Supporting Information

Ensemble-based Virtual Screening Reveals Novel Antiviral Compounds for Avian Influenza Neuraminidase

Lily S. Cheng, Rommie E. Amaro, Dong Xu, Wilfred W. Li, Peter W. Arzberger, and J. Andrew McCammon

Table of Contents:

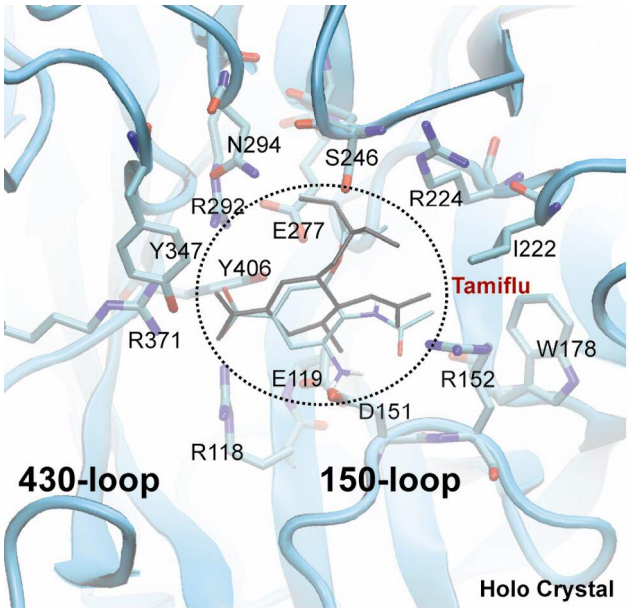

Figure S1. The successful re-docking of Tamiflu to the holo crystal structure (2HU4) is shown. The crystallized position of Tamiflu is shown in black without hydrogens. The docked position of Tamiflu is shown colored by atom type with polar hydrogens added. The docked pose agrees closely with the crystal pose, with minor deviations.

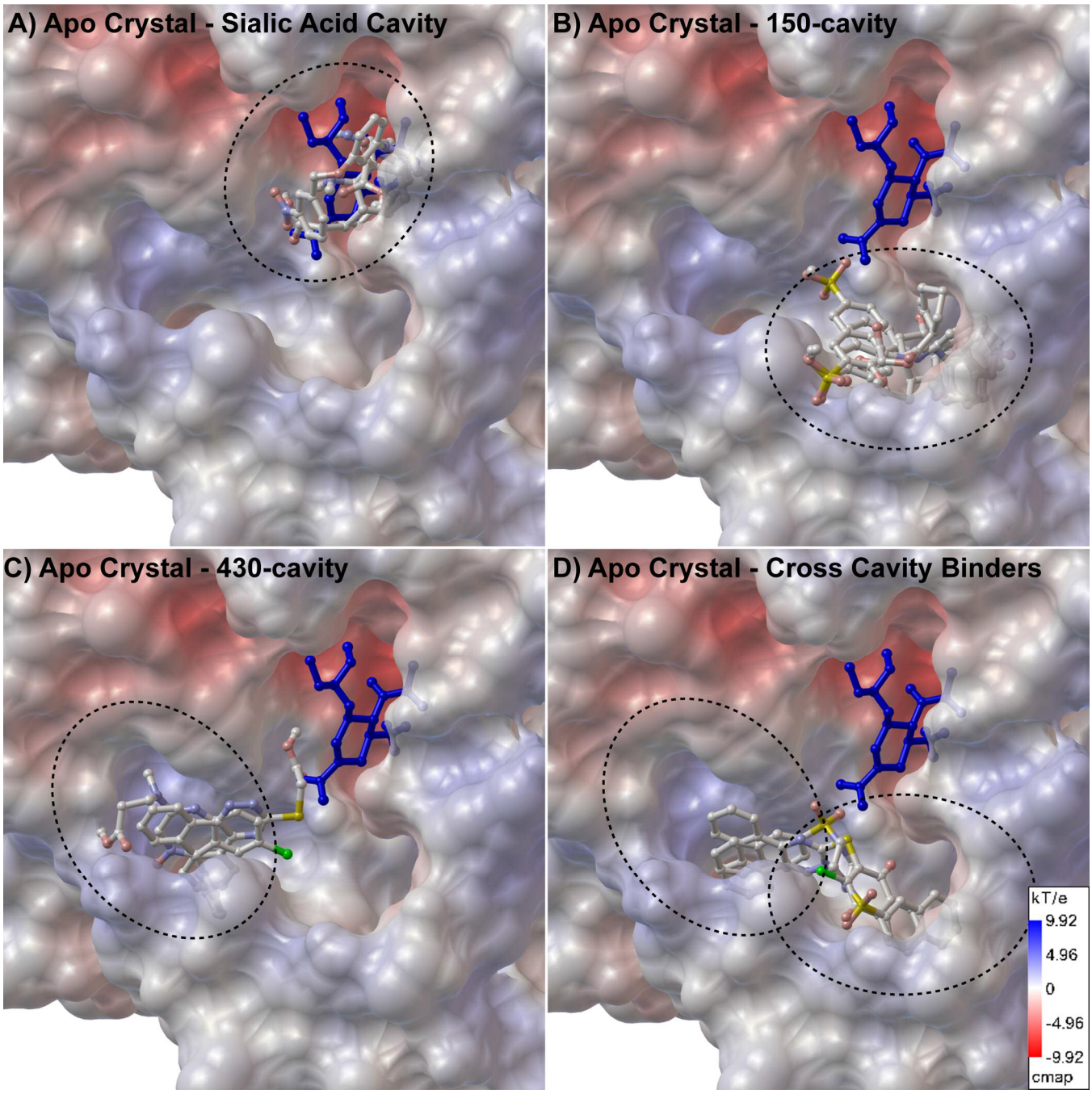

Figure S2. The molecular surface of the apo crystal structure with selected hits docked illustrates the position of (A) the sialic acid cavity, (B) the 150-cavity, and (C) the 430-cavity. Two hits that span the 150- and 430-cavities are also shown (D). Electrostatic surface potential scale shown in panel D.

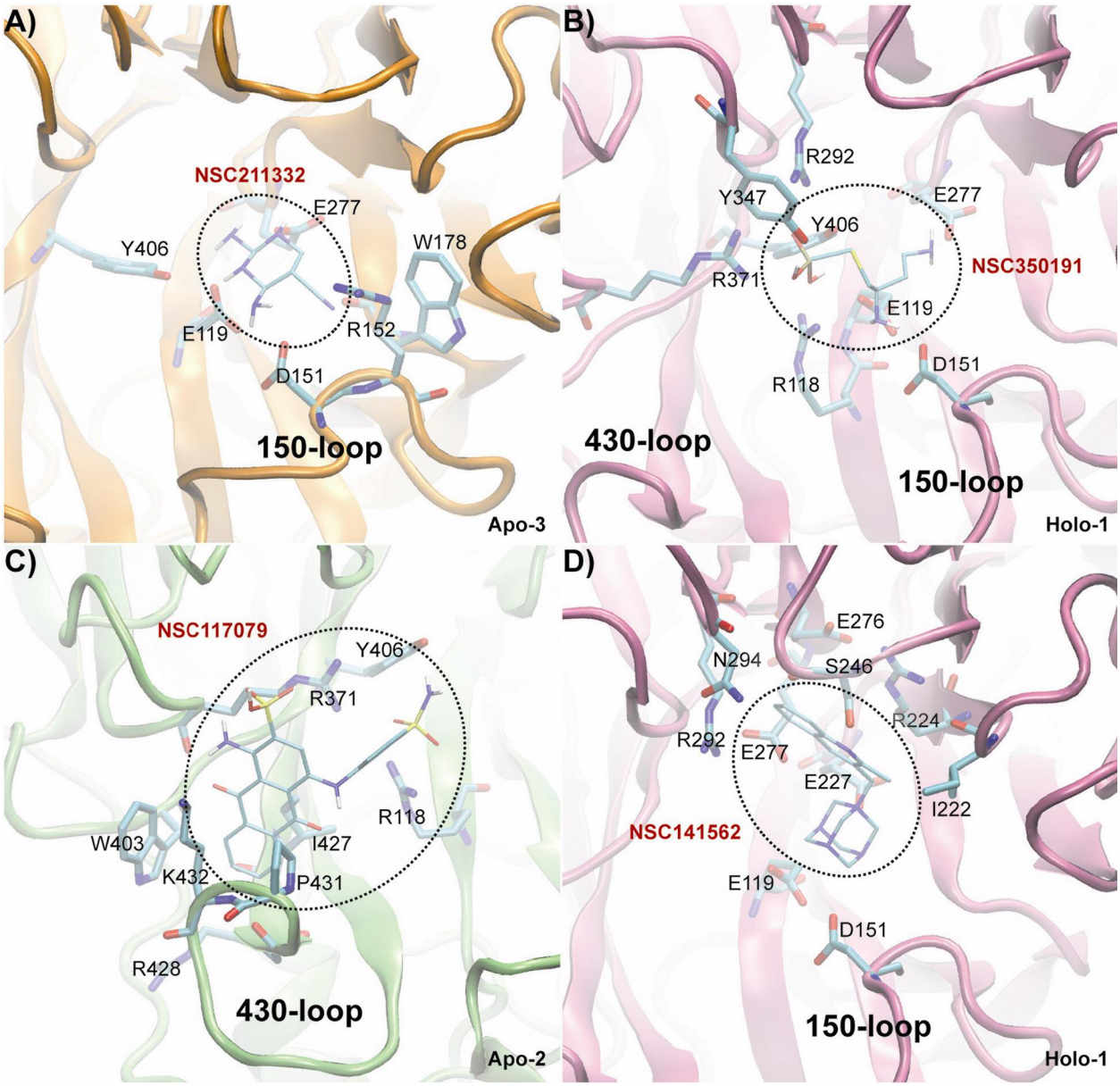

Figure S3. Four hits that bind in the sialic acid cavity are shown, each docked to the structure to which they exhibit the best binding affinity. (C) NSC117079 simultaneously docks in the sialic acid cavity and in the adjacent open 430-cavity.

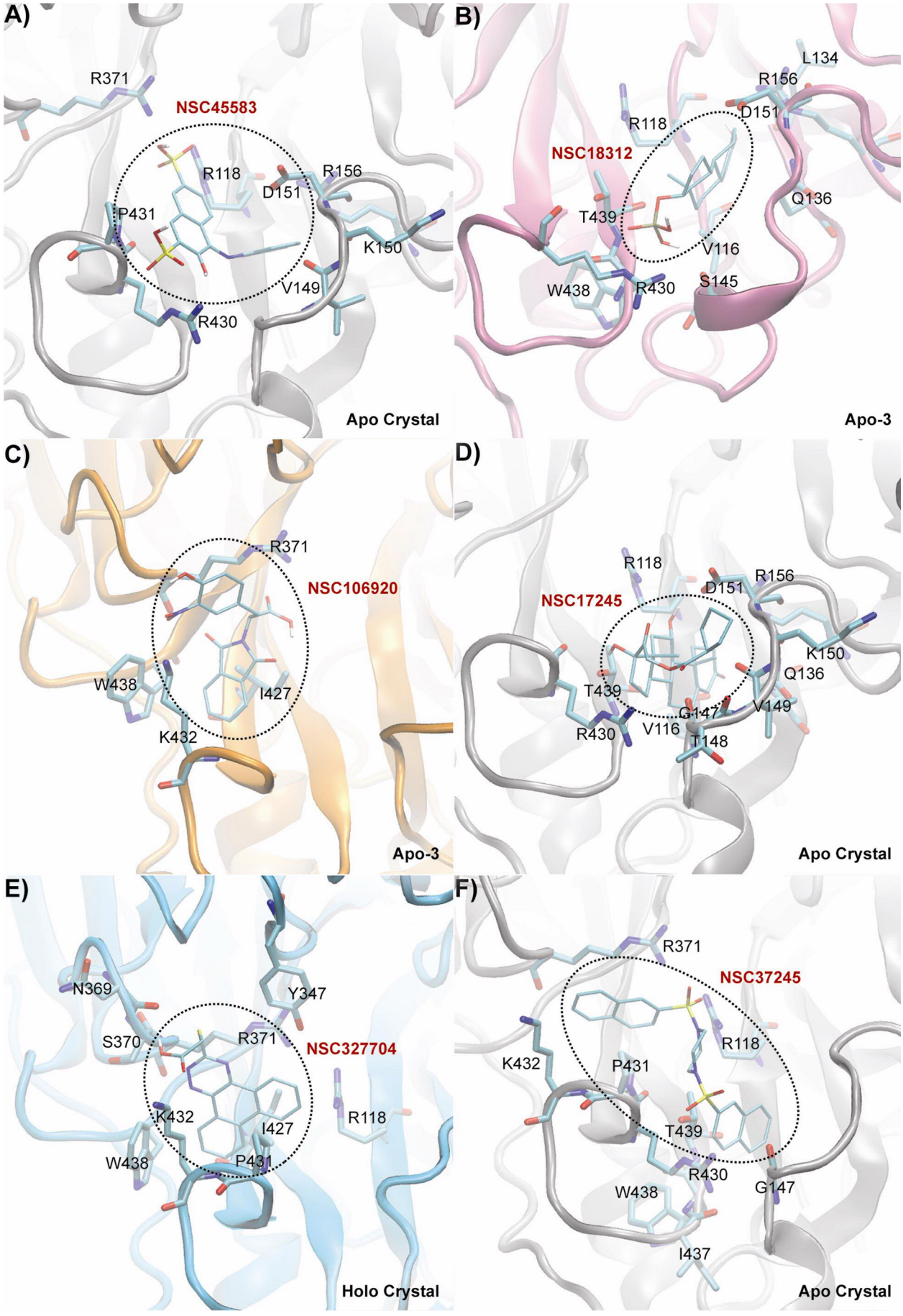
A)
R371
NSC45583
R118
R156
D151
P431
K150
V149
R430
Apo Crystal
B)
L134
R156
D151
R118
NSC18312
Q136
T439
V116
W438
R430
S145
Apo-3
C)
R371
NSC106920
W438
I427
K432
Apo-3
D)
R118
R156
D151
NSC17245
K150
Q136
T439
G147
V149
R430
V116
T148
Apo Crystal
E)
N369
Y347
S370
R371
NSC327704
K432
I427
R118
W438
P431
Holo Crystal
F)
R371
NSC37245
R118
K432
P431
T439
R430
W438
G147
I437
Apo Crystal

Figure S4. Six hits that bind in the 150-cavity and/or the 430-lcavity are shown, each docked to the structure to which they exhibit the best binding affinity. (A) NSC45583, (B) NSC18312, and (D) NSC17245 bind in the open or wide-open 150-cavity. (C) NSC106920 and NSC327704 bind in the open and closed 430-cavity, respectively. (F) NSC37245 is a compound that binds across the 150- and 430-cavities.

Table S1. Top hits to the receptor active site (sialic acid binding site)

Ligand	Apo Crystal	Apo-1	Apo-2	Apo-3	Holo Crystal	Holo-1	Holo 2	Holo-3
211332	-7.99 (63)	**-9.74 (94)**	-8.91 (100)	**-11.6 (100)**	**-9.96 (100)**	**-10.89 (74)**	**-9.96 (100)**	**-9.35 (98)**
109836	**-9.83 (31)**	-9.50 (19)	-8.99 (18)	-9.92 (17)	**-11.14 (43)**	-9.17 (12)	-10.41 (22)	**-9.89 (32)**
Oseltamivir	-7.91 (14)	-7.22 (28)	-7.82 (39)	-7.88 (28)	**-10.20 (43)**	**-10.24 (53)**	**-10.06 (60)**	-8.57 (41)
350191	-7.54 (8)	-8.24 (15)	-6.89 (20)	-8.70 (20)	-8.52 (26)	**-10.0 (34)**	-8.73 (35)	-6.79 (33)
117079	-8.29 (11)	-10.6 (22)	**-9.81 (25)**	-9.99 (10)	-7.99 (11)	-9.17 (7)	-10.89 (7)	-9.29 (6)
Zanamivir	-7.93 (12)	-6.71 (6)	-5.49 (8)	-8.35 (6)	**-9.79 (29)**	**-9.41 (26)**	-8.45 (23)	-6.66 (7)
141562	-6.98 (35)	-8.23 (52)	-8.90 (77)	-9.05 (32)	-8.28 (20)	**-9.71 (46)**	-7.81 (29)	-7.77 (23)
5069	-8.35 (30)	-7.31 (25)	**-9.60 (63)**	-6.39 (18)	-8.60 (28)	-7.21 (39)	-8.97 (25)	-7.50 (29)
131612	**-9.54 (34)**	-8.05 (10)	-8.11 (30)	-7.90 (15)	**-9.31 (29)**	-8.36 (19)	-9.33 (32)	-8.09 (49)
70194	-7.55 (29)	-8.73 (36)	**-9.43 (50)**	-7.97 (26)	-6.40 (18)	-7.61 (19)	-7.86 (41)	-7.36 (34)
135371	-7.28 (27)	-7.81 (16)	-6.65 (25)	-8.69 (66)	**-9.24 (62)**	-8.22 (48)	-8.15 (22)	-7.77 (20)
164640	-6.60 (16)	-7.42 (12)	-7.04 (19)	-8.21 (24)	-8.66 (25)	-8.69 (21)	-7.38 (30)	**-9.06 (34)**
Peramivir	-6.11 (23)	-5.09 (18)	-5.77 (16)	-7.81 (15)	-9.00 (16)	-8.08 (17)	-7.04 (9)	-6.98 (10)
SA	-7.54 (9)	-6.86 (7)	-7.31 (11)	-7.98 (9)	-7.89 (18)	-7.35 (11)	-7.96 (25)	-7.93 (13)
DANA	-6.77 (9)	-6.66 (19)	-6.86 (9)	-7.15 (12)	-7.85 (12)	-7.00 (11)	-7.69 (15)	-7.77 (20)

Top hits are shown sorted by lowest binding energy. Each compound is identified by its NSC classification number. Known inhibitors are included for reference. The table expresses the binding affinity of each ligand as the Gibbs free energy (kcal/mol) of the most populated cluster followed by the percentage of docked conformations with that energy in parenthesis. Bold font denotes that the compound is among the top 30 lowest energy hits for the given receptor. Gray font denotes that the clustering of the docking did not meet the 25 percent cutoff. Compounds identified from the ensemble screens that were not identified by the crystal structure screens are underlined.

Table S2. Top hits to other locales in the receptor binding site.

Ligand	Apo Crystal	Apo-1	Apo-2	Apo-3	Holo Crystal	Holo-1	Holo 2	Holo-3
46080	**-10.60 (35)**	-6.34 (8)	-8.98 (40)	-7.98 (13)	**-10.35 (29)**	-8.70 (23)	-7.56 (14)	-7.24 (14)
45583	**-10.57 (32)**	**-10.24 (25)**	-9.36 (18)	-8.46 (19)	-9.57 (14)	**-9.55 (34)**	**-9.95 (37)**	-10.53 (13)
18312	-8.49 (47)	**-10.56 (63)**	-8.29 (76)	-7.44 (50)	-7.79 (29)	-7.43 (49)	-7.88 (49)	-7.48 (30)
106920	-8.62 (10)	-7.03 (11)	-9.13 (19)	**-10.08 (45)**	-8.80 (11)	-7.83 (18)	-9.32 (34)	-8.02 (9)
17245	**-10.07 (33)**	**-9.60 (33)**	-9.76 (20)	-8.34 (13)	-8.83 (8)	-9.53 (16)	-9.58 (11)	-8.15 (8)
45576	**-9.79 (29)**	**-10.05 (54)**	-9.20 (42)	-9.00 (20)	-7.63 (13)	-7.43 (20)	-8.33 (18)	-8.12 (26)
327704	**-9.52 (27)**	-8.52 (28)	**-9.75 (69)**	-9.01 (66)	**-9.97 (48)**	-8.96 (84)	-8.39 (64)	-7.98 (38)
37245	**-9.80 (28)**	-7.20 (8)	-7.47 (7)	-7.98 (17)	-7.83 (24)	-8.31 (11)	-8.84 (14)	-7.20 (5)
59620	-8.41 (72)	**-9.79 (60)**	-8.18 (52)	-7.45 (62)	-7.88 (70)	-7.87 (34)	-8.25 (51)	-7.33 (69)
16163	-7.15 (12)	**-9.72 (28)**	-8.45 (14)	-7.23 (10)	-7.06 (26)	-5.79 (9)	-6.25 (26)	-7.16 (15)
90630	**-7.53 (25)**	-7.92 (19)	-8.56 (36)	-8.33 (14)	-7.86 (16)	-7.27 (22)	**-9.71 (42)**	-7.74 (14)
148354	-7.95 (11)	-8.78 (7)	-8.48 (9)	-9.49 (23)	-8.15 (7)	-7.82 (11)	**-9.70 (29)**	-8.73 (22)
327705	**-9.53 (36)**	-7.32 (17)	-8.82 (26)	-8.37 (37)	-9.76 (24)	-9.30 (44)	-9.39 (30)	-7.69 (15)
372294	-8.09 (12)	-7.17 (22)	-8.86 (7)	**-9.45 (31)**	-9.02 (17)	-8.36 (21)	-7.54 (36)	-8.01 (17)
45582	-7.44 (8)	-8.03 (9)	**-9.37 (28)**	-8.27 (9)	-7.62 (10)	-8.50 (17)	-7.44 (9)	-7.79 (14)
371688	-7.87 (22)	-8.94 (21)	-8.94 (17)	-7.74 (18)	**-9.35 (26)**	-8.71 (19)	-7.72 (10)	-7.28 (14)
72254	-8.30 (13)	-8.59 (41)	**-9.25 (25)**	-8.48 (37)	-8.76 (14)	-8.43 (17)	-8.61 (28)	-7.88 (17)

Same annotation used as Table S1.

Table 3. Top hits re-ranked by RCS

Ligand	W.H. RC Mean	W.A. RC Mean	RC STD	RC Mean	RC Min.	E. Mean	E. Min.	IUPAC Name
109836	-10.63	-10.42	0.55	-10.00	-11.08	-9.82	-10.41	1-(3-((3-(hydroxy(oxido)amino)benzyl)oxy)phenyl)-4,6-diimino-2,2-dimethyl-1,3,5-triazinane
211332	-10.34	-10.13	0.55	-10.05	-12.34	-10.07	-10.89	2,4-diaminohexahydro-5-pyrimidinecarbonitrile
45583	-10.09	-9.78	0.86	-9.09	-10.58	-10.01	-10.53	3-hydroxy-4-(1-naphthyldiazenyl)-2,7-naphthalenedisulfonic acid
Oseltamivir	-9.82	-9.60	0.66	-9.05	-10.63	-9.62	-10.24	(3R,4R,5S)-4-acetamido-5-amino-3-pentan-3-yloxycyclohexene-1-carboxylic acid
Zanamivir	-9.38	-8.89	0.92	-8.00	-10.22	-8.17	-9.41	(4S,5R,6R)-5-acetamido-4-(diaminomethylideneamino)-6-[(1R,2R)-1,2,3-trihydroxypropyl]-5,6-dihydro-4H-pyran-2-carboxylic acid
106920	-9.20	-8.82	0.77	-7.11	-10.23	-8.39	-9.32	N/A
17245	-9.18	-8.92	0.57	-8.64	-9.74	-9.09	-9.58	11,17-dihydroxy-3,20-dioxopregn-4-en-21-yl cyclopentanecarboxylate
350191	-9.14	-8.80	0.70	-8.36	-10.29	-8.51	-10.00	S-(2,4-diaminobutyl) dihydrogen thiophosphate
148354	-9.12	-8.74	0.90	-8.00	-9.62	-8.75	-9.70	8-hydroxy-7-(1-naphthyldiazenyl)-5-quinolinesulfonic acid
131612	-9.10	-8.82	0.60	-8.44	-10.00	-8.59	-9.33	2-(1,3-dioxo-1,3-dihydro-2H-isoindol-2-yl)-3-(4-hydroxy-3-(hydroxy(oxido)amino)phenyl)propanoic acid
37245	-9.02	-8.61	0.90	-8.34	-9.61	-8.12	-8.84	1,4-bis(2-naphthylsulfonyl)piperazine
45582	-9.00	-8.45	0.78	-8.07	-9.77	-7.91	-8.50	7-hydroxy-8-(1-naphthyldiazenyl)-1,3-naphthalenedisulfonic acid
90630	-8.89	-7.97	0.97	-7.70	-9.73	-8.24	-9.71	1-dibenzo[b,d]furan-3-yl-2-(2-piperidinyl)ethanol
5069	-8.83	-8.73	0.36	-8.31	-9.85	-7.89	-8.97	4-nitro-2,3-dioxo-2,3-dihydro-9,10-secostrychnidin-10-oic acid
327705	-8.83	-8.59	0.62	-8.27	-9.44	-8.79	-9.39	2-chloro-3-phenanthro[9,10-e][1,2,4]triazin-3-ylsulfanylnaphthalene-1,4-dione
141562	-8.78	-8.12	1.08	-8.44	-10.65	-8.43	-9.71	1-((2-methyl-3-quinolinyl)methyl)-1(lambda)5,3,5,7-tetraazatricyclo [3.3.1.1~3,7~] decane
117079	-8.72	-7.90	0.78	-8.67	-10.41	-9.78	-10.89	1-amino-4-(3-(aminosulfonyl)anilino)-9,10-dioxo-9,10-dihydro-2-anthracenesulfonic acid
327704	-8.60	-8.29	0.70	-7.86	-10.04	-8.44	-8.96	(phenanthro[9,10-e][1,2,4]triazin-3-ylthio)acetic acid
72254	-8.47	-8.17	0.64	-8.13	-9.40	-8.31	-8.61	2-hydroxy-2-((3-oxoestr-4-en-17-yl)oxy)-1H-indene-1,3(2H)-dione
164640	-8.37	-8.14	0.59	-8.05	-9.54	-8.38	-9.06	2-(4-chloro-2-methylphenoxy)-N-(4,6-diamino-1,3,5-triazin-2-yl)acetamide
135371	-8.28	-8.05	0.51	-7.87	-9.80	-8.05	-8.22	N/A
45576	-8.20	-8.16	0.21	-8.33	-9.32	-7.96	-8.33	5-((2-hydroxy-1-naphthyl)diazenyl)-2-naphthalenesulfonic acid
46080	-8.05	-7.43	0.93	-7.02	-9.92	-7.83	-8.70	N-(7-chloro-4-hydroxy-5-phenyl-3H-1,4(lambda)^5-benzodiazepin-2-yl)-N-methyl-beta-alanine
372294	-7.92	-7.81	0.35	-7.86	-9.32	-7.97	-8.36	(phenanthro[9,10-e][1,2,4]triazin-3-ylthio)acetic acid

59620	-7.85	-7.55	0.53	-7.76	-9.42	-7.82	-8.25	4a,6a,9-trimethyl-6,7-dioxo-2,3,4,4a,4b,5,6,6a,6b,7,10,10a,11,11a, 11b,12,13,13a-octadecahydro-1H-indeno[2,1-a]phenanthren-2-yl acetate
Peramivir	-7.79	-7.43	0.80	-6.99	-8.74	-7.37	-8.08	(1S,2S,3S,4R)-3-[(1S)-1-acetamido-2-ethylbutyl]-4-(diaminomethylidene-amino)-2-hydroxycyclopentane-1-carboxylic acid trihydrate
70194	-7.74	-7.62	0.38	-8.00	-9.92	-7.61	-7.86	N/A
18312	-7.72	-7.69	0.18	-7.66	-9.88	-7.60	-7.88	3-hydroxyestra-1,3,5(10)-trien-17-yl dihydrogen phosphate
SA	-7.47	-7.32	0.42	-7.37	-9.04	-7.75	-7.96	N-acetylneuraminic acid
371688	-7.44	-7.18	0.37	-7.53	-9.21	-7.90	-8.71	7-methoxy-3-(7-methoxy-4,5-dihydrobenzo[g][1,2]benzoxazol-3-yl)-4,5-dihydrobenzo[g][1,2]benzoxazole
DANA	-7.41	-7.27	0.45	-7.11	-8.54	-7.49	-7.77	(4S,5R,6R)-5-acetamido-4-hydroxy-6-[(1R,2R)-1,2,3-trihydroxypropyl]-5,6-dihydro-4H-pyran-2-carboxylic acid
16163	-7.07	-6.56	0.65	-6.84	-8.28	-6.40	-7.16	((1-amino-6-sulfo-2-naphthyl)oxy)acetic acid

The 27 top hits are listed in the table above, re-ranked by their weighted harmonic RC mean. The columns are binding energies in kcal/mol in the following order: weighted harmonic RC mean (WH RC mean), weighted arithmetic mean (WA RC mean), RC standard deviation (RC std), RC mean, RC minimum (RC min.), Ensemble mean (E mean), and Ensemble minimum (E. min.).