Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 12 (2004) 5819–5831
Bioorganic & Medicinal Chemistry

A mechanistic study of 3-aminoindazole cyclic urea HIV-1 protease inhibitors using comparative QSAR

Rajni Garg* and Barun Bhhatarai

Chemistry Department, Clarkson University, Potsdam, NY 13699-5810, USA

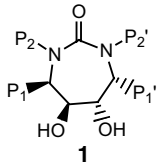
Received 18 June 2004; accepted 26 August 2004
Available online 28 September 2004

Abstract—Comparative QSAR studies on P2/P2′ and P1/P1′ substituted symmetrical and nonsymmetrical 3-aminoindazole cyclic urea HIV-1 protease inhibitors were performed. The protease inhibitory activity of these compounds was found to decrease with larger and more hydrophobic molecules, whereas the antiviral potency and translation across the cell membrane increases with increase in hydrophobicity and size. These results provide mechanistic insight about the mode of interaction of these compounds with HIV-1 protease receptor and would help in further improving the biological activity.

1. Introduction

Acquired immunodeficiency syndrome (AIDS) and its related disorders are major health concerns worldwide, and are caused by retrovirus human immunodeficiency virus type 1 (HIV-1).[1,2] One of the key enzymes encoded by HIV is an homodimeric aspartyl protease that cleaves the *gag* and *gag-pol* viral polyproteins.[3] This processing is essential for the maturation of viral particles and production of infectious virions.[4,5] Inhibition of HIV-1 protease (HIVPR) results in the production of immature and noninfectious virions.[6–9] Currently, there are many HIVPR inhibitor drugs in the market, approved by the US Food and Drug Administration (FDA).[10] These drugs have shown promising results when used in combination with reverse transcriptase (RT) inhibitors.[11,12] However, the use of current drugs regimens is compounded by several issues such as—adherence, tolerability, long-term toxicity, and drug- and cross-resistance.[11,12] The development of new protease drugs that are less toxic, tolerable, convenient, and are active against drug-resistant viruses is highly desirable.

All the HIVPR drugs approved, so far, are peptidic in nature, and their prolonged use is resulting in the emergence of mutant virus, therefore, several authors have paid attention to nonpeptidic inhibitors in the hope of getting new better drugs.[13–17] Compounds possessing seven-member cyclic urea scaffold (**1**) were found to be potent inhibitors of the enzyme.[13,14] These compounds create an effective hydrogen bond network between the aspartic residues and the flap region of the enzyme without the intervention of a water molecule commonly found in linear inhibitor.[13] To identify the full potential of the various classes of HIVPR inhibitors, numerous quantitative structure–activity relationship (QSAR) studies have been performed.[18–25] QSAR studies help in understanding structural features of both the inhibitor and the target receptor responsible for biological activity and guide in the design of more effective inhibitors.

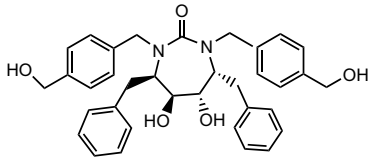


1

Two of the nonpeptidic cyclic urea derivatives (**2**, **3**) reached advance stage of clinical trial. However, DMP 323 (**2**) has been discontinued due to poor bioavailability and metabolic instability. DMP 450 (**3**) showed good aqueous solubility and oral bioavailability in humans, but only moderate potency and resistance profile as compared to other protease inhibitors.[26,27]

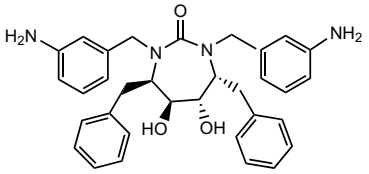
Keywords: HIV-1-protease; Comparative QSAR; 3-Aminoindazole cyclic urea; Hydrophobicity.
* Corresponding author. Tel.: +1 315 268 4421; fax: +1 315 268 6610; e-mail: rgarg@clarkson.edu

0968-0896/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2004.08.036



2

DMP 323
K_i = 0.31 nM
IC_{90} = 100 nM
ClogP = 4.66



3

DMP 450
K_i = 0.30 nM
IC_{90} = 140 nM
ClogP = 4.20

To investigate further other structurally diverse class of nonpeptidic cyclic urea inhibitors, some authors[27–31] synthesized 3-aminoindazole cyclic urea (**4**–**7**), and reported their enzyme inhibition, antiviral, and translation activity. The current study reports the comparative QSAR analysis of 3-aminoindazole cyclic urea HIV-1 protease inhibitors. It is envisioned that these studies will reveal a deeper insight into the mechanism of interaction of these compounds with the receptor and provide guidelines for designing better analogs with superior pharmacokinetic and efficacy profiles.

QSAR studies were performed by us on the compounds listed in Tables 1–5, where the biological activity IC_{90} is a measure of antiviral potency, K_i is the enzyme inhibition constant, and T (K_i/IC_{90}) defines translation of enzyme activity to antiviral activity accounting for penetration across cell membrane. All the QSAR reported here are derived by us and were not given with the original data sets taken from the literature as referenced.

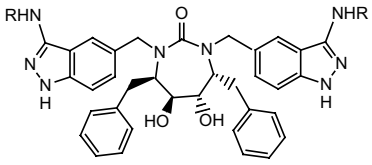
2. Results and discussion

In all the QSAR reported here, n is the number of data points, r is the correlation coefficient, s is the standard deviation, q is the quality of fit and calculated as described by Cramer et al.[32] and the data within parentheses are for the 95% confidence intervals. The QSAR multiple linear regression analyses were executed with the CQSAR program.[33] C log P is the calculated partition coefficient in octanol/water and is a measure of hydrophobicity.[34] CMR is the calculated molar refractivity for the whole molecule, and is a measure of volume and polarizability. C log P and CMR are normally for the neutral form of partially ionized compounds. L is the Verloop's sterimol parameter that defines substituents' length.[35] More details are given in Section 5. Outliers, if any, in each QSAR indicate the molecules of the data set that were not fit in the QSAR model. A discussion about the presence of outliers in QSAR is given later.

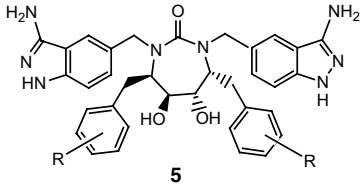
2.1. Symmetrical P2/P2′ substituted 3-aminoindazole cyclic urea derivatives (4)

Rodgers et al.[28] reported the structure–activity data for potent cyclic urea HIV-1 protease inhibitors with alkyl substituted P2/P2′ 3-aminoindazole (**4**). QSAR 1 and 2 were developed for the enzyme inhibition (K_i) data of **4** (Table 1, panel a).

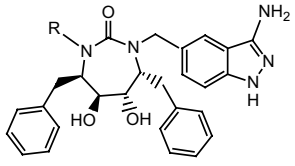
$$\begin{aligned} &\mathrm{Log}\,1/K_i = -0.39(\pm 0.22)\mathrm{C}\log P + 13.33(\pm 1.79)\\ &n = 6, \quad r^2 = 0.858, \quad q^2 = 0.475, \quad s = 0.256, \\ &\text{outlier: } -(CH_2)_4- \end{aligned} \tag{1}$$



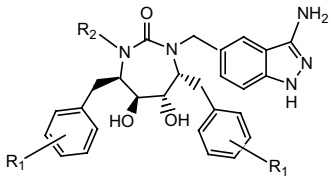
4



5



6



7

Table 1. Symmetrical P2/P2′ substituted 3-aminoindazole cyclic urea derivatives, their physico-chemical parameters and HIV-protease inhibition activity studied by Rodgers et al.[28] (panel a, **4**) and their HIV-antiviral potency and translation data of compounds of panel a (panel b, **4**[28])

S. No.	R	C log *P*	CMR	L_4	I_1	Log(1/K_i)			
						Obsd.[28]	Calcd		
							QSAR 1	QSAR 2	QSAR 25
Panel (a)									
1	H	5.45	17.85	2.06	0.0	11.00	11.18	11.18	10.97
2	CH_3	7.04	18.78	2.06	0.0	10.74	10.56	10.68	10.43
3	C_2H_5	8.10	19.70	2.06	0.0	10.39	10.14	10.19	10.07
4	$CH(CH_3)_2$	8.71	20.63	2.06	0.0	10.05	9.90	9.69	9.86
5	CH_2-cy-C_3H_5	8.98	21.28	2.06	0.0	9.47	9.79	9.34	9.77
6	C_3H_7	9.15	20.63	2.06	0.0	9.64	9.72	9.69	9.71
7[a]	$(CH_2)_4$-	8.86	21.20	2.06	0.0	8.89	9.84	9.39	9.81

	Log(1/IC_{90})				Log(1/*T*)			
	Obsd.[28]	Calcd			Obsd.[28]	Calcd		
		QSAR 3	QSAR 4	QSAR 26		QSAR 5	QSAR 6	QSAR 27
Panel (b)								
1	5.56	5.68	5.86	5.70	−5.40	−5.52	−5.20	−5.32
2	6.55	6.46	6.38	6.42	−4.20	−4.02	−4.29	−4.09
3	7.19	6.98	6.90	6.90	−3.20	−3.02	−3.37	−3.27
4	7.62	7.28	7.41	7.19	−2.40	−2.44	−2.46	−2.79
5	7.42	7.41	7.78	7.31	−2.10	−2.18	−1.82	−2.58
6	7.40	7.50	7.41	7.39	−2.20	−2.02	−2.46	−2.44
7[b]	6.92	7.35	7.73	7.25	−2.00	−2.30	−1.90	−2.67

[a] Not included in deriving QSAR 1 and 25.
[b] Not included in deriving QSAR 4 and 26.

$$\text{Log}\,1/K_i = -0.53(\pm 0.25)\text{CMR} + 20.72(\pm 4.97)$$
$$n = 7, \quad r^2 = 0.860, \quad q^2 = 0.708, \quad s = 0.306 \tag{2}$$

The negative coefficient of C log *P* in QSAR 1 indicates that more hydrophobic substituents have negative influence on the enzyme inhibition. A high negative correlation with CMR was also observed (QSAR 2), that shows that larger P2/P2′ alkyl groups may be involved in unfavorable steric interactions with the receptor.

The same authors[28] also studied the antiviral activity (IC_{90}) of analogs of **4** (Table 1, panel b), the analyses of the data gave QSAR 3 and 4. From these models, it seems that more hydrophobic and larger molecules are required for good antiviral activity.

$$\text{Log}\,1/\text{IC}_{90} = 0.49(\pm 0.21)\text{C log}\,P + 3.01(\pm 1.74)$$
$$n = 7, \quad r^2 = 0.874, \quad q^2 = 0.733, \quad s = 0.276 \tag{3}$$

$$\text{Log}\,1/\text{IC}_{90} = 0.56(\pm 0.29)\text{CMR} + 4.12(\pm 5.82)$$
$$n = 6, \quad r^2 = 0.875, \quad q^2 = 0.576, \quad s = 0.307, \quad \text{outlier}: -(\text{CH}_2)_4- \tag{4}$$

In order to have good HIV protease inhibitory activity, the compound need to have good antiviral activity, to account for the translation of enzyme activity to antiviral activity, the translation data of derivatives of **4** ($T = K_i/\text{IC}_{90}$) was correlated, as given in QSAR 5 and 6.

$$\text{Log}\,1/T = 0.95(\pm 0.16)\text{C log}\,P - 10.67(\pm 1.28)$$
$$n = 7, \quad r^2 = 0.979, \quad q^2 = 0.945, \quad s = 0.203 \tag{5}$$

$$\text{Log}\,1/T = 0.98(\pm 0.17)\text{CMR} - 22.75(\pm 3.55)$$
$$n = 7, \quad r^2 = 0.976, \quad q^2 = 0.939, \quad s = 0.219 \tag{6}$$

For QSAR 1–6

$$\text{C log}\,P \text{ range} = 5.45\text{–}9.15,$$
$$\text{CMR range} = 17.85\text{–}21.28,$$
$$\text{C log}\,P \text{ versus CMR } (r^2 = 0.922)$$

It appears to us that if a compound is more hydrophobic and bulky, its enzymatic activity translate well into antiviral activity. Overall, if we study the QSAR 1–6, it is clear that while the good enzyme inhibition require a compound to be less hydrophobic and less bulky, the antiviral activity and translation across the cell membrane depends more on larger and hydrophobic compound.

Analysis of the activity data in Table 1 reveals that there is a progressive decrease in the enzyme inhibition activity of these compounds with larger, more hydrophobic

Table 2. Symmetrical P1/P1′ substituted 3-aminoindazole cyclic urea derivatives, their physico-chemical parameters and HIV-protease inhibition activity studied by Kaltenbach et al.[29] (panel a, **5**) and their HIV-antiviral potency and translation data of compounds of panel a (panel b, **5**[29])

S. No.	R	C log *P*	CMR	L_4	I_1	Log(1/K_i)			
						Obsd.[29]	Calcd		
							QSAR 7	QSAR 8	QSAR 25
Panel (a)									
1	H	5.45	17.85	2.06	0.0	11.00	11.04	11.04	10.97
2[a]	4-CH_3	6.45	18.78	2.87	0.0	10.21	10.69	10.70	10.63
3	3-CH_3	6.45	18.78	2.06	0.0	10.96	10.69	10.70	10.63
4	3,5-Di-CH_3	7.45	19.70	2.06	0.0	10.17	10.34	10.35	10.29
5	2,5-Di-CH_3	7.35	19.70	2.06	0.0	10.47	10.37	10.35	10.33
6	4-C_2H_5	7.51	19.70	4.11	0.0	10.17	10.32	10.35	10.27
7	4-C_3H_7	8.56	20.63	4.92	0.0	10.17	9.95	10.01	9.91
8	4-$CH(CH_3)_2$	8.30	20.63	4.11	0.0	9.72	10.04	10.01	10.00
9	4-$C(CH_3)_3$	9.10	21.56	4.11	0.0	9.72	9.76	9.67	9.73
10	3-$C(CH_3)_3$	9.10	21.56	2.06	0.0	9.68	9.76	9.67	9.73
11	4-C_4H_9	9.62	21.56	6.17	0.0	9.77	9.57	9.67	9.55

	Log(1/IC_{90})				Log(1/*T*)			
	Obsd.[29]	Calcd			Obsd.[29]	Calcd		
		QSAR 9	QSAR 10	QSAR 26		QSAR 11	QSAR 12	QSAR 27
Panel (b)								
1	5.56	5.78	5.77	5.70	−5.44	−5.10	−5.08	−5.32
2	6.11	6.06	6.05	6.00	−4.11	−4.54	−4.54	−4.55
3	6.13	6.06	6.05	6.16	−4.83	−4.54	−4.54	−4.55
4	6.27	6.33	6.32	6.61	−3.91	−3.98	−4.00	−3.77
5	6.52	6.30	6.32	6.56	−3.95	−4.04	−4.00	−3.85
6	6.52	6.35	6.32	6.25	−3.64	−3.95	−4.00	−3.73
7	6.54	6.64	6.60	6.58	−3.63	−3.35	−3.46	−2.90
8	6.39	6.57	6.60	6.61	−3.33	−3.50	−3.46	−3.11
9	7.00	6.78	6.87	6.98	−2.72	−3.05	−2.92	−2.48
10[b]	6.41	6.78	6.87	7.36	−3.27	−3.05	−2.92	−2.48
11[c]	6.74	6.93	6.87	6.83	−3.02	−2.76	−2.92	−2.08

[a] Not included in deriving QSAR 7, 8, and 25.
[b] Not included in deriving QSAR 9, 10, 26, and 27.
[c] Not included in deriving QSAR 27.

compounds. However, the antiviral activity first increases and then decreases with the increase in hydrophobicity and size. We were unable to see this change in the relationship in QSAR 3 and 4, as this is a very small dataset. To have a statistically valid meaningful QSAR, a minimum of five data points per parameter are recommended.[36] The choice of substituents and the range of C log *P* did not allow the QSAR to reveal the dependence of biological activity (first increase and then decrease) on physicochemical parameters C log *P* and/or CMR in these QSAR. The range of C log *P*, CMR and mutual correlation between C log *P* and CMR ($r^2 = 0.922$) is noted for QSAR 1–6.

2.2. Symmetrical P1/P1′ substituted 3-aminoindazole cyclic urea derivatives (5)

In an attempt to further increase the antiviral activity and maintain the good enzyme inhibition of 3-aminoindazole cyclic urea HIV-1 protease inhibitors, Kaltenbach et al.,[29] synthesized the P1/P1′ substituted symmetrical aminoindazole analogs (**5**) of cyclic urea, and reported their structure–activity data. We derived QSAR 7–12 for the enzyme inhibition (K_i), antiviral (IC_{90}) and translation ($T = K_i/IC_{90}$) data (Table 2).

$$\text{Log}\,1/K_i = -0.35(\pm0.12)\text{C}\log P + 12.95(\pm0.95)$$
$$n = 10, \quad r^2 = 0.851, \quad q^2 = 0.780, \quad s = 0.202, \quad (7)$$
$$\text{outlier: 4-CH}_3$$

$$\text{Log}\,1/K_i = -0.37(\pm0.11)\text{CMR} + 17.62(\pm2.28)$$
$$n = 10, \quad r^2 = 0.877, \quad q^2 = 0.825, \quad s = 0.184, \quad (8)$$
$$\text{outlier: 4-CH}_3$$

$$\text{Log}\,1/\text{IC}_{90} = 0.28(\pm0.11)\text{C}\log P + 4.28(\pm0.83)$$
$$n = 10, \quad r^2 = 0.813, \quad q^2 = 0.659, \quad s = 0.181, \quad (9)$$
$$\text{outlier: 3-C(CH}_3)_3$$

$$\text{Log}\,1/\text{IC}_{90} = 0.29(\pm0.10)\text{CMR} + 0.47(\pm2.08)$$
$$n = 10, \quad r^2 = 0.844, \quad q^2 = 0.728, \quad s = 0.166, \quad (10)$$
$$\text{outlier: 3-C(CH}_3)_3$$

$$\text{Log}\,1/T = 0.56(\pm0.16)\text{C}\log P - 8.15(\pm1.30)$$
$$n = 11, \quad r^2 = 0.867, \quad q^2 = 0.776, \quad s = 0.302 \quad (11)$$

Table 3. Nonsymmetrical P2/P2′ substituted 3-aminoindazole cyclic urea derivatives, their physico-chemical parameters and HIV-protease inhibition activity studied by Rodgers et al.[30] (panel a, **6**) and their HIV-antiviral potency and translation data of compounds of panel a (panel b, **6**[30])

S. No.	R	C log *P*	CMR	L_4	I_1	Log(1/K_i)			
						Obsd.[30]	Calcd		
							QSAR 13	QSAR 14	QSAR 25
Panel (a)									
1[a]	CH_2-cy-C_3H_5	5.63	15.32	2.06	1.00	10.70	11.14	10.84	10.91
2	CH_2-cy-C_4H_7	6.19	15.74	2.06	1.00	10.80	10.73	10.51	10.72
3[b]	CH_2-2-naphthyl	7.52	18.26	2.06	1.00	10.64	9.76	8.54	10.27
4[c]	$CH_2C_6H_5$	6.34	16.57	2.06	1.00	10.51	10.62	9.86	10.67
5	C_4H_9	6.25	15.45	2.06	1.00	10.68	10.69	10.73	10.70
6	C_5H_{11}	6.78	15.92	2.06	1.00	10.39	10.30	10.37	10.52
7[b]	C_6H_{13}	7.31	16.38	2.06	1.00	9.89	9.92	10.01	10.34

	Log(1/IC_{90})				Log(1/*T*)			
	Obsd.[30]	Calcd			Obsd.[30]	Calcd		
		QSAR 15	QSAR 16	QSAR 26		QSAR 17	QSAR 18	QSAR 27
Panel (b)								
1[d]	6.94	8.20	6.93	7.01	−3.76	−3.72	−3.73	−4.02
2[e]	7.74	7.49	7.02	7.27	−3.06	−3.66	−3.55	−3.58
3[d]	7.64	5.81	7.58	7.87	−3.00	−3.28	−3.14	−2.55
4	7.21	7.30	7.21	7.34	−3.30	−3.53	−3.51	−3.46
5	7.25	7.42	6.96	7.29	−3.43	−3.70	−3.54	−3.53
6[f]	6.69	6.75	7.06	7.53	−3.70	−3.63	−3.37	−3.12
7[g]	6.12	6.07	7.16	7.77	−3.77	−3.56	−3.20	−2.71

[a] Not included in deriving QSAR 13.
[b] Not included in deriving QSAR 13, 14, and 25.
[c] Not included in deriving QSAR 14 and 25.
[d] Not included in deriving QSAR 15.
[e] Not included in deriving QSAR 16, 17, and 18.
[f] Not included in deriving QSAR 18 and 26.
[g] Not included in deriving QSAR 16, 17, 26, and 27.

$$\text{Log}\,1/T = 0.59(\pm 0.16)\text{CMR} - 15.50(\pm 3.29)$$
$$n = 11, \quad r^2 = 0.878, \quad q^2 = 0.793, \quad s = 0.290 \qquad (12)$$

For QSAR 7–12

$$\text{C log}\,P \text{ range} = 5.45\text{–}9.62,$$
$$\text{CMR range} = 17.85\text{–}21.56,$$
$$\text{C log}\,P \text{ versus CMR}(r^2 = 0.986)$$

Similar to the observations in QSAR 1–6 derived for P2/P2′ substituted symmetrical 3-aminoindazole (**4**), for P1/P1′ symmetrical 3-aminoindazoles (**5**) also, good enzyme inhibition seems to require a compound to be less hydrophobic and less bulky. However, the antiviral activity and translation of enzymatic activity to antiviral activity were found to be favored by a larger and hydrophobic compound. Investigation of Table 2 shows that out of 11 compounds, one compound has C log *P* 5.45, two compounds have C log *P* 6.45 and all others were >7.45.

2.3. Nonsymmetrical P2/P2′ substituted 3-aminoindazole cyclic urea derivatives (6)

The symmetrical indazoles were reported to have poor bioavailability,[28–30] due to the high molecular weight and presence of large number of hydrogen bond donor/acceptor groups. Thus, the group reporting structure–activity data on **4**, also synthesized and studied the nonsymmetrical 3-aminoindazole (**6**).[30] QSAR analyses on their data revealed QSAR 13–18 for the enzyme inhibition (K_i), antiviral (IC_{90}) and translation ($T = K_i/IC_{90}$) (Table 3).

$$\text{Log}\,1/K_i = -0.73(\pm 0.31)\text{C log}\,P + 15.26(\pm 2.04)$$
$$n = 5, \quad r^2 = 0.950, \quad q^2 = 0.859, \quad s = 0.091 \qquad (13)$$
$$\text{outliers: } CH_2\text{-2-naphthyl}, \quad CH_2\text{-cy-}C_3H_5$$

$$\text{Log}\,1/K_i = -0.78(\pm 0.76)\text{CMR} + 22.79(\pm 12.0)$$
$$n = 5, \quad r^2 = 0.780, \quad q^2 = 0.210, \quad s = 0.200 \qquad (14)$$
$$\text{outliers: } CH_2\text{-2-naphthyl}, \quad CH_2C_6H_5$$

$$\text{Log}\,1/IC_{90} = 1.27(\pm 0.64)\text{C log}\,P + 15.35(\pm 4.21)$$
$$n = 5, \quad r^2 = 0.930, \quad q^2 = 0.800, \quad s = 0.189 \qquad (15)$$
$$\text{outliers: } CH_2\text{-2-naphthyl}, \quad CH_2\text{-cy-}C_3H_5$$

$$\text{Log}\,1/IC_{90} = 0.22(\pm 0.36)\text{CMR} + 3.56(\pm 5.95)$$
$$n = 5, \quad r^2 = 0.553, \quad q^2 = -0.234, \quad s = 0.274 \qquad (16)$$
$$\text{outliers: } CH_2\text{-cy-}C_4H_7, \quad C_6H_{13}$$

Table 4. Nonsymmetrical P1/P1′ substituted 3-aminoindazole cyclic urea derivatives, their physico-chemical parameters and HIV-protease inhibition activity studied by Kaltenbach et al.[31] (panel a, **7**) and their HIV-antiviral potency and translation data of compounds of panel a (panel b, **7**[31])

S. No.	Substituents		C log *P*	CMR	L_4	I_1	Log($1/K_i$)			
	R_1	R_2					Obsd.[31]	Calcd		
								QSAR 19	QSAR 20	QSAR 25
1	H	$CH_2C_6H_5$	6.34	16.57	2.06	1.0	10.51	10.66	10.39	10.67
2	4-CH_3	$CH_2C_6H_5$	7.34	17.50	2.87	0.0	10.33	10.22	10.11	10.33
3	3-CH_3	$CH_2C_6H_5$	7.34	17.50	2.06	0.0	10.21	10.22	10.11	10.33
4[a]	4-C_2H_5	$CH_2C_6H_5$	8.40	18.43	4.11	0.0	10.14	9.75	9.84	9.97
5	3,5-Di-CH_3	$CH_2C_6H_5$	8.34	18.43	2.06	0.0	9.70	9.77	9.84	9.99
6	H	C_4H_9	6.25	15.45	2.06	1.0	10.68	10.71	10.72	10.70
7	4-CH_3	C_4H_9	7.25	16.38	2.87	0.0	10.43	10.26	10.44	10.36
8	3-CH_3	C_4H_9	7.25	16.38	2.06	0.0	10.35	10.26	10.44	10.36
9[b]	4-C_2H_5	C_4H_9	8.30	17.31	4.11	0.0	10.03	9.79	10.17	10.00
10[c]	3,5-Di-CH_3	C_4H_9	8.24	17.31	2.06	0.0	9.72	9.82	10.17	10.02

	Log($1/IC_{90}$)				Log($1/T$)			
	Obsd.[31]	Calcd			Obsd.[31]	Calcd		
		QSAR 21	QSAR 22	QSAR 26		QSAR 23	QSAR 24	QSAR 27
1	6.30	6.19	6.34	—	−3.30	−4.40	−4.05	—
2[d]	6.72	6.42	6.63	6.41	−3.61	−3.89	−3.64	−3.85
3	6.52	6.42	6.63	6.56	−3.68	−3.89	−3.64	−3.85
4[e]	6.61	6.67	6.93	6.66	−3.54	−3.34	−3.23	−3.03
5[f]	6.46	6.65	6.93	—	−3.24	−3.37	−3.23	−3.08
6	6.02	6.17	5.98	7.29	−3.43	−4.45	−4.54	—
7	6.38	6.40	6.28	6.37	−4.06	−3.94	−4.13	−3.93
8[d]	6.11	6.40	6.28	6.52	−4.24	−3.94	−4.13	−3.93
9	6.66	6.64	6.57	6.62	−3.37	−3.39	−3.72	−3.10
10	—	—	—	—	—	—	—	—

[a] Not included in deriving QSAR 19 and 20.
[b] Not included in deriving QSAR 19.
[c] Not included in deriving QSAR 20.
[d] Not included in deriving QSAR 21.
[e] Not included in deriving QSAR 22 and 27.
[f] Not included in deriving QSAR 21 and 22.

Table 5. Symmetrical P1/P1′ substituted 3-aminoindazole cyclic urea derivatives (**7**), their physico-chemical parameters and HIV-antiviral potency studied by Kaltenbach et al.[31]

S. No.	Substituents		C log *P*	CMR	QSAR 28*		QSAR 29#	
	R_1	R_2			Obsd.[31]	Calcd	Obsd.[31]	Calcd
1	H	$CH_2C_6H_5$	6.34	16.57	7.04	7.16	6.43	6.42
2	4-CH_3	$CH_2C_6H_5$	7.34	17.50	7.77	7.71	6.51	6.51
3[a]	3-CH_3	$CH_2C_6H_5$	7.34	17.50	7.59	7.71	6.98	6.51
4[a]	4-C_2H_5	$CH_2C_6\ H_5$	8.40	18.43	7.09	7.03	6.39	6.59
5	3,5-Di-CH_3	$CH_2C_6H_5$	8.34	18.43	7.00	7.10	—	—
6	H	C_4H_9	6.25	15.45	7.16	7.05	6.41	6.42
7	4-CH_3	C_4H_9	7.25	16.38	7.80	7.71	—	—
8[b]	3-CH_3	C_4H_9	7.25	16.38	7.27	7.71	—	—
9[b]	4-C_2H_5	C_4H_9	8.30	17.31	7.66	7.14	6.59	6.59
10	3,5-Di-CH_3	C_4H_9	8.24	17.31	7.25	7.21	—	—

[a] Not included in deriving QSAR 29.
[b] Not included in deriving QSAR 28.
[*] IC_{90} data of wild type viral isolate HXB_2.
[#] IC_{90} data of HIV-1 mutant I84V.

$$\text{Log}\,1/T = 0.15(\pm 0.13)\text{C log}\,P - 6.05(\pm 2.18)$$
$$n = 5, \quad r^2 = 0.814, \quad q^2 = 0.516, \quad s = 0.100 \qquad (17)$$
$$\text{outliers: } CH_2\text{-cy-}C_4H_7, \quad C_6H_{13}$$

$$\text{Log}\,1/T = 0.31(\pm 0.37)\text{CMR} - 5.50(\pm 2.46)$$
$$n = 5, \quad r^2 = 0.708, \quad q^2 = 0.136, \quad s = 0.182 \qquad (18)$$
$$\text{outliers: } CH_2\text{-cy-}C_4H_7, \quad C_5H_{11}$$

For QSAR 13–18

$$\begin{aligned}&\text{C}\log P \text{ range} = 5.63\text{–}7.52,\\ &\text{CMR range} = 15.32\text{–}18.26,\\ &\text{C}\log P \text{ versus CMR}(r^2 = 0.669)\end{aligned}$$

QSAR 13 and 14 were derived for the protease inhibitory activity (Table 3, panel a). From QSAR 13 it appears that if the compounds are very hydrophobic, the enzyme activity decreases. To study whether this decrease is due to hydrophobic and/or steric interactions (as observed in QSAR 1, 2, 7, and 8), we also studied the correlation with CMR (QSAR 14), which do show a negative correlation with CMR, however, it is statistically not very significant (note poor q^2 0.21). The QSAR with CMR are given for comparison. So we believe that in unsymmetrical 3-aminoindazole cyclic urea the interactions between P2/P2′ substituents and receptor are more hydrophobic than steric.

The analyses of antiviral activity (IC_{90}) of analogs of **7** (Table 3, panel b), gave QSAR 15 and 16. More hydrophobic molecules were found conducive to good antiviral activity. The correlation with CMR were not that significant ($r^2 = 0.964$, $q^2 = 0.800$ versus $r^2 = 0.553$, $q^2 = -0.234$ for QSAR 15 and 16, respectively). The translation of enzyme activity to antiviral activity depended well on hydrophobicity (QSAR 17). QSAR 18 was not very significant ($q^2 = 0.136$). There were only seven compounds in the dataset reported in Table 3, for which QSAR 13–18 were developed, only one of them have C$\log P$ 5.63, another 6.19, and all others >7.52.

Overall analyses of QSAR 1–18 show that whereas the substituted symmetrical 3-aminoindazole cyclic urea (**4** and **5**) have hydrophobic and steric interactions with the receptor and both play an important role in their biological activities, for nonsymmetrical P2/P2′ cyclic urea (**6**) the hydrophobic interactions are more important. May be the removal of one of the P2/P2′ 3-aminoindazole moiety causes some conformational change in the molecule making it interact more favorably with the hydrophobic site.

2.4. Nonsymmetrical P1/P1′ substituted 3-aminoindazole cyclic urea derivatives (7)

Nonsymmetrical P2/P2′ substituted 3-aminoindazole cyclic urea compounds were found to possess excellent enzyme inhibition activity but poor antiviral activity. In order to further increase the antiviral potency of **6**, Kaltenbach et al.[31] designed P1/P1′ substituted analogs (**7**) and reported their activity data. We derived QSAR 19–24 for their enzyme inhibition (K_i), antiviral (IC_{90}), and translation ($T = K_i/IC_{90}$) data (Table 4).

$$\begin{aligned}&\text{Log}\,1/K_i = -0.47(\pm 0.15)\text{C}\log P + 13.50(\pm 1.09)\\ &n = 8,\quad r^2 = 0.900,\quad q^2 = 0.819,\quad s = 0.121\\ &\text{outliers: } R_1 = \text{4-C}_2\text{H}_5,\quad R_2 = \text{CH}_2\text{C}_6\text{H}_5;\\ &\qquad R_1 = \text{4-C}_2\text{H}_5,\quad R_2 = \text{C}_4\text{H}_9\end{aligned} \tag{19}$$

$$\begin{aligned}&\text{Log}\,1/K_i = -0.30(\pm 0.14)\text{CMR} + 15.32(\pm 2.42)\\ &n = 8,\quad r^2 = 0.813,\quad q^2 = 0.651,\quad s = 0.142\\ &\text{outliers: } R_1 = \text{4-C}_2\text{H}_5,\quad R_2 = \text{CH}_2\text{C}_6\text{H}_5;\\ &\qquad R_1 = \text{3, 5-di-CH}_3,\quad R_2 = \text{C}_4\text{H}_9\end{aligned} \tag{20}$$

$$\begin{aligned}&\text{Log}\,1/\text{IC}_{90} = 0.23(\pm 0.15)\text{C}\log P + 4.72(\pm 1.10)\\ &n = 6,\quad r^2 = 0.822,\quad q^2 = 0.530,\quad s = 0.111\\ &\text{outliers: } R_1 = \text{4-CH}_3,\quad R_2 = \text{CH}_2\text{C}_6\text{H}_5;\\ &\qquad R_1 = \text{3, 5-di-CH}_3,\quad R_2 = \text{CH}_2\text{C}_6\text{H}_5;\\ &\qquad R_1 = \text{3-CH}_3,\quad R_2 = \text{C}_4\text{H}_9\end{aligned} \tag{21}$$

$$\begin{aligned}&\text{Log}\,1/\text{IC}_{90} = 0.32(\pm 0.16)\text{CMR} + 1.03(\pm 2.69)\\ &n = 7,\quad r^2 = 0.840,\quad q^2 = -0.705,\quad s = 0.116\\ &\text{outliers: } R_1 = \text{4-C}_2\text{H}_5,\quad R_2 = \text{CH}_2\text{C}_6\text{H}_5;\\ &\qquad R_1 = \text{3, 5-di-CH}_3,\quad R_2 = \text{CH}_2\text{C}_6\text{H}_5\end{aligned} \tag{22}$$

$$\begin{aligned}&\text{Log}\,1/T = 0.52(\pm 0.21)\text{C}\log P - 7.69(\pm 1.55)\\ &n = 9,\quad r^2 = 0.808,\quad q^2 = 0.695,\quad s = 0.215\end{aligned} \tag{23}$$

$$\begin{aligned}&\text{Log}\,1/T = 0.44(\pm 0.18)\text{CMR} - 11.37(\pm 3.02)\\ &n = 9,\quad r^2 = 0.807,\quad q^2 = 0.697,\quad s = 0.215\end{aligned} \tag{24}$$

For QSAR 19–24

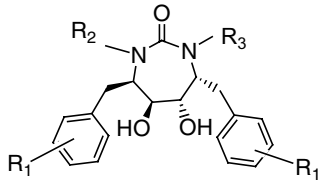
$$\begin{aligned}&\text{C}\log P \text{ range} = 6.34\text{–}8.40,\\ &\text{CMR range} = 15.45\text{–}18.43,\\ &\text{C}\log P \text{ versus CMR}(r^2 = 0.665)\end{aligned}$$

The QSAR models derived for biological activities of nonsymmetrical P1/P1′-substituted 3-aminoindazole cyclic urea followed the same pattern as visible in QSAR 1–18. However, the antiviral activity used in QSAR 21 and 22 were for protein binding adjusted IC_{90}.[31] The molar concentration of the compound, required to reduce the concentration of HIV viral RNA by 90% from the level measured in an infected culture did not reveal statistically significant QSAR. We again noticed the importance of hydrophobic and steric interactions in the action of these compounds. However, these QSAR compared more with QSAR 1–12 than QSAR 13–18, in that, both the C$\log P$ and CMR seems to be almost equally well correlated with the biological activities except for QSAR 22. The substitution at P1/P1′ position of nonsymmetrical indazole (**7**) compounds seems to increase the steric as well as hydrophobic interactions with the receptor, similar to the derivatives of **4** and **5**. The binding mode of these compounds may be different from P2/P2′ substituted nonsymmetrical indazoles (**6**), which primarily seems to be interacting hydrophobically

(see QSAR 13–18). Except one compound (C log P = 6.34), all the compounds in Table 4, have C log P more than 7.34.

The different modes of binding of symmetrical and non-symmetrical 3-aminoindazole cyclic urea are clear from these studies. The mutual correlations between C log P and CMR for QSAR 13–18 and 19–24 were 0.669 and 0.665, as compared to 0.922 and 0.982 for QSAR 1–6 and 7–12, respectively.

To get deeper insight about the mode of interaction of these compounds with the HIV-1 protease receptor, we combined the data of Tables 1–4 and developed QSAR 25–27 (**8**).[28–31]



8

R_1 = H / alkyl
R_2 and/or R_3 = 3-x-aminoindazole / alkyl / benzyl

2.4.1. K_i data of combined dataset of 3-aminoindazole cyclic urea derivatives (8).[28–31]

$$\text{Log}\,1/K_i = -0.34(\pm 0.07)\text{C log}\,P + 12.82(0.53)$$
$$n = 28, \quad r^2 = 0.801, \quad q^2 = 0.771, \quad s = 0.195$$
outliers: R_1 = H, R_2/R_3 = 3-NH-$(CH_2)_4$-indazole;
R_1 = 4-CH_3, R_2/R_3 = 3-NH_2-indazole;
R_1 = H, R_2 = CH_2-2-naphthyl,
R_3 = 3-NH_2-indazole;
R_1 = H, R_2 = C_6H_{13},
R_3 = 3-NH_2-indazole (25)

2.4.2. IC_{90} data of combined dataset of 3-aminoindazole cyclic urea derivatives (8).[28–31]

$$\text{Log}\,1/\text{IC}_{90} = 0.46(\pm 0.11)\text{C log}\,P + 1.23(\pm 0.28)I_1 - 0.19(\pm 0.10)L_4 + 3.61(\pm 0.78)$$
$$n = 27, \quad r^2 = 0.833, \quad q^2 = 0.773, \quad s = 0.234$$
outliers: R_1 = 3-CMe_3, R_2/R_3 = 3-NH_2-indazole;
R_1 = H, R_2 = C_5H_{11},
R_3 = 3-NH_2-indazole;
R_1 = H, R_2 = C_6H_{13},
R_3 = 3-NH_2-indazole;
R_1 = 3,5-di-Me, R_2 = $CH_2C_6H_5$,
R_3 = 3-NH_2-indazole (26)

2.4.3. Translation data of combined dataset of 3-aminoindazole cyclic urea derivatives (8).[28–31]

$$\text{Log}\,1/T = 0.78(\pm 0.15)\text{C log}\,P + 1.17(\pm 0.40)I_1 - 9.56(\pm 1.25)$$
$$n = 27, \quad r^2 = 0.808, \quad q^2 = 0.750, \quad s = 0.360$$
outliers: R_1 = 4-C_4H_9, R_2/R_3 = 3-NH_2-indazole;
R_1 = 3-$(CH_3)_3$, R_2/R_3 = 3-NH_2-indazole;
R_1 = H, R_2 = C_6H_{13},
R_3 = 3-NH_2-indazole;
R_1 = 4-C_2H_5, R_2 = C_6H_5,
R_3 = 3-NH_2-indazole (27)

QSAR 25–27 reflects the pattern shown by QSAR 1–24. The best correlations were observed with C log P. It seems that overall the hydrophobicity of these compounds is the most important parameter governing their biological activity. The enzyme inhibition decreases with increase in hydrophobicity (QSAR 25), whereas the antiviral activity and translation data correlated positively with C log P (QSAR 26 and 27). It is to be noted that the antiviral data for derivatives of **7**, used for developing QSAR 26 and 27 was for protein binding adjusted IC_{90}[31] (similar to used for QSAR 21 and 22). The RF IC_{90} did not reveal significant correlations.[31] We used this value in the combined data set because the individual QSAR 21 and 22 derived for the antiviral activity of analogs of **7** were consistent with all the other correlations. QSAR 26 also have a negative coefficient of parameter L_4, indicating that length of the fourth position substituent at P1/P1′-benzyl is detrimental to the activity. This effect was not seen in the QSAR 9 and 21 (subsets listed in Table 2(panel a) and Table 4(panel a)), may be due to small size of dataset. It appears that in P1/P1′ substituted compounds overall bulk may be good for the antiviral activity (QSAR 10 and 22), but the longer substituent at fourth position of P1/P1′ benzyl has negative interactions with the receptor and is detrimental to the activity. It may be involved in steric interactions at the binding site. The limited number of substituents and narrow range in their parameter value, at other positions of the benzyl precluded their more systematic study.

In QSAR 26 and 27 additional indicator parameter I_1 was used with a value of unity for nonsymmetrical P2/P2′ substituted analogs (**6**). The positive coefficient of this parameter indicates that these analogs of 3-aminoindazole cyclic urea are more suitable for achieving good antiviral activity and good translation against HIV-1 infected cells. However, two of the compounds (**6** and **7**, Table 3, panel a) belonging to this group were found outlier in the derivation of QSAR 26. These compounds are highly hydrophobic and long straight chain alkyls; it is possible that their larger size causes steric hindrance at the binding site. Compound 7, Table 3(panel a) was also found misfit in QSAR 27 derived for translation. An indicator used for these straight chain alkyl substituent showed negative effect of these substituents on the activity, however, the QSAR model did not improve

much by adding an additional indicator for this subset. The inclusion of the other series of nonsymmetrical analogs (**7**) in the I_1 parameter was detrimental to the activity.

Most of the compounds found outlier in QSAR 25–27 were also found 'misfit' when we studied these series independently (QSAR 1–24). The calculated value of these compounds was either too high or too low than the corresponding observed value (see individual QSAR and corresponding tables for the data). This problem of 'misfit' of the congeners in the final QSAR could be associated with any one of the following reasons:

1. Outliers due to what seem to be 'congeners' but are not.
2. Mathematical form of the equation may be off the mark.
3. Different rates of metabolism of the members of a set.
4. The quality of the experimental data.
5. Finally, the parameters used may not be the best. Sometimes, experimentally obtained parameters are better than those calculated and vice versa.

In fact they are the leads to new understanding, and covering them up by including them in a QSAR, at the cost of lower r^2, can be more confusing than helpful.

3. Overview

The three important factors, which describe the physicochemical properties of the molecules and are used in developing QSAR, are hydrophobic, steric, and electronic. One needs variation in these properties of the substituents at each position of the parent structure to be sure that these properties are considered, In addition, the test sets should be large enough to be able to include the three factors to see their influence on activity. Very often either of the two aspects is not considered while designing a series for investigation.

The most important conclusion is the importance of hydrophobicity and steric effect. Overall, study of QSAR 1–27, show that the HIV-1 protease inhibitory activity decreases with larger and more hydrophobic molecules. The antiviral activity and translation (K_i/IC_{90}) across the cell membrane was found to increase with increase in hydrophobicity and size. Presence of parameter CMR in a QSAR has been associated with size as well as polarizability and also points toward the involvement of electronic effects beside steric interactions.[37] It is of note that in the design of all these series, the major attention was paid to study only the effect of change in hydrophobicity and size. Not much attempt was reported to study the electronic effect of the substituent. Most of the substituents studied have almost similar Hammett σ values, used to study most commonly the electronic effects in a QSAR.[38] We have included range of Clog*P* and CMR values of the congeners in each series with each QSAR for the convenience of the readers. A brief summary is given in Table 6.

Table 6. Clog*P* range for QSAR models

QSAR model	Clog*P* range	Comment
1–6	5.45–9.15	One compound Clog*P* 5.45, all others > 7.04
7–12	5.45–9.62	One compound Clog*P* 5.45, two 6.45, all others > 7.45
1–18	5.63–7.52	One compound Clog*P* 5.63, second 6.19, all others > 7.52
19–24	6.34–8.40	One compound Clog*P* 6.34, all others > 7.34

It is important to note that in our previous reports,[18,25] the optimum Clog*P* range for protease inhibitors was found to be from 4.49 to 6.96. To observe the optimum value of a parameter, a sufficient spread in the data is required. For the compounds of Table 1, there is only one compound (compound **1**), that has a Clog*P* of 5.45. All others were >7.04. Investigation of Table 1(panel a) shows that there is insufficient spread in the range of Clog*P* values to establish the optimum point for antiviral potency or enzyme inhibition. Most of the data points fall either on positive side or negative side of the optimum that is why we observe either a positive or negative term in the QSAR. Previous work also showed that normally the optimum Clog*P* for antiviral potency is higher than for enzyme inhibition for HIV-1 protease inhibitors.[18,25] That could be one of the reason for the presence of only a positive term in QSAR for antiviral potency and negative term in QSAR for enzyme inhibition.

QSAR 25–27 developed for the combined data of all the four series were consistent with the individual QSAR and brought out the importance of Clog*P*. The QSAR with CMR were not found significant in the combined dataset. QSAR 25 supported individual QSAR 1, 7, 13, and 19 for enzyme inhibition data. QSAR 26 derived for antiviral potency and QSAR 27 for translation data reflected the observations noticed in QSAR 3, 9, 15, 21, and QSAR 5, 11, 17, 23, respectively. A proposed model for the interaction of nonsymmetrical 3-aminoindazole cyclic urea with HIV-1 protease is shown in Figure 1. Plots of observed activity versus Clog*P* value of the compounds used for developing QSAR 25–27, and observed versus calculated activity for these models are shown in Figures 2 and 3.

None of the QSAR derived for wild type HIV-1, showed the optimum value of Clog*P*. However, the analysis of the antiviral data (IC_{90})[31] of second wild type viral isolate HXB_2 determined by measuring the accumulation of viral RNA transcript after infection of cells with HIV-1, gave QSAR 28 for derivatives of **7** (Table 5).

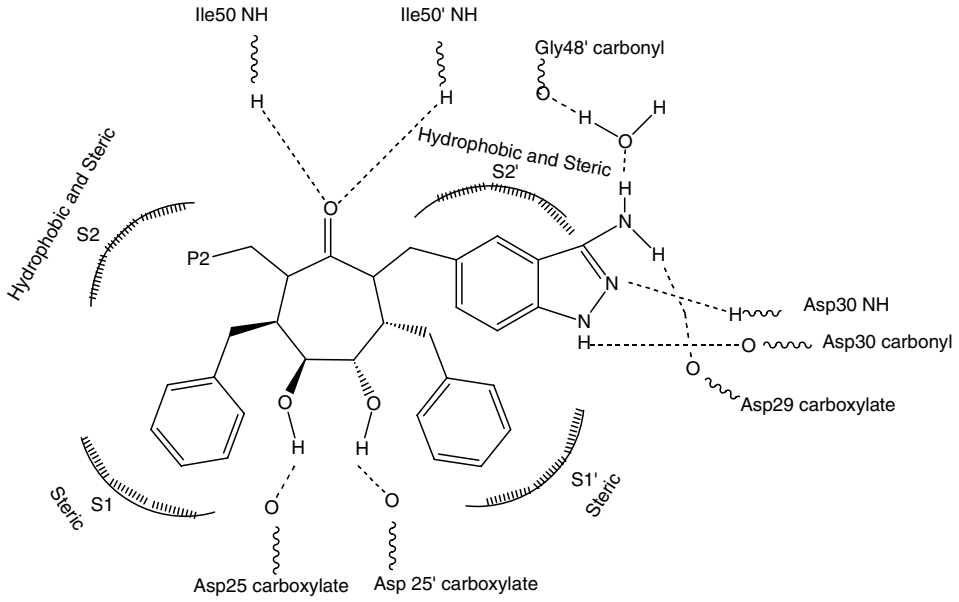


Figure 1. The proposed model for the interaction of nonsymmetrical 3-aminoindazole cyclic urea with HIV-1 protease.

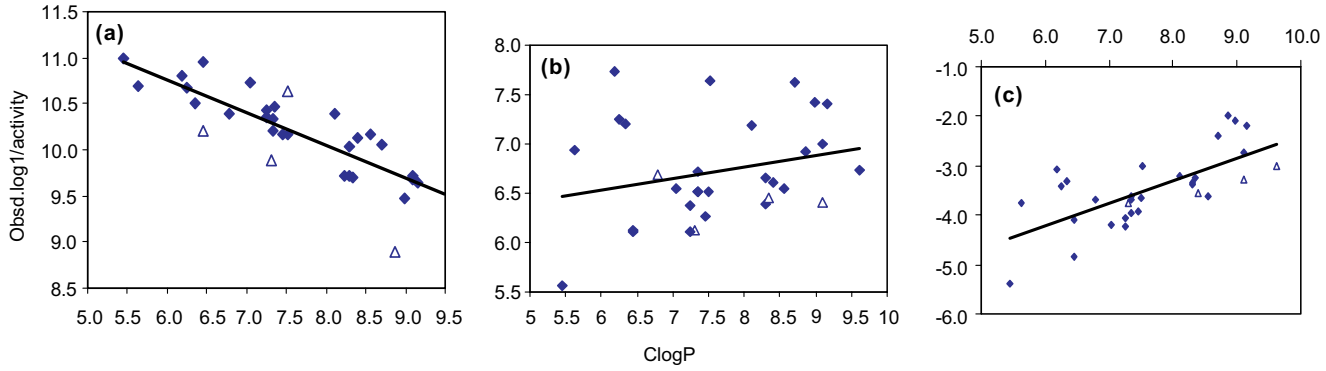


Figure 2. Plots of observed biological activity versus C log *P* of the compounds used for deriving (a) QSAR 25 (K_i); (b) QSAR 26 (IC_{90}) and (c) QSAR 27 (*T*) (Δ represents outliers).

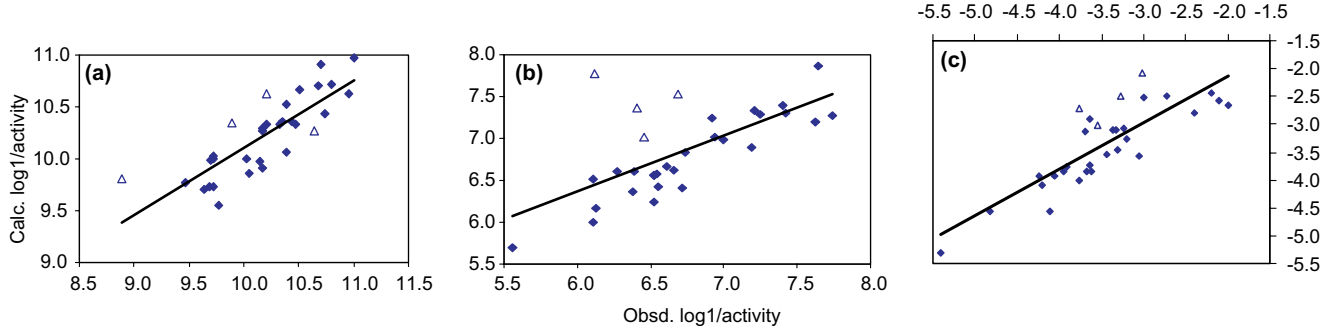


Figure 3. Plots of observed versus calculated protease inhibition (K_i), antiviral (IC_{90}) and translational data ($T = K_i/IC_{90}$) for QSAR 25–27 (Δ represents outliers).

$$\text{Log}\,1/\text{IC}_{90} = 8.46(\pm 3.17)\text{C}\log P - 0.58(\pm 0.22)\text{C}\log P^2 - 23.23(\pm 11.54)$$

$$n = 8, \quad r^2 = 0.906, \quad q^2 = 0.721, \quad s = 0.119$$

optimum $\text{C}\log P = 7.315$ (7.181–7.437)

outliers: $R_1 = 3\text{-}CH_3$, $R_2 = C_4H_9$; $R_1 = 4\text{-}C_2H_5$, $R_2 = C_4H_9$ (28)

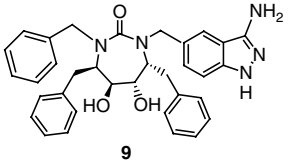
Interestingly, QSAR 28 showed a parabolic dependence of antiviral activity on hydrophobicity. It seems that the activity first increases with increasing hydrophobicity up to an optimum of 7.315, and then decreases with further increase. We were surprised to observe this QSAR, because QSAR 1–27 discussed so far in this report did not reveal any optimum value of C log *P* to aid in defining the size of the hydrophobic cavity at the binding site

with the receptor more clearly. As discussed earlier that may be due to the limitations of insufficient spread in the data and parameter values to clearly establish the optimum. In our earlier reports we have found the optimum C log *P* range for protease inhibitors to be from 4.49 to 6.96.[18,25] The optimum observed in QSAR 28 is slightly higher than observed so far in our earlier reports, it may be attributed to the type of viral isolate HXB2 that was used for studies. This second wild type viral isolate HXB2 may have a larger hydrophobic binding pocket as compared to the more common wild type HIV-1. More biological data on this isolate is required to learn more about this observation. The HIV-1 protease inhibitors already approved by US Food and Drug Administration (FDA) and in the market for use have the following C log *P* value:

FDA approved protease inhibitors*	C log *P*#
(1) Saquinavir (Invirase®)	4.73
(2) Ritanovir (Norvir®)	4.94
(3) Indinavir (Crixivan®)	3.68
(4) Nelfinavir (Viracept®)	5.84
(5) Amprenavir (Agenerase®)	3.29
(6) Lopinavir (Aluviran®)	5.54

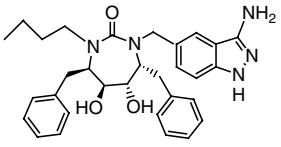
Recently approved Kaletra® is a combination of Lopinavir and Ritonavir. # Calculated using CQSAR program, Biobyte Corp., Claremont, CA. * All are peptidic in nature.

There is a close agreement in the C log *P* values of these drug molecules and the optimum log P observed in our QSAR models. The C log *P* value of the nonsymmetrical 3-aminoindazole cyclic urea protease inhibitors DMP 850 (**9**, compound **4** Table 3, panel a), and DMP 851 (**10**, compound **5** Table 3, panel a) that are currently in advanced stage of clinical trial is given below.



9

DMP 850
K_i = 0.031 nM
IC_{90}=62 nM
ClogP = 6.34



10

DMP 851
K_i = 0.021 nM
IC_{90}=56 nM
ClogP = 6.25

The closeness in the C log *P* values of drugs in the market, the nonpeptidic drugs in clinical trials, and optimum log *P* observed in our QSAR models[18,25] was surprising and established that QSAR is an important tool in drug design. Thus an individual QSAR is considered to be weak on its own and significant when strengthened by other similar QSAR models. The availability of CQSAR database[33] that allows for lateral validation of a newly derived biological QSAR makes this approach of comparative QSAR useful and amenable to mechanistic interpretation that provides an important lead in drug development.[39]

The majority of HIV research is done with cells and these studies tend to over estimate log P_0 for animal systems. From a study of CQSAR database (omitting QSAR based on charged molecules), it has been found that optimum log *P* (log P_0) for cells is about 2 log unit higher than for whole organisms.[33,34]

We also analyzed the antiviral data (IC_{90}) of HIV-1 mutant I84V for the derivatives of **7**.[31] The most significant correlation was found as given in QSAR 29 (Table 5).

$$\begin{aligned} &\text{Log}1/\text{IC}_{90} = 0.082(\pm 0.020)\text{C}\log P + 5.90(\pm 0.14) \\ &n = 4, \quad r^2 = 0.994, \quad q^2 = 0.977, \quad s = 0.008 \\ &\text{outliers: } \text{R}_1 = \text{3-CH}_3, \quad \text{R}_2 = \text{CH}_2\text{C}_6\text{H}_5; \\ &\qquad \text{R}_1 = \text{4-C}_2\text{H}_5, \quad \text{R}_2 = \text{CH}_2\text{C}_6\text{H}_5 \end{aligned} \tag{29}$$

QSAR 29 outlines the importance of hydrophobicity of substituents in the design of new drugs active against mutant viruses. The crsytal structure of multidrug-resistance HIV type-1 revealed an expanded active site cavity.[40] More such QSAR will be reported in future reports.

It is very well known that the HIV-1 protease is a C_2 symmetrical homodimer.[41,42] Each monomer has 99 residues. The C_2 axis of the enzyme lies between and perpendicular to catalytic aspartates (Asp 25 and Asp 25′) in the active site. As per the standard nomenclature[43] the S_1 and S_1' (S_2 and S_2' etc.) subsites are structurally identical.[44] The two equivalent S_1 subsites are also mostly hydrophobic except for the Asp29, Asp29′ and Asp30, and Asp30′. Wang et al.,[45] observed that at least two additional factors are also important in the binding of a compound to HIV-1 PR. The first is the conformational flexibility of the inhibitor molecule and the second is the hydrophobic interactions between an inhibitor and the enzyme.

The main interactions of cyclic urea inhibitors with HIVPR involve hydrogen bonds. The high resolution structural studies of HIV-1 PR complexed with peptidomimetic inhibitors showed the presence of a structural water molecule hydrogen bonded to the mobile flaps of the enzyme and the two carbonyls flanking the transition state mimic of the inhibitors.[14] Cyclic urea oxygen accepts two hydrogen bonds from the backbone Ile50/Ile50′ NH and the diols form multiple hydrogen bonds with catalytic Asp25/25′. The X-ray crystal structure of symmetrical 3-aminoindazole cyclic urea (compound **1**, Table 1) bound to HIV-1 protease showed same hydrogen-bond interactions with the urea core.[30] In addition, the indazole N–H and nitrogen atom forms H-bond to the Asp30 carbonyl and Asp30 N–H, respectively. 3-NH_2 of indazole forms two hydrogen bonds to the enzyme—one from NH_2 group to Asp30 carboxylate

side chain and another to a water molecule, which further forms a bond to Gly48 carbonyl. These additional hydrogen bonds to Asp29, Asp 30, and Gly48, are responsible for the increased binding of these compounds with the enzyme.[30] On the basis of this information the series of compounds belonging to **4**–**7** were synthesized. However, compounds belonging to **4** and **5** were found to have poor absorption and poor bioavailability, which was related to their high molecular weight and large number of donor/acceptor groups. In order to better address the problems associated with symmetrical 3-aminoindazoles, the nonsymmetrical derivatives (**6** and **7**) were then studied. It was believed that the removal of one of the P2/P2′ 3-aminoindazole should reduce the molecular weight, hydrophobicity, eliminate three hydrogen bonding groups and break the crystal packing, thus improving the solubility and translation of enzyme activity to antiviral activity.[28–31] The modifications studied so far have helped in improving the antiviral activity while retaining good enzyme inhibition. However, it did not lead to the desired pharmacokinetics.

4. Conclusion

So far mostly alkyl groups at 3-NH_2 of P2/P2′-indazole are tried in symmetrical 3-aminoindazole cyclic urea derivatives (**4** and **5**) and alkyl/benzyl tried in nonsymmetrical derivatives (**6** and **7**). At P1/P1′ only alkyl substituents at 3,4,5-positions of benzyl are reported (**5** and **7**). This class of cyclic urea protease inhibitors have shown very good enzyme inhibitory, antiviral, and translation activity. However, the choice of substituent and the range of parameter did not allow one to firmly reveal the optimum value of a physicochemical parameter. Most of the time, most of the data points fall on one side of the optimum, that is why we either observe a positive term or a negative term for the parameter of an importance for biological activity and sometimes come to misleading conclusions.

Other substituent that would allow the investigation of all the three physicochemical parameters namely hydrophobic, steric, and electronic at each position of the parent structure may reveal the optimum effect, and help in retaining the excellent enzyme inhibitory activity of these compounds as well as improve antiviral potency and translation further. It is hoped that the results reported herein would help in enhancing the activity and resistance profile of these 3-aminoindazole cyclic urea protease inhibitors.

One constant concern in formulating a new QSAR is to find support for it in as many ways as possible. A single QSAR standing alone can not be taken seriously until it is laterally validated by comparative QSAR (CQSAR).[18,25,39] Hydrophobic interactions play an important role in the inhibition of HIV-1 protease enzyme. To study them in detail it is necessary to design molecules of series with enough wide range of C$\log P$ values to firmly establish the optimum point. These results and our earlier reports[18,25] clearly show that the QSAR models provide valuable insight in understanding complex biological interactions.

5. Experimental

All the HIV-1 inhibitory data has been collected from the literature.[28–31] The biological activity K_i is the HIV-1 protease enzyme inhibition constant, which was measured by assaying the cleavage of a fluorescent peptide substrate using high performance liquid chromatography. Antiviral potency IC_{90} was assayed by measuring the effect of the compounds on the accumulation of viral RNA transcripts 3 days after infection of MT-2 cells with HIV-1 RF, except for QSAR 21 and 22 where it is for protein binding adjusted IC_{90}.[31] Translation data ($T = K_i/IC_{90}$) defines translation of enzyme activity to antiviral activity accounting for penetration across cell membrane.

The QSAR multiple linear regression analyses were executed with the CQSAR program and all the physicochemical parameters were auto loaded.[33] For details about the utility of CQSAR program in comparative correlation analyses see references.[36,37] The physicochemical parameter—C$\log P$ is the calculated partition coefficient in octanol/water and is a measure of hydrophobicity,[34] CMR is the calculated molar refractivity for the whole molecule, and is a measure of volume and polarizability.[36] MR is calculated as follows: $(n^2 - 1/n^2+2)$ (MW/d), where n is the refractive index, MW is the molecular weight, and d is the density of a substance. Since there is very little variation in n, MR is largely a measure of volume with a small correction for polarizability. MR values have been scaled by 0.1 and can be used for a substituent or for the whole molecule. C$\log P$ and CMR are normally for the neutral form of partially ionized compounds. L is the Verloop's sterimol parameter that defines substituents' length.[35]

Acknowledgements

This study was supported by grant 1 R15 GM 069323-01 from NIGMS. The authors are indebted to Biobyte Corp., Claremont, CA for the use of CQSAR program.

References and notes

1. Gallo, R. C.; Salahuddin, S. Z.; Popovic, M.; Shearer, G. M.; Kaplan, M.; Haynes, B. F.; Palker, T. J.; Redfield, R.; Oleske, J.; Safai, B.; White, G.; Foster, P.; Markham, P. D. *Science* **1984**, *224*, 500.
2. Barre-Sinoussi, F.; Chermann, J.-C.; Rey, F.; Nugeyre, M. T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axler-Blin, C.; Vezinet-Brun, F.; Rouzioux, C.; Rozenbaum, W.; Montagnier, L. *Science* **1983**, *220*, 868.
3. Kohl, N. E.; Emini, N. A.; Schleif, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A. F.; Scolnick, E. M.; Sigal, I. S. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4686.
4. Henderson, L. E.; Sowder, R. C.; Copeland, T. D.; Oroszlan, S.; Benveniste, R. E. *J. Med. Primatol.* **1990**, *19*, 411.

5. Vogt, V. M. *Curr. Top. Microbiol. Immunol.* **1996**, *219*, 95.
6. Gottlinger, H. G.; Sodroski, J. G.; Haseltine, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5781.
7. Peng, C.; Ho, B. K.; Chang, T. W.; Chang, N. T. *J. Virol.* **1989**, *63*, 2550.
8. Seelmier, S.; Schmidt, H.; Turk, V.; von der Helm, K. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6612.
9. McQuade, T. J.; Tomasselli, A. G.; Liu, L.; Karacostas, V.; Moss, B. *Science* **1990**, *247*, 454.
10. http://www.aids.org/atn/a-393-04.html (accessed 03/16/2004).
11. De Clercq, E. *Med. Res. Rev.* **2002**, *22*, 531.
12. Gulick, R. M. *Clin. Microbiol. Infect.* **2003**, *9*, 186.
13. Lam, P. Y.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; Chang, C. H.; Wever, P. C.; Jackson, D. A.; Sharpe, T. R.; Erickson-Vittanen, S. *Science* **1994**, *263*, 380.
14. Lam, P. Y. S.; Ru, Y.; Jadhav, P. K.; Aldrich, P. E.; De Lucca, G. V.; Eyerman, C. J.; Chang, C.-H.; Emmett, G.; Holler, E. R.; Danekar, W. F.; Li, L.; Confalone, P. N.; McHugh, R. J.; Han, Q.; Li, R.; Markwalder, J. A.; Seitz, S. P.; Sharpe, T. R.; Bacheler, L. T.; Rayner, M. M.; Klabe, R. M.; Shum, L.; Winslow, D. L.; Kornhauser, D. M.; Jackson, D. A.; Erickson-Viitanen, S.; Hodge, C. N. *J. Med. Chem.* **1996**, *39*, 3514.
15. Nugiel, D. A.; Jacobs, K.; Worley, T.; Patel, M.; Kaltenbach, R. F., III; Meyer, D. T.; Jadhav, P. K.; De Lucca, G. V.; Smyser, T. E.; Klabe, R. M.; Bacheler, L. T.; Rayner, M. M.; Seitz, S. P. *J. Med. Chem.* **1996**, *39*, 2156.
16. Wilkerson, W. W.; Dax, S.; Cheatham, W. W. *J. Med. Chem.* **1997**, *40*, 4079.
17. De Lucca, G. V.; Liang, J.; De Lucca, I. *J. Med Chem.* **1999**, *42*, 135.
18. Garg, R.; Gupta, S. P.; Gao, H.; Mekapati, S. B.; Debnath, A. K.; Hansch, C. *Chem. Rev.* **1999**, *99*, 3525.
19. Gupta, S. P.; Babu, M. S.; Garg, R.; Sowmya, S. *J. Enzym. Inhib.* **1998**, *13*, 399.
20. Gupta, S. P.; Babu, M. S.; Kaw, N. *J. Enzym. Inhib.* **1999**, *14*, 109.
21. Gupta, S. P.; Babu, M. S.; Sowmya, S. *Bioorg. Med. Chem.* **1998**, *6*, 2185.
22. Gayathri, P.; Pande, V.; Sivakumar, R.; Gupta, S. P. *Bioorg. Med. Chem.* **2001**, *9*, 3059.
23. Debnath, A. K. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 761.
24. Debnath, A. K. *J. Med. Chem.* **1999**, *42*, 249.
25. Kurup, A.; Mekapati, S. B.; Garg, R.; Hansch, C. *Curr. Med. Chem.* **2003**, *10*, 1819.
26. Wong, Y. N.; Burcham, D. L.; Saxton, P. L.; Erickson-Viitanen, S.; Grubb, M. F.; Quon, C. Y.; Huang, S. M. *Biopharm. Drug Dispos.* **1994**, *15*, 535.
27. Rodgers, J. D.; Johnson, B. L.; Wang, H.; Greenberg, R. A.; Erickson-Viitanen, S.; Klabe, R. M.; Cordova, B. C.; Rayner, M. M.; Lam, G. N.; Chang, C.-H. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2919.
28. Rodgers, J. D.; Lam, P. Y. S.; Johnson, B. L.; Wang, H.; Ko, S. S.; Seitz, S. S.; Trainor, G. L.; Anderson, P. S.; Klabe, R. M.; Bachelor, L.; Cordova, B. C.; Garber, S.; Reid, C.; Wright, M. R.; Chang, C.-H.; Erickson-Viitanen, S. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 715.
29. Kaltenbach, R. F., III; Klabe, R. M.; Cordova, B. C.; Seitz, S. P. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2259.
30. Rodgers, J. D.; Johnson, B. L.; Wang, H.; Erickson-Viitanen, S.; Klabe, R. M.; Bachelor, L.; Cordova, B. C.; Chang, C.-H. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 715.
31. Kaltenbach, R. F., III; Patel, M.; Waltermire, R. E.; Harris, G. D.; Stone, B. R. P.; Klabe, R. M.; Garber, S.; Bacheler, L. T.; Cordova, B. C.; Logue, K.; Wright, M. R.; Erickson-Viitanen, S.; Trainor, G. L. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 605.
32. Cramer, R. D., III; Bunce, J. D.; Patterson, D. E.; Frank, I. E. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18.
33. CQSAR program, Biobyte Corp., 201 W. 4th street, Claremont, CA 91711.
34. Leo, A. *Chem. Rev.* **1993**, *93*, 1281.
35. Verloop, A. In *The Sterimol Approach to Drug Design*; Marcel Dekker: New York, 1987.
36. Hansch, C.; Leo, A. In *Exploring QSAR, Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington DC, 1995.
37. Hansch, C.; Steinmetz, W. E.; Leo, A. J.; Mekapati, S. B.; Kurup, A.; Hoekman, D. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 120.
38. Hansch, C.; Leo, A. In *Exploring QSAR*; American Chemical Society: Washington DC, 1995.
39. Selassie, C. D.; Garg, R.; Kapur, S.; Kurup, A.; Verma, R. P.; Mekapati, S. B.; Hansch, C. *Chem. Rev.* **2002**, *102*, 2585.
40. Logsdon, B. C.; Vickrey, J. F.; Martin, P.; Proteasa, G.; Koepke, J. I.; Terlecky, S. R.; Wawrzak, Z.; Winters, M. A.; Merigan, T. C.; Kovari, L. C. *J. Virol.* **2004**, *78*, 3123.
41. Toh, H.; Ono, M.; Saigo, K.; Miyata, T. *Nature* **1985**, *315*, 691.
42. Miller, M.; Jaskolski, M.; Rao, J. K. M.; Leis, J.; Wlodawer, A. *Nature* **1989**, *337*, 576.
43. Schechtee, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157.
44. Babine, R. E.; Bender, S. L. *Chem. Rev.* **1997**, *97*, 1359.
45. Wang, S.; Milne, G. W.; Yan, X.; Posey, I. J.; Nicklaus, M. C.; Graham, L.; Rice, W. G. *J. Med. Chem.* **1996**, *39*, 2047.