)C(CCCNC(N)=N)NC(=O)c2ccccc2)c(cc1)OC)OC

Table S11. Compounds inactive in a phenotypic assay and active in at least one target-based assay

ChEMBL ID	SMILES
CHEMBL324842	OC(=O)C(Cc1ccc(O)c(O)c1)OC(=O)C=Cc1ccc(O)c(O)c1
CHEMBL288114	OC(=O)c1cc(O)c(O)c(O)c1
CHEMBL4288319	O=C(OC1=CS(=O)(=O)Nc2ccc(cc12)Oc1ccccc1Br)c1ccc(F)c(F)c1
CHEMBL2440341	CCCCC(NC(=O)C(CCCCN)NC(=O)C(CCCNC(N)=N)NC(=O)c1ccc(cc1)C=C(C#N)C(=O)NC1CC1)C(N)=O
CHEMBL4460508	Cc1ccc(cc1)S(=O)(=O)N1CCCC1C(=O)Nc1[n]c2cc(O)c(O)cc2[s]1
CHEMBL4560058	Cc1ccc(cc1)S(=O)(=O)N1CCCC1C(=O)Nc1[n]c2cc(O)c(O)cc2[s]1
CHEMBL4437288	C[n]1c2ccc(cc2c2cc(ccc12)-c1cc(ccc1)C(N)=N)-c1cc(ccc1)C#N
CHEMBL4518903	NC(=N)c1cc(ccc1)S(=O)(=O)[n]1c2ccc(cc2c2cc(ccc12)-c1cc(ccc1)C#N)-c1cc(ccc1)C#N
CHEMBL4548788	COc1cc(ccc1)S(=O)(=O)[n]1c2ccc(Br)cc2c2cc(ccc12)-c1cc(ccc1)C(N)=N
CHEMBL4591196	OCC1Cc2ccccc2N1C(=O)OCc1ccccc1
CHEMBL1214186	Nc1[n]c[n]c2c1[n]c[n]2C1OC(CC(N)CCC(N)C(O)=O)C(O)C1O
CHEMBL221722	NC(=O)C1=NC(F)=CNC1=O
CHEMBL55400	Nc1cc2[n]c3cc(N)ccc3cc2cc1
CHEMBL4080284	O=C1C2CCCN2C(c2ccc(cc2)N(=O)=O)N1c1ccc(cc1)N(=O)=O
CHEMBL4436496	Nc1cc(N)c(cc1N=Nc1ccc(cc1)S(=O)(=O)Nc1[n]cc[s]1)N=Nc1ccc(cc1)S(=O)(=O)Nc1[n]cc[s]1
CHEMBL2440353	CCCCC(NC(=O)C(CCCCN)NC(=O)C(CCCNC(N)=N)NC(=O)c1ccc(cc1)C=C1SC(=O)N(C2CCCC2)C1=O)C(N)=O
CHEMBL4568823	COc1cc(ccc1)S(=O)(=O)[n]1c2ccc(cc2c2cc(ccc12)-c1cc(ccc1)C(N)=N)-c1cc(ccc1)C(N)=N
CHEMBL4558642	COc1cc(ccc1)S(=O)(=O)[n]1c2ccc(cc2c2cc(ccc12)-c1cc(ccc1)C#N)-c1cc(ccc1)C(N)=N
CHEMBL485818	OC(=O)C1OC(OC2C=C3OC(=CC(O)=C3C(=O)C=2O)c2ccccc2)C(O)C(O)C1O
CHEMBL8260	OC1C=C2OC(=CC(O)=C2C(=O)C=1O)c1ccccc1
CHEMBL660	NC12CC3CC(C1)CC(C2)C3
CHEMBL3989759	Cc1c(F)cc2[n]c3C4=CC5=C(COC(=O)C5(O)CC)C(=O)N4Cc3c3C(N)CCc1c23
CHEMBL4297188	Nc1[n]cccc1-c1[n]c2ccc([n]c2[n]1-c1ccc(cc1)C1(N)CCC1)-c1ccccc1
CHEMBL69863	Oc1ccc(cc1O)C=Cc1cc(O)cc(O)c1
CHEMBL1555751	CN(C)C1C2C(O)C3C(=C)c4cccc(O)c4C(=O)C3=C(O)C2(O)C(=O)C(C(N)=O)C1=O
CHEMBL4303331	CC1OC(OC2C(=O)c3c(cc(O)cc3O)OC=2c2cc(O)c(O)c(O)c2)C(O)C(O)C1O
CHEMBL1201073	OC(=O)C(S)C(S)C(O)=O
CHEMBL4303376	CN(c1ccccc1C)C(=O)c1cc(c(cc1)S(=O)c1[n]cc[s]1)N(=O)=O
CHEMBL391997	Oc1cc2CN(CCCc2cc1O)C(=S)NCCc1ccc(Cl)cc1
CHEMBL4303445	COC(=O)C1=C(NC(=S)N1C(CC)c1cc(F)c(F)cc1)c1cc[n]o1
CHEMBL94841	CCN1C(=O)NN=C1c1ccc(Cl)cc1
CHEMBL964	CCN(CC)C(=S)SSC(=S)N(CC)CC
CHEMBL1414870	NC(=S)NNc1ccc(cc1)N=NC(N)=N
CHEMBL3769507	CCN1C=C(c2[s]c(cc2C1=O)C(=N)NC1CCS(=O)(=O)CC1)c1cc(ccc1)C(F)(F)F
CHEMBL178459	CC1C(=S)SSC=1c1c[n]cc[n]1
CHEMBL485549	Oc1c(cc(cc1N(=O)=O)N(=O)=O)C(=O)NN=Cc1ccc(o1)N(=O)=O
CHEMBL4303154	COC(=O)C1=COC(OC2OC(CO)C(O)C(O)C2O)C(=CC)C1CC(=O)OCCc1cc(O)c(O)cc1
CHEMBL4303622	Nc1[n]c(N)c(cc1SC#N)SC#N
CHEMBL353335	Oc1c(Cl)c2CCNCC(c2cc1O)c1ccccc1
CHEMBL2107818	CC1OC(CC(N)C1O)OC1CC(O)(Cc2c(O)c3C(=O)c4cccc(OC)c4C(=O)c3c(O)c21)C(CO)=NNC(=O)CCCCCN1C(=O)C=CC1=O
CHEMBL1256915	CN1CCC2(C)C1N(C)c1ccc(O)cc21

CHEMBL286738	NS(=O)(=O)Oc1ccc2C3CCCCCC=3C(=O)Oc2c1
CHEMBL1256752	CCOC(=O)Nc1ccc(NCc2ccc(F)cc2)[n]c1N
CHEMBL1380480	COc1cc(ccc1O)C(=S)N1CCOCC1
CHEMBL1516068	Clc1ccc(cc1)C(OCCN1CCCCC1)c1ccccc1
CHEMBL66953	OC1C=C2C=CC=C(O)C(O)=C2C(=O)C=1O
CHEMBL224120	C[S+](CC1OC(C(O)C1O)[n]1c[n]c2c1[n]c[n]c2N)CCC(N)C(O)=O
CHEMBL1213887	CC1NC(=S)NC(C=1C(=O)OCC)c1cc(O)ccc1
CHEMBL1356238	O[n+]1ccccc1S
CHEMBL3392011	OC(Cc1cc(O)c(O)cc1)C(O)=O
CHEMBL4303595	O=C1C=Cc2cc(Br)ccc2C1=O
CHEMBL1271764	CN1CCN(CC1)C1=C(Cl)C(=O)N(c2ccc(Cl)c(Cl)c2)C1=O
CHEMBL297453	Oc1cc2OC(C(Cc2c(O)c1)OC(=O)c1cc(O)c(O)c(O)c1)c1cc(O)c(O)c(O)c1
CHEMBL421	OC(=O)c1cc(ccc1O)N=Nc1ccc(cc1)S(=O)(=O)Nc1cccc[n]1
CHEMBL60718	Cc1ccc(cc1)S(=O)(=O)NC(Cc1ccccc1)C(=O)CCl
CHEMBL3348861	OC1C(Oc2cc(O)cc(O)c2C1=O)c1cc(O)c(O)c(O)c1
CHEMBL571700	CN(C)C(=S)SC(=S)N(C)C
CHEMBL120563	CN(C)C(=S)SSC(=S)N(C)C
CHEMBL1092926	Fc1cc(cc(F)c1CN1CCOCC1)-c1cccc2[n]cc([n]c21)-c1c[n]([n]c1)C1CCNCC1
CHEMBL1230122	CN1C(=O)C(=Cc2c[n]c(NC3CCOCC3)[n]c12)Oc1ccc(F)cc1F
CHEMBL3963349	CC(C)C(NC(=O)C(CCC(=O)OC)NC(=O)C(CC(=O)OC)NC(=O)OCc1ccccc1)C(=O)NC(CC(=O)OC)C(=O)CF
CHEMBL3797437	Oc1cc2OCOc2cc1C(c1ccc(cc1)C(F)(F)F)N1CCOCC1
CHEMBL596674	CC(C)N(Cc1ccccc1)CCC(=O)c1cc2ccccc2cc1
CHEMBL1725120	Cc1cc(c(C)cc1)N1Sc2cc(F)ccc2C1=O
CHEMBL1382627	Nc1ccc(cc1)S(=O)(=O)Nc1[n]ccc[n]1
CHEMBL1408862	OCC(Br)(CO)N(=O)=O
CHEMBL1271993	O=C1C(=C(Cl)C(=O)N1c1ccc(Cl)c(Cl)c1)N1CCOCC1
CHEMBL1161936	OC(=O)c1c[n]c(cc1)SSc1ccc(c[n]1)C(O)=O
CHEMBL1374426	CCC(=O)OCN1C(=O)C=CC1=O
CHEMBL3545157	O=C1SN(c2cccc3ccccc32)C(=O)N1Cc1ccccc1
CHEMBL198997	Cc1[s]c([n]c1OCCOCCCC)-c1ccc(C(O)=O)c(F)c1
CHEMBL3989514	CCc1c2cc(ccc2[n]c2C3=CC4=C(COC(=O)C4(O)CC)C(=O)N3Cc21)OC(=O)N1CCC(CC1)N1CCCCC1
CHEMBL12208	CC1=CC(=O)Oc2cc(O)ccc21
CHEMBL4303503	CC1(C)CCC(O)C23COC(O)(CC12)C12C3CCC(C1O)C(=C)C2=O
CHEMBL1530581	C1COCCN1Sc1[n]c2ccccc2[s]1
CHEMBL1201236	CC(Cc1cc(O)c(O)cc1)(NN)C(O)=O
CHEMBL536282	NCC1=CNC(=S)N1C1Cc2cc(F)cc(F)c2CC1
CHEMBL270299	CC(C)(C)c1ccc(cc1)S(=O)(=O)C=CC#N
CHEMBL403183	Cc1ccc(cc1)S(=O)(=O)C=CC#N
CHEMBL17329	COC(=O)C=Cc1cc(O)ccc1O
CHEMBL2106486	CON=C(c1c[s]c(N)[n]1)C(=O)NC1C2SCC(CSC(=O)c3ccco3)=C(C(O)=O)N2C1=O
CHEMBL510038	Nc1[n][n]c(Sc2[n]cc([s]2)N(=O)=O)[s]1
CHEMBL28	Oc1cc(O)cc2OC(=CC(=O)c12)c1ccc(O)cc1

CHEMBL1671898	Cc1c[n]2c([n]1)c(cc(Cl)c2N)C(=O)NCC1CCN(CC(C)C)CC1
CHEMBL1200712	Cc1cc(ccc1NN=C1C=Cc2c(C1=O)c(N)c(cc2S(O)(=O)=O)S(O)(=O)=O)-c1cc(C)c(cc1)NN=C1C=Cc2c(C1=O)c(N)c(cc2S(O)(=O)=O)S(O)(=O)=O
CHEMBL680	NC(c1ccccc1)C(=O)NC1C2SCC(Cl)=C(C(O)=O)N2C1=O
CHEMBL3747259	OC(=O)C(Cc1cc(O)c(O)cc1)OC(=O)C=Cc1ccc(O)c2OC(C(c12)C(=O)OC(Cc1cc(O)c(O)cc1)C(O)=O)c1cc(O)c(O)cc1
CHEMBL51628	Oc1cc2OC(=O)C=Cc2cc1
CHEMBL596380	O=N(=O)C=Cc1cc2OCOc2cc1
CHEMBL961	CCN(CC)C(S)=S
CHEMBL1616	CN1CCc2cccc3c2C1Cc1ccc(O)c(O)c1-3
CHEMBL582355	CCOc1cc2c(cc1)[n]c1cc(N)ccc1c2N
CHEMBL243652	O=C1C(CC(=O)Nc2ccc(Br)cc2)Sc2[n]c[n][n]12
CHEMBL481	CCc1c2cc(ccc2[n]c2C3=CC4=C(COC(=O)C4(O)CC)C(=O)N3Cc21)OC(=O)N1CCC(CC1)N1CCCCC1
CHEMBL107217	CCN(CC)C(S)=S
CHEMBL1015	Cc1cc(ccc1NN=C1C=Cc2c(C1=O)c(N)c(cc2S(O)(=O)=O)S(O)(=O)=O)-c1cc(C)c(cc1)NN=C1C=Cc2c(C1=O)c(N)c(cc2S(O)(=O)=O)S(O)(=O)=O
CHEMBL1449651	SC(=S)N1CCCC1
CHEMBL2359093	OC(=O)c1c(C2c3cc(I)c(O)c(I)c3OC3C=2C=C(I)C(=O)C=3I)c(Cl)c(Cl)c(Cl)c1Cl
CHEMBL444186	CC(CN1CC(=O)NC(=O)C1)N1CC(=O)NC(=O)C1
CHEMBL1923234	CC(NC(=O)C(=Cc1cccc(Br)[n]1)C#N)c1ccccc1
CHEMBL244948	Oc1ccc2C=CC(=O)Oc2c1O
CHEMBL1886408	CCOC(=O)Cc1ccc(cc1)-c1ccccc1
CHEMBL9116	CC(=O)N(OC(=O)NC)C(=O)NC
CHEMBL487253	C[n]1c2ccc(cc2[n]c1CCCC(O)=O)N(CCCl)CCCl
CHEMBL1683636	Cc1[n](Cc2ccccc2)cc[nH+]1CCCCCCCCCCCCCCCC
CHEMBL1628564	CN(CC(O)=O)CCOc1ccc(cc1C(C)(C)C)-c1cc2OCOc2cc1
CHEMBL1964991	NC(=N)c1ccc(cc1)-c1ccc(o1)-c1ccc(cc1)C(N)=N
CHEMBL2354521	CC(NCCC(=O)c1cc(ccc1)OC)C(O)c1ccccc1
CHEMBL1917204	Cc1ccc(cc1)N1SC(=O)N(Cc2ccc(F)cc2)C1=O
CHEMBL1422849	CCN1CCN(CC1)C(c1ccc2ccc[n]c2c1O)c1ccc(cc1)C(F)(F)F
CHEMBL128251	OC(=O)CCCCCCCCCCCCCCCc1ccc(I)cc1
CHEMBL1026	Oc1c(Cl)c2CCNCC(c2cc1O)c1ccc(O)cc1
CHEMBL188434	CC[n]1c2c(c[n]cc2CNC2CCNCC2)[n]c1-c1[n]o[n]c1N
CHEMBL2093109	COc1c(cc(CC2NCCc3cc(O)c(O)cc32)cc1OC)OC
CHEMBL1517651	CN1CCc2cc3OCOc3cc2C1O
CHEMBL286120	C=CCN1CC(c2cc(O)c(O)cc2CC1)c1ccccc1
CHEMBL451197	NCCc1cc(O)c(O)cc1O
CHEMBL1370630	COc1ccc(cc1)C1CC(=NN1c1ccc(cc1)S(N)(=O)=O)c1ccccc1
CHEMBL4303312	Oc1c(I)cc(I)c(O)c1I
CHEMBL2204502	Cc1cc(C(N)=O)c(cc1C(=O)NC1CC2CCC(C1)N2c1ccc(c[n]1)C(=O)C1CC1)NC(C)CC
CHEMBL472940	Oc1ccc2ccccc2c1SSc1c2ccccc2ccc1O
CHEMBL3894518	CC[n]1[n]c([n]c1-c1c[n]c(N)c([n]1)-c1[n][n]c(o1)C(C)(C)C)C1CCN(CC1)C(=O)CCO
CHEMBL477197	Oc1cc2C3c4ccc(O)c(O)c4OCC3(O)Cc2cc1O
CHEMBL1738731	C[n]1cc(c[n]1)-c1ccc2[n]cc(Cc3ccc4[n]cccc4c3)[n]2[n]1

CHEMBL3392049	ON1C=CC=CC1=S
CHEMBL329522	Oc1ccc(C(=O)c2cc(O)c(O)c(O)c2)c(O)c1O
CHEMBL959	CC(N)C12CC3CC(C1)CC(C2)C3
CHEMBL505670	O[n+]1ccccc1Sc1ccc(c2[n]o[n]c12)N(=O)=O
CHEMBL2368547	OC(=O)c1cc(O)c2c(c1)c(-c1c3cc(cc(O)c3c(O)c3c1cccc3OC1OC(CO)C(O)C(O)C1O)C(O)=O)c1cccc(OC3OC(CO)C(O)C(O)C3O)c1c2O
CHEMBL1553873	CN(C)c1cc(OC)c(O)c(CN(C)CCc2ccc(cc2)N(=O)=O)c1
CHEMBL284861	CN1SC(=O)N(Cc2ccccc2)C1=O
CHEMBL87708	CN(CCCN1c2ccccc2CCc2ccccc12)CC(=O)c1ccc(Cl)cc1
CHEMBL4303361	Oc1c(O)cc(CNC(=S)C=Cc2cc(Br)c(O)c(O)c2)cc1O
CHEMBL1474701	O=C1C=CC(=O)C2C3c4ccccc4C(C1=2)c1ccccc13
CHEMBL111654	S=C1Nc2ccccc2S1
CHEMBL1096	CC(C)c1cc2c(cc1)Oc1[n]c(N)c(cc1C2=O)C(O)=O
CHEMBL1790039	CCN(CC)c1ccc(cc1)[C+](c1ccc(cc1)N(CC)CC)c1ccccc1
CHEMBL585	Nc1[n]c(N)[n]c2[n]c(N)c([n]c12)-c1ccccc1
CHEMBL105712	O=C(CCl)c1ccccc1
CHEMBL291143	NCC1OC(Cc2c1ccc(O)c2O)C12CC3CC(C1)CC(C2)C3
CHEMBL118109	COc1cc(cc(CSc2[n]c3ccccc3[s]2)c1O)C=C(C#N)C(N)=O
CHEMBL1610810	Cc1cc(ccc1N)-c1[n]c2[n]cccc2o1
CHEMBL388676	O=C1C2CC=CCC2C(=O)N1SC(Cl)(Cl)Cl
CHEMBL1188627	CN=C1N=C(c2ccccc2)N(S1)c1ccccc1
CHEMBL508280	O=N(=O)c1c[n]c(SC2=NNC(=O)N2c2cc3OCCOc3cc2)[s]1
CHEMBL2106153	CN1CC2CC1CN2c1cc2c(cc1F)C(=O)C(=CN2C1CC1)C(O)=O
CHEMBL331743	CCOC(=S)SSC(=S)OCC
CHEMBL566136	Cc1c([s]c(N)c1C(=O)OCC)-c1ccccc1
CHEMBL16312	COc1ccc(cc1)C1=CC(=O)c2ccccc2O1
CHEMBL3191466	NC(=S)NN=CC1=COc2ccccc2C1=O
CHEMBL154580	C=CC(=O)c1cc2ccccc2cc1
CHEMBL505308	Oc1cc2CCNCC(c2cc1O)c1ccccc1
CHEMBL4061483	CCCCS(=O)c1[s]c2[n]c(cc(c2c1N)-c1ccccc1)-c1ccc[s]1
CHEMBL1397952	Cc1cc(NC(c2ccc3ccc[n]c3c2O)c2ccc(OC)c(c2)OCC)[n]cc1
CHEMBL286080	Oc1cc2CCNCC(c2cc1O)c1ccccc1
CHEMBL418052	Nc1[n]c[n]c2c1[n]c[n]2C1OC(CSCCC(N)C(O)=O)C(O)C1O
CHEMBL3770443	Cc1cccc([n]1)-c1ccc(c(Cl)c1)C1=Cc2c[n]c(NC)[n]c2N(CC2OCC(N)CO2)C1=O
CHEMBL1255778	NC(CO)C(=O)NNCc1ccc(O)c(O)c1O
CHEMBL1612112	Cc1ccc(cc1)S(=O)(=O)c1c(NS(=O)(=O)c2ccccc2)[n](CC=C)c2[n]c3ccccc3[n]c21
CHEMBL119171	NC1Nc2c([n]c[n]2C2OC(CO)C(O)C2O)C(=S)N=1
CHEMBL1208422	OC(=O)c1c(C2c3cc(I)c(O)c(I)c3OC3C=2C=C(I)C(=O)C=3I)c(Cl)c(Cl)c(Cl)c1Cl
CHEMBL1256832	Oc1cc2CCC3NCc4ccccc4C3c2cc1O
CHEMBL4559134	COc1ccc(CN2C3CN(CC2C3)c2ccc(c[n]2)-c2cc(c[n]3[n]cc(C#N)c32)OCC(C)(C)O)c[n]1
CHEMBL155103	CCOC(=O)C1=Cc2ccc(O)cc2OC1=O
CHEMBL2269009	S=C1NN=C(O1)c1ccccc1Cl
CHEMBL243644	COc1ccc(cc1)NC(=O)CC1Sc2[n]c[n][n]2C1=O

CHEMBL508112	c1cccc2[s]c([n]c21)SSc1[n]c2ccccc2[s]1
CHEMBL50588	CCC1CN2CCc3cc(OC)c(cc3C2CC1CC1NCCc2cc(OC)c(cc21)OC)OC
CHEMBL168067	C[N+]1(C)CCC(CC1)OC(=O)C(c1ccccc1)c1ccccc1
CHEMBL192627	O=C1C(Cl)=C(Cl)C(=O)C(Cl)=C1Cl
CHEMBL4303240	Oc1cc2OC(=O)C=Cc2cc1OC1OC(CO)C(O)C(O)C1O
CHEMBL449298	Nc1ccc(O)c2[n]cccc21
CHEMBL3665060	CN1CCN(CC1)C(=O)c1ccc([s]1)-c1cc2ccc(O)c(C=O)c2cc1
CHEMBL45152	O=C1c2ccccc2SN1c1ccccc1
CHEMBL1339149	CN(C)S(=O)(=O)c1cc(ccc1)N1Sc2ccccc2C1=O
CHEMBL570345	O[n+]1o[n]c(c1C(=O)c1ccc[s]1)C(=O)c1ccc[s]1
CHEMBL481049	COc1cc(OC)c2C=CC(=O)Oc2c1
CHEMBL156592	CN1CC2CC1CN2c1cc2c(cc1F)C(=O)C(=CN2C1CC1)C(O)=O
CHEMBL1096979	NC(CO)C(=O)NNCc1ccc(O)c(O)c1O
CHEMBL275938	OC(=O)C1=CC(C=CC1=O)=C(c1cc(c(O)cc1)C(O)=O)c1cc(c(O)cc1)C(O)=O
CHEMBL299784	O=C1c2ccccc2S(=O)(=O)N1CCl
CHEMBL940	NCC1(CC(O)=O)CCCCC1
CHEMBL289356	S=C1NN=C(Nc2ccccc2)S1
CHEMBL1713178	CCN(CC)S(=O)(=O)c1cc(ccc1)N1Sc2ccccc2C1=O
CHEMBL311909	COc1cc(ccc1NN=C1C=Cc2c(C1=O)c(N)c(cc2S(O)(=O)=O)S(O)(=O)=O)-c1cc(OC)c(cc1)NN=C1C=Cc2c(C1=O)c(N)c(cc2S(O)(=O)=O)S(O)(=O)=O
CHEMBL1509639	COc1cc2NC(=S)Nc2cc1
CHEMBL360055	CC[N+](CCOc1cccc(OCC[N+](CC)(CC)CC)c1OCC[N+](CC)(CC)CC)(CC)CC
CHEMBL1336959	Cc1ccc(cc1)N1Sc2ccccc2C1=O
CHEMBL3628274	NCCCCC(NC(=O)C(Cc1ccc(cc1)C(N)=N)NC(=O)c1ccc(cc1)C(F)(F)F)C(=O)NC(c1ccccc1)C(N)=O
CHEMBL3628261	NCCCCC(NC(=O)C(Cc1ccc(cc1)C(N)=N)NC(=O)c1ccccc1)C(=O)NC(c1ccccc1)C(N)=O
CHEMBL3335494	NCCCCC(NC(=O)C(CCCNC(N)=N)NC(=O)c1ccccc1)C(=O)NC(c1ccccc1)C(N)=O

We have described above protocols for curating data of interest to the antiviral drug discovery from ChEMBL database. The resulting SMACC database currently exists as a searchable Excel spreadsheet that we include for the public with this paper. This spreadsheet allows multiple approaches to identify compounds of interest. The approach described here in Methods, i.e., removing compounds with conflicting activity calls from our final BSA activity analysis, was stringent and resulted in a concise list of potential BSA compounds that we had the highest confidence in. While the filtering criteria described above were appropriate to achieve our project's objective, we acknowledge the value of extracting different subsets of the SMACC database using different criteria depending on the study objectives, and SMACC database (even in the form of an Excel spreadsheet) enables multiple analyses.

For instance, in contrast with the approach described above, another method for identifying BSA compounds would be to consider all compounds with at least one active assay result against two or more viruses. This would increase the number of compounds considered in the analysis because inconclusive entries would not be removed as described above. To do this, we created a subset of the SMACC database with all compound entries reported as active in phenotypic assays, enumerated the number of viruses the compounds were reported active against and removed all compounds reported active against only one virus. This approach

resulted in 21 new hit compounds identified from phenotypic assays (Table S12) and 10 new hit compounds from target based assays that were not identified in the previous approach (Table S13). One could also search the database for actives using some threshold of activity calls, setting a percentage of active assay results compared to inactive at a specific virus or viral target.

As mentioned above, there are currently only 90 antiviral drugs approved for treating nine human infectious diseases (Erik and Guangdi, 2016). We utilized SMACCs filtering tools to enumerate their presence in our database. Currently, SMACC includes chemogenomics data for RSV and two strains of human influenza virus (H1N2 and H7N7), which covers only two of nine diseases with approved drugs. Despite this, we identified 53 of 90 approved drugs in our phenotypic dataset and 57 of 90 in our target-based dataset (Table S14). The compounds with reported active assay results are summarized in Table 3. Clearly, these drugs have broader activity than they are approved for. Further experimental testing based on hypotheses from this table will be extremely valuable to understanding their broad-spectrum potential.

Table S12. Additional compounds identified as potential hits due to activity in at least one phenotypic assay, regardless of inactivity in another assay.

Molecule ChEMBL ID	SMILES	Viruses-active (#data points)
CHEMBL866	COc1c(C)c2c(c(O)c1CC=C(C)CCC(=O)O)C(=O)OC2	SARS-CoV-2, RSV(3), HPIV-3(2), Dengue Virus(4), West Nile Virus(2), Yellow Fever Virus(2)
CHEMBL1643	NC(=O)c1ncn(C2OC(CO)C(O)C2O)n1	HPIV-3(3), Sandfly_Fever(1), Dengue Virus(3), Yellow Fever Virus(1), RSV(19)
CHEMBL160	CC=CCC(C)C(O)C1C(=O)NC(CC)C(=O)N(C)CC(=O)N(C)C(CC(C)C)C(=O)NC(C(C)C)C(=O)N(C)C(CC(C)C)C(=O)NC(C)C(=O)NC(C)C(=O)N(C)C(CC(C)C)C(=O)N(C)C(CC(C)C)C(=O)N(C)C(C(C)C)C(=O)N1C	Dengue Virus(4), West Nile Virus(4), Yellow Fever Virus(2), SARS-CoV-2(1)
CHEMBL76	CCN(CC)CCCC(C)Nc1ccnc2cc(Cl)ccc12	SARS-CoV-2(3), Zika Virus(3), Dengue Virus(1)
CHEMBL1397	CCC(C(C)O)n1ncn(-c2ccc(N3CCN(c4ccc(OCC5COC(Cn6cncn6)(c6ccc(F)cc6F)C5)cc4)CC3)cc2)c1=O	SARS-CoV-2(2), Dengue Virus(1)
CHEMBL1421	Cc1nc(Nc2ncc(C(=O)Nc3c(C)cccc3Cl)s2)cc(N2CCN(CCO)CC2)n1	Dengue Virus(14), SARS-CoV-2(1)
CHEMBL1448	O=C(Nc1ccc([N+](=O)[O-])cc1Cl)c1cc(Cl)ccc1O	SARS-CoV-2(2), Dengue Virus(1)
CHEMBL147806	CC(C)(C)NCc1ccc(Nc2ccnc3cc(Cl)ccc23)cc1O	SARS-CoV-2(2), Zika Virus(1)
CHEMBL1535	CCN(CCO)CCCC(C)Nc1ccnc2cc(Cl)ccc12	SARS-CoV-2(2), Dengue Virus(5)
CHEMBL1631361	Clc1ccc(-c2ccc(-c3nc(NCc4cccnc4)nc4ccccc34)s2)cc1	Dengue Virus(10), West Nile Virus(1)
CHEMBL1650595	NCC(=O)Nc1ccc(-n2nc(C(F)(F)F)cc2-c2ccc3c(ccc4ccccc43)c2)cc1	Dengue Virus(4), SARS-CoV-2(1)
CHEMBL1652118	CC1CCC(C(C)C)C(OCCCCCCN2CC(O)C(O)C(O)C2CO)C1	Dengue Virus(2), West Nile Virus(1)
CHEMBL1652119	Cc1ccc(C)c(OCCCCCCN2CC(O)C(O)C(O)C2CO)c1	Dengue Virus(2), West Nile Virus(1)
CHEMBL1652120	CC1CCCCC1OCCCCCCN1CC(O)C(O)C(O)C1CO	Dengue Virus(2), West Nile Virus(1)
CHEMBL3578033	Cc1cn(C2CC(n3nncc3-c3ccccc3)C(CO[Si](C)(C)C(C)(C)C)O2)c(=O)[nH]c1=O	West Nile Virus(2), Dengue Virus(1)

CHEMBL36	CCc1nc(N)nc(N)c1-c1ccc(Cl)cc1	RSV(1), SARS-CoV-2(1)
CHEMBL4065616	CCC(CC)COC(=O)C(C)NP(=O)(OCC1OC(C#N)(c2ccc3c(N)ncnn23)C(O)C1O)Oc1ccccc1	RSV(1), SARS-CoV-2(4)
CHEMBL4438019	O=C(C=Cc1ccccc1)NC(CC1CCCC1)C(=O)NC(CC1CCNC1=O)C(=O)C(=O)NCc1ccccc1	MERS-CoV(1), HCoV-229E(1)
CHEMBL4442435	O=C(C=Cc1ccccc1)NC(CC1CC1)C(=O)NC(CC1CCNC1=O)C(=O)C(=O)NCc1ccccc1	MERS-CoV(1), HCoV-229E(1)
CHEMBL453392	O=C(O)c1c[nH]c2c(c1=O)c(=O)n(Cc1ccccc1)c(=O)n2Cc1ccccc1	Yellow Fever Virus(1), Dengue Virus(1)
CHEMBL4556306	COc1ccc(S(=O)(=O)N(CCCC(=O)C2CCCCCC2)c2ccccc2)cc1	Dengue Virus(17), Zika Virus(16)

Table S13. Additional compounds identified as potential hits due to activity in at least one target-based assay, regardless of inactivity in another assay.

Molecule ChEMBL ID	SMILES	Viruses-active	Target-active (#data points)
CHEMBL4452943	COc1ccc(CC2C(=O)NC(CCCCN)C(=O)N(C)C(Cc3ccccc3)C(=O)NC(CCCCN)C(=O)N(C)C(CO)C(=O)N(C)C(Cc3ccc(OC)cc3)C(=O)NC(CCCCN)C(=O)N(C)C(Cc3ccc(OC)cc3)C(=O)N(C)C(Cc3ccc(OC)cc3)C(=O)NC(CCCCN)C(=O)NC(C(N)=O)CSCC(=O)NC(Cc3ccc(O)cc3)C(=O)N2C)cc1	Dengue, Zika, West Nile	NS2B-NS3 Protease(7)
CHEMBL4483351	CCC(C)C1NC(=O)C(CC(N)=O)NC(=O)C(C(C)C)NC(=O)C(CC(C)C)NC(=O)C(C(C)O)NC(=O)C(CC(N)=O)NC(=O)C(Cc2ccc(OC)cc2)N(C)C(=O)C(C(C)CC)NC(=O)C(CCCCN)NC(=O)C(Cc2c[nH]c3ccccc23)NC(=O)C(Cc2ccc(O)cc2)NC(=O)CSCC(C(N)=O)NC1=O	Dengue, Zika, West Nile	NS2B-NS3 Protease(6)
CHEMBL164	O=c1c(O)c(-c2cc(O)c(O)c(O)c2)oc2cc(O)cc(O)c12	Zika, SARS-CoV-2	NS2B-NS3 Protease(2), Main Protease (3CLpro, Mpro)(4)
CHEMBL4435541	N=C(N)NCCCC(NC(=O)c1ccccc1)C(=O)NC(CCCCN)C(=O)NC(C(=O)NC1CN(c2ccccc2)C1=O)c1ccc(OCc2ccccc2)cc1	Dengue, West Nile	NS2B-NS3 Protease(5)
CHEMBL4437757	N=C(N)Nc1ccc(NC(=O)Oc2ccc(N)cc2)cc1	Dengue, West Nile	NS2B-NS3 Protease(4)
CHEMBL4463470	N=C(N)NCCCC(NC(=O)c1ccccc1)C(=O)NC(CC	Dengue, West Nile	NS2B-NS3 Protease(2)

	CCN)C(=O)NC1CN(c2ccccc2)C1=O			
CHEMBL4464884	CC(NC(=O)C(CCCCN)NC(=O)C(CCCNC(=N)N)NC(=O)c1ccccc1)(C(N)=O)c1ccccc1	Dengue, West Nile	NS2B-NS3 Protease(3)	
CHEMBL4469272	CC(NC(=O)C(CCCCN)NC(=O)C(CCCNC(=N)N)NC(=O)c1ccccc1)(C(N)=O)c1ccccc1	Dengue, West Nile	NS2B-NS3 Protease(3)	
CHEMBL4545602	Cc1ccc(N2CC(NC(=O)C(NC(=O)C(CCCCN)NC(=O)C(CCCNC(=N)N)NC(=O)c3ccccc3)c3ccc(OCc4ccccc4)cc3)C2=O)cc1	Dengue, West Nile	NS2B-NS3 Protease(2)	
CHEMBL4561419	N=C(N)NCCCC(NC(=O)c1ccccc1)C(=O)NC(CCCCN)C(=O)NC(C(=O)NC1CN(c2cccc(F)c2)C1=O)c1ccc(OCc2ccccc2)cc1	Dengue, West Nile	NS2B-NS3 Protease(2)	

Table S14. List of approved antiviral drugs present in the SMACC database, including the viruses and targets the compounds were reported active against.

Compound Information			**Presence in SMACC**				
Drug name	**Approved Use**	**ChEMBL ID**	**Phenotypic Data**	**Active Virus**	**Target Based**	**Virus**	**Active Target**
Cidofovir	CMV	CHEMBL152	Yes		Yes		
Adefovir dipivoxil	HBV	CHEMBL922	Yes		Yes		
Telbivudine	HBV	CHEMBL374731	Yes		Yes		
Ganciclovir	HCMV	CHEMBL182	Yes		Yes		
Simeprevir	HCV	CHEMBL501849	Yes		Yes	H7N7	Matrix Protein 2
Paritaprevir	HCV	CHEMBL3391662	Yes		Yes		
Dasabuvir	HCV	CHEMBL3137312	Yes		Yes		
Ombitasvir	HCV	CHEMBL3127326	Yes		Yes		
Elbasvir	HCV	CHEMBL3039514			Yes		
Ledipasvir	HCV	CHEMBL2374220			Yes		
Telaprevir	HCV	CHEMBL231813			Yes		
Boceprevir	HCV	CHEMBL218394	Yes		Yes		
Asunaprevir	HCV	CHEMBL2105735	Yes		Yes	H7N7	Matrix Protein 2
Grazoprevir	HCV	CHEMBL2063090	Yes		Yes		
Daclatasvir	HCV	CHEMBL2023898			Yes		
Sofosbuvir	HCV	CHEMBL1259059	Yes	Dengue	Yes		

				HPIV-3, Sandfly Fever, Dengue, Yellow Fever,			
Ribavirin	HCV, RSV	CHEMBL1643	Yes	RSV	Yes		
Stavudine	HIV	CHEMBL991	Yes		Yes		
Emtricitabine	HIV	CHEMBL885	Yes		Yes		
Zalcitabine	HIV	CHEMBL853	Yes		Yes		
				SARS-			
Lopinavir	HIV	CHEMBL729	Yes	CoV-2	Yes		
Nelfinavir	HIV	CHEMBL584	Yes	Dengue			
Raltegravir	HIV	CHEMBL254316	Yes		Yes	H7N7	Matrix Protein 2
						SARS-	
Elvitegravir*	HIV	CHEMBL204656	Yes		Yes	CoV-2	3CLpro
Ritonavir	HIV	CHEMBL163	Yes		Yes		
Didanosine	HIV	CHEMBL1460			Yes		
Darunavir	HIV	CHEMBL1323	Yes				
Zidovudine	HIV	CHEMBL129	Yes				
Dolutegravir	HIV	CHEMBL1229211	Yes		Yes		
Maraviroc	HIV	CHEMBL1201187	Yes		Yes		
						SARS-	
Atazanavir	HIV	CHEMBL1163	Yes		Yes	CoV-2	3CLPro
Indinavir	HIV	CHEMBL115	Yes		Yes		
Delavirdine	HIV-1	CHEMBL593			Yes		
Nevirapine	HIV-1	CHEMBL57	Yes		Yes		
Etravirine	HIV-1	CHEMBL308954	Yes		Yes		
Efavirenz	HIV-1	CHEMBL223228			Yes		
Tipranavir	HIV-1	CHEMBL222559	Yes		Yes		
				SARS-			
Rilpivirine	HIV-1	CHEMBL175691	Yes	CoV-2	Yes		
Fosamprenavir	HIV-1	CHEMBL1664	Yes		Yes		
Amprenavir	HIV-1	CHEMBL116	Yes		Yes		
TDF	HIV, HBV	CHEMBL1538	Yes		Yes		
Lamivudine	HIV, HBV	CHEMBL141	Yes		Yes		
	HPV-related						
Imiquimod	diseases	CHEMBL1282	Yes		Yes		

Podofilox	HPV-related diseases	CHEMBL61	Yes	SARS-CoV-2	Yes		
Penciclovir	HSV	CHEMBL1540	Yes		Yes		
Docosanol	HSV	CHEMBL1200453	Yes		Yes		
Trifluridine	HSV	CHEMBL1129	Yes	HPIV-3	Yes		
Idoxuridine	HSV-1	CHEMBL788	Yes		Yes	Zika	NS5
Brivudine	HSV-1,VZV	CHEMBL31634	Yes		Yes		
Acyclovir	HSV, VZV	CHEMBL184	Yes		Yes	H7N7	Neuraminidase
Valaciclovir	HSV, VZV	CHEMBL1349			Yes		
Vidarabine	HSV, VZV	CHEMBL1090	Yes		Yes		
Famciclovir	HSV,VZV	CHEMBL880	Yes		Yes		
Rimantadine	Influenza A	CHEMBL959	Yes		Yes		
Amantadine	Influenza A	CHEMBL660	Yes		Yes		
Laninamivir octanoate	Influenza A and B	CHEMBL467058			Yes		
Zanamivir	Influenza A and B	CHEMBL222813	Yes	H7N7	Yes		
Peramivir	Influenza A and B	CHEMBL139367	Yes		Yes		
Oseltamivir	Influenza A and B	CHEMBL1229	Yes		Yes		
Favipiravir	Influenza A, B and C	CHEMBL221722	Yes		Yes		

Comment on Viral ChEMBL

Efforts similar to ours have been made to collect antiviral data prior to the SARS-CoV-2 outbreak. There is a collection of antiviral activity data from ChEMBL with enhanced taxonomy annotations as a tool for studying the antiviral chemical space that the authors dubbed “Viral ChEMBL” (Nikitina et al., 2019). Viral ChEMBL was compiled using information collected on compounds related to many virus types (human, animal, plant) and additional curation was performed to the data by mapping lists for assay and target organism data and using a dictionary of virus-related terms. While this collection is quite valuable to the field, it is based on an old version of ChEMBL20 (released in 2015, the current release is ChEMBL29) and some data are not relevant to human disease. Thus, our database was collected and manually curated to provide

a structured, annotated repository of all data available in the most current version of ChEMBL for viruses that hold the greatest risk for human contraction and pandemic potential.

Table S15. Additional resources for SARS-CoV-2 drug discovery

Source	**Link**	**Description**
PubMed	https://www.ncbi.nlm.nih.gov/sars-cov-2/	SARS-CoV-2 Data Resource
European Bioinformatics Institute (EBI)	https://www.covid19dataportal.org	Comprehensive COVID-19 Data Portal
ChEMBL database	https://www.ebi.ac.uk/chembl/	A collection of over 20,000 screening results for compounds tested against SARS-CoV-2
NCATS at the NIH	https://opendata.ncats.nih.gov/covid19/index.html	Target specific and phenotypic screening results of chemical libraries in SARS-CoV-2
Kaggle	https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge	CORD-19
MML UNC	https://coke.mml.unc.edu	CoKE: COVID-19 Knowledge Extractor