Identification of selective nanomolar inhibitors of the human neuraminidase, NEU4

Amgad Albohy[1], Yi Zhang[1], Victoria Smutova[2], Alexey V. Pshezhetsky[2], and Christopher W. Cairo[1]*

[1]Alberta Glycomics Center, Department of Chemistry, University of Alberta, Edmonton Alberta T6G 2G2, Canada

[2]Division of Medical Genetics, Centre Hospitaliere Universitaire Sainte-Justine, University of Montreal, Montreal, Quebec, Canada. Department of Anatomy; Cell Biology, Faculty of Medicine, McGill University, Montreal, Quebec, Canada

*Corresponding author. Tel.: 780 492 0377; fax: 780 492 8231; e-mail: ccairo@ualberta.ca

General protocols

All reagents used were purchased from commercial sources, such as Sigma-Aldrich and Acros Organics, and used without further purification unless otherwise noted. All reactions were carried out under a positive pressure of argon at room temperature unless otherwise indicated. Reactions were monitored by analytical TLC on silica gel 60-F_{254} (0.25mm, Silicycle, Quebec, Canada) and the spots were visualized under UV light (254 nm) or stained by charring with ceric ammonium molybdate (CAM). Organic solvents were evaporated under reduced pressure, and organic products were purified by flash column chromatography on silica gel (230-400 mesh, Silicycle, Quebec, Canada) or by HPLC (Waters). ^{1}H NMR spectra were acquired on Varian 400, 500 or 600 MHz instruments at room temperature as noted. ^{13}C NMR spectra were recorded at 125 or 150 MHz. Electrospray-ionization mass spectra were recorded on an Agilent Technologies 6220 TOF instrument. All compounds were synthesized following reported protocols.[1]

To determine the purity of all tested compounds, samples were analyzed using HPLC (Waters® XTerra® RP C18 analytical column), particle size: 3.5µm, column dimensions: 4.8*150 mm. A gradient elution was used following the program given in **Table SI1**. Solvent A was 0.05% trifluoroacetic acid in milliQ water, and solvent B was acetonitrile.

Table SI1: HPLC Gradient elution program for testing compound purity.

Step	Time	Flow rate	% A	% B
1	-	0.8	100	0
2	4	0.8	100	0
3	8	0.8	35	65
4	16	0.8	35	65
5	17	0.8	0	100
6	20	0.8	0	100
7	24	0.8	100	0
8	25	0	-	-

The compounds were detected using a Waters 2996 photodiode array detector. The traces below (**Figure SI1)** show response at 254 nm.

Table SI2: HPLC and HRMS results for compounds (1-11).

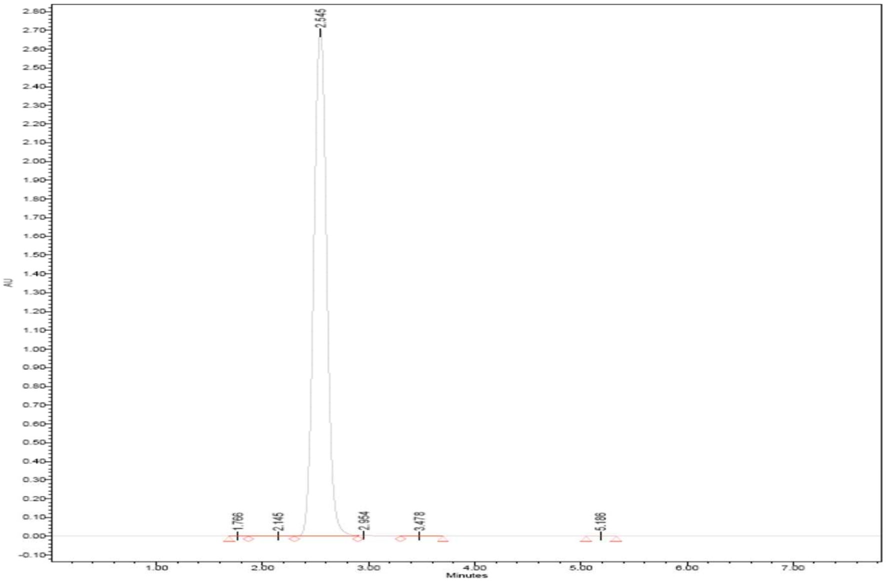
To confirm the identity of the all tested compounds, samples were submitted for high resolution mass spectrometry (HRMS). The results of HPLC and HRMS are summarized in **Table SI2.** HPLC traces and mass spectra are shown in **Figure SI1** and **Figure SI2**, respectively.

Compound Number	HPLC trace		HRMS $(M-H)^-$	
	Retention Time (min)	% Area	Calculated M/z	Found M/z
DANA (1)	2.545	99.53	290.0881	290.0887
2	3.947	99.62	315.0946	315.0954
3	13.809	95.51	417.1416	417.1421
4	2.529	95.55	385.1365	385.1363
5	13.979	97.40	447.1521	447.1527
6	5.507	99.52	371.1208	371.1208
7	4.754	96.78	385.1365	385.1362
8	9.744	95.04	399.1521	399.1518
9	1.943	100	399.1521	399.1517
10	3.403	98.23	331.0895	331.0893
11	13.636	100	433.1365	433.1366

The IC_{50} curves for all the tested compounds **(1-11)** with the human neuraminidases (NEU1, NEU2, NEU3 and NEU4) are shown in **Figure SI3.** We also measured the inhibition constant (K_i) for **6** and **7** with their target enzyme NEU4 (**Figure SI4**). For each inhibitor, the double reciprocal curve (1/product formation rate versus 1/[S]) for several inhibitor concentrations were plotted and fit by linear regression. The slopes of these lines were plotted versus inhibitor concentration and fit to a standard equation ($y = ax+b$). K_i was given by $-b/a$ and the error was calculated from the linear fit.

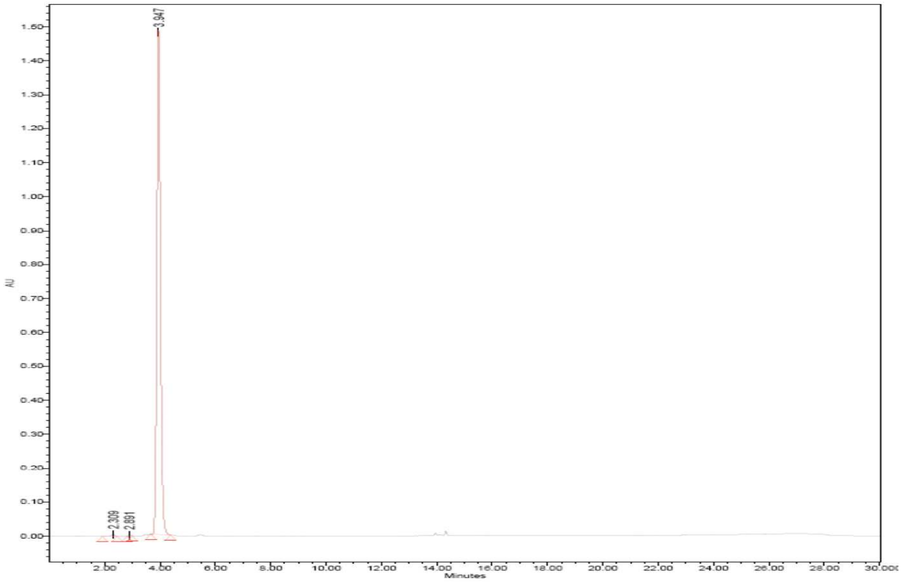
Figure SI1: HPLC traces of the tested compounds (1-11).

a- HPLC trace of DANA (**1**)



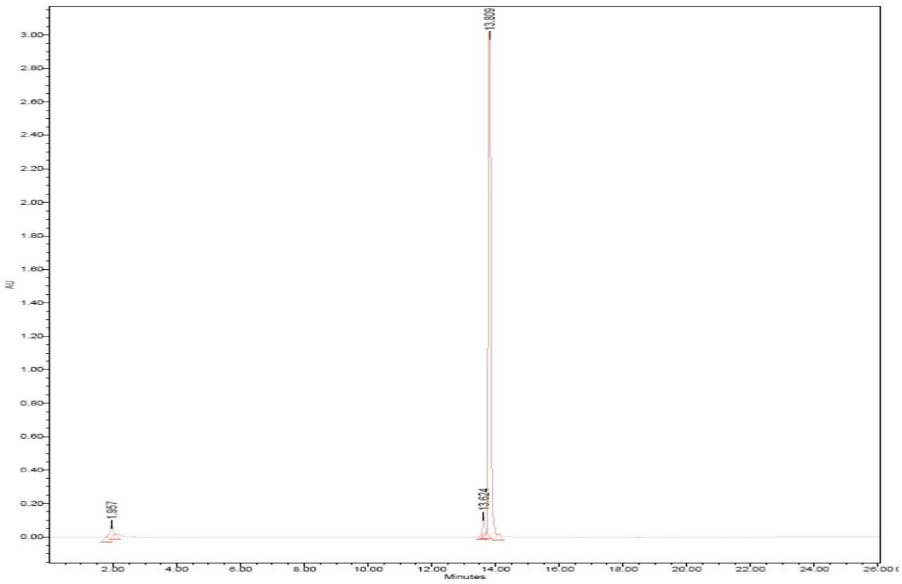
	Retention Time (min)	Area (µV*sec)	%Area	Hight (µV)	Int. Type
1	1.766	1434	0.01	240	BV
2	2.145	15441	0.07	1041	VV
3	2.545	22972011	99.53	2711004	VV
4	2.954	65053	0.28	5414	VV
5	3.478	22964	0.10	1717	VB
6	5.188	3343	0.01	393	BB

b- HPLC trace of **2**



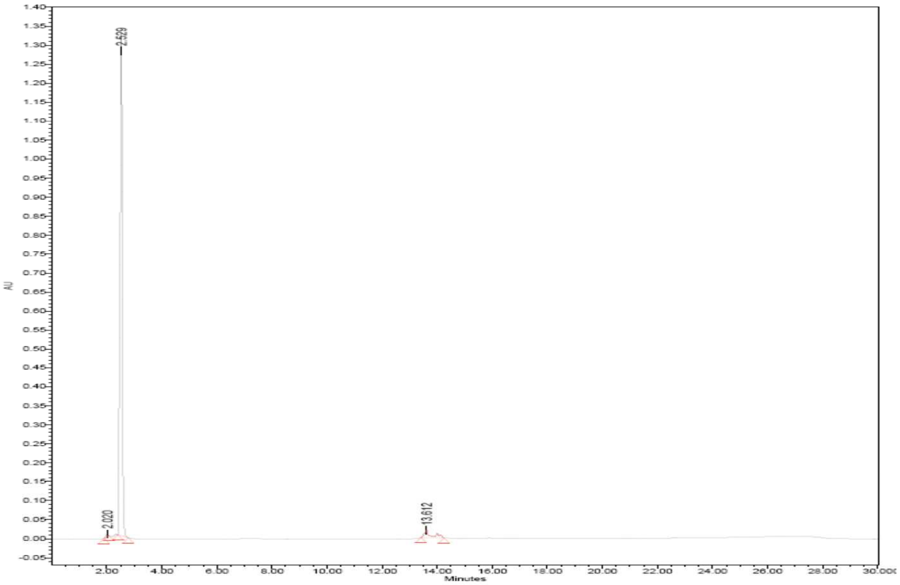
	Retention Time (min)	Area (μV*sec)	%Area	Hight (μV)	Int. Type
1	2.309	39733	0.29	5074	BB
2	2.891	11637	0.09	2223	BB
3	3.947	13551725	99.62	1484883	BB

c- HPLC trace of **3**



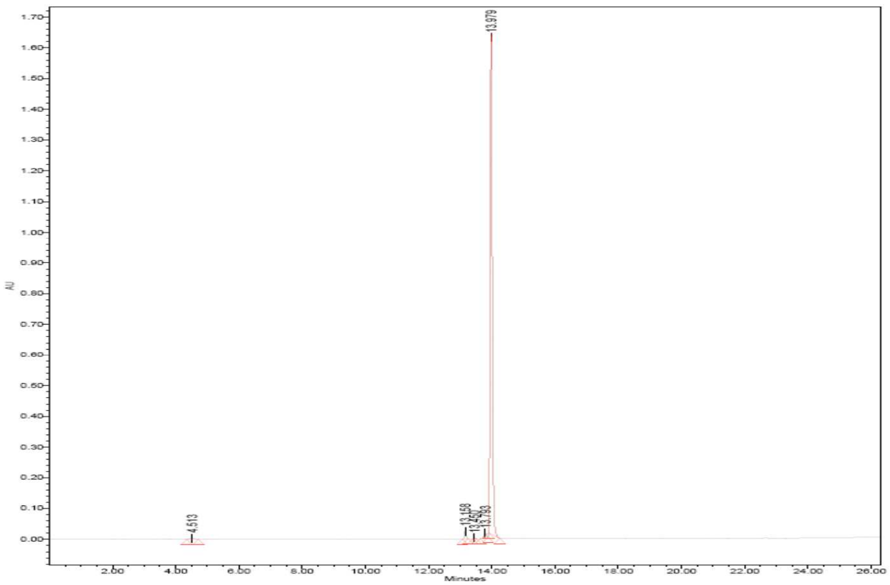
	Retention Time (min)	**Area (µV*sec)**	**%Area**	**Hight (µV)**	**Int. Type**
1	1.957	393536	2.12	64682	BB
2	13.624	439705	2.37	106983	BB
3	13.809	17722666	95.51	3119332	BB

d- HPLC trace of **4**



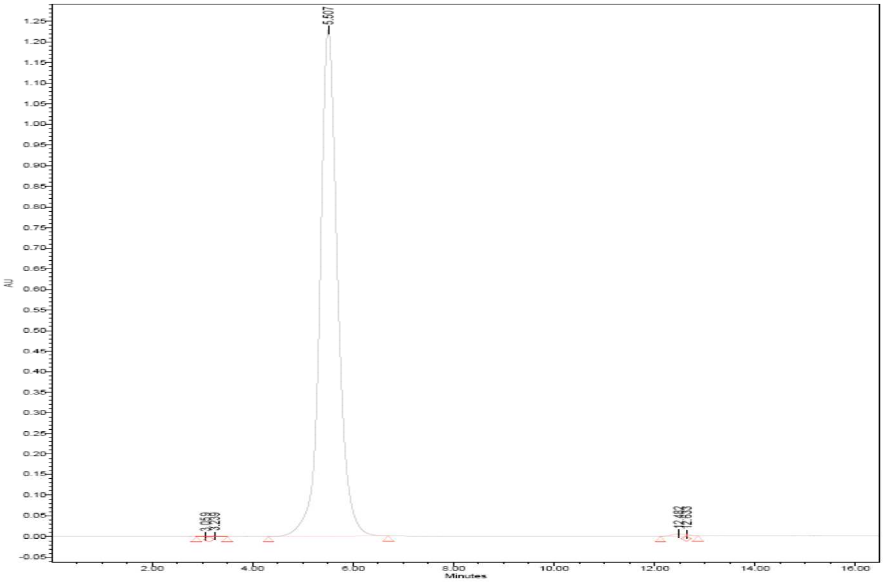
	Retention Time (min)	Area (µV*sec)	%Area	Hight (µV)	Int. Type
1	2.020	27419	0.34	5789	BB
2	2.529	7623335	95.55	1267933	BB
3	13.612	327262	4.10	19026	BB

e- HPLC trace of **5**



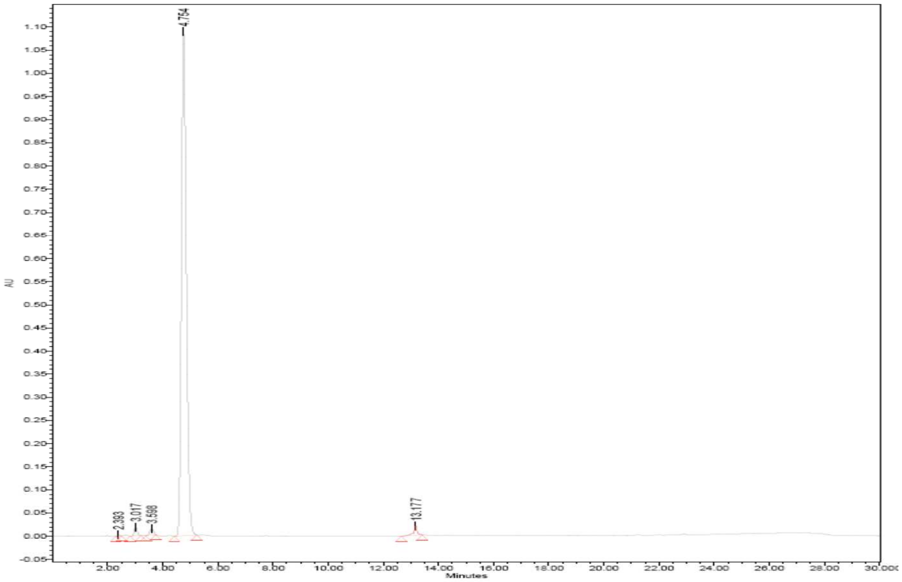
	Retention Time (min)	Area (µV*sec)	%Area	Hight (µV)	Int. Type
1	4.513	33200	0.42	3401	BB
2	13.158	92336	1.16	24174	BB
3	13.450	7984	0.10	2183	BB
4	13.793	73418	0.92	14344	BB
5	13.979	7755670	97.40	1618956	BB

f- HPLC trace of **6**



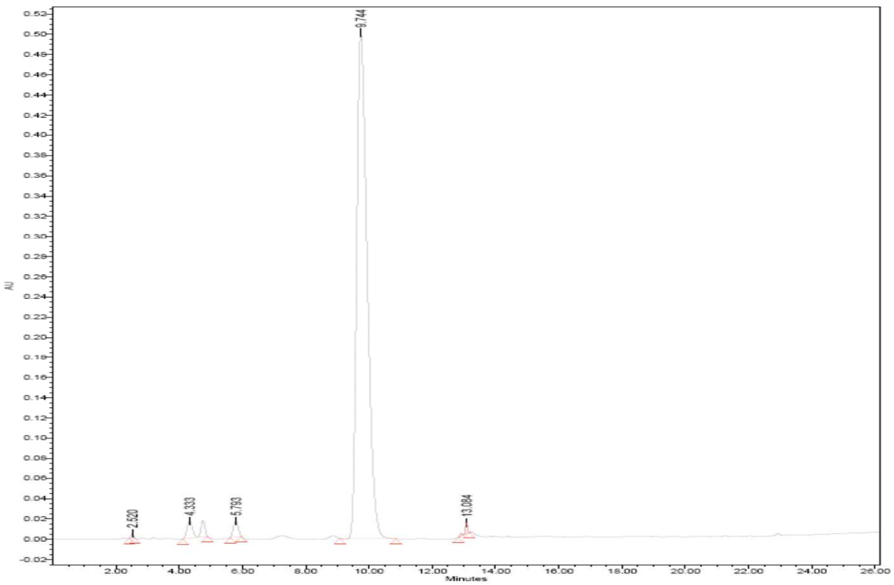
	Retention Time (min)	**Area (µV*sec)**	**%Area**	**Hight (µV)**	**Int. Type**
1	3.059	6678	0.02	744	BV
2	3.239	17092	0.06	1524	VB
3	5.507	30162534	99.52	1229252	BB
4	12.482	103272	0.34	5997	BV
5	12.633	19365	0.06	4086	VB

g- HPLC trace of **7**



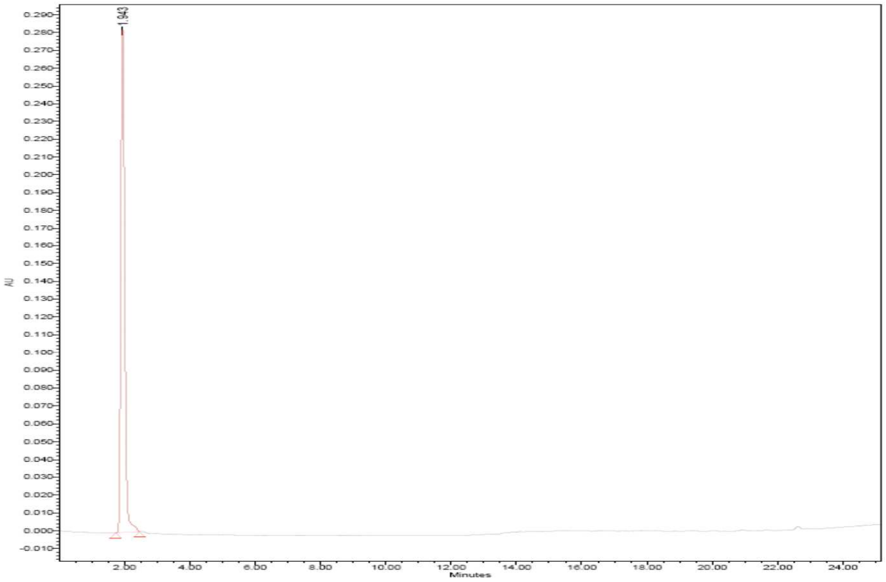
	Retention Time (min)	Area (µV*sec)	%Area	Hight (µV)	Int. Type
1	2.393	13191	0.09	2574	BB
2	3.017	138138	0.99	18289	BB
3	3.598	124768	0.89	13230	BB
4	4.754	13533015	96.78	1093283	BB
5	13.177	173816	1.24	20052	BB

h- HPLC trace of **8**



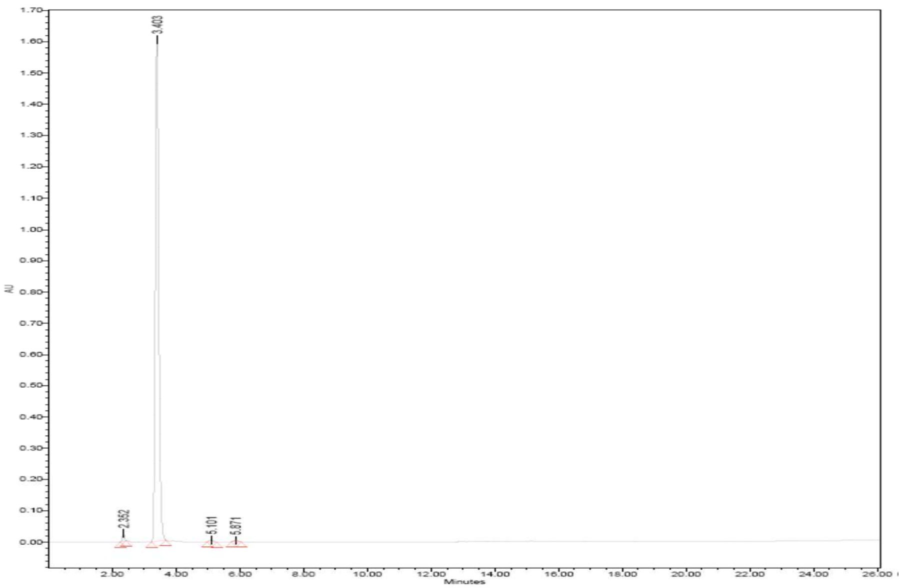
	Retention Time (min)	Area (µV*sec)	%Area	Hight (µV)	Int. Type
1	2.520	24032	0.20	4554	BB
2	4.333	347679	2.90	16945	BB
3	5.793	171629	1.34	15768	BB
4	9.744	11404544	95.04	501348	BB
5	13.084	52279	0.44	10488	BB

i- HPLC trace of **9**



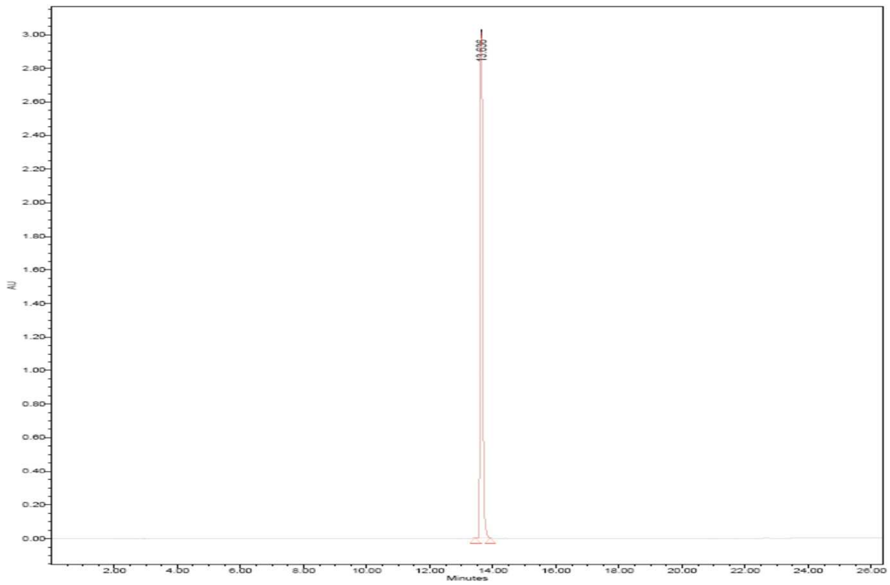
	Retention Time (min)	Area (µV*sec)	%Area	Hight (µV)	Int. Type
1	1.943	2290792	100.00	283664	BB

j- HPLC trace of **10**



	Retention Time (min)	Area (µV*sec)	%Area	Hight (µV)	Int. Type
1	2.352	111889	0.92	24986	BB
2	3.403	11934962	98.23	1614598	BB
3	5.101	56756	0.47	6185	BB
4	5.871	45875	0.38	4516	BB

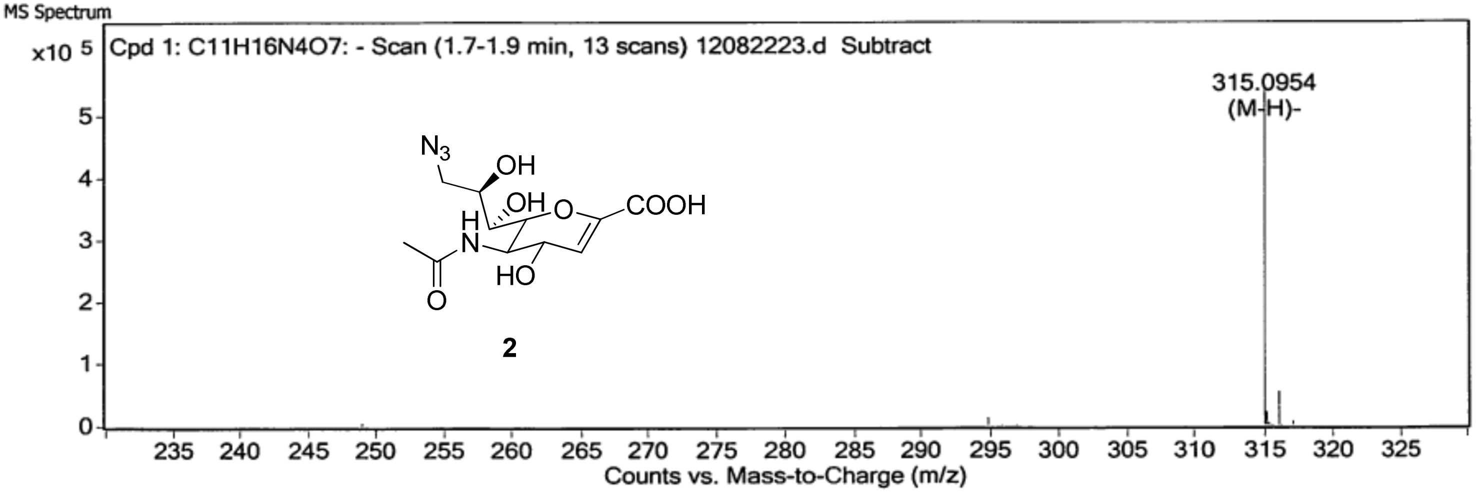
k- HPLC trace of **11**



	Retention Time (min)	**Area (µV*sec)**	**%Area**	**Hight (µV)**	**Int. Type**
1	13.636	16618000	100.00	3107171	BB

Figure SI2: HRMS spectra of the tested compounds (1-11).

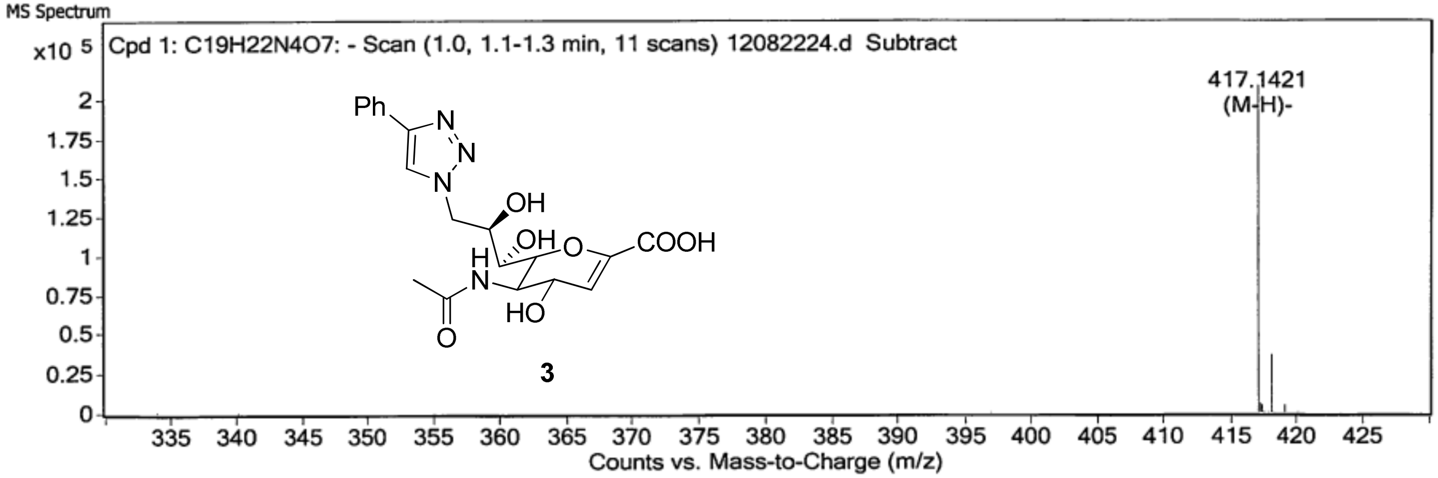
a- HRMS of **2**: $(M-H)^-$ found 315.0954, calculated 315.0946



MS Spectrum Peak List

m/z	Calc m/z	Diff(ppm)	Abund	Formula	Ion
315.0954	315.0946	2.41	544510	C11 H15 N4 O7	(M-H)-

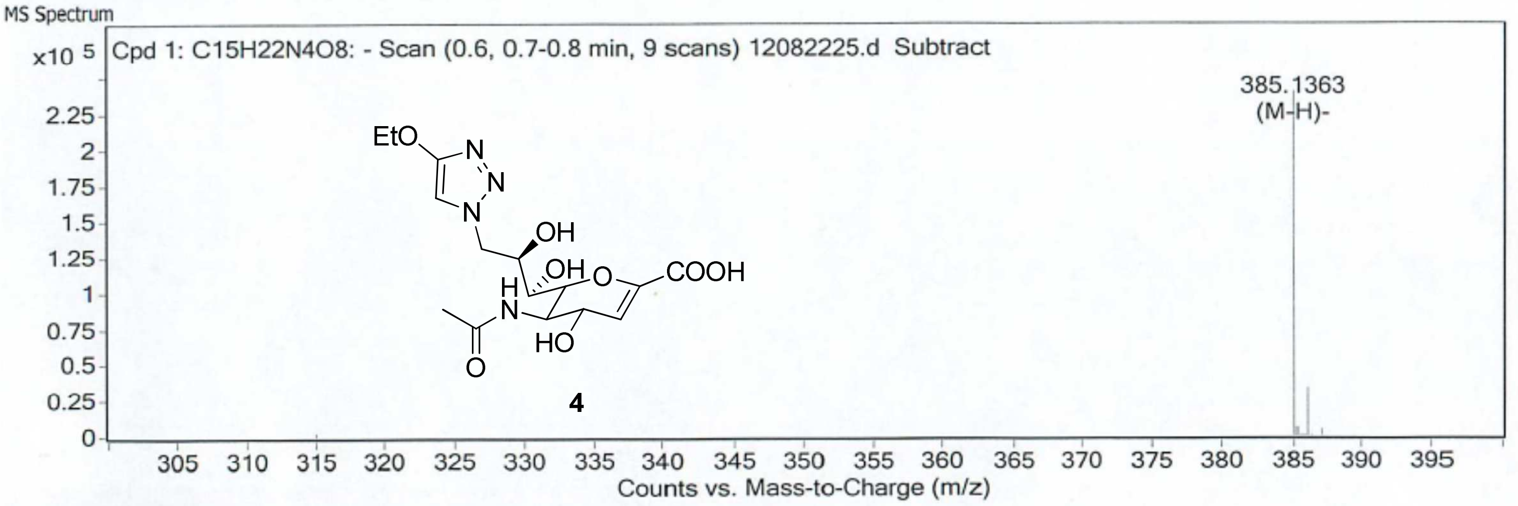
b- HRMS of **3**: (M-H)$^-$ found 417.1421, calculated 417.1416



MS Spectrum Peak List

m/z	*Calc m/z*	Diff(ppm)	z	Abund	Formula	Ion
417.1421	417.1416	1.16	1	209037	C19 H21 N4 O7	(M-H)-

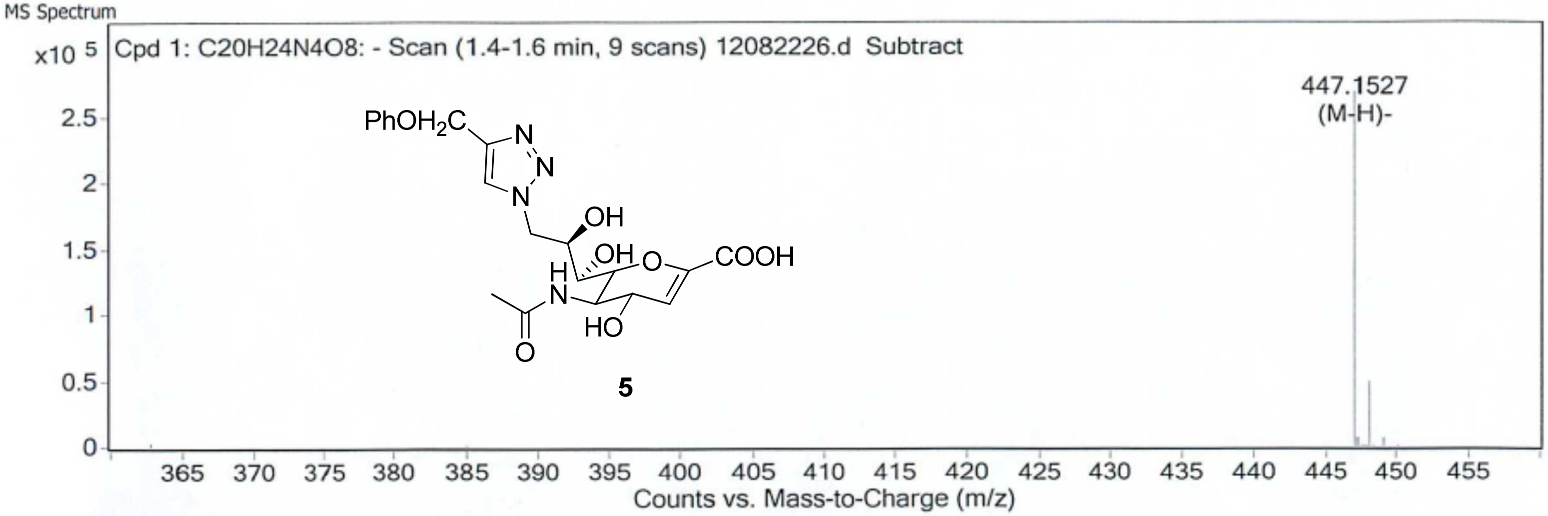
c- HRMS of **4**: (M-H)$^-$ found 385.1363, calculated 385.1365



MS Spectrum Peak List

m/z	*Calc m/z*	Diff(ppm)	Abund	Formula	Ion
385.1363	385.1365	-0.44	240633	C15 H21 N4 O8	(M-H)-

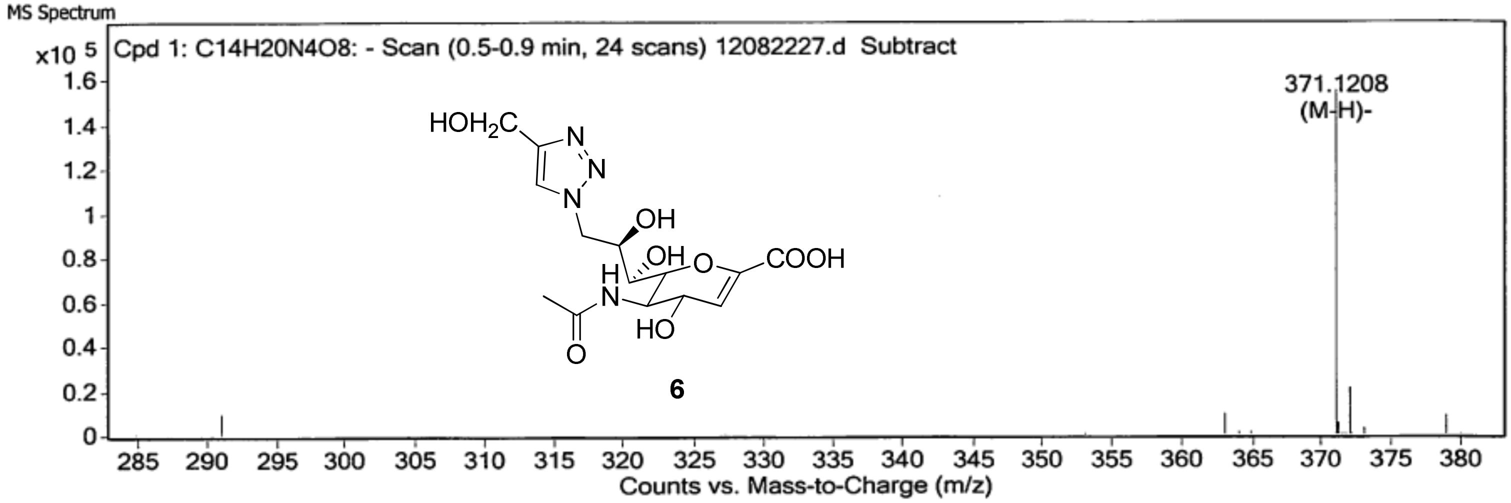
d- HRMS of **5**: (M-H)$^-$ found 447.1527, calculated 447.1521



MS Spectrum Peak List

m/z	*Calc m/z*	Diff(ppm)	Abund	Formula	Ion
447.1527	447.1521	1.18	269017	C20 H23 N4 O8	(M-H)-

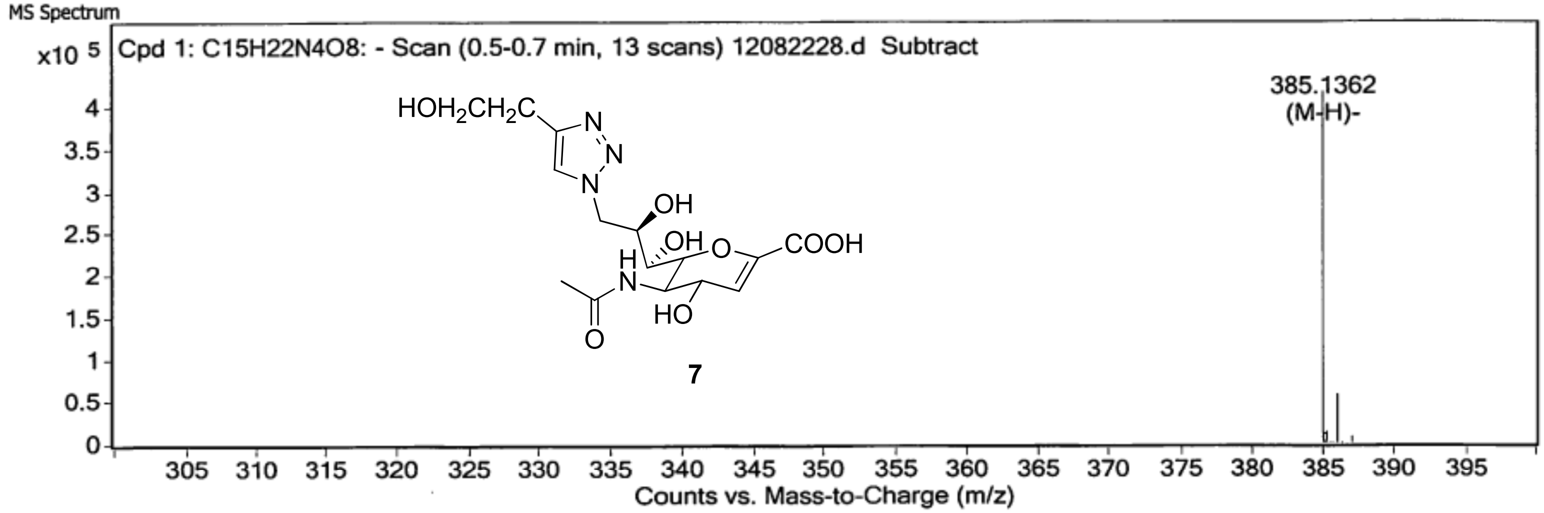
e- HRMS of **6**: (M-H)$^-$ found 371.1208, calculated 371.1208



MS Spectrum Peak List

m/z	*Calc m/z*	Diff(ppm)	z	Abund	Formula	Ion
371.1208	371.1208	-0.17	-1	155741	C14 H19 N4 O8	(M-H)-

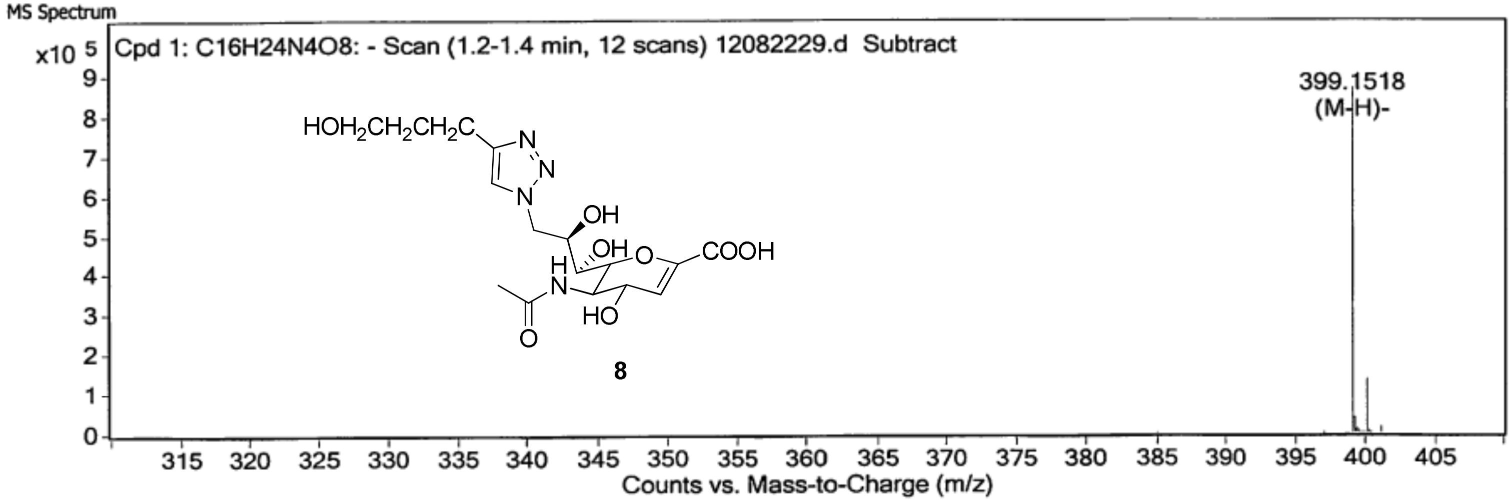
f- HRMS of **7**: $(M\text{-}H)^-$ found 385.1362, calculated 385.1365



MS Spectrum Peak List

m/z	*Calc m/z*	Diff(ppm)	Abund	Formula	Ion
385.1362	385.1365	-0.66	419004	C15 H21 N4 O8	(M-H)-

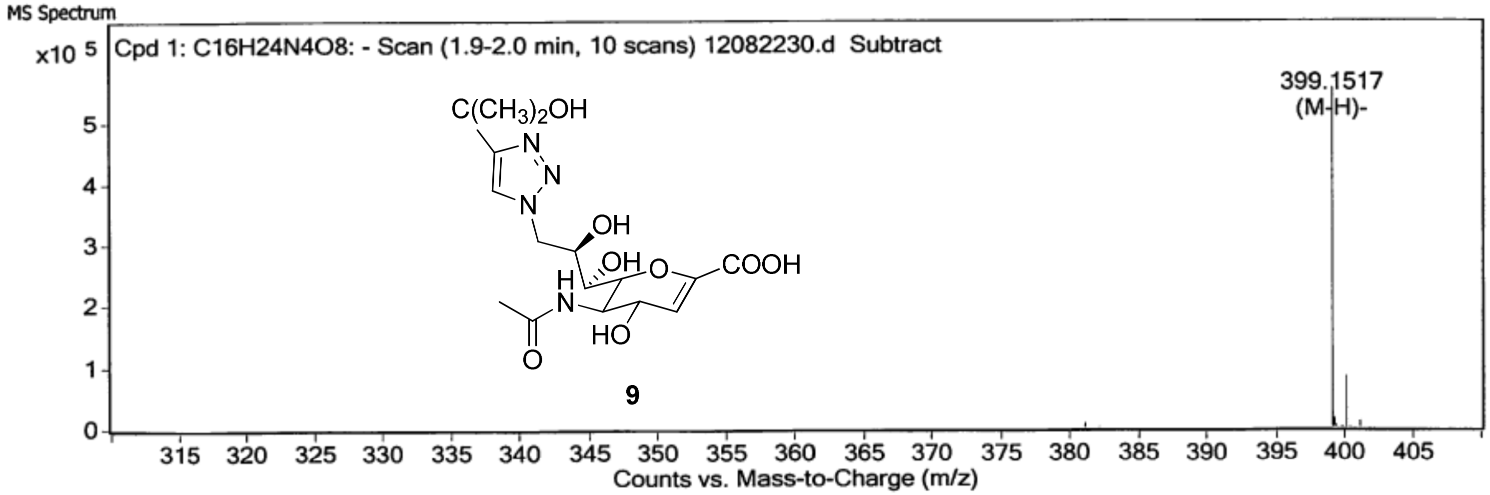
g- HRMS of **8**: $(M\text{-}H)^-$ found 399.1518, calculated 399.1521



MS Spectrum Peak List

m/z	*Calc m/z*	Diff(ppm)	Abund	Formula	Ion
399.1518	399.1521	-0.96	870595	C16 H23 N4 O8	(M-H)-

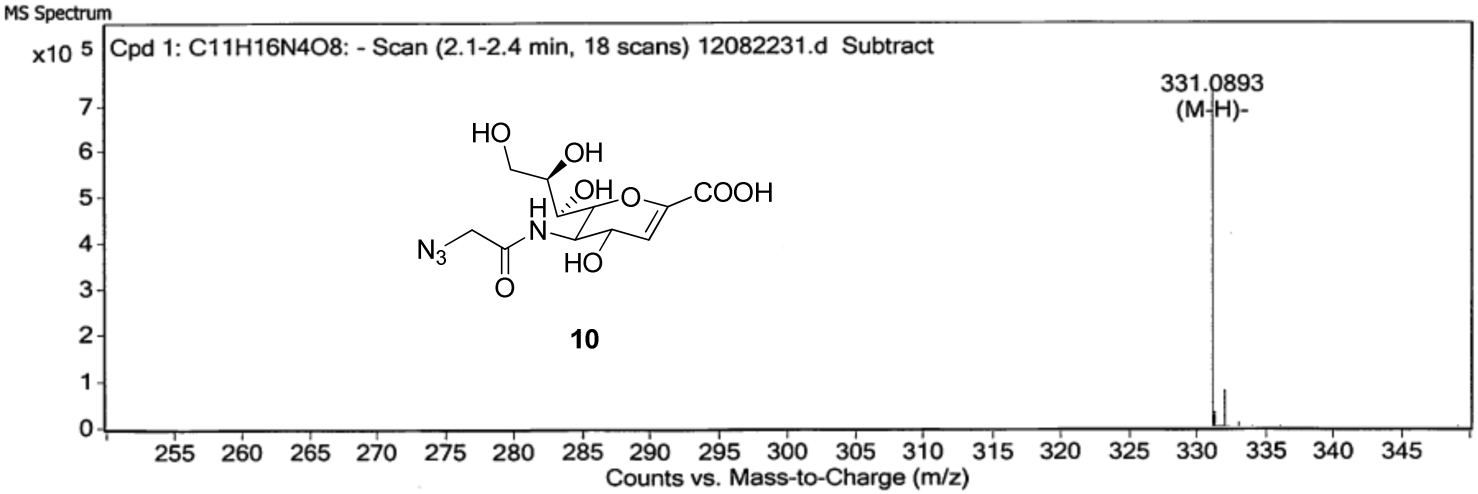
h- HRMS of **9**: $(M-H)^-$ found 399.1517, calculated 399.1521



MS Spectrum Peak List

m/z	*Calc m/z*	Diff(ppm)	Abund	Formula	Ion
399.1517	399.1521	-1.13	555721	C16 H23 N4 O8	(M-H)-

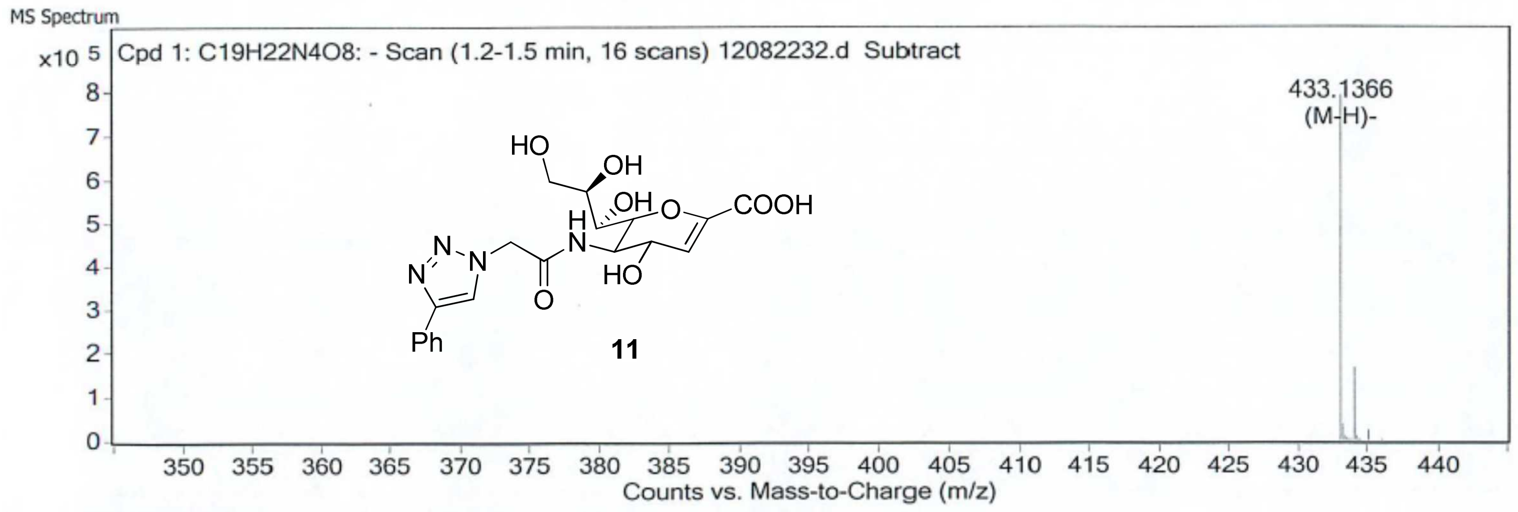
i- HRMS of **10**: $(M\text{-}H)^-$ found 331.0893, calculated 331.0895



MS Spectrum Peak List

m/z	*Calc m/z*	Diff(ppm)	Abund	Formula	Ion
331.0893	331.0895	-0.64	734783	C11 H15 N4 O8	(M-H)-

j- HRMS of **11**: (M-H)$^{-}$ found 433.1366, calculated 433.1365



MS Spectrum Peak List

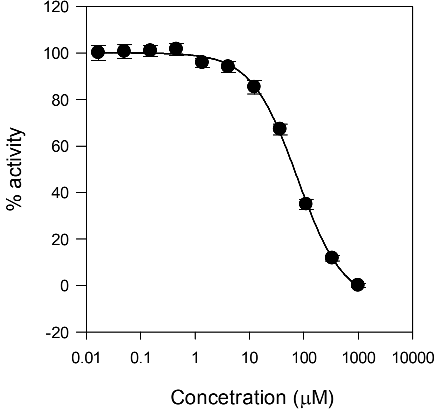
m/z	*Calc m/z*	Diff(ppm)	Abund	Formula	Ion
433.1366	433.1365	0.36	793553	C19 H21 N4 O8	(M-H)-

Figure SI3: IC50 curves of compounds 1-11 with NEU1, NEU2, NEU3 and NEU4.

a - IC_{50} curves of compounds **1-11** with NEU1

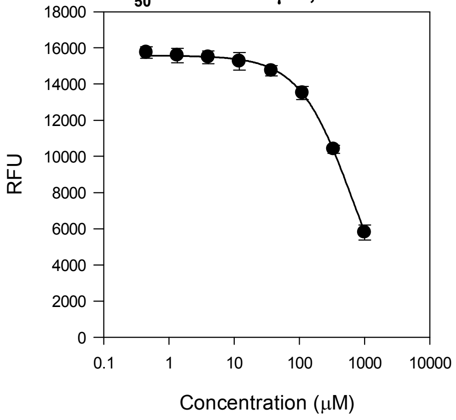
IC_{50} of DANA (1) with NEU1

IC_{50} = 76 ± 6 μM, r^2= 0.998



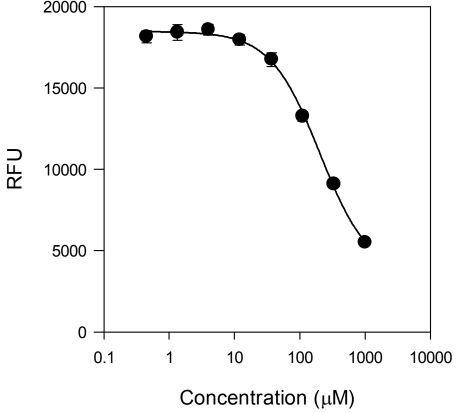
IC_{50} of 2 with NEU1

IC_{50} = 620 ± 10 μM, r^2= 0.992



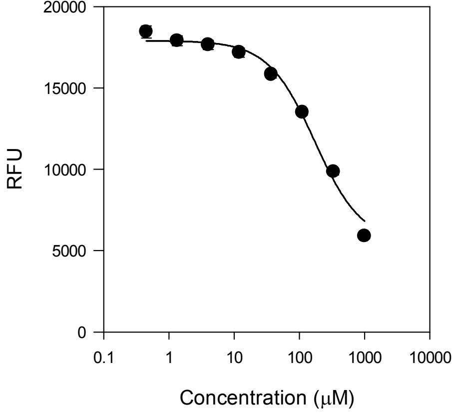
IC_{50} of 3 with NEU1

IC_{50} = 200 ± 30 μM, r^2= 0.998



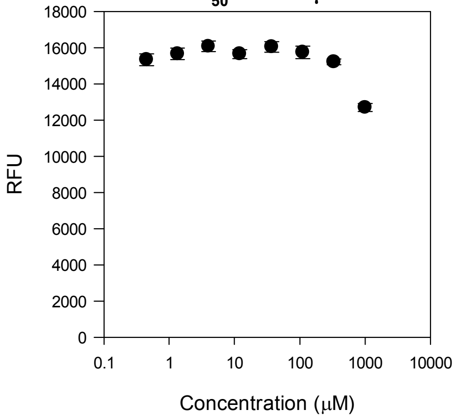
IC_{50} of 4 with NEU1

IC_{50} = 170 ± 20 μM, r^2= 0.988



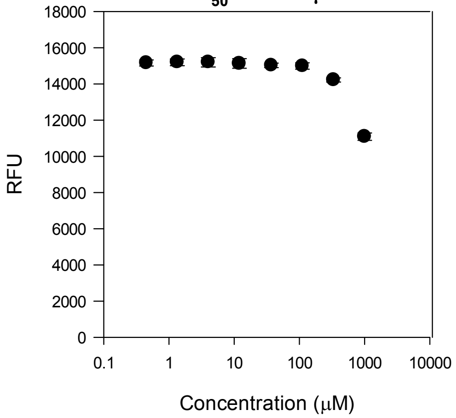
IC_{50} of 5 with NEU1

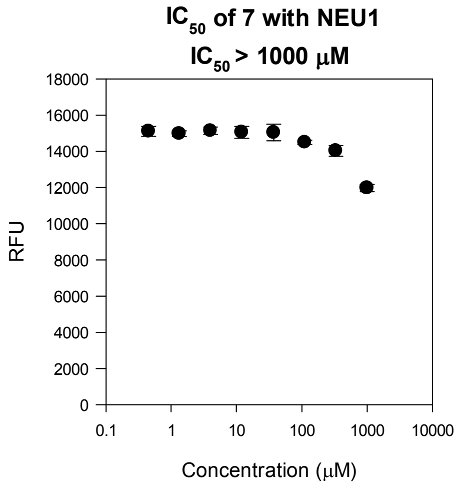
IC_{50} > 1000 μM



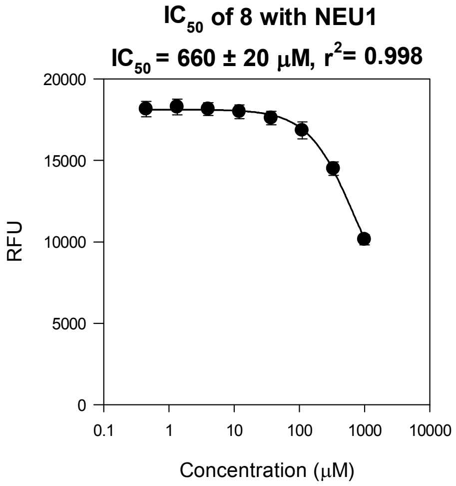
IC_{50} of 6 with NEU1

IC_{50} > 1000 μM

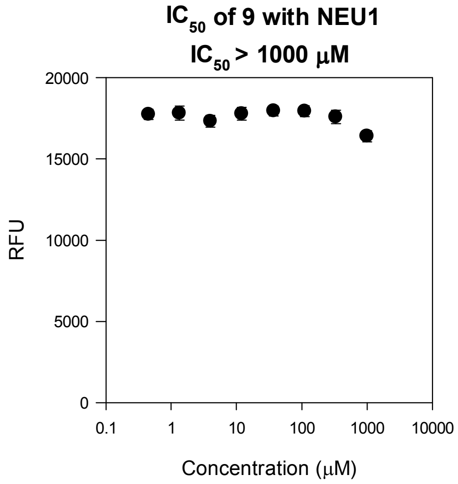



IC_{50} of 7 with NEU1

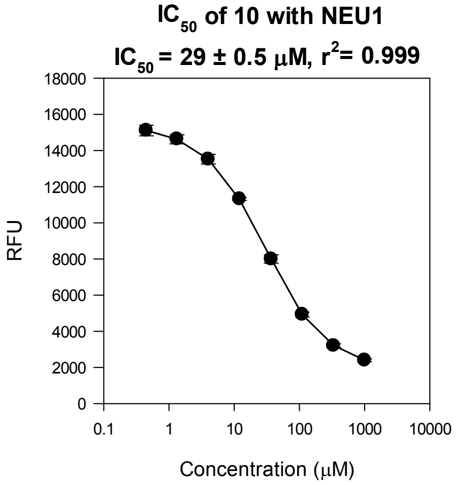
IC_{50} > 1000 μM


IC_{50} of 8 with NEU1

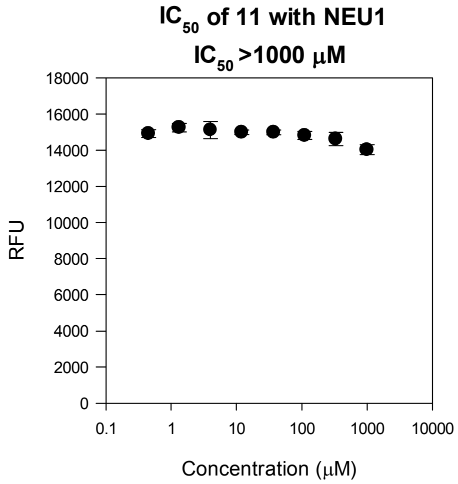
IC_{50} = 660 ± 20 μM, r^2= 0.998


IC_{50} of 9 with NEU1

IC_{50} > 1000 μM


IC_{50} of 10 with NEU1

IC_{50} = 29 ± 0.5 μM, r^2= 0.999

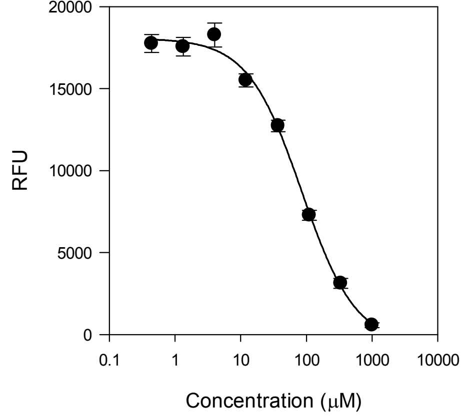

IC_{50} of 11 with NEU1

IC_{50} >1000 μM

b - IC_{50} curves of compounds **1-11** with NEU2

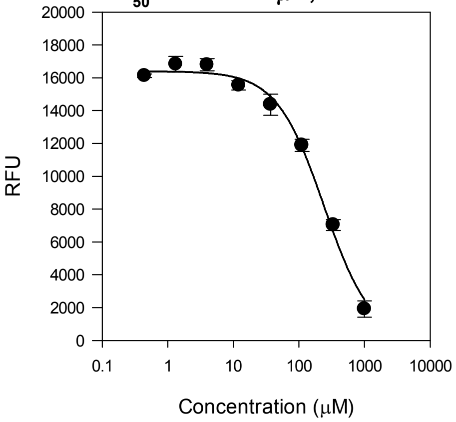
IC_{50} of DANA (1) with NEU2

IC_{50} = 85 ± 13 μM, r^2= 0.996



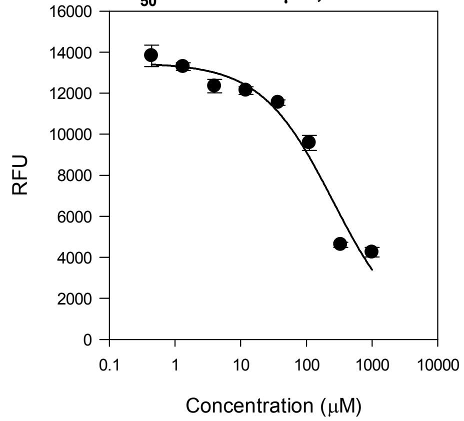
IC_{50} of 2 with NEU2

IC_{50} = 240 ± 20 μM, r^2= 0.992



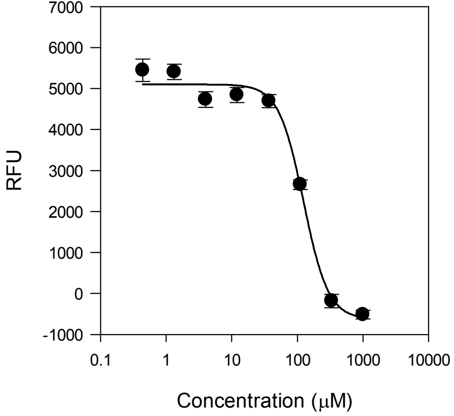
IC_{50} of 3 with NEU2

IC_{50} = 250 ± 70 μM, r^2= 0.959



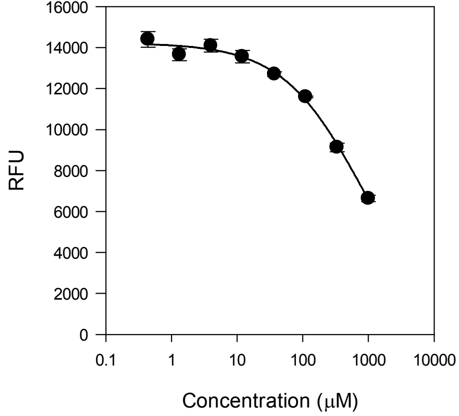
IC_{50} of 4 with NEU2

IC_{50} = 130 ± 20 μM, r^2= 0.990



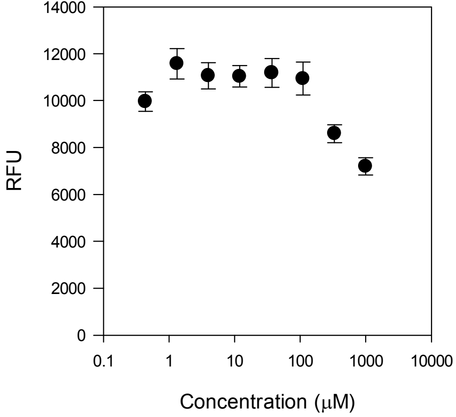
IC_{50} of 5 with NEU2

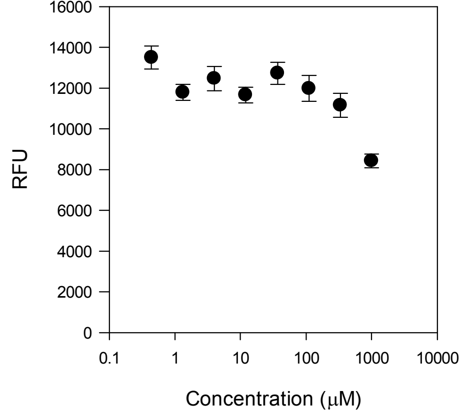
IC_{50} = 800 ± 70 μM, r^2= 0.994

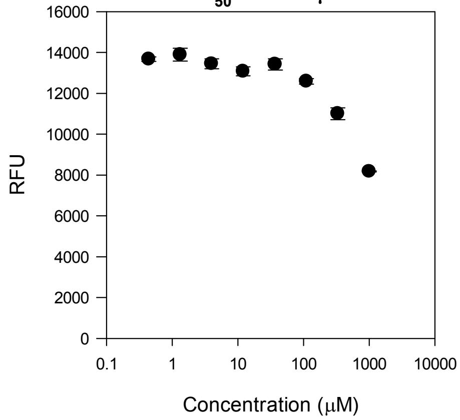


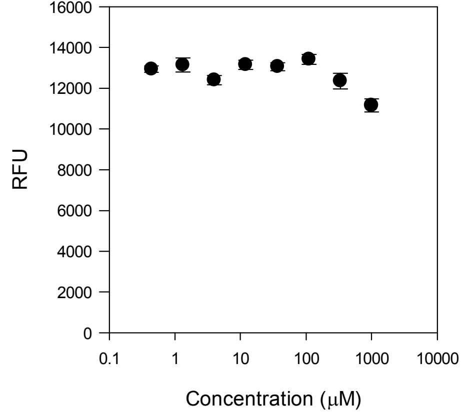
IC_{50} of 6 with NEU2

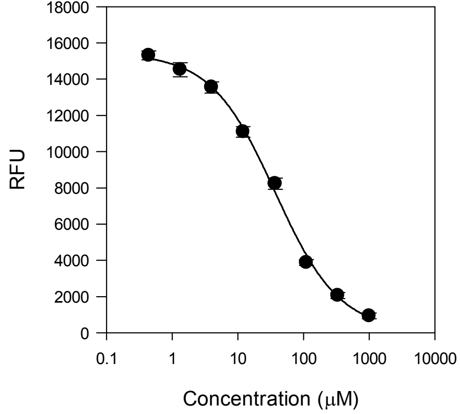
IC_{50} >1000 μM

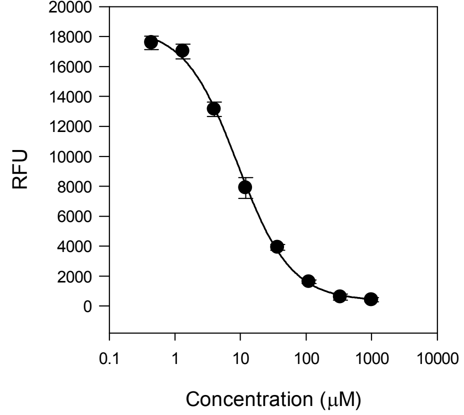



IC_{50} of 7 with NEU2
IC_{50} > 1000μM


IC_{50} of 8 with NEU2
IC_{50} > 1000 μM


IC_{50} of 9 with NEU2
IC_{50} > 1000 μM

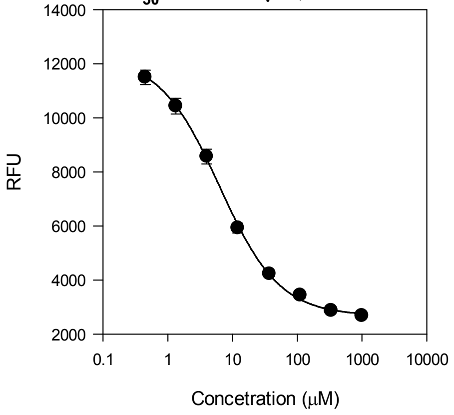

IC_{50} of 10 with NEU2
IC_{50} = 37 ± 5 μM, r^2= 0.998


IC_{50} of 11 with NEU2
IC_{50} = 9 ± 1 μM, r^2= 0.999

c - IC_{50} curves of compounds **1-11** with NEU3

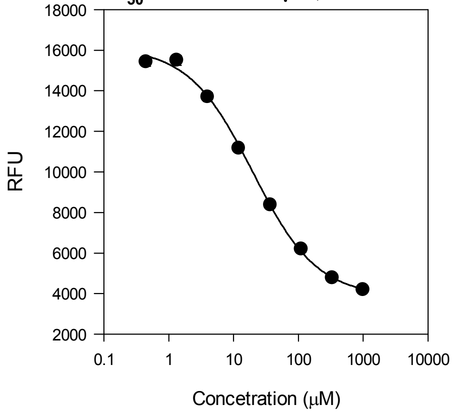
IC_{50} of DANA (1) with NEU3

IC_{50} = 6 ± 0.5 μM, r^2= 0.999



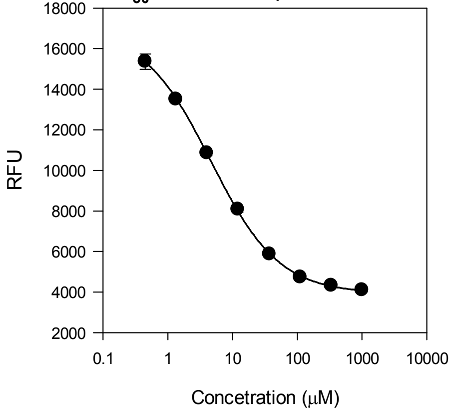
IC_{50} of 2 with NEU3

IC_{50} = 19.7 ± 2.3 μM, r^2= 0.998



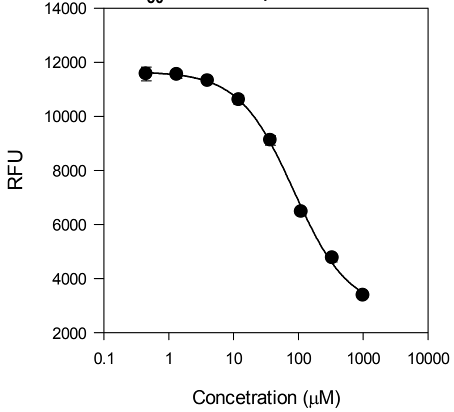
IC_{50} of 3 with NEU3

IC_{50} = 4.6 ± 0.2 μM, r^2= 0.999



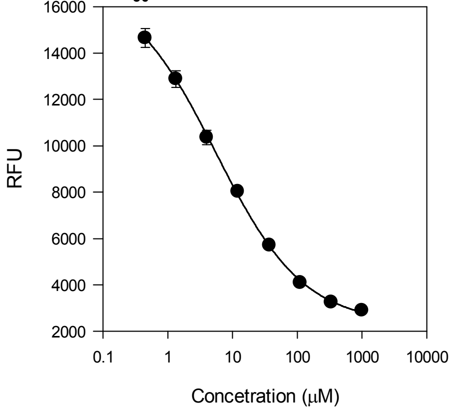
IC_{50} of 4 with NEU3

IC_{50} = 85 ± 8 μM, r^2= 0.999



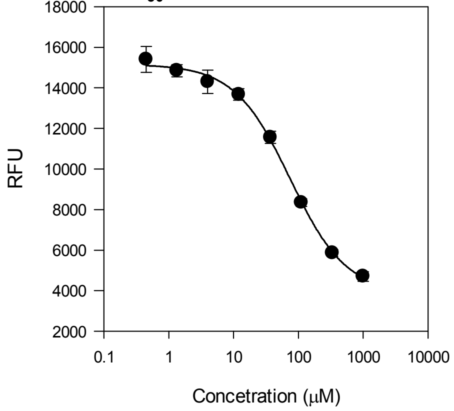
IC_{50} of 5 with NEU3

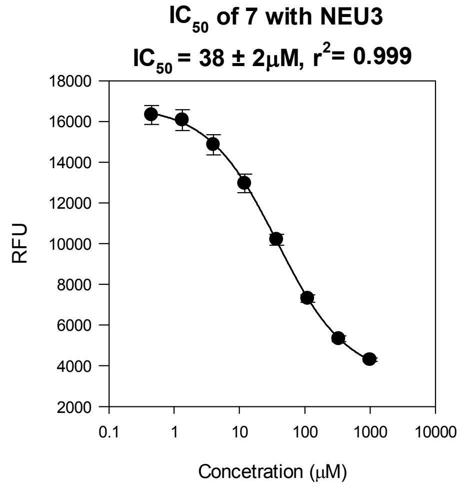
IC_{50} = 5.5 ± 0.6 μM, r^2= 0.999



IC_{50} of 6 with NEU3

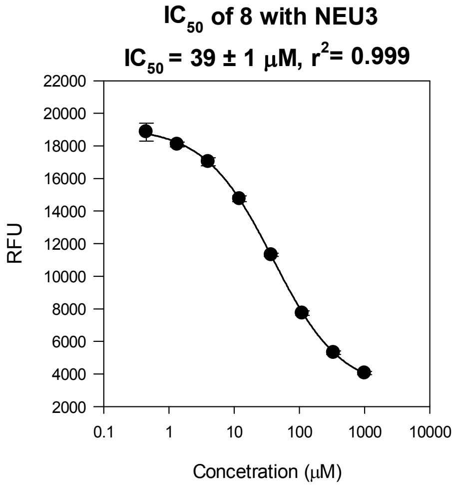
IC_{50} = 75 ± 9 μM, r^2= 0.998





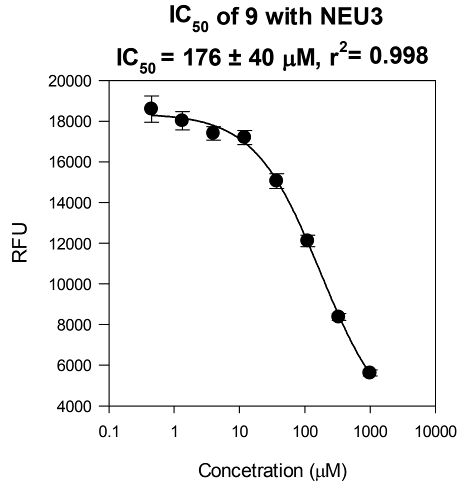
IC$_{50}$ of 7 with NEU3

IC$_{50}$ = 38 ± 2μM, r^2= 0.999



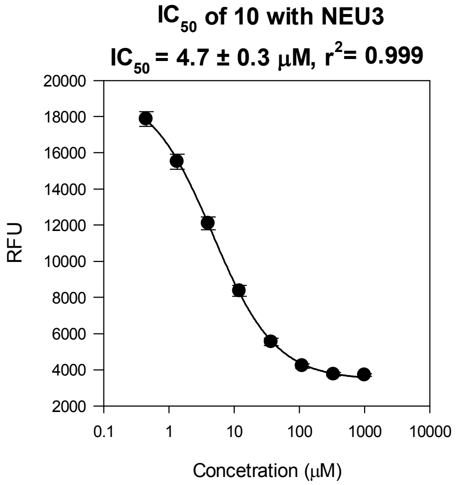
IC$_{50}$ of 8 with NEU3

IC$_{50}$ = 39 ± 1 μM, r^2= 0.999



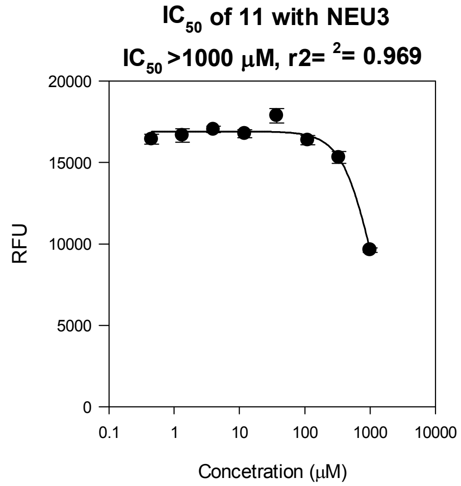
IC$_{50}$ of 9 with NEU3

IC$_{50}$ = 176 ± 40 μM, r^2= 0.998



IC$_{50}$ of 10 with NEU3

IC$_{50}$ = 4.7 ± 0.3 μM, r^2= 0.999



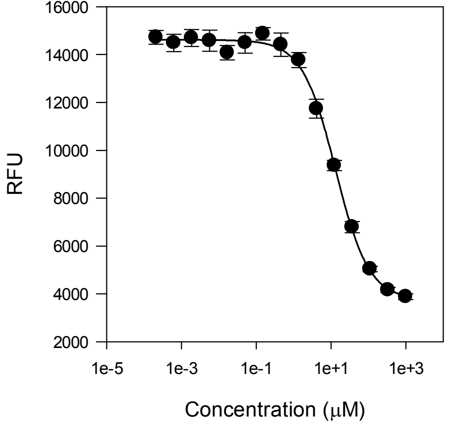
IC$_{50}$ of 11 with NEU3

IC$_{50}$ >1000 μM, r2= ^{2}= 0.969

d - IC_{50} curves of compounds **1-11** with NEU4

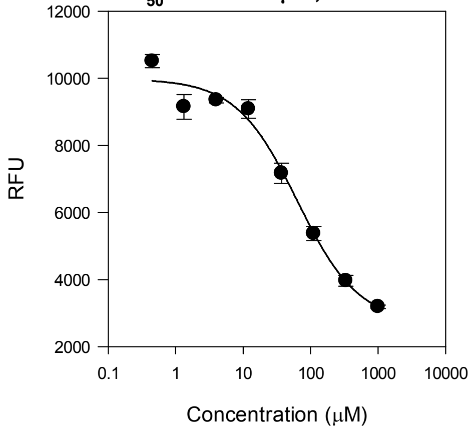
IC_{50} of DANA (1) with NEU4

IC_{50} = 13 ± 1μM, r^2= 0.998



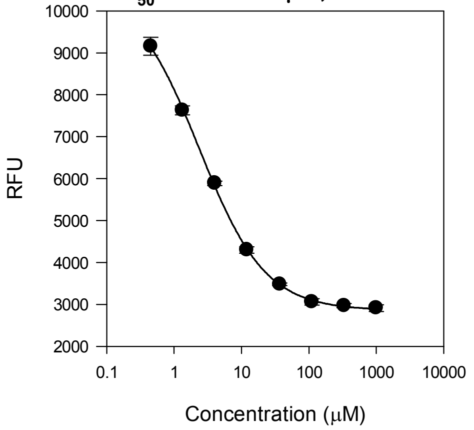
IC_{50} of 2 with NEU4

IC_{50} = 63 ± 20 μM, r^2= 0.983



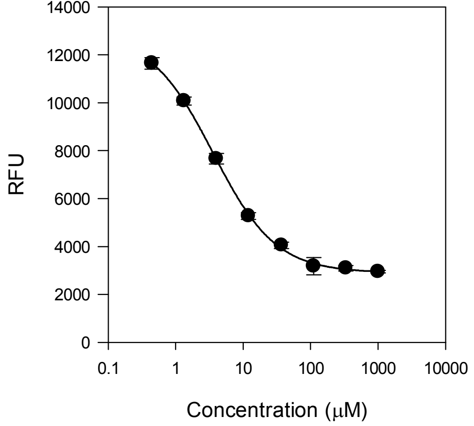
IC_{50} of 3 with NEU4

IC_{50} = 2.3 ± 0.2 μM, r^2= 0.999



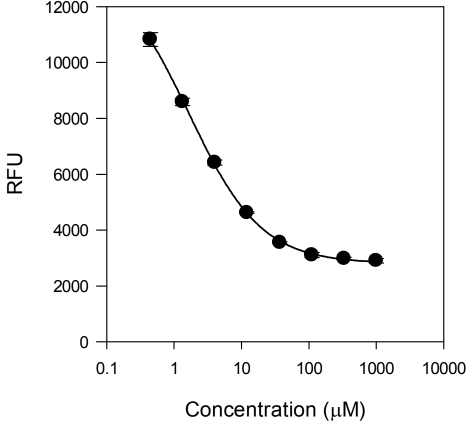
IC_{50} of 4 with NEU4

IC_{50} = 3.6 ± 0.3 μM, r^2= 0.999



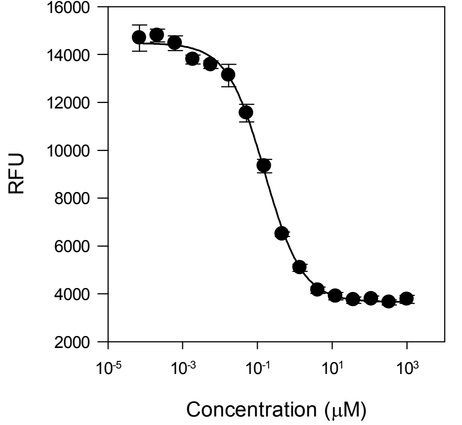
IC_{50} of 5 with NEU4

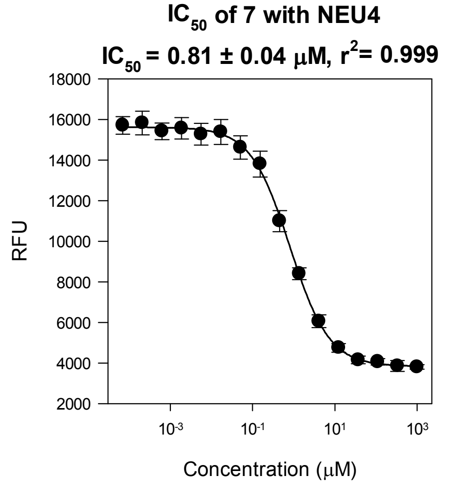
IC_{50} = 1.7 ± 0.2 μM, r^2= 0.999



IC_{50} of 6 with NEU4

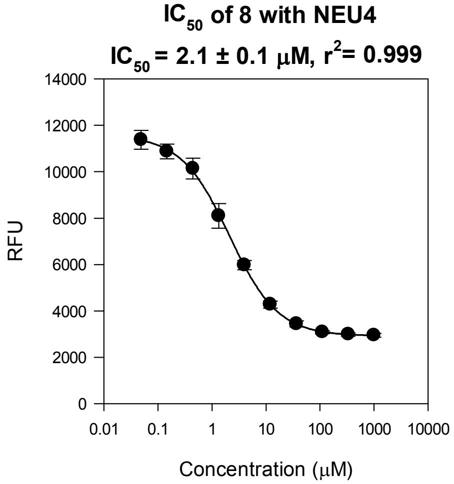
IC_{50} = 0.16 ± 0.01 μM, r^2= 0.998





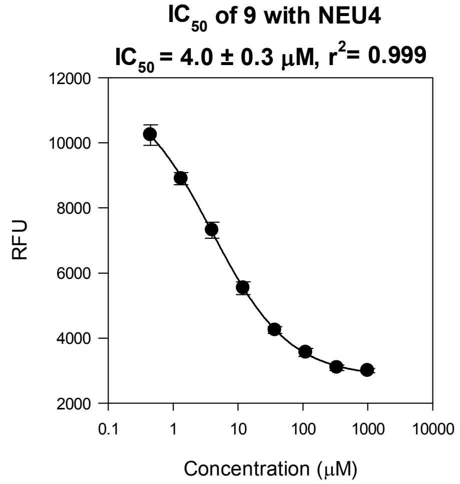
IC$_{50}$ of 7 with NEU4

IC$_{50}$ = 0.81 ± 0.04 μM, r^2= 0.999



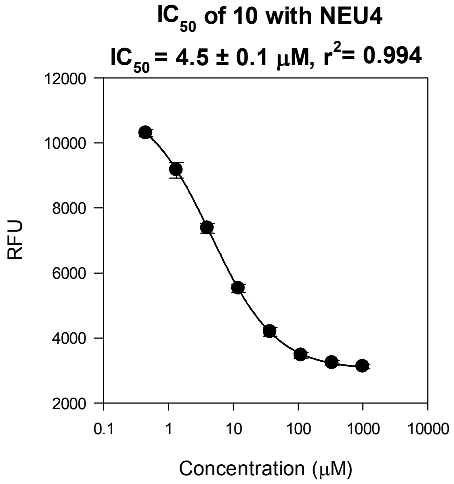
IC$_{50}$ of 8 with NEU4

IC$_{50}$ = 2.1 ± 0.1 μM, r^2= 0.999



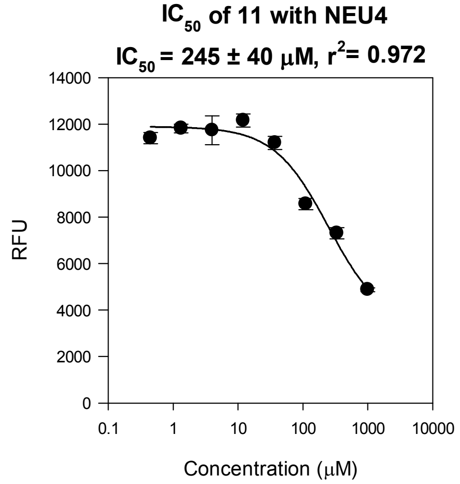
IC$_{50}$ of 9 with NEU4

IC$_{50}$ = 4.0 ± 0.3 μM, r^2= 0.999



IC$_{50}$ of 10 with NEU4

IC$_{50}$ = 4.5 ± 0.1 μM, r^2= 0.994



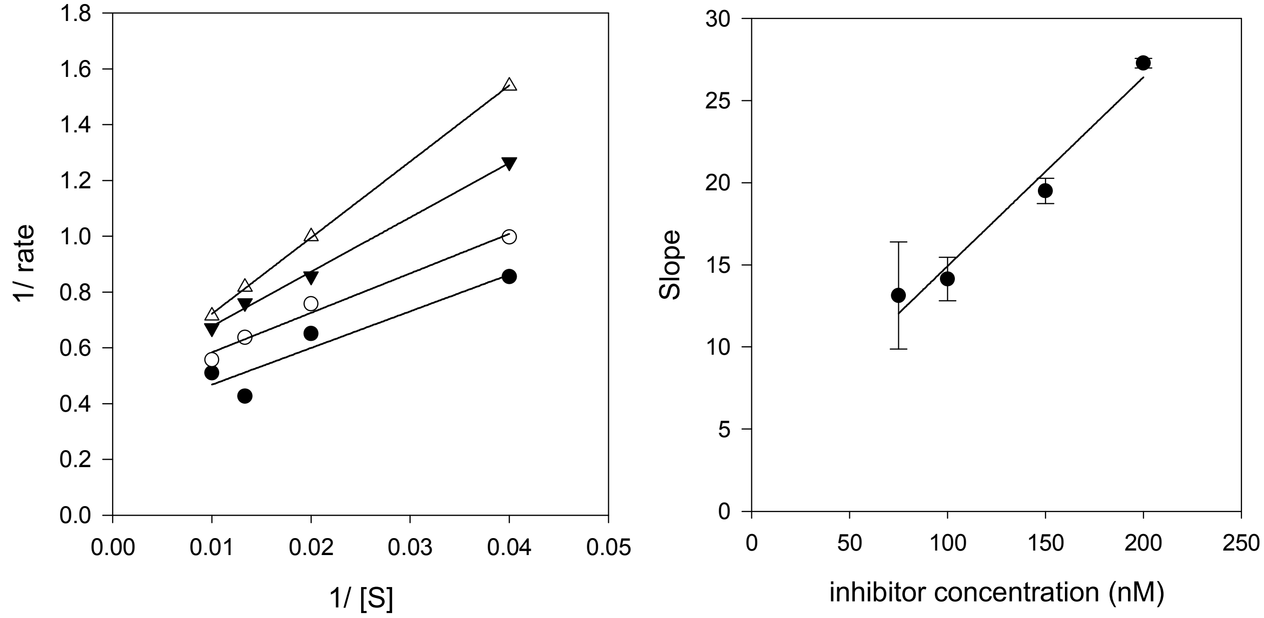
IC$_{50}$ of 11 with NEU4

IC$_{50}$ = 245 ± 40 μM, r^2= 0.972

Figure SI4: K_i curves for compounds 6 and 7 with NEU4.

Measurements were done according to protocol in Materials and Methods. The top left graph is the double reciprocal curve of rate and the substrate concentration at serial inhibitor concentrations (See table insert for legend and fit results). Plots of the slopes versus inhibitor concentration provide a linear curve (y= ab + c) where K_i is found as –b/a.

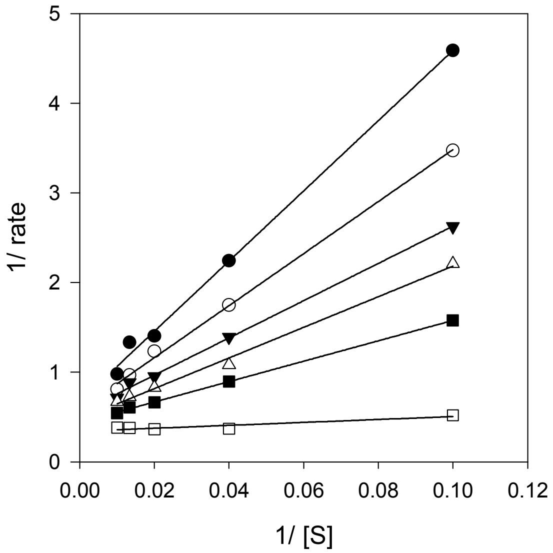
a- K_i of 6 with NEU4, K_i= 30 ± 19 nM



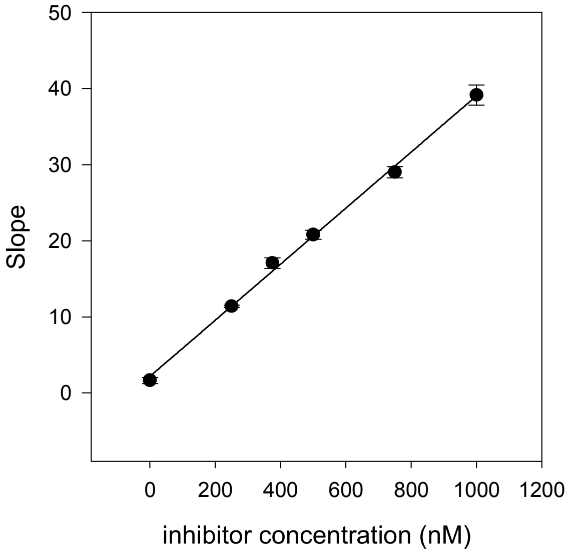
	[I] (nM)	Slope	±	r^2
△	200	27.28	0.297	0.999
▼	150	19.5	0.764	0.997
○	100	14.15	1.332	0.983
●	75	13.13	3.269	0.890

y= 0.115 x + 3.4135
r^2= 0.9691

b- K_i of 7 with NEU4, K_i= 60 ± 16 nM



	[I] nM	Slope	±	r^2
●	1000	39.147	1.321	0.997
○	750	29.006	0.739	0.998
▼	500	20.771	0.584	0.998
△	375	17.077	0.694	0.995
■	250	11.372	0.129	0.999
□	0	1.632	0.392	0.852



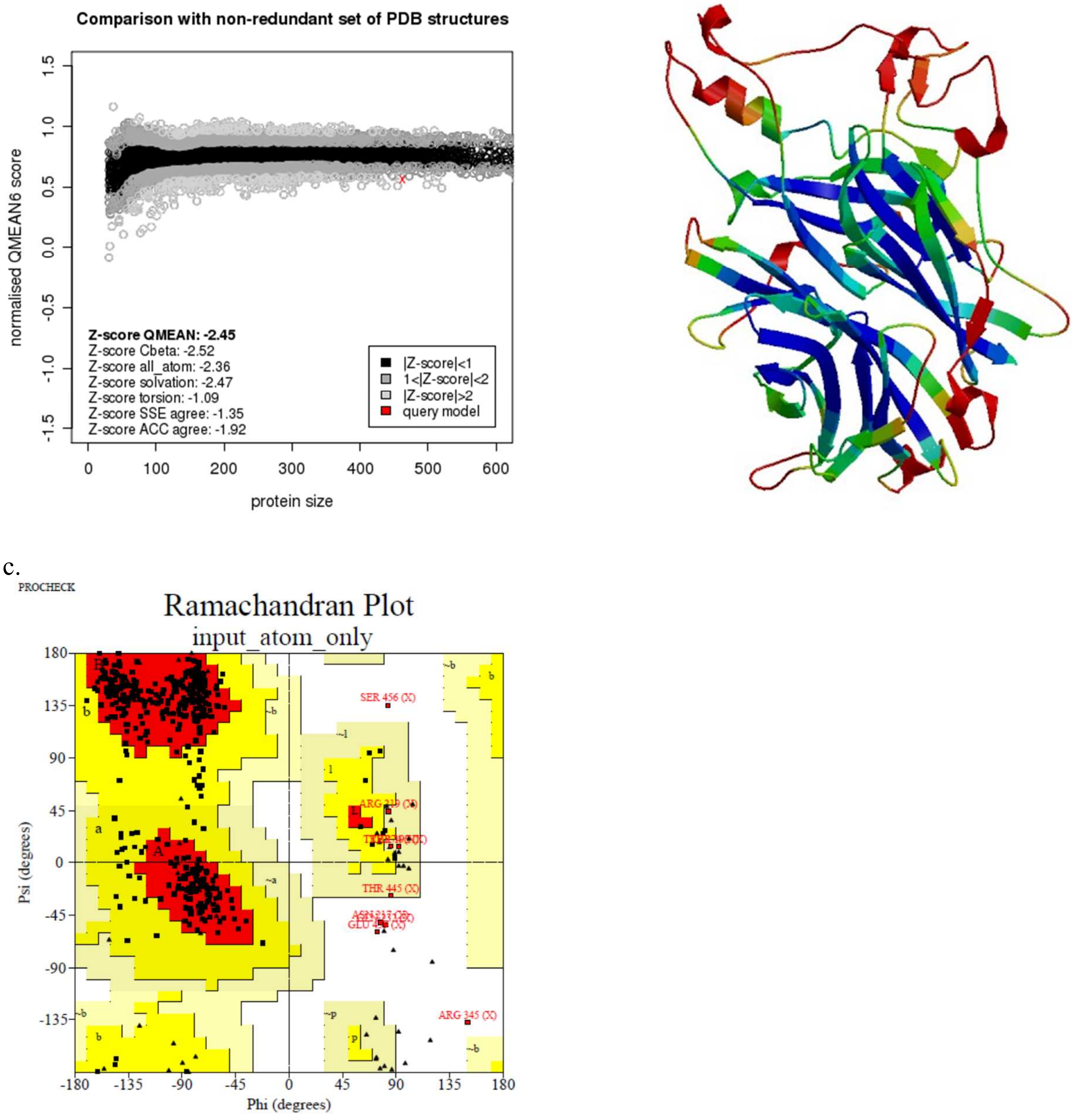
y= 0.0368 x + 2.1772
r^2= 0.9975

Figure SI5: Quality of NEU4 homology model

The NEU4L sequence has an extended sequence (residues 287-374) that was could not be aligned to the NEU2 template. This region formed a disorganized loop in the final model. Evaluation of the model using Swiss-model assessment tools shows that the quality of the model as whole is moderate with a Z-score of -2.451, which is close to extreme values which are found in protein data bank (panel a). In terms of local score, most of the unreliable amino acids (shown in red) are part of the unaligned region, and the C-terminus of the protein (panel b). The Ramachandran plot shows that 97.4 % of the amino acids are found in favoured or allowed conformational regions, while 1.1% is in the generously allowed region, and 1.4% (5 amino acids) are in the disallowed region (panel c).[2]

a.

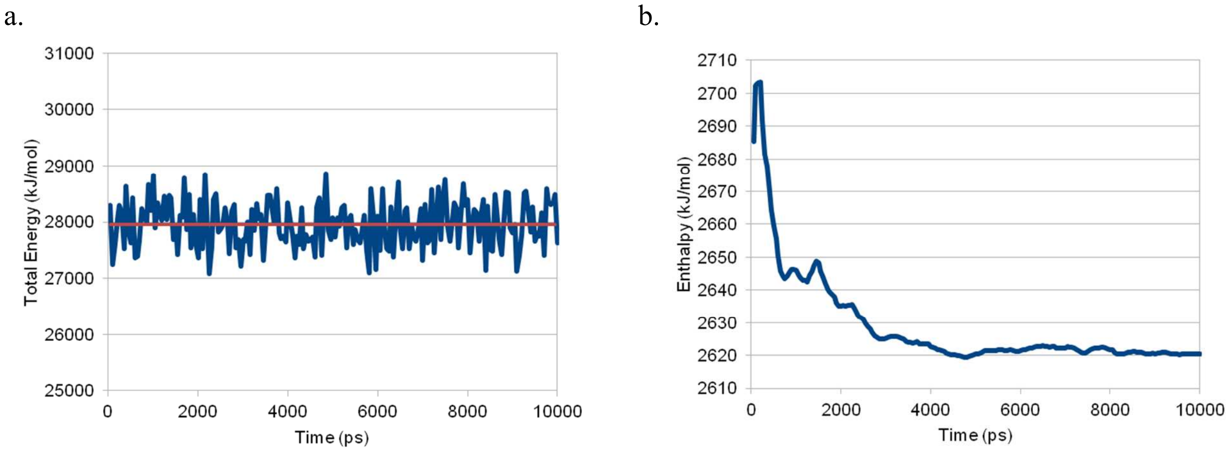
b.



c.

Figure SI6: Convergence and total energy during molecular dynamics calculation

The total energy of the system was conserved during the calculation, indicating the stability of the system (panel a). Molecular dynamics calculations were run to convergence, as illustrated below (panel b).

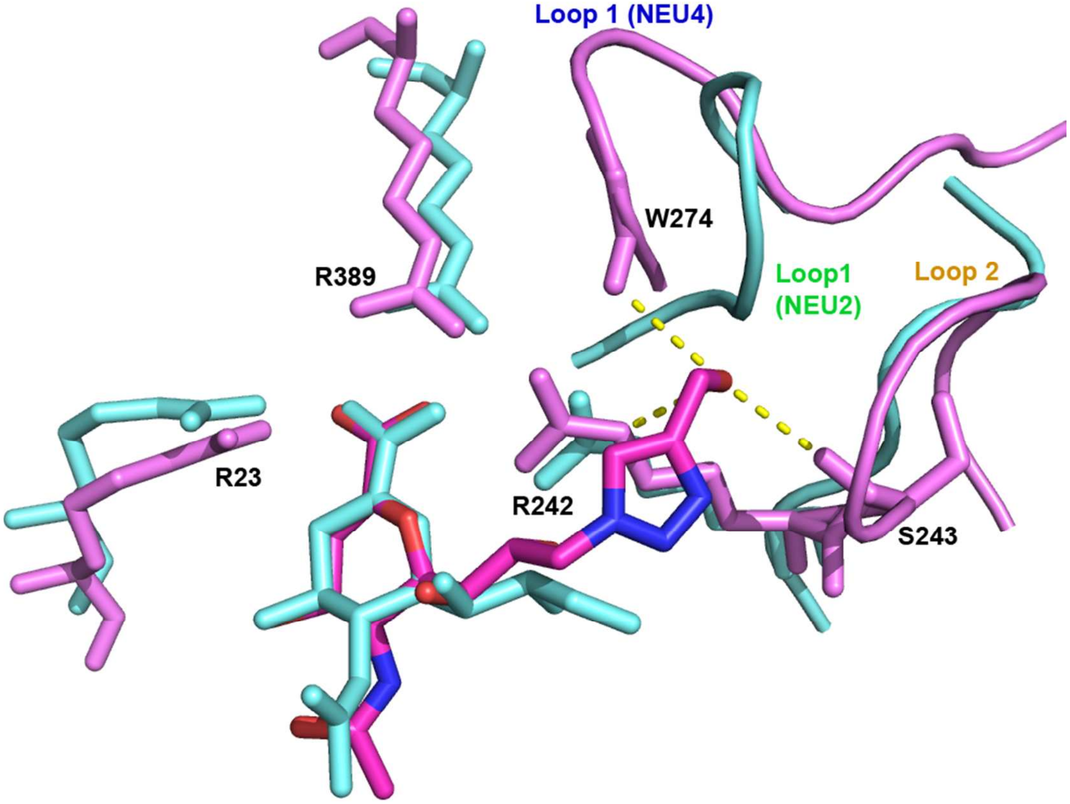


a.

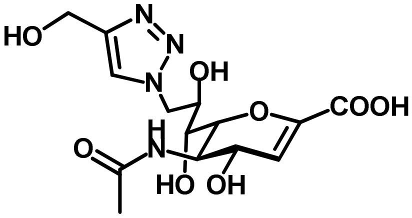
b.

Figure SI7: Selectivity of compound 6 for NEU4 over NEU2

An overlay of our NEU4 homology model with a reported co-crystal of NEU2 illustrates one potential source of selectivity between these two isoenzymes. The proposed contacts between the *C*4'-sidechain and NEU4 are between R242, S243, and W274. Although the first two of these residues have identity in the NEU2 structure, the third is found on a loop (Loop 1, NEU4 R242-R248), which adopts a radically different orientation in the NEU2 structure. Loop 1 of the NEU2 structure obscures the pocket occupied by the *C*4'-sidechain in the NEU4 model.



Characterization of compound 6 (^{1}H NMR and HRMS).



(6) 5-Acetamido-9-[4-hydroxymethyl-[1,2,3]triazol-1-yl]-2,6-anhydro-5,9-dideoxy-D-*glycero*-D-*galacto*-non-2-enonic acid

The triazole, **6**, was synthesized as previously reported.[1] ^{1}H NMR (600 MHz, D_2O) δ 8.10 (s, 1H, $H_{5'}$), 6.038-6.033 (d, J = 3.0 Hz, 1H, H_3), 4.87-4.75 (dd, J = 14.4 Hz, 2.4 Hz, 1H, H_9), 4.74 (s, 2H, CH_2), 4.57-4.53 (dd, J = 14.4 Hz, 4.2 Hz, 1H, H_9), 4.50-4.48 (dd, J = 9.0 Hz, 2.4 Hz, 1H, H_4), 4.31-4.26 (m, 1H, H_8), 4.26-4.24 (dd, J = 10.8 Hz, 0.6 Hz, 1H, H_6), 4.07-4.04 (dd, J = 10.8 Hz, 3.0 Hz, 1H, H_5), 3.50-3.49 (dd, J = 9.6 Hz, 1.2 Hz, 1H, H_7), 2.02 (s, 3H, CH_3). HRMS (ESI) calculated for $C_{14}H_{19}N_4O_8$ $[M-H]^-$, 371.1208, found 371.1208.

References

1. Zou, Y.; Albohy, A.; Sandbhor, M.; Cairo, C. W. Inhibition of the human neuraminidase 3 (NEU3) by *C*9-triazole derivatives of 2,3-didehydro-*N*-acetylneuraminic acid. *Bioorg. Med. Chem. Lett.* **2010,** *20*, 7529-7533.

2. Benkert, P., Schwede, T. and Tosatto, S.C.E. QMEANclust: Estimation of protein model quality by combining a composite scoring function with structural density information. *BMC Struct Biol.* **2009,** 20, 9:35