Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 13 (2005) 4078–4084
Bioorganic & Medicinal Chemistry

From SAR to comparative QSAR: role of hydrophobicity in the design of 4-hydroxy-5,6-dihydropyran-2-ones HIV-1 protease inhibitors

Barun Bhhatarai and Rajni Garg*

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

Received 3 February 2005; revised 24 March 2005; accepted 24 March 2005
Available online 25 April 2005

Abstract—Role of hydrophobicity in the design of 4-hydroxy-5,6-dihydropyran-2-ones—a new class of emerging HIV-1 protease inhibitors (HIV-PI) was investigated by using comparative QSAR. These studies show that most of the data points in the individual dataset studied fall either on positive or negative side of the optimum value of Clog *P*. This is why, we observe either a positive or negative Clog *P* term in the QSAR. To observe the optimum value of Clog *P* for these inhibitors, a sufficient spread in the data is required. It is hoped that the results of this study would help in optimizing substituents for better binding at enzyme pockets and guide in the design of more effective HIV-PI of this class.

1. Introduction

HIV-1 (Human Immunodeficiency Virus Type-1) is the pathogenic retrovirus and causative agent of AIDS or AIDS-related complex (ARC).[1,2] When viral RNA is translated into a polypeptide sequence, it is assembled in a long polypeptide chain, which includes several individual proteins namely reverse transcriptase, protease, integrase, etc. Before these enzymes become functional, they must be cut from the longer polypeptide chain. Viral protease cuts the long chain into its individual enzyme components, which then facilitate the production of new viruses.[3–5] By blocking the ability of the protease to cleave the viral polypeptide into functional enzymes, protease inhibitors interfere with continued infection.[6–9] Mutations enable HIV-1 to resist currently available treatments, therefore, there is a growing need for the development of new improved chemotherapeutics.[10–12]

The structure–activity relationship (SAR) relates the effect of a drug or toxic chemical on animal, plant, or the environment to its molecular structure. This relationship may be assessed by considering a series of biologically active molecules and making gradual changes to them, and noting the effect of these changes upon the biological activity. Alternatively, it may be possible to assess a large body of data using intelligent tools to establish such relationships that are formulated as quantitative structure–activity relationship (QSAR). The QSAR models lead to the judgment of the specific effects of substituents and modifications in the lead structure, and thus reduce trial-and-error factors.[13] The QSAR studies have been found useful in the design and development of HIV-1 protease inhibitors (HIV-PI).[14–20] 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 cannot be taken seriously until it is laterally validated by Comparative QSAR.[14,19–21]

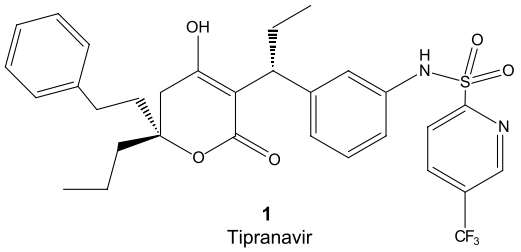
Hydrophobic interactions play important role in the inhibition of HIV-1 protease enzyme (HIV-PR).[22] They are not observed in many of the QSAR. Recent QSAR studies have revealed that the insufficient spread in the substituents' hydrophobicity could be one of the reasons for absence of the hydrophobic term in QSAR.[20] To study them in detail, it is necessary to design molecules of series with enough wide range of Clog *P* values to firmly establish the optimum point.

In the present study, we have used Comparative QSAR technique to investigate the role of hydrophobicity in

Keywords: HIV-1-protease inhibitors; Comparative QSAR; 4-Hydroxy-5,6-dihydropyran-2-ones; Hydrophobicity; Clog *P*.
* Corresponding author. Tel.: +1 1352684421; fax: +1 1352686610; e-mail: rgarg@clarkson.edu

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.03.049

the design of 4-hydroxy-5,6-dihydropyran-2-ones—a new emerging class of HIV-PI.[23] One of the drugs, Tipranavir (**1**), of this class is already in clinical trial, but has not yet achieved the desired pharmacological profile and efficacy.[24] It is envisioned that the results of this study would 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.

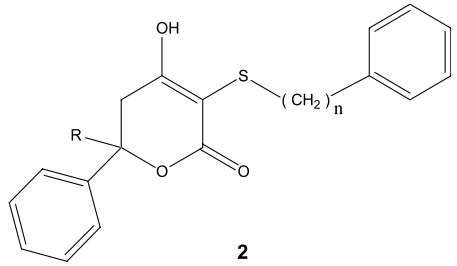


1
Tipranavir

2. Results and discussion

Tait et al. reported the antiviral activity data (IC_{50}) of 4-hydroxy-5,6-dihydropyran-2-ones HIV-PI.[23] The authors tried different hydrophobic/hydrophilic or polar/nonpolar substituents on the parent dihydropyranone and reported several series. QSAR studies were performed on these compounds by us (Tables 1–3). 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.,[25] and the data within parentheses are for the 95% confidence intervals. The QSAR multiple linear regression analyses were executed with the CQSAR program.[26] Clog$P$ is the calculated log of octanol/water partition coefficient and is a measure of hydrophobicity.[27] CMR is the calculated molar refractivity for the whole molecule and is a measure of volume dependent polarizability.[13,28] $B5$ is the Verloop's sterimol parameter that defines maximum width of the substituents.[29] Clog$P$ and CMR are normally for the neutral form of partially ionized compounds. See Experimental Section for more details.

2.1. 3-(S-$(CH_2)_n$-C_6H_5), 6-C_6H_5, 6′-R (hydrophobic) substituted 4-hydroxy-5,6-dihydro pyran-2-ones (2) (Table 1)[23]



2

Initial SAR studies by Tait et al.[23] focused on preparing alkyl derivatives (**2**) designed to fill the S_2 pocket of the enzyme and readily adjust to its surface to maximize the hydrophobic interactions. We developed QSAR 1 for the antiviral activity data (IC_{50}) of these derivatives.

$$\begin{aligned}\mathrm{Log}(1/\mathrm{IC}_{50}) &= 1.41(\pm0.36)\mathrm{C}\log P - 0.62(\pm0.24)B5_{\mathrm{R}} \\ &\quad - 1.16(\pm0.51)I_0 + 1.38(\pm1.08) \\ n = 28,\ r &= 0.901,\ r^2 = 0.812,\ q^2 = 0.715,\ s = 0.287\end{aligned} \quad (1)$$

QSAR 1 showed linear dependence of the antiviral activity on hydrophobicity as evidenced by the positive Clog$P$ term. It strengthens the fact that the biological activity increases linearly with increasing hydrophobicity. The presence of a negative $B5$ steric term for R-substituents in the QSAR suggests that the bulky groups such as phenyl, cyclo-pentyl, and cyclo-hexyl at 6-position of pyranone may cause steric hindrance. As mentioned earlier, $B5$ is Verloop's Sterimol parameter and defines the maximum width of the substituents.[29] The indicator parameter I_0 was used with a value of unity for $n = 0$ and zero for $n = 1$ and 2. Its negative coefficient indicates that $–CH_2–$ linker between sulfur and phenyl group at C3-position is required for substituent at this position to reach the binding pocket of the enzyme.

2.2. 3-(S-$(CH_2)_n$-C_6H_5), 6-C_6H_5, 6′-R (polar) substituted 4-hydroxy-5,6-dihydro-pyran-2-ones (2) (Table 2)[23]

The analyses of a small dataset where polar groups were substituted at the 6th position of (**2**) gave QSAR 2.

$$\begin{aligned}\mathrm{Log}(1/\mathrm{IC}_{50}) &= 2.47(\pm1.21)\mathrm{CMR} \\ &\quad + 1.58(\pm0.84)I_{\mathrm{COOH}} - 23.03(\pm14.33) \\ n = 6,\ r &= 0.972,\ r^2 = 0.945,\ q^2 = 0.661,\ s = 0.276\end{aligned} \quad (2)$$

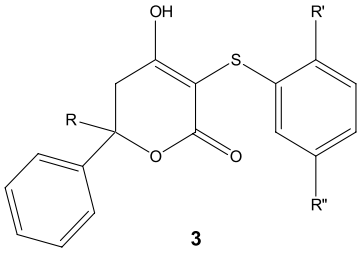
As mentioned earlier CMR is a measure of volume dependent polarizability of the compounds. The positive coefficient of CMR indicates that the bulky polar R groups may be involved in charge interaction with the binding pocket. The ionization of the polar groups in cellular assays and at physiological concentration is different at the binding site of the enzyme. The indicator parameter I_{COOH} was used with a value of one for R = $(CH_2)_n$COOH and zero for others. Its positive coefficient emphasizes the role of acidic substituents in the polar interactions. This QSAR model is statistically less robust because of its low q^2 and use of two terms for 6 data points. Normally use of 5 data points per parameter is recommended for deriving a good QSAR model.[13] However, we believe this QSAR gives important information about receptor–ligand interaction because all these compounds fitted well in the final QSAR (shown later) derived for the combined dataset. The small number of data points in the dataset and insufficient range in the parameter value precluded study of other parameters.

Table 1. The physicochemical parameters and the antiviral activity data used for deriving QSAR 1 and 4

S. No.	Substituents		Clog *P*	$B5_R$	I_0	Log ($1/IC_{50}$)		
	n	R				Obsd.[23]	Calcd.	
							QSAR 1	QSAR 4
1	1	H	3.32	1.00	0	5.05	5.45	4.92
2	1	C_3H_7	4.90	3.49	0	6.54	6.14	6.02
3	1	C_4H_9	5.43	4.54	0	6.57	6.23	6.31
4	1	C_5H_{11}	5.96	4.94	0	6.91	6.73	6.58
5	1	C_6H_{13}	6.49	5.96	0	6.74	6.85	6.80
6	1	CH_2CHMe_2	5.30	4.45	0	6.36	6.11	6.30
7	1	$CH_2CH_2CHMe_2$	5.83	4.54	0	7.01	6.80	6.57
8	1	$CH_2CH_2CH_2CHMe_2$	6.36	5.59	0	6.74	6.89	6.80
9[b]	1	CH_2-*c*-pentyl	5.93	5.42	0	7.06	6.40	6.71
10	1	CH_2-*c*-hexyl	6.49	5.42	0	6.86	7.19	6.93
11	1	C_6H_5	4.88	3.11	0	6.58	6.34	6.54
12	1	$(CH_2)_2C_6H_5$	5.79	3.58	0	7.22	7.33	7.09
13	1	$(CH_2)_2$–	4.20	3.24	0	4.93	5.30	5.79
14	1	$(CH_2)_3$–	4.76	3.82	0	5.64	5.73	6.12
15	1	$(CH_2)_4Me$	5.96	4.94	0	6.49	6.73	6.58
16	0	$CH_2CH_2CH_2CHMe_2$	6.16	4.54	1	5.85	6.10	6.37
17[a]	0	C_6H_5	5.21	3.11	1	6.96	5.65	6.38
18	0	$(CH_2)_2C_6H_5$	6.12	3.58	1	6.89	6.64	6.89
19	2	H	3.44	1.00	0	5.68	5.61	5.18
20[a]	2	C_3H_7	5.01	3.49	0	5.29	6.29	6.25
21	2	C_4H_9	5.54	4.54	0	6.40	6.39	6.54
22	2	C_5H_{11}	6.07	4.94	0	7.08	6.89	6.80
23	2	C_6H_{13}	6.60	5.96	0	6.95	7.01	7.01
24	2	CH_2CHMe_2	5.41	4.45	0	5.92	6.26	6.53
25	2	$CH_2CH_2CHMe_2$	5.94	4.54	0	7.02	6.95	6.79
26	2	$CH_2CH_2CH_2CHMe_2$	6.47	5.59	0	6.81	7.05	7.02
27[b]	2	CH_2-*c*-pentyl	6.05	5.42	0	6.29	6.56	6.93
28	2	C_6H_5	4.99	3.11	0	6.82	6.50	6.77
29	2	–$(CH_2)_2C_6H_5$	5.90	3.58	0	7.29	7.49	7.31
30	2	–$(CH_2)_3$–	4.87	3.82	0	5.89	5.89	6.36

[a] Not included in deriving QSAR 1.
[b] *B*5 value is estimated to be equal to CH_2-*c*-hexyl.

2.3. 3-(S-(2R′,5R″-C_6H_5)), 6-C_6H_5, 6′-R (hydrophobic) substituted 4-hydroxy-5,6-dihydro-pyran-2-ones (3) (Table 3)[23]



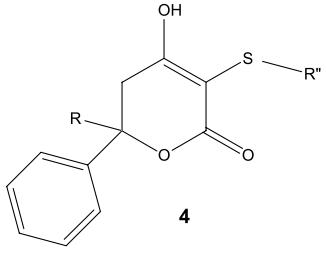
3

The same authors reported anti-HIV-1 data for another series of dihydropyranones, where several alkyl groups were tried at 2′- and 5′-position of thiophenyl ring at C-3 of pyranone (**3**) in order to further improve the antiviral activity. QSAR 3 was developed for these compounds.

$$\mathrm{Log}(1/\mathrm{IC}_{50}) = -0.28(\pm 0.23)\mathrm{Clog}P + 0.73(\pm 0.24)B5_{R'} - 0.92(0.40)I_R + 7.67(\pm 1.45)$$
$$n = 17,\ r = 0.892,\ r^2 = 0.797,\ q^2 = 0.654,\ s = 0.255 \quad (3)$$

The QSAR shows a linear relationship of antiviral activity with the Clog *P*. Its negative sign indicates that highly hydrophobic groups are not good for improving the activity of this series. A positive *B*5 term for R′ substituent was also found significant in our model. The substituent at this position seems to be involved in favorable steric interactions. Indicator variable I_R was used with a value of unity for R = C_6H_5 and zero for R = $(CH_2)_2C_6H_5$. Its negative coefficient suggests that the C_6H_5 group without linker –$(CH_2)_2$– at the 6th position of pyranone would not be able to achieve good antiviral potency. The linker –$(CH_2)_2$– for R-substituent seems to be important for facilitating the access to the binding pocket of the enzyme.

2.4. 3-S-R″, 6-C_6H_5, 6′-R substituted 4-hydroxy-5,6-dihydro-pyran-2-ones (4)[23]



4

Table 2. The physicochemical parameters and the antiviral activity data used for deriving QSAR 2 and 4

S. No.	Substituents		CMR	I_{COOH}	Log ($1/IC_{50}$)		
	n	R			Obsd.[23]	Calcd.	
						QSAR 2	QSAR 4
1	2	$(CH_2)_2COOH$	11.05	1	5.92	5.79	5.93
2	2	$(CH_2)_3COOH$	11.51	1	6.57	6.93	6.19
3[b]	2	$(CH_2)_4COOH$	11.98	1	8.30	8.07	6.57
4	2	$(CH_2)_3CONH_2$	11.73	0	5.89	5.89	5.91
5[a]	2	$(CH_2)_4CONH_2$	12.19	0	6.47	7.03	6.35
6	2	4-Pyridyl	11.77	0	6.10	5.98	6.27
7[a,b]	2	$CH_2N(Me)C_6H_5$	13.28	0	5.64	9.70	7.47
8	1	$CH_2OC_6H_5$	12.13	0	6.76	6.88	6.88
9[a]	1	$(CH_2)_4OH$	11.01	0	5.90	4.12	5.90

[a] Not included in deriving QSAR 2.
[b] Not included in deriving QSAR 4.

Table 3. The physicochemical parameters and the antiviral activity data used for deriving QSAR 3 and 4

S. No.	Substituents			Clog *P*	I_R	$B5_{R'}$	Log ($1/IC_{50}$)		
	R	R′	R″				Obsd.[23]	Calcd.	
								QSAR 3	QSAR 4
1[a]	C_6H_5	H	H	5.21	1	1.00	6.96	6.04	6.38
2	C_6H_5	Me	H	5.71	1	2.04	6.41	6.66	6.65
3	C_6H_5	$CHMe_2$	H	5.94	1	3.17	7.64	7.42	7.09
4	C_6H_5	CHMeEt	H	6.47	1	3.49	7.77	7.50	7.32
5	C_6H_5	*c*-Pentyl	H	6.57	1	4.09	7.85	7.91	7.45
6	C_6H_5	*c*-Hexyl	H	7.13	1	3.49	6.90	7.32	7.63
7	C_6H_5	CMe_3	H	7.03	1	3.17	7.31	7.11	7.29
8	C_6H_5	$CHMe_2$	Me	6.43	1	3.17	7.42	7.28	7.32
9	C_6H_5	$CHMe_2$	$CHMe_2$	7.36	1	3.17	6.92	7.02	7.69
10[a]	C_6H_5	CMe_3	Me	7.53	1	3.17	7.92	6.98	7.45
11	$CH_2C_6H_5$	H	H	6.12	0	1.00	6.89	6.71	6.89
12	$CH_2C_6H_5$	Me	H	6.62	0	2.04	7.14	7.33	7.10
13	$CH_2C_6H_5$	$CHMe_2$	H	6.84	0	3.17	7.85	8.09	7.52
14	$CH_2C_6H_5$	CHMeEt	H	7.37	0	3.49	8.00	8.17	7.68
15[a]	$CH_2C_6H_5$	*c*-Hexyl	H	8.04	0	3.49	7.49	7.99	7.92
16[a]	$CH_2C_6H_5$	CMe_3	H	7.94	0	3.17	8.52	7.78	7.59
17	$CH_2C_6H_5$	$CHMe_2$	Me	7.34	0	3.17	8.15	7.95	7.69
18	$CH_2C_6H_5$	$CHMe_2$	$CHMe_2$	8.27	0	3.17	7.85	7.69	7.95
19	$CH_2C_6H_5$	CMe_3	Me	8.44	0	3.17	8.00	7.65	7.69
20	$CH_2C_6H_5$	CMe_3	$CHMe_2$	8.67	0	3.17	7.33	7.58	8.05
21[a]	$CH_2C_6H_5$(S)	$CHMe_2$	Me	8.04	0	2.04	8.22	6.93	7.69
22	$CH_2C_6H_5$(R)	$CHMe_2$	Me	8.04	0	2.04	6.89	6.93	7.69

[a] Not included in deriving QSAR 3.

The QSAR analyses of individual datasets and study of the hydrophobic parameter resulted in three QSAR models. QSAR 1 has a positive Clog *P* term, QSAR 2 has no Clog *P* term whereas QSAR 3 showed a negative Clog *P* term. We combined all the three datasets having general structure (**4**) and obtained the following QSAR:

$$\begin{aligned}\mathrm{Log}(1/\mathrm{IC}_{50}) &= 0.82(\pm 0.60)\mathrm{Clog}P \\ &\quad - 0.07(\pm 0.06)\mathrm{Clog}P^2 + 0.47(\pm 0.19)\mathrm{CMR} \\ &\quad - 1.28(\pm 2.93)\end{aligned}$$
$$n = 57,\ r = 0.850,\ r^2 = 0.722,\ q^2 = 0.676,\ s = 0.436$$
$$\text{Optimum Clog}P\ (\log P_0) = 6.345 \qquad (4)$$

This QSAR showed a parabolic dependence of antiviral activity on hydrophobicity. The activity first increases with increasing hydrophobicity upto an optimum value ($\log P_0 = 6.345$), and then decreases with further increase. The presence of CMR in the final QSAR 4 (of all the three combined datasets) shows the importance of size dependent polarizability of these molecules for achieving good antiviral activity. QSAR 4 explains 85% variance in the activity. Further jackknifing did not improve the QSAR. Plot of observed versus calculated activity for this QSAR is shown in Figure 1.

Most of the compounds found outliers (the number of the molecules of the dataset that did not fit in the QSAR model) in QSAR 4 were also found misfit when we studied these series independently. The calculated value of these compounds was either too high or too low than the corresponding observed value (see Tables 1–3). This problem of 'misfit' of the congeners in the final QSAR could be associated with any one of the following reasons:

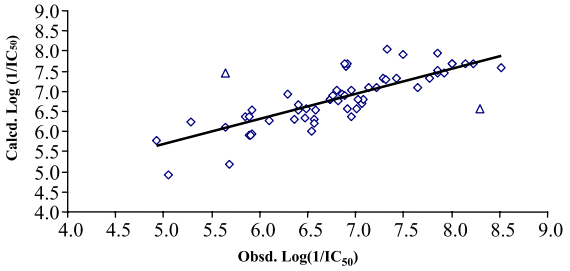


Figure 1. Plot of observed versus calculated antiviral activity for QSAR 4 (△ represents outliers).

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

Therefore, one has to expect outliers that must not be forgotten because many times they can lead to new understanding. To cover them up by including in a QSAR, at the cost of a lower correlation coefficient (r), 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. Variations in these properties of the substituents at each position of the parent structure are required to make sure that all of 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 dataset for investigation.

Overall study of QSAR 1–4, shows that the antiviral activity of dihydropyranones under study depends on its hydrophobic and volume dependent polar interactions with the receptor. Calculated molecular refractivity (CMR) and the Verloop's width parameter ($B5$) along with the hydrophobic parameter Clog$P$ were found the most important parameters in our QSAR models.

Presence of CMR parameter in a QSAR has been associated with size as well as polarizability and also points toward the involvement of electronic effects beside steric interactions.[30] This parameter was also found significant in QSAR models of other structurally diverse classes of HIV-PR inhibitors.[14,19,20] Interaction of the inhibitor with the enzyme pocket of protease is critical for potent binding. The HIV-PR is a dimer with an elliptical open-ended cylindrical active site formed as well-defined sub-sites (S_1, S'_1, S_2, S'_2, S_3, S'_3) and not as clear sub-sites (S_4, S'_4 and S_5, S'_5) containing mostly hydrophobic residues.[31–34] The active site of HIV-1 protease can accommodate 6–8 amino acids with its 18 potential H-bond donors/acceptors.[31–34] The lone pair electrons on the oxygen and nitrogen atom of the polar groups in the congeners will have an appreciable effect on H-bonding with surrounding amino acids, and positive coefficient of CMR is a good indicator of volume dependent polarizability. The positive value of CMR also gives hint that the bulk of substituents may assist in holding the ligand in place or cause a conformational change in receptor structure.[35] This is also supported by the presence of a negative $B5$ term in QSAR 1, which suggests steric hindrance in binding with the receptor of bigger substituents at 6th position of pyranone. The positive $B5$ term in QSAR 3 suggests that the bigger groups may be better substituents in the S-aryl group at C-3 of pyranone.

HIV-PR binding site has a hydrophobic binding domain.[22] However, the site seems to have an optimum size as indicated by the presence of parabolic Clog$P$ in several QSAR models.[14,19] In the present study, HIV protease inhibitory activity of the smaller set of compounds showed linear relationship with Clog$P$ (QSAR 1 & 3) whereas the combined dataset of all compounds exhibited a parabolic Clog$P$ term (QSAR 4; optimum log$P$ = 6.345). A brief summary is given in Table 4.

Closer inspection of all the datasets used for developing QSAR 1–3 showed that there was insufficient spread in the parameter value of Clog$P$ to reveal the optimum value (6.345) observed in QSAR 4. The first dataset which gave QSAR 1 (with a positive Clog$P$ term) has most of the compounds with a Clog$P$ value <6.345 (Table 1). The second dataset used for developing QSAR 2 (with no Clog$P$ term) has mostly polar substituents (Table 2). The third dataset that showed QSAR 3 (with a negative Clog$P$ term) has most of the compounds with a Clog$P$ value >6.345 (Table 3). As shown in Figure 2, only the combined dataset has sufficient spread in the range of parameter values of Clog$P$ to reveal the optimum value.

Table 4. Summary of QSAR 1–4

QSAR models	Compounds	Log (1/IC_{50}) range	Clog$P$ range	Comments
QSAR 1	Hydrophobic groups (**2**)	4.93–7.29	3.32–6.60	Linear positive Clog$P$ term
QSAR 2	Polar groups (**2**)	5.64–8.30	2.68–5.70	No Clog$P$
QSAR 3	Hydrophobic groups (**3**)	6.41–8.52	5.21–8.67	Linear negative Clog$P$ term
QSAR 4	Combined dataset (**4**)	4.93–8.52	2.68–8.67	Parabolic Clog$P$ term (Opt. log$P$ = 6.345)

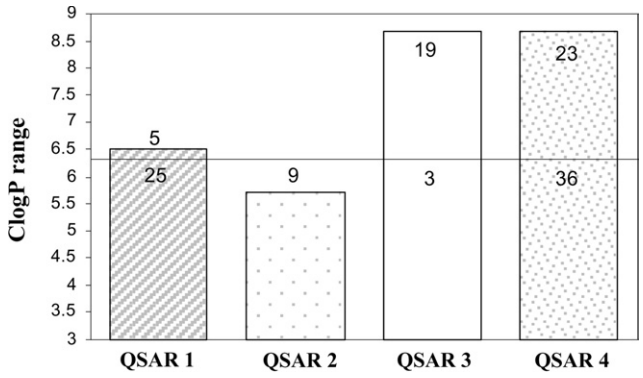


Figure 2. The range of Clog *P* in different QSAR models. The number of data points above and below the optimum Clog *P* (6.345) line in each dataset is given.

These new comparative QSAR studies show that most of the data points in a dataset often fall either on positive or negative side of the optimum Clog *P* value. This is why, we do not observe an optimum, but either a positive or negative term in the QSAR. To observe the optimum value of a parameter, a sufficient spread in the data is required. More such studies are in progress to fully establish this hypothesis and will be reported in separate communications.

4. Conclusion

Several studies have proved that the HIV-PR binding pockets have hydrophobic residues.[31–34] Earlier QSAR studies on HIV-PIs revealed that there is an optimum value of Clog *P* ($\log P_0$), which an inhibitor is required to have for potent inhibition.[14,19] It is important to note that in our previous reports[14,19] the $\log P_0$ range for protease inhibitors was found to be from 4.49 to 6.96. The optimum (6.345) observed in QSAR 4 in the present study also fall in this range. The majority of HIV research is done with cells and isolated receptors and these studies tend to overestimate the value of $\log P_0$ (i.e., over 1 or 2 log unit higher) for the whole organism.[26] Nonetheless, we found a close agreement in the Clog *P* values of US Food and Drug Administration (FDA) approved HIV-PI[36] and the optimum log *P* observed in our QSAR models as shown below:

FDA approved protease inhibitors[a]	Clog *P*[c]
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
7. Tipranavir[b]	7.76

Recently approved Kaletra® is a combination of Lopinavir and Ritonavir.

[a] All FDA approved HIV-PI are peptidic in nature.

[b] First nonpeptidic dihydropyranone HIV-PI in clinical trial.

[c] Calculated using CQSAR program, Biobyte Corp., Claremont, CA.

These results clearly show that the comparative QSAR studies provide valuable insight in understanding complex biological interactions. We hope that this study would help in optimizing substituents for better binding at HIV-PR pockets and guide in the design of more effective inhibitors.

5. Experimental

All the HIV-1 inhibitory activity data have been collected from the literature.[23] Antiviral potency IC_{50} is the effect of the compounds on the 50% inhibition of accumulation of viral RNA transcripts 7 days after infection of H-9 cells by measuring RT assay.[23,37] The QSAR multiple linear regression analyses were executed with the CQSAR program and all the physicochemical parameters were auto loaded.[26] For details about the utility of CQSAR program in comparative QSAR analyses, see Refs. 13 and 30. The physicochemical parameter—Clog *P* is the calculated partition coefficient in octanol/water and is a measure of the hydrophobicity of the molecule.[27] It explains: (a) hydrophobic interactions between ligand and receptor, and (b) random walk process in movement of the drug molecule in the organism from site of injection to sites of action. Clog *P* is for the *neutral form* of acids and bases that may be partially ionized. If the degree of ionization is about the same for a set of congeners, the ionization factor can be neglected. Otherwise, good correlations can be obtained using electronic terms. CMR is the calculated molar refractivity for the whole molecule and is a measure of volume and polarizability.[13] 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. *B*5 is the Verloop's sterimol parameter that defines maximum width of the substituents.[29] The indicator variable *I* is assigned the value of 1 or 0 for special affects that cannot be parameterized and has been explained wherever used.

Acknowledgements

The authors are indebted to Clarkson University, Chemistry Department for the use of their facilities and Biobyte Corp., Claremont, CA for the use of CQSAR program. This study was supported by grant 1 R15 GM 069323-01 from NIH/NIGMS.

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. De Clercq, E. *Med. Res. Rev.* **2002**, *22*, 531.
11. Gulick, R. M. *Clin. Microbiol. Infect.* **2003**, *9*, 186.
12. Strohbach, J. W.; Chrusciel, R. A. *Curr. Top. Med. Chem.* **2004**, *4*, 1097.
13. Hansch, C.; Leo, A. *Exploring QSAR Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, DC, 1995.
14. Garg, R.; Gupta, S. P.; Gao, H.; Mekapati, S. B.; Debnath, A. K.; Hansch, C. *Chem. Rev.* **1999**, *99*, 3525.
15. Gupta, S. P.; Babu, M. S.; Garg, R.; Sowmya, S. *J. Enzyme Inhib.* **1998**, *13*, 399.
16. Gupta, S. P.; Babu, M. S.; Kaw, N. *J. Enzyme Inhib.* **1999**, *14*, 109.
17. Gupta, S. P.; Babu, M. S.; Sowmya, S. *Bioorg. Med. Chem.* **1998**, *6*, 2185.
18. Gayathri, P.; Pande, V.; Sivakumar, R.; Gupta, S. P. *Bioorg. Med. Chem.* **2001**, *9*, 3059.
19. Kurup, A.; Mekapati, S. B.; Garg, R.; Hansch, C. *Curr. Med. Chem.* **2003**, *10*, 1819.
20. Garg, R.; Bhhatarai, B. *Bioorg. Med. Chem.* **2004**, *12*, 5819.
21. Selassie, C. D.; Garg, R.; Kapur, S.; Kurup, A.; Verma, R. P.; Mekapati, S. B.; Hansch, C. *Chem. Rev.* **2002**, *102*, 2585.
22. Babine, R. E.; Bender, S. L. *Chem. Rev.* **1997**, *97*, 1359.
23. Tait, B. D.; Hagen, S.; Domagala, J.; Ellsworth, E. L.; Gajda, C.; Hamilton, H. W.; Vara Prasad, J. V. N.; Ferguson, D.; Graham, N.; Hupe, D.; Nouhan, C.; Tummino, P. J.; Humblet, C.; Lunney, E. A.; Pavlovsky, A.; Rubin, J.; Gracheck, S. J.; Baldwin, E. T.; Bhat, T. N.; Erickson, J. W.; Gulnik, S. V.; Liu, B. *J. Med. Chem.* **1997**, *40*, 3781.
24. Mehandru, S.; Markowitz, M. *Expert Opin. Invest. Drugs* **2003**, *12*, 1821.
25. Cramer, R. D., III; Bunce, J. D.; Patterson, D. E.; Frank, I. E. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18.
26. CQSAR program, Biobyte Corp., 201 W. 4th street, Claremont, CA 91711.
27. Leo, A. *Chem. Rev.* **1993**, *93*, 1281.
28. Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR Hydrophobic, Electronic, and Steric Constants*; American Chemical Society: Washington, DC, 1995.
29. Verloop, A. *The Sterimol Approach to Drug Design*; Marcel Dekker: New York, NY, 1987.
30. Hansch, C.; Steinmetz, W. E.; Leo, A. J.; Mekapati, S. B.; Kurup, A.; Hoekman, D. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 120.
31. Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157.
32. Graves, B. J.; Hatada, M. H.; Miller, J. K.; Graves, M. C.; Roy, S.; Cook, C. M.; Krohn, A.; Martin, J. A.; Roberts, N. A. In *Structure and Function of the Aspartic Proteinases*; Dunn, B., Ed.; Plenum: New York, 1992; p 455.
33. Thompson, W. J.; Fitzgerald, P. M. D.; Holloway, M. K.; Emini, E. A.; Darke, P. L.; McKeever, B. M.; Schleif, W. A.; Quintero, J. C.; Zugay, J. A.; Tucker, T. J.; Schwering, J. E.; Homnick, C. F.; Nunberg, J.; Springer, J. P.; Huff, J. R. *J. Med. Chem.* **1992**, *35*, 1685.
34. Wang, S.; Milne, G. W.; Yan, X.; Posey, I. J.; Nicklaus, M. C.; Graham, L.; Rice, W. G. *J. Med. Chem.* **1996**, *39*, 2047.
35. Hansch, C.; Klein, T. *Acc. Chem. Res.* **1986**, *19*, 392.
36. http://www.niadi.nih.gov/daids/dtpdb/FDADRUG.HTM (accessed 12/10/2004).
37. Vara Prasad, J. V. N.; Para, K. S.; Tummino, P. J.; Ferguson, D.; McQuade, T. J.; Lunney, E. A.; Rapundalo, S. T.; Batley, B. L.; Hingorani, G.; Domagala, J. M.; Gracheck, S. J.; Bhat, T. N.; Liu, B.; Baldwin, E. T.; Erickson, J. W.; Sawyer, T. K. *J. Med. Chem.* **1995**, *38*, 898.