Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 14 (2006) 982–996
Bioorganic & Medicinal Chemistry

A QSAR study on influenza neuraminidase inhibitors

Rajeshwar P. Verma and Corwin Hansch*

Department of Chemistry, Pomona College, Claremont, CA 91711, USA

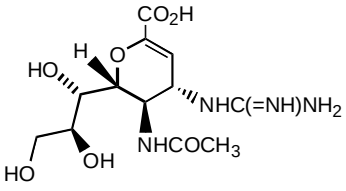
Received 3 August 2005; revised 2 September 2005; accepted 6 September 2005
Available online 6 October 2005

Abstract—Influenza is a major respiratory infection associated with significant morbidity in the general population and mortality in elderly and high-risk patients. It is an RNA virus that contains two major surface glycoproteins, neuraminidase and hemagglutinin. These proteins are essential for infection. Neuraminidase has been found to be a potential target to control influenza virus. In this paper, we have developed 17 quantitative structure–activity relationships (QSAR) for different sets of compounds to understand chemical–biological interactions governing their activities toward influenza neuraminidase.

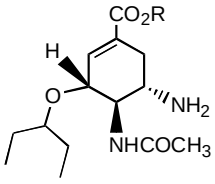
1. Introduction

Understanding the molecular and cell biology of influenza is highly important for suppressing this widely spread disease. Influenza is an RNA virus that contains two major surface glycoproteins, neuraminidase and hemaggutinin. Neuraminidase has been found to be a potential target to control influenza virus. Neuraminidase can cleave the α-ketosidic connections of sialic acid and nearby sugar residues. The removal of sialic acid lowers the viscosity of the virus, thus permitting the entry of the virus into epithelial cells. Neuraminidase also destroys hemagglutinin on the virus surface allowing the emergence of progeny virus units from infected cells. Chemicals that inhibit neuraminidase can protect the host from viral infection.[1]

Recently, two influenza neuraminidase inhibitors, zanamivir (from GlaxoSmithKline and Biota) and oseltamivir (from Hoffman La Roche and Gilead Sciences), have been approved by the FDA for the treatment and prevention of influenza.[2–4] Zanamivir is administered by oral inhalation due to high polar compounds, and oseltamivir is a prodrug that is converted after oral intake to its active form, the carboxylic acid (GS 4071). Zanamivir and oseltamivir are effective inhibitors of both A and B forms of neuraminidase. In the present paper, we have discussed the QSAR studies on benzoic acids, carbocyclic derivatives, cyclopentanes, isoquinolines, pyrrolidines, and miscellaneous compounds for their influence on neuraminidase inhibiting activities.



Zanamivir

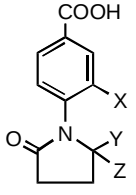


R = Et (Oseltamivir)
R = H (GS 4071)

2. Results and discussion

2.1. Benzoic acids

2.1.1. Inhibition of influenza neuraminidase A by I. Data from Atigadda et al.[5] (Table 1)



I

Keywords: Hydrophobicity; Molar refractivity; Molar volume; Number of valence electrons; Neuraminidase; QSAR.
* Corresponding author. Tel.: +1 909 621 8445; fax: +1 909 607 7726; e-mail: atessier@pomona.edu

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.09.016

Table 1. Biological and physicochemical constants used to derive QSAR Eq. 1 for the inhibition of influenza neuraminidase A by I

	X	Y	Z	log 1/C (Eq. 1)			NVE
				Observed	Predicted	Δ	
1	$NHC(=NH)NH_2$	H	H	3.60	3.08	0.52	100
2	$NHC(=NH)NH_2$	H	CH_2OH	4.70	4.31	0.39	112
3[a]	$NHC(=NH)NH_2$	H	CH_2NH_2	2.59	4.31	−1.72	112
4	H	CH_2OH	CH_2OH	3.12	3.29	−0.17	102
5	$NHC(=NH)NH_2$	CH_2OH	CH_2OH	5.30	5.54	−0.24	124
6	$NHCH(C_2H_5)_2$	H	H	3.65	4.51	−0.86	114
7	$NHCH(C_2H_5)_2$	CH_2OH	CH_2OH	7.32	6.97	0.35	138

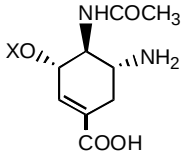
[a] Outlier.

$$\log 1/C = 0.10(\pm 0.05)\mathrm{NVE} - 7.15(\pm 5.90)$$
$$n = 6, \quad r^2 = 0.886, \quad s = 0.585,$$
$$q^2 = 0.713 \quad outlier: \mathrm{X} = \mathrm{NHC}(=\mathrm{NH})\mathrm{NH}_2,$$
$$\mathrm{Y} = \mathrm{H}, \ \mathrm{Z} = \mathrm{CH}_2\mathrm{NH}_2 \qquad (1)$$

With respect to Eq. 1, there is a high mutual correlation between two polarizability parameters, NVE and CMR ($r^2 = 0.930$, $q^2 = 0.880$). We preferred the correlation with NVE term because it gave better statistics than that with CMR.

2.2. Carbocyclic derivatives

2.2.1. Inhibition of influenza neuraminidase A by II. Data from Williams et al.[6] (Table 2)



II

$$\log 1/\mathrm{C} = 8.84(\pm 3.14)\mathrm{MgVol} - 11.29(\pm 6.49)$$
$$n = 6, \quad r^2 = 0.938, \quad s = 0.352,$$
$$q^2 = 0.875 \quad outliers: \mathrm{X} = \text{cyclohexyl; phenyl} \qquad (2)$$

The authors note that while the inhibition of neuraminidase has been shown to be useful to treat influenza infection, the exact role of neuraminidase in the life cycle of the virus remains a subject of discussion. The equation is linear in MgVol. One wonders if a wider range of values of MgVol were studied if an allosteric type QSAR would have been obtained.

Table 2. Biological and physicochemical constants used to derive QSAR Eq. 2 for the inhibition of influenza neuraminidase A by II

	X	log 1/C (Eq. 2)			MgVol
		Observed	Predicted	Δ	
1	$CH(C_2H_5)C_2H_5$	9.00	8.85	0.15	2.28
2	$CH_2CH_2CH_3$	6.89	6.36	0.53	2.00
3	CH_2OCH_3	5.70	5.63	0.07	1.91
4	$CH_2CH_2CF_3$	6.65	6.83	−0.18	2.05
5	$CH_2CH{=}CH_2$	5.66	5.98	−0.32	1.95
6	Cyclopentyl	7.66	7.89	−0.23	2.17
7[a]	Cyclohexyl	7.22	9.14	−1.92	2.31
8[a]	Phenyl	6.28	8.00	−1.72	2.18

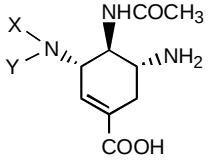
[a] Outlier.

2.2.2. Inhibition of influenza neuraminidase A by II. Data from Kim et al.[7] (Table 3)

$$\log 1/C = 0.63(\pm 0.21)C\log P - 0.74(\pm 0.39)\text{bilin } C\log P + 2.15(\pm 0.27)I + 7.15(\pm 0.40)$$
$$n = 20, \quad r^2 = 0.965, \quad s = 0.263,$$
$$q^2 = 0.945 \quad optimum\ value = 0.384$$
$$outliers: \text{see Table 3} \qquad (3)$$

This equation contains a bilinear term that indicates that instead of a parabola, one has a ∧-shaped curve. This model was developed by Kubinyi.[8] I is for 11 examples where X is branch chained alkyl groups. Its positive coefficient indicates that these compounds were more active, regardless of whether X is a straight-chained alkyl. It should also be noted that all of the compounds contain a COOH group that will be partially ionized.

2.2.3. Inhibition of influenza neuraminidase A by III. Data from Lew et al.[9] (Table 4)



III

$$\log 1/C = 0.72(\pm 0.42)C\pi_X - 1.20(\pm 0.44)I + 6.59(\pm 0.33)$$
$$n = 11, \quad r^2 = 0.944, \quad s = 0.225, \quad q^2 = 0.861$$
$$outliers: \mathrm{X} = \mathrm{Me}, \mathrm{Y} = \mathrm{CH}(C_2H_5)_2; \quad \mathrm{X} = \mathrm{H}, \mathrm{Y} = \mathrm{CH}(C_2H_5)_2 \qquad (4)$$

The authors point out that the hydrophobic side chain ($(-\mathrm{N}\langle^{\mathrm{Y}}_{\mathrm{X}})$) is most important. However, we find that only X is important. This substituent, with one exception, is always the most hydrophobic. $I = 1$ for Y = CO-alkyl groups.

Table 3. Biological and physicochemical constants used to derive QSAR Eq. 3 for the inhibition of influenza neuraminidase A by II

	X	log 1/*C* (Eq. 3)			*C*log *P*	*I*
		Observed	Predicted	*Δ*		
1	H	5.20	5.29	−0.09	−2.96	0
2	Me	5.43	5.30	0.13	−2.93	0
3	C_2H_5	5.70	5.63	0.07	−2.40	0
4[a]	C_3H_7	6.74	4.58	2.17	−1.88	0
5	C_4H_9	6.52	6.27	0.25	−1.35	0
6	C_5H_{11}	6.70	6.54	0.16	−0.82	0
7	C_6H_{13}	6.82	6.71	0.11	−0.29	0
8	C_7H_{15}	6.57	6.78	−0.21	0.24	0
9	C_8H_{17}	6.74	6.77	−0.02	0.77	0
10	C_9H_{17}	6.68	6.72	−0.05	1.30	0
11	$C_{10}H_{21}$	6.22	6.67	−0.45	1.83	0
12[a]	CH_2CHMe_2	6.70	7.28	−0.58	−0.48	1
13	$CH(Me)C_2H_5$ (*R*)	8.00	8.30	−0.30	−1.57	1
14	$CH(Me)C_2H_5$ (*S*)	8.05	8.30	−0.25	−1.57	1
15	$CH(C_2H_5)_2$	9.00	8.59	0.41	−1.04	1
16	$CH(C_2H_5)CH_2CH{=}CH_2$ (*R*)	9.00	8.85	0.15	−0.37	1
17	$CH(C_2H_5)CH_2CH{=}CH_2$ (*S*)	8.52	8.85	−0.32	−0.37	1
18	$CH(C_2H_5)C_7H_{15}$	9.00	8.85	0.15	1.61	1
19[a]	Cy-C_6H_{11}	7.22	8.07	−0.85	−0.90	0
20[a]	$CH(C_2H_5)CH_2$-Cy-C_6H_{11}	7.80	9.18	−1.39	1.08	1
21	$CH(C_2H_5)CH_2CH_2$-Cy-C_6H_{11}	9.00	8.85	0.15	1.61	1
22	C_6H_5	6.28	6.33	−0.05	−1.24	0
23	$CH_2C_6H_5$	6.21	6.41	−0.20	−1.07	0
24[a]	$CH(C_2H_5)CH_2CH_2C_6H_5$ (*R*)	9.52	7.11	2.41	0.38	1
25[a]	$CH(C_2H_5)CH_2CH_2C_6H_5$ (*S*)	7.92	7.11	0.81	0.38	1
26	$CH_2CH_2CH_2$–C_6H_4(4-C_6H_5)	7.05	6.71	0.34	1.43	0

[a] Outlier.

Table 4. Biological and physicochemical constants used to derive QSAR Eqs. 4 and 5 for the inhibition of influenza neuraminidase A and B, respectively by III

	X	Y	log 1/*C* (Eq. 4)			log 1/*C* (Eq. 5)			$C\pi_X$	CMR-Y	*I*
			Observed	Predicted	*Δ*	Observed	Predicted	*Δ*			
1	Me	$(CH_2)_2CH_3$	7.19	6.89	0.29	7.19	7.26	−0.07	0.43	1.39	0
2	Me	$(CH_2)_3CH_3$	6.74	6.89	−0.15	—	—	—	0.43	1.86	0
3[a]	Me	$CH(CH_2CH_3)_2$	8.22	6.89	1.33	7.22	6.85	0.37	0.43	2.32	0
4	Me	$(CH_2)_2$-C_6H_5	7.00	6.89	0.11	6.25	6.36	−0.12	0.43	3.44	0
5	Me	Cy-C_6H_{11}	6.70	6.89	−0.19	—	—	—	0.43	2.61	0
6	CH_2CH_3	$(CH_2)_2CH_3$	7.05	7.28	−0.23	—	—	—	0.96	1.39	0
7	CH_2CH_3	$(CH_2)_3CH_3$	7.07	7.28	−0.20	6.76	7.05	−0.30	0.96	1.86	0
8	$(CH_2)_2CH_3$	$(CH_2)_2CH_3$	7.92	7.66	0.26	7.22	7.26	−0.03	1.48	1.39	0
9[b]	H	$(CH_2)_3CH_3$	6.70	6.59	0.11	6.62	7.26	−0.64	0.00	1.39	0
10[a]	H	$CH(CH_2CH_3)_2$	7.96	6.59	1.37	7.00	6.85	0.15	0.00	2.32	0
11	H	$COCH_2CH_3$	5.57	5.39	0.18	6.30	6.12	0.18	0.00	1.43	1
12	H	$COCHMe_2$	5.19	5.39	−0.19	5.77	5.92	−0.15	0.00	1.89	1
13	H	$COCH(CH_2CH_3)_2$	5.40	5.39	0.01	5.50	5.52	−0.02	0.00	2.82	1

[a] Outliers for Eq. 4.
[b] Outlier for Eq. 5.

2.2.4. Inhibition of influenza neuraminidase B by III. Data from Lew et al.[9] (Table 4)

$$\log 1/C = -0.44(\pm 0.28)\text{CMR-Y} - 1.12(\pm 0.40)I + 7.86(\pm 0.64)$$
$$n = 9, \quad r^2 = 0.906, \quad s = 0.231, \quad q^2 = 0.819$$
$$outlier: \text{X} = \text{H}, \text{Y} = C_4H_9 \qquad (5)$$

$I = 1$ for Y = CO-alkyl groups. The authors found in their study that the hydrophobic substituent ((–N<Y/X>)) promoted activity. It is of interest that we did not find a hydrophobic parameter to be of value in Eq. 5.

2.2.5. Inhibition of influenza neuraminidase by II. Data from Kim et al.[3] (Table 5)

$$\log 1/C = 0.84(\pm 0.22)C\log P - 1.48(\pm 0.46)\text{IS} + 9.34(\pm 0.53)$$
$$n = 9, \quad r^2 = 0.967, \quad s = 0.231, \quad q^2 = 0.946$$
$$outlier: CH(C_2H_5)_2 \qquad (6)$$

IS = 1 for R/S isomers. This represents early work in the search for influenza drugs.

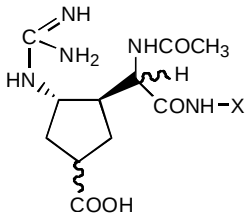
Table 5. Biological and physicochemical constants used to derive QSAR Eq. 6 for the inhibition of influenza neuraminidase by II

	X	log 1/C (Eq. 6)			*C*log *P*	IS
		Observed	Predicted	Δ		
1	H	5.20	5.37	−0.17	−2.96	1
2	Me	5.43	5.39	0.04	−2.93	1
3	C_2H_5	5.70	5.84	−0.14	−2.40	1
4	C_3H_7	6.74	6.28	0.46	−1.88	1
5	C_4H_9	6.52	6.72	−0.20	−1.35	1
6	CH_2CHMe_2	6.70	6.62	0.08	−1.48	1
7	$CH(Me)CH_2CH_3$ (*R*)	8.00	8.02	−0.02	−1.57	0
8	$CH(Me)CH_2CH_3$, (*S*)	8.05	8.02	0.02	−1.57	0
9[a]	$CH(C_2H_5)_2$	9.00	6.98	2.02	−1.04	1
10	$CH(C_3H_7)_2$	7.80	7.87	−0.08	0.02	1

[a] Outlier.

2.3. Cyclopentanes

2.3.1. Inhibition of influenza neuraminidase A by cyclopentane amide derivatives IV. Data from Chand et al.[4] (Table 6)



IV

$$\log 1/C = -42.35(\pm 12.30)\mathrm{MgVol} + 7.46(\pm 2.17)\mathrm{MgVol}^2 + 65.21(\pm 17.20)$$
$$n = 11, \quad r^2 = 0.889, \quad s = 0.216,$$
$$q^2 = 0.815 \quad \textit{inversion point} : 2.84(2.80\text{–}2.88)$$
$$\textit{outliers} : CH(CH_3)C_2H_5; CH(C_2H_5)_2; CH(C_2H_5)CH_2OCH_3 \quad (7)$$

Eq. (7) defines an allosteric reaction. We found our first example of such a reaction in 2001.[10] The word 'allosteric' comes from Greek and means 'another shape.' That is, at the inversion point, the shape of the receptor changes. In the case of MgVol, we have 30 allosteric QSAR. For *C*log *P* we have found 126.[11] MgVol was defined by Abraham and McGowan.[12] It is the sum of the volume of the atoms in a molecule. At present, we have 918 QSAR based on this term.[11] The general formula for allosteric reactions is $\log 1/C = -aX + bX^2 + \text{constant}$. In the present example at first as MgVol increases, activity decreases until at the inversion point (MgVol = 2.84), and then the X^2 term takes over and activity begins to increase. This is taken to mean that a structural change in the receptor occurs at the inversion point. For in-depth knowledge about the use of QSAR in allosteric reaction, the interested reader is to refer to earlier publications.[13–20]

2.3.2. Inhibition of influenza neuraminidase B by cyclopentane amide derivatives IV. Data from Chand et al.[4] (Table 6)

$$\log 1/C = -17.58(\pm 9.42)\mathrm{MgVol} + 2.96(\pm 1.67)\mathrm{MgVol}^2 + 30.80(\pm 13.25)$$
$$n = 10, \quad r^2 = 0.838, \quad s = 0.163,$$
$$q^2 = 0.708 \quad \textit{inversion point} : 2.97(2.88\text{–}3.18)$$
$$\textit{outliers} : C_4H_9, CH(C_2H_5)_2, CH(CH_3)CH_2CH(CH_3)_2; (CH_2)_2C_6H_5 \quad (8)$$

Table 6. Biological and physicochemical constants used to derive QSAR Eqs. 7 and 8 for the inhibition of influenza neuraminidase A and B, respectively by IV

	X	log 1/C (Eq. 7)			log 1/C (Eq. 8)			MgVol
		Observed	Predicted	Δ	Observed	Predicted	Δ	
1	CH_2CH_3	6.48	6.59	−0.11	5.51	5.60	−0.09	2.39
2	$CH(CH_3)_2$	5.87	5.81	0.07	5.34	5.18	0.15	2.53
3	$CH_2CH{=}CH_2$	6.14	6.02	0.13	5.37	5.30	0.07	2.49
4[b]	$(CH_2)_3CH_3$	5.00	5.32	−0.31	4.33	4.88	−0.55	2.68
5[a]	$CH(CH_3)CH_2CH_3$	6.39	5.32	1.07	4.78	4.88	−0.10	2.68
6[a,b]	$CH(CH_2CH_3)_2$	7.10	5.12	1.97	5.52	4.70	0.83	2.82
7	$CH(CH_3)(CH_2)_2CH_3$	5.49	5.12	0.37	4.91	4.70	0.22	2.82
8[b]	$CH(CH_3)CH_2CH(CH_3)_2$	5.08	5.23	−0.15	4.94	4.63	0.31	2.96
9	$CH(CH_3)(CH_2)_3CH_3$	5.06	5.23	−0.17	4.43	4.63	−0.20	2.96
10	$CH(CH_2CH_3)(CH_2)_3CH_3$	5.74	5.63	0.12	4.80	4.68	0.12	3.10
11[b]	$(CH_2)_2C_6H_5$	5.19	5.32	−0.13	4.15	4.63	−0.48	3.00
12	$CH(CH_3)(CH_2)_2C_6H_5$	6.64	6.60	0.04	4.90	4.93	−0.02	3.28
13	$CH(CH_3)CH_2OCH_3$	5.36	5.20	0.16	4.60	4.79	−0.19	2.73
14[a]	$CH(CH_2CH_3)CH_2OCH_3$	5.66	5.13	0.53	4.70	4.65	0.04	2.87

[a] Outliers for Eq. 7.
[b] Outliers for Eq. 8.

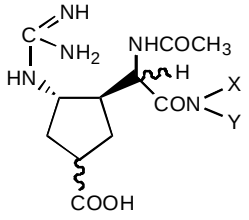
Table 7. Biological and physicochemical constants used to derive QSAR Eq. 9 for the inhibition of influenza neuraminidase A by V

	X	Y	log 1/*C* (Eq. 9)			NVE	$C\pi_X$
			Observed	Predicted	Δ		
1	CH_3	$(CH_2)_2CH_3$	6.03	6.05	−0.02	136	0.27
2[a]	CH_3	$CH(CH_3)_2$	5.49	6.05	−0.55	136	0.27
3	CH_3	$CH_2CH{=}CH_2$	6.19	6.14	0.05	134	0.27
4	CH_3	$(CH_2)_3CH_3$	5.62	5.76	−0.14	142	0.27
5	CH_3	$(CH_2)_5CH_3$	4.91	5.19	−0.28	154	0.27
6	CH_3	$(CH_2)_2C_6H_5$	5.10	5.00	0.10	158	0.27
7[a]	CH_2CH_3	CH_2CH_3	7.82	6.91	0.91	136	0.80
8	CH_2CH_3	$(CH_2)_2CH_3$	6.89	6.63	0.26	142	0.80
9	CH_2CH_3	$(CH_2)_3CH_3$	6.37	6.34	0.03	148	0.80
10	CH_2CH_3	$CH_2C_6H_5$	6.17	5.86	0.31	158	0.80
11[a]	CH_2CH_3	CH_2CH_2OH	6.14	6.63	−0.48	142	0.80
12	$CH_2CH_2CH_3$	$CH_2CH(CH_2)_2$	6.70	7.01	−0.32	152	1.32
13	$CH_2CH_2CH_3$	$CH_2CH_2CH_3$	7.22	7.20	0.02	148	1.32

[a] Outliers.

Again we obtained an allosteric QSAR in terms of MgVol. This suggests that, at first the activity declines as MgVol increases up to the inversion point for MgVol = 2.97 and then increases.

2.3.3. Inhibition of influenza neuraminidase A by cyclopentane amide derivatives V. Data from Chand et al.[4] (Table 7)



V

Table 8. Biological and physicochemical constants used to derive QSAR Eq. 10 for the inhibition of influenza neuraminidase by VI

	X		log 1/*C* (Eq. 10)			π	σ^+
			Observed	Predicted	Δ		
1	4-NO_2		2.90	2.81	0.09	0.24	0.79
2[a]	4-Br		2.77	2.93	−0.16	1.02	0.15
3[a]	4-CN		2.84	2.64	0.20	−0.32	0.66
4	4-Cl		2.81	2.83	−0.02	0.70	0.11
5	4-F		2.63	2.65	−0.02	0.15	−0.07
6	H		2.58	2.62	−0.04	0.00	0.00
7	4-CH_3		2.68	2.71	−0.03	0.52	−0.31
8	4-OCH_3		2.62	2.48	0.14	−0.04	−0.78
9	4-OH		2.24	2.30	−0.06	−0.61	−0.92
10	4-OC_2H_5		2.65	2.62	0.03	0.46	−0.78
11	4-OC_3H_7		2.79	2.76	0.03	0.96	−0.78
12	4-OC_4H_9		2.78	2.90	−0.12	1.46	−0.78
13	4-$C(CH_3)_3$		3.15	3.05	0.10	1.68	−0.26
14	3-CH_3		2.78	2.75	0.03	0.51	−0.07
15	3-F		2.67	2.71	−0.04	0.13	0.35
16	3-Cl		2.82	2.90	−0.08	0.76	0.40

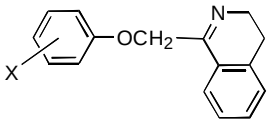
[a] Outliers.

$$\log 1/C = -0.05(\pm 0.02)\mathrm{NVE} + 1.64(\pm 0.44)C\pi_X + 12.09(\pm 3.22)$$
$$n = 10, \quad r^2 = 0.924, \quad s = 0.232, \quad q^2 = 0.843$$
$$outliers: X = CH_3, Y = CH(CH_3)_2; \; X = C_2H_5, Y = C_2H_5; X = C_2H_5, Y = CH_2CH_2OH \quad (9)$$

The NVE parameter is defined as the sum of the number of valence electrons. We have 747 QSAR in our database with this term.[11] Generally they have positive values for NVE. $C\pi_X$ is the calculated hydrophobicity for X-substituents.

2.4. Isoquinolines

2.4.1. Inhibition of influenza neuroaminidase by 1-phenoxymethyl-3,4-dihydroisoquinolines VI. Data from Tute.[21] (Table 8)



VI

$$\log 1/C = 0.28(\pm 0.08)\pi + 0.16(\pm 0.09)\sigma^+ + 2.62(\pm 0.06)$$
$$n = 14, \quad r^2 = 0.868, \quad s = 0.079, \quad q^2 = 0.728$$
$$outliers: \text{4-Br; 4-CN} \quad (10)$$

This is a rather strange equation with a very weak σ^+ term and a hydrophobic term π for the substituents. The spread in log 1/*C* is low 2.24–3.15. This simple equation does show that there can be a hydrophobic region in the virus.

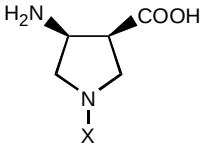
Table 9. Biological and physicochemical constants used to derive QSAR Eq. 11 for the inhibition of influenza neuraminidase A by VII

	X	log 1/*C* (Eq. 11)			MgVol
		Observed	Predicted	*Δ*	
1	$OCH(Et)_2$	4.66	4.62	0.04	1.84
2	$CON(Et)_2$	4.60	4.73	−0.13	1.80
3	$CON(N\text{-}Pr)_2$	4.49	4.44	0.06	2.08
4[a]	$CON(Et)CH(Me)_2$	5.80	4.45	1.35	1.94
5[a]	$CON[CH(Me)_2]_2$	5.40	4.44	0.96	2.08
6	$CON(CH_2CH_2OH)CH(Me)_2$	4.68	4.41	0.27	2.00
7[a]	$CON[(CH_2)_3OH]CH(Me)_2$	5.68	4.51	−1.17	2.14
8	$CON[(CH_2)_5OH]CH(Me)_2$	5.70	5.54	0.16	2.42
9	$CON[(CH_2)_3COOH]CH(Me)_2$	4.72	4.95	−0.23	2.30
10	$CON[(CH_2)_4COOH]CH(Me)_2$	5.89	4.63	0.26	2.44
11	$CON[(CH_2)_3NH_2]CH(CH_3)_2$	4.34	4.59	−0.26	2.18
12	$CON[(CH_2)_2$-2-Pyridyl$]CH(Me)_2$	5.89	6.06	−0.18	2.51

[a] Outliers.

2.5. Pyrrolidines

2.5.1. Inhibition of influenza neuraminidase A by pyrrolidines VII. Data from Wang et al.[22] (Table 9)



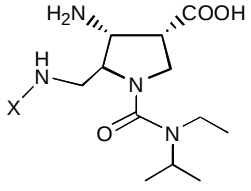
VII

$$\log 1/C = -27.72(\pm 17.70)\mathrm{MgVol} + 6.88(\pm 4.10)\mathrm{MgVol}^2 + 32.33(\pm 18.82)$$
$$n = 9, \quad r^2 = 0.896, \quad s = 0.236, \quad q^2 = 0.768$$
$$\textit{inversion point} : 2.02(1.80\text{–}2.09)$$
$$\textit{outliers} : CON(C_2H_5)CH(CH_3)_2;\ CON[CH(CH_3)_2]_2;\ CON[(CH_2)_3OH]CH(CH_3)_2 \quad (11)$$

It is another example of the propensity of neuraminidase A to undergo an allosteric reaction. Although there are only nine data points and three outliers, the inversion point is very well defined. It is much lower than the other two examples Eqs. 7 and 8 that are based on highly polar molecules. This might explain the low inversion point. The authors were interested in developing antimalarial agents.

It is also interesting that MgVol is involved in the allosteric reactions. This has important implications for drug design where researchers often use intuition to make the next congener in a series. The ∪ curve for allosteric reactions is in contrast to the ∩ curve often seen for thousands of QSAR.

2.5.2. Inhibition of influenza neuraminidase A by pyrrolidines VIII. Data from Wang et al.[22] (Table 10)



VIII

$$\log 1/C = 2.81(\pm 1.11)C\log P + 8.54(\pm 1.53)$$
$$n = 4, \quad r^2 = 0.983, \quad s = 0.194,$$
$$q^2 = 0.888 \quad \textit{outlier} : X = COCH_3 \quad (12)$$

Linear *C*log *P* is the most significant model and suggests that the highly hydrophobic molecules will be more active.

Table 10. Biological and physicochemical constants used to derive QSAR Eq. 12 for the inhibition of influenza neuraminidase A by VIII

	X	log 1/*C* (Eq. 12)			*C*log *P*
		Observed	Predicted	*Δ*	
1[a]	$COCH_3$	5.12	3.37	1.75	−1.84
2	$COCH_2CH_3$	4.80	4.86	−0.07	−1.31
3	$COCH\text{-}CH_2$	4.02	4.20	−0.18	−1.54
4	$COCF_3$	6.55	6.49	0.06	−0.73
5	SO_2CH_3	3.89	3.70	0.19	−1.72

[a] Outlier.

Table 11. Biological and physicochemical constants used to derive QSAR Eq. 13 for the inhibition of influenza neuraminidase by miscellaneous compounds

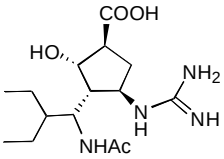
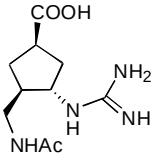
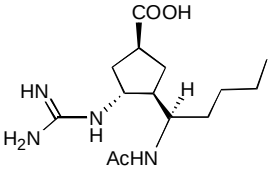
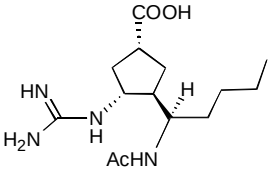
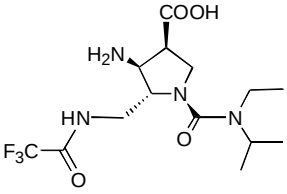
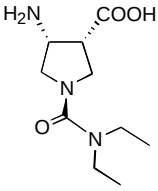
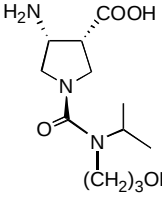
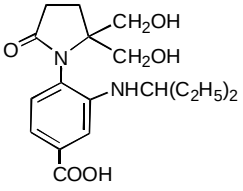
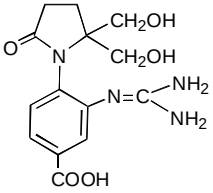
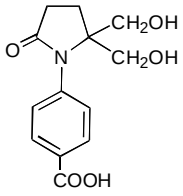
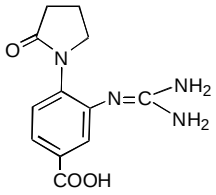
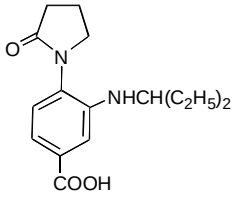
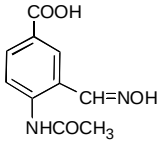
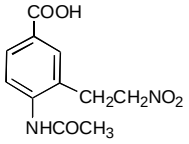
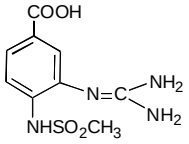
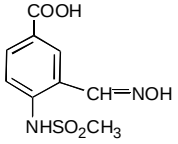
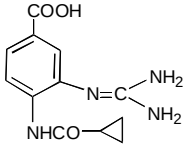
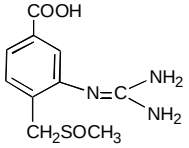
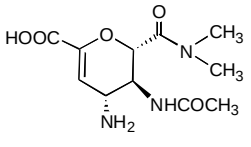
	Compound	log 1/*C* (Eq. 13)			*C*log *P*	MgVol
		Observed	Predicted	*Δ*		
1		10.08	8.13	1.95	−1.31	2.62
2		3.97	5.18	−1.21	−3.89	1.86
3		6.44	7.37	−0.93	−2.00	2.42
4		8.89	7.37	1.52	−2.00	2.42
5		6.13	7.39	−1.26	−0.65	2.53
6		4.36	3.88	0.48	−1.52	1.80
7		5.31	5.85	−0.54	−1.90	2.14
8		7.38	6.90	0.48	2.70	2.71
9		5.34	5.91	−0.57	−0.06	2.30

Table 11 (*continued*)

	Compound	log 1/*C* (Eq. 13)			*C*log *P*	MgVol
		Observed	Predicted	*Δ*		
10		3.14	3.38	−0.24	0.96	1.90
11		3.63	3.42	0.21	0.83	1.90
12		3.68	4.41	−0.73	3.59	2.31
13		2.33	1.73	0.60	0.89	1.59
14		2.68	3.17	−0.49	0.0	1.79
15		4.12	3.62	0.50	−0.26	1.85
16		2.78	2.51	0.27	0.61	1.71
17		3.09	3.54	−0.45	0.56	1.90
18		3.97	3.80	0.17	−0.89	1.83
19		5.89	6.01	−0.12	−4.40	1.97

(*continued on next page*)

Table 11 (*continued*)

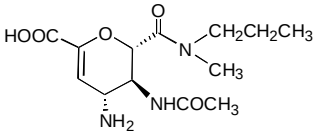
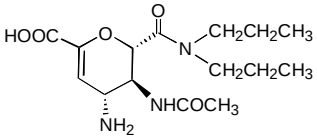
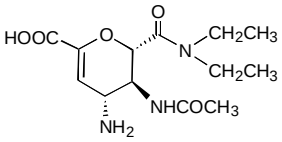
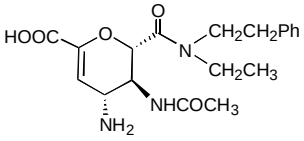
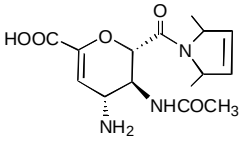
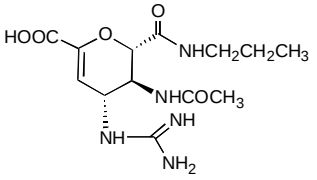
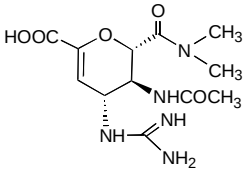
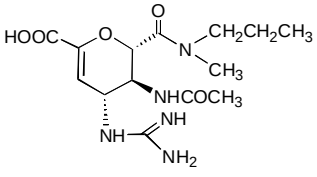
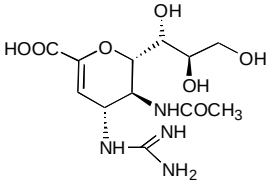
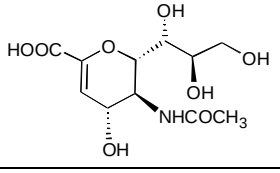
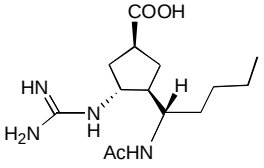
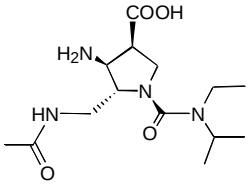
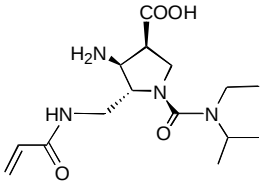
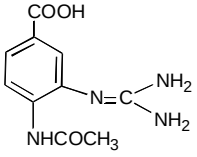
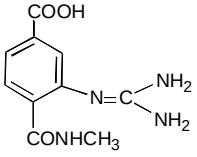
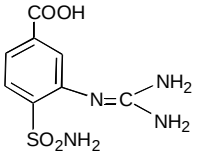
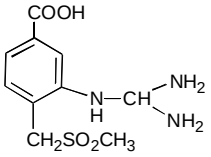
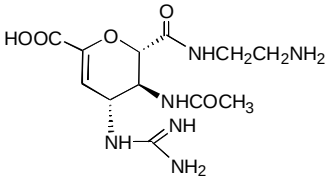
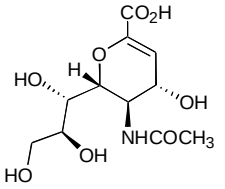
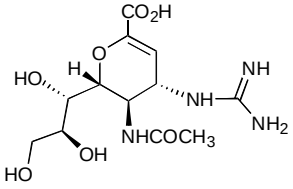
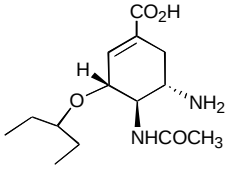
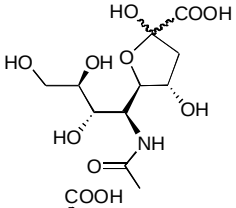
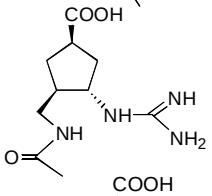
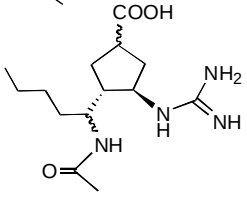
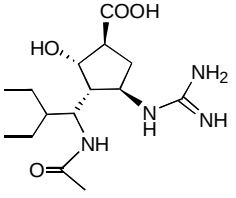
	Compound	log 1/*C* (Eq. 13)			*C*log *P*	MgVol
		Observed	Predicted	*Δ*		
20		7.06	7.06	0.00	−3.34	2.25
21		8.30	8.10	0.20	−2.28	2.53
22		8.93	7.06	1.87	−3.34	2.25
23		8.70	9.61	−0.91	−1.77	2.86
24		8.70	7.70	1.00	−3.24	2.38
25		6.60	8.36	−1.76	−4.45	2.41
26		7.97	7.85	0.12	−5.02	2.27
27		8.80	8.89	−0.09	−3.96	2.55
28		8.70	8.40	0.30	−5.56	2.33
29		5.04	5.41	−0.37	−2.74	1.99

Table 12. Biological and physicochemical constants used to derive QSAR Eq. 14 for the inhibition of influenza neuraminidase by miscellaneous compounds

	Compound	log 1/*C* (Eq. 14)			*C*log *P*
		Observed	Predicted	*Δ*	
1		4.83	4.76	0.07	−2.00
2		4.79	4.85	−0.06	−1.76
3[a]		3.76	5.01	−1.25	−1.37
4		5.77	5.57	0.20	−0.03
5		5.46	5.54	−0.08	−0.10
6		5.19	5.34	−0.15	−0.58
7		3.63	3.62	0.01	−4.76
8[a]		4.93	3.21	1.72	−5.76

[a] Outliers.

Table 13. Biological and physicochemical constants used to derive QSAR Eq. 15 for the inhibition of influenza neuraminidase A by miscellaneous compounds

	Compound	log 1/*C* (Eq. 15)			CMR
		Observed	Predicted	*Δ*	
1		5.00	4.56	0.44	6.51
2[a]		9.52	7.33	2.19	7.75
3[a]		11.00	7.12	3.88	7.66
4		4.40	4.79	−0.39	6.61
5		3.94	3.77	0.17	6.16
6		7.00	7.92	−0.92	8.02
7		10.00	9.30	0.70	8.63

[a] Outliers.

2.6. Miscellaneous compounds

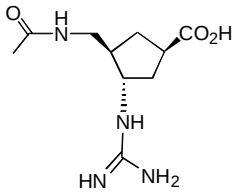
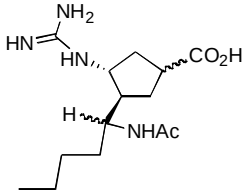
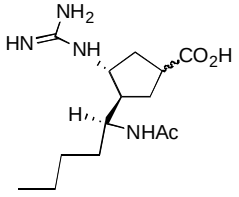
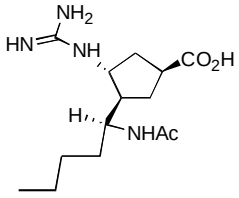
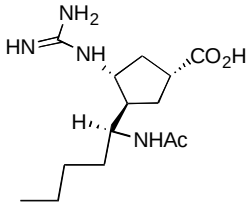
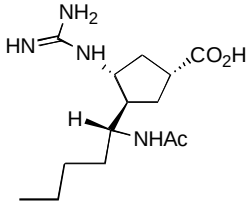
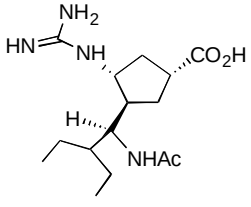
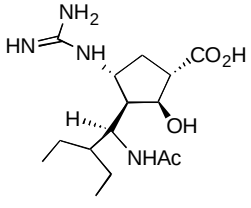
2.6.1. Inhibition of influenza neuraminidase by miscellaneous compounds. Data from Yi et al.[23] (Tables 11 and 12)

$$\log 1/C = -0.42(\pm 0.16)C\log P + 5.30(\pm 1.08)\mathrm{MgVol} - 6.29(\pm 2.36)$$
$$n = 29, \quad r^2 = 0.860, \quad s = 0.912, \quad q^2 = 0.829 \tag{13}$$

$$\log 1/C = 0.41(\pm 0.10)C\log P + 5.58(\pm 0.22)$$
$$n = 6, \quad r^2 = 0.972, \quad s = 0.141, \quad q^2 = 0.945$$
$$\textit{outliers}: 3 \text{ and } 8 \text{ in Table } 12 \tag{14}$$

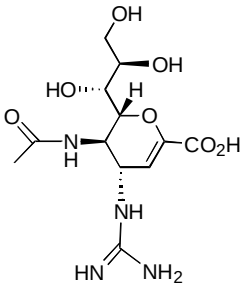
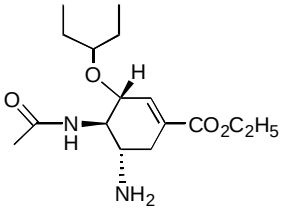
From the 37 data points of this data set, we were unable to derive a good QSAR due to the presence of eight outliers. The presence of a large number of outliers may be due to their interaction with influenza neuraminidase in

Table 14. Biological and physicochemical constants used to derive QSAR Eqs. 16 and 17 for the inhibition of influenza neuraminidase A and B, respectively by miscellaneous compounds

	Compound	log 1/*C* (Eq. 16)			log 1/*C* (Eq. 17)			NVE
		Observed	Predicted	*Δ*	Observed	Predicted	*Δ*	
1		4.56	4.60	−0.04	4.57	4.56	0.02	96
2[b]		7.30	7.55	−0.25	6.05	7.19	−1.14	120
3		7.60	7.55	0.05	7.37	7.19	0.18	120
4[a,b]		6.44	7.55	−1.11	5.28	7.19	−1.91	120
5		7.89	7.55	0.33	6.39	7.19	−0.80	120
6[a,b]		4.84	7.55	−2.72	3.52	7.19	−3.66	120
7		—	—	—	8.70	7.84	0.86	126
8		—	—	—	8.30	8.50	−0.20	132

(*continued on next page*)

Table 14 (*continued*)

	Compound	log 1/*C* (Eq. 16)			log 1/*C* (Eq. 17)			NVE
		Observed	Predicted	Δ	Observed	Predicted	Δ	
9		8.70	8.79	−0.09	7.77	8.28	−0.51	130
10		—	—	—	8.30	7.84	0.46	126

[a] Outliers for Eq. 16.
[b] Outliers for Eq. 17.

a different mode. Thus, this data set was divided into two subsets on the basis of outliers and derived two Eqs. 13 and 14 with good statistics.[18–20] The authors explored docking techniques for this data set in an attempt to understand the SAR.

2.6.2. Inhibition of influenza neuraminidase A by miscellaneous compounds. Data from Babu et al.[24] (Table 13)

$$\log 1/C = 2.24(\pm 1.12)\mathrm{CMR} - 10.00(\pm 8.1)$$
$$n = 5, \quad r^2 = 0.931, \quad s = 0.755,$$
$$q^2 = 0.724 \quad \textit{outliers} : 2 \text{ and } 3 \text{ in Table 13} \quad (15)$$

CMR is the most important parameter for this equation. No correlation with a hydrophobic factor was found. The positive sign of the coefficient associated with the CMR term indicates that an increase in overall molar refractivity should result in stronger inhibition of influenza neuraminidase A.

2.6.3. Inhibition of influenza neuraminidase A by miscellaneous compounds. Data from Chand et al.[25] (Table 14)

$$\log 1/C = 0.12(\pm 0.03)\mathrm{NVE} - 7.23(\pm 3.69)$$
$$n = 5, \quad r^2 = 0.981, \quad s = 0.249, \quad q^2 = 0.953$$
$$\textit{outliers} : 4 \text{ and } 6 \text{ in Table 14} \quad (16)$$

2.6.4. Inhibition of influenza neuraminidase B by miscellaneous chemicals. Data from Chand et al.[25] (Table 14)

$$\log 1/C = 0.11(\pm 0.05)\mathrm{NVE} - 5.96(\pm 6.55)$$
$$n = 7, \quad r^2 = 0.846, \quad s = 0.619, \quad q^2 = 0.774$$
$$\textit{outliers} : 2, 4, \text{ and } 6 \text{ in Table 14} \quad (17)$$

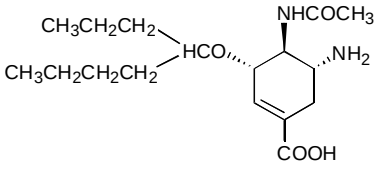
Eqs. 16 and 17 show a good correlation with polarizability parameter (NVE).

3. Summary

The above QSAR give an overview of the studies of neuraminidase starting with the work of Tute in 1970. However, most of the studies from which it was possible to derive a QSAR were published after 1995. These studies illustrate how QSAR can be used to control a virus. In a previous report,[26] we demonstrated how QSAR could be employed to elucidate the action of HIV virus.

An analysis of our QSAR reveals a number of points of interest. The most important of these is hydrophobicity. Out of 17 QSAR, eight contain a correlation between activity and hydrophobicity. A positive linear correlation is found in six Eqs. 4, 6, 9, 10, 12, and 14. The coefficient with the hydrophobic parameter varies considerably, from a low value of 0.28 (Eq. 10) to a high value of 2.81 (Eq. 12). These data suggest that although activity might be improved by increasing compound substituent hydrophobicity, this will produce considerably different results for different compounds. Bilinear correlation with hydrophobicity is found in one Eq. 3, which shows that activity is optimal for a particular value of $\log P(0.384)$. The other important parameter is MgVol. Parabolic correlations with MgVol are found in three Eqs. 7, 8, and 11, where activity declines with increasing MgVol and then changes direction and increases. These may correspond to allosteric reactions. Other parameters, electronic, molar refractivity, and polarizability, also appear in several QSAR but, they do not seem to play as important a role as hydrophobicity and molar volume for the data sets that we have examined.

QSAR (3) is an encouraging example, where the optimal hydrophobicity is well defined that is $\log P = 0.384$. We believe that this will be the predictive model to narrow the synthetic challenges in order to yield very specific influenza neuraminidase A inhibitors. On the basis of this model, we predict one compound (1X) that may be the next synthetic target.



Clog P = 0.35

log 1/C = 9.46
(predicted from equation 3)

IX

4. Materials and methods

All the data have been collected from the literature (see individual QSAR for respective references). C is the molar concentration of a compound and $\log 1/C$ is the dependent variable that defines the biological parameter for QSAR equations. Physicochemical descriptors are auto-loaded, and multiregression analyses (MRA) used to derive the QSAR were executed with the C-QSAR program.[11] The parameters used in this report have already been discussed.[27] Briefly, $C\log P$ is the calculated octanol/water partition coefficient. $C\pi$ and π are the calculated and experimentally obtained hydrophobic parameters for the substituents. MR is the molar refractivity and is defined by the Lorentz–Lorenz equation: $MR = n^2 - 1/n^2 + 2(MW/\delta)$, where n = refractive index, δ = density, and MW = molecular weight. CMR is the calculated molar refractivity. Molar refractivity (MR) has a strong correlation with the molecular polarizability. Thus, it has also been used as a measure of polarizability. MR can be used for a substituent or for the whole molecule. Number of valence electrons (NVE) is a parameter[28–30] that was found to be another approach to understanding polarizability and calculated by simply summing up the valence electrons in a molecule, for example, H = 1, C = 4, Si = 4, N = 5, P = 5, O = 6, S = 6, and halogens = 7. There are three commonly encountered electronic parameters: σ, σ^-, and σ^+ that account for specific electronic effects of substituents on aromatic systems. These parameters are known as the Hammett parameters and their application has been illustrated.[27] MgVol is the molar volume calculated by C-QSAR Program using the method of Abraham and McGowan.[12] I is an indicator variable that takes the value of 1 or 0 for structural features that cannot be defined by the normal parameters. Each regression equation includes 95% confidence limits for each term in parentheses.

In QSAR equation, n is the number of data points, r is the correlation coefficient, r^2 is the goodness of fit, q^2 is the goodness of prediction, and s is the standard deviation. All the QSAR reported here are derived by us and were not given with the original data sets taken from the literature as referenced.

References and notes

1. Palese, P.; Tobita, K.; Ueda, M. *Virology* **1974**, *61*, 397.
2. Von Itzstein, M.; Wu, W.-Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Phan, T. V.; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, P. M.; Varghese, J. M.; Ryan, D. M.; Woods, J. M.; Bethell, R. C.; Hotham, V. J.; Cameron, J. M.; Penn, C. R. *Nature* **1993**, *363*, 418.
3. Kim, C. U.; Lew, W.; Williams, M. A.; Liu, H.; Zhang, L.; Swaminathan, S.; Bischofberger, N.; Chen, M. S.; Mendel, D. B.; Tai, C. Y.; Laver, W. G.; Stevens, R. C. *J. Am. Chem. Soc.* **1997**, *119*, 681.
4. Chand, P.; Babu, Y. S.; Bantia, S.; Rowland, S.; Dehghani, A.; Kotian, P. L.; Hutchison, T. L.; Ali, S.; Brouillette, W.; El-Kattan, Y.; Lin, T.-H. *J. Med. Chem.* **2004**, *47*, 1919.
5. Atigadda, V. R.; Brouillette, W. J.; Duarte, F.; Ali, S. M.; Babu, Y. S.; Bantia, S.; Chand, P.; Chu, N.; Montgomery, J. A.; Walsh, D. A.; Sudbeck, E. A.; Finley, J.; Luo, M.; Air, G. M.; Laver, G. M. *J. Med. Chem.* **1999**, *42*, 2332.
6. Williams, M. A.; Lew, W.; Mendel, D. B.; Tai, C. Y.; Escarpe, P. A.; Laver, W. G.; Stevens, R. C.; Kim, C. U. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1837.
7. Kim, C. U.; Lew, W.; Williams, M. S.; Wu, H.; Zhang, L.; Chen, X.; Escarpe, P. A.; Mendel, D. B.; Laver, W. G.; Stevens, R. C. *J. Med. Chem.* **1998**, *41*, 2451.
8. Kubinyi, H. *Arzmein-Forsch.* **1976**, *26*, 1991.
9. Lew, W.; Wu, H.; Mendel, D. B.; Escarpe, P. A.; Chen, X.; Laver, G.; Graves, B. J.; Kim, C. U. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 3321.
10. Hansch, C.; Garg, R.; Kurup, A. *Bioorg. Med. Chem.* **2001**, *9*, 283.
11. C-QSAR Program, BioByte Corp., 201 W. 4th Street, Suite 204, Claremont, CA 91711, USA. www.biobyte.com.
12. Abraham, M. H.; McGowan, J. C. *Chromatographia* **1987**, *23*, 243.
13. Garg, R.; Kurup, A.; Hansch, C. *Bioorg. Med. Chem.* **2001**, *9*, 3161.
14. Hansch, C.; Hoekman, D.; Leo, A.; Weininger, D.; Selassie, C. D. *Chem. Rev.* **2002**, *102*, 783.
15. Hansch, C.; Garg, R.; Kurup, A.; Mekapati, S. B. *Bioorg. Med. Chem.* **2003**, *11*, 2075.
16. Mekapati, S. B.; Kurup, A.; Verma, R. P.; Hansch, C. *Bioorg. Med. Chem.* **2005**, *13*, 3737.
17. Verma, R. P. *Lett. Drug Des. Discov.* **2005**, *2*, 205.
18. Verma, R. P. *Bioorg. Med. Chem.* **2005**, *13*, 1059.
19. Hansch, C.; Verma, R. P.; Kurup, A.; Mekapati, S. B. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 2149.
20. Verma, R. P.; Hansch, C. *Bioorg. Med. Chem.* **2005**, *13*, 4597.
21. Tute, M. S. *J. Med. Chem.* **1970**, *13*, 48.
22. Wang, G. T.; Chen, Y.; Wang, S.; Gentles, R.; Sowin, T.; Kati, W.; Muchmore, S.; Giranda, V.; Stewart, K.; Sham, H.; Kempf, D.; Laver, W. G. *J. Med. Chem.* **2001**, *44*, 1192.
23. Yi, X.; Guo, Z.; Chu, F. M. *Bioorg. Med. Chem.* **2003**, *11*, 1465.
24. Babu, Y. S.; Chand, P.; Bantia, S.; Kotian, P.; Dehghani, A.; El-Kattan, Y.; Lin, T.-H.; Hutchison, T. L.; Elliott, A. J.; Parker, C. D.; Ananth, S. L.; Horn, L. L.; Laver, G. W.; Montgomery, J. A. *J. Med. Chem.* **2000**, *43*, 3482.
25. Chand, P.; Kotian, P. L.; Dehghani, A.; El-Kattan, Y.; Lin, T.-H.; Hutchison, T. L.; Babu, Y. S.; Bantia, S.; Elliott, A. J.; Montgomery, J. A. *J. Med. Chem.* **2001**, *44*, 4379.
26. Garg, R.; Gupta, S. P.; Gao, H.; Babu, M. S.; Debnath, A. K.; Hansch, C. *Chem. Rev.* **1999**, *99*, 3525.
27. Hansch, C.; Leo, A. *Exploring QSAR Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, DC, 1995.

28. Hansch, C.; Steinmetz, W. E.; Leo, A. J.; Mekapati, S. B.; Kurup, A.; Hoekman, D. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 120.
29. Verma, R. P.; Kurup, A.; Hansch, C. *Bioorg. Med. Chem.* **2005**, *13*, 237.
30. Verma, R. P.; Hansch, C. *Bioorg. Med. Chem.* **2005**, *13*, 2355.