Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 15 (2005) 3767–3770
Bioorganic & Medicinal Chemistry Letters

Hydrophobicity in the design of P2/P2′ tetrahydropyrimidinone HIV protease inhibitors

Rajni Garg* and Disha Patel

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

Received 29 April 2005; revised 13 May 2005; accepted 16 May 2005
Available online 1 July 2005

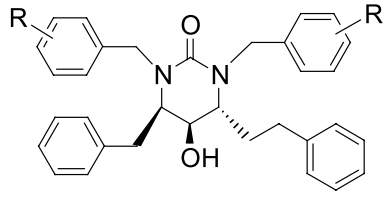
Abstract—As part of an ongoing effort in understanding the role of hydrophobicity in the design of nonpeptidic HIV protease inhibitors, the QSAR study on P2/P2′ tetrahydropyrimidinone is presented in this report. Our results suggest that the balance of hydrophobicity and a volume dependent polarizability term plays a key role in the inhibition of the viral protease by these inhibitors. The size of the substituent of ligands at particular positions which induce steric fit is crucial. The role of hydrophobicity in the design of tetrahydropyrimidinone is discussed. It has been found that a sufficient spread in the data is required to observe the optimum value of C log *P* for these inhibitors.

Highly active anti-retroviral therapy (HAART) has made dramatic impacts on the mortality and morbidity associated with human immunodeficiency virus (HIV) infection. However, HIV is still a major concern worldwide because the use of current drugs regimens is compounded by many problems such as long-term toxicity, drug-resistance, and mutations that enable HIV to resist currently available treatments.[1] Therefore, there is a growing need for the development of new chemotherapeutics with improved antiviral potency and favorable pharmacokinetic profiles. HIV protease (HIVPR) inhibitors are a critical component of this therapy.

Several in silico techniques are utilized in the process of drug design and development of HIV protease inhibitors (HIVPI). One such technique is quantitative structure–activity relationship (QSAR). QSAR models reveal significant correlations between the biological activity and physicochemical parameters. This quantitative technology can be utilized to improve the structure of the inhibitor molecule and to interpret the improved structure in terms of favorable biological interactions.[2]

Robust models and their relevance to possible modes of action provide a better understanding of protein–inhibitor interactions and structure-based drug design. QSAR models lead to assessment of the specific effects of various types of substituents reducing trial experiments and investments. It also helps in identifying possible expensive failures in the early stage of the drug design and discovery process and eliminates them as outliers.

QSAR studies have provided valuable insight in the design and development of HIVPI.[3–6] Recently, we have been interested in understanding the role of hydrophobicity in the design of nonpeptidic HIVPI using this technique.[7,8] As part of an ongoing effort the QSAR study on P2/P2′ tetrahydropyrimidinone (**1**) HIVPI reported by De Lucca et al.[9] is presented.



1

All the HIV inhibitory data have been taken from the literature.[9] The biological activity K_i is the HIV 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

Keywords: Hydrophobicity; QSAR; P2/P2′ tetrahydropyrimidinone; C log *P*; MR.
* Corresponding author. Tel.: +1 315 268 4421; fax: +1 315 268 6610; e-mail: rgarg@clarkson.edu

0960-894X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2005.05.087

transcripts 3 days after infection of MT-2 cells with HIV-1 RF (a heterologous strain with mutation observed in the V3 hypervariable region of HIV envelop glycoprotein gp120).[9]

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 is calculated as described by Cramer et al.[10] The data within parentheses are for the 95% confidence intervals. The QSAR multiple linear regression analyses were executed with the CQSAR program and all the physicochemical parameters were auto loaded.[11] For details about the utility of the CQSAR program see Refs. 12,13. The physicochemical parameter—C$\log P$ is the calculated partition coefficient in octanol/water and is a measure of the hydrophobicity of the molecule.[14] MR is the calculated molar refractivity and 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. The indicator variable I is assigned the value of 1 or 0 for special effects that cannot be parameterized and has been explained wherever used.

De Lucca et al.[9] reported the structure–activity data for N-substituted P2/P2′ tetrahydropyrimidinone (**1**) HIVPI that can interact with the lipophilic S2/S2′ enzyme sites. Both the biological activity K_i (enzyme inhibition) and IC_{90} (antiviral activity) were found to be significantly correlated with C$\log P$ and MR_4 as shown in QSAR 1 and 2 (Table 1).

$$\begin{aligned}\mathrm{Log}(1/K_i) &= -0.738(\pm 0.416)\mathrm{C}\log P \\ &\quad - 1.997(\pm 1.114)\mathrm{MR}_4 \\ &\quad + 14.029(\pm 2.633) \\ n &= 12,\ r = 0.911,\ r^2 = 0.830, \\ q^2 &= 0.716,\ s = 0.595. \end{aligned} \quad (1)$$

$$\begin{aligned}\mathrm{Log}(1/\mathrm{IC}_{90}) &= -0.726(\pm 0.357)\mathrm{C}\log P \\ &\quad - 2.371(\pm 1.281)\mathrm{MR}_4 \\ &\quad + 11.500(\pm 2.274) \\ n &= 10,\ r = 0.912,\ r^2 = 0.832, \\ q^2 &= 0.606,\ s = 0.359. \end{aligned} \quad (2)$$

The negative coefficient of C$\log P$ in QSAR 1 and 2 indicates that the more hydrophobic molecule will have a detrimental effect on the biological activity. As mentioned earlier, C$\log P$ is the calculated partition coefficient of the molecule in octanol/water and measures its hydrophobicity.[14] The substituent at the 4th position of the P2/P2′ benzyl group seems to be involved in unfavorable polarizability dependent steric interactions with the receptor binding site as shown by the presence of a negative MR_4 term. There was no mutual correlation between the two parameters C$\log P$ and MR_4 ($r^2 = 0.099$).

Compounds **9** and **11** (Table 1) were not included in the derivation of QSAR 1 and 2 because they exhibited aberrant behavior. The calculated value of these compounds was either too high or too low than the corresponding observed value. This problem of 'misfit' of the congeners in the final QSAR could be associated with any one of the following reasons:

- 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.

In order to have a good HIV antiviral activity, the compound needs to have good protease inhibitory activity. QSAR 1 and 2 are quiet parallel and suggest that the enzyme inhibitory activity (K_i) translates well to antiviral activity (IC_{90}). The regression analysis was found

Table 1. The physicochemical parameters and the biological data used for deriving QSAR 1 and 2

S. No.	R	C$\log P$	MR_4	$\log(1/K_i)$ QSAR 1		$\log(1/IC_{90})$ QSAR 2	
				Obsd[9]	Calcd	Obsd[9]	Calcd
1	H	7.70	0.10	7.82	8.13	5.55	5.67
2	3-CN	6.89	0.10	7.96	8.73	5.80	6.25
3	3-$COOCH_3$	7.90	0.10	8.82	7.99	—	—
4	3-OH	6.37	0.10	9.52	9.12	6.94	6.63
5	3-CH_2OH	5.63	0.10	9.31	9.67	6.96	7.17
6	3-NH_2, 4-F	5.98	0.09	9.60	9.43	7.44	6.95
7	3-$CONH_2$	4.89	0.10	10.06	10.21	7.46	7.71
8	3-C(=NOH)NH_2	5.47	0.10	10.74	9.79	7.31	7.29
9[a]	3-NH_2	5.25	0.10	8.77	9.95	6.44	7.45
10	3-$NHCH_3$	6.70	0.10	8.31	8.87	6.48	6.39
11[a]	3-$CONH_2$, 4-F	4.71	0.09	8.85	10.37	6.31	7.87
12	3-F, 4-NH_2	5.98	0.54	8.10	8.53	6.27	5.88
13	4-CH_2OH	5.63	0.72	8.60	8.44	5.44	5.71
14	4-$COOCH_3$	7.90	1.29	5.71	5.63	—	—

[a] Not included in deriving QSAR 1 and 2.

to reveal a significant correlation between these two biological parameters (QSAR 3).

$$\begin{aligned}\mathrm{Log}(1/\mathrm{IC}_{90}) &= 0.666(\pm 0.292)\mathrm{Log}(1/K_i) \\ &\quad + 0.555(\pm 2.633) \\ n &= 12,\ r = 0.849,\ r^2 = 0.721, \\ q^2 &= 0.582,\ s = 0.390. \end{aligned} \quad (3)$$

De Lucca et al.[9] also studied the effect of stereochemistry on the HIVPR inhibitory activity and reported that the RRR isomer of (**1**) is more active than SSS isomer. Analysis of the data gave QSAR 4 (Table 2). A highly significant positive coefficient of the indicator variable I that was used with a value of unity for the RRR isomer also shows that this isomer is more potent than the SSS isomer.

$$\begin{aligned}\mathrm{Log}(1/K_i) &= 3.726(\pm 1.229)I + 5.653(\pm 0.869) \\ n &= 8,\ r = 0.927,\ r^2 = 0.859,\ q^2 = 0.780, \\ s &= 0.843. \end{aligned} \quad (4)$$

Inclusion of an electronic term for the substituent at the 3rd position of P2/P2′ benzyl improved the correlation significantly (QSAR 5). It indicates that the electron donating substituents at this position will improve the enzyme inhibitory activity as shown by the presence of a negative σ (Hammett electronic constant) term. These substituents may be involved in some charge transfer phenomenon with the amino acid residues at the active site. It is of note that although the σ term is very weak in QSAR 5, still it is important for enhancing the potency of the inhibitor as evident by the improved correlation. A small number of data in the set and narrow range in parameter values precluded a more detailed analysis.

$$\begin{aligned}\mathrm{Log}(1/K_i) &= 3.498(\pm 0.982)I - 2.245(\pm 2.442)\sigma_{R,3} \\ &\quad + 6.213(\pm 1.009) \\ n &= 8,\ r = 0.973,\ r^2 = 0.947,\ q^2 = 0.874, \\ s &= 0.540. \end{aligned} \quad (5)$$

Analysis of these results suggests that the balance of hydrophobicity (C$\log P$) and the volume dependent polarizability (MR) term plays a key role in the inhibition of the viral protease. The size of the substituent of ligands at particular positions which induce steric fit is crucial. Our results indicate that the various substituents used to target individual pockets need to be designed cautiously to achieve better inhibition and an improved pharmacokinetic profile.

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 substituent 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.

HIV protease binding sites have a hydrophobic binding domain.[15] However, the site seems to have an optimum size as indicated by the presence of parabolic and bilinear correlations of C$\log P$ in earlier reports[5,6] and our recent publications.[7,8] These studies have revealed that there is an optimum value of C$\log P$ ($\log P_0$) which HIV-PI is required to have for potent inhibition.[5–8] QSAR 1 and 2 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.

It is important to note that in our previous reports,[5–8] the $\log P_0$ range for protease inhibitors was found to be from 4.49 to 6.96. Investigation of the data in Table 1 shows that compound **8** with a C$\log P$ value 5.47 is one of the most potent ($\log 1/K_i = 10.74$; $\log 1/\mathrm{IC}_{90} = 7.31$). A closer look at the C$\log P$ value of all the compounds studied in Table 1 shows that except for three compounds all others have a C$\log P$ value >5.47. It is clear that there is insufficient spread in the range of C$\log P$ values to establish the optimum point for antiviral potency or enzyme inhibition for this data set. To observe the optimum value of a parameter, a sufficient spread with a wide range in the data is required.

Table 2. The physicochemical parameters & the biological data used for deriving QSAR 5

S. No.	R	$\sigma_{R,3}$	I	$\log(1/K_i)$	
				Obsd[9]	Calcd QSAR 5
1	3-CN (A)	0.56	1	7.96	8.45
2	3-$COOCH_3$ (A)	0.36	1	8.82	8.90
3	3-CH_2OH (A)	0.00	1	9.31	9.71
4	3-$CONH_2$ (A)	0.28	1	10.06	9.08
5[a]	3-C(=NOH)NH_2 (A)	—	1	10.74	—
6	3-CN (B)	0.56	0	5.10	4.96
7	3-$COOCH_3$ (B)	0.36	0	5.48	5.40
8	3-CH_2OH (B)	0.00	0	6.26	6.21
9	3-$CONH_2$ (B)	0.28	0	5.32	5.58
10[a]	3-C(=NOH)NH_2 (B)	—	0	6.11	—

(A) R,R,R isomer.
(B) S,S,S isomer.
[a] Not included in deriving QSAR 5 because σ values are not available.

Most of the data points in this set fall on the negative side of the optimum and that is why we may be observing a negative Clog P term in the QSAR 1 and 2.

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 units higher) as compared with the whole organism.[11] Nonetheless, we found a close agreement in the Clog$P$ values of US Food and Drug Administration (FDA) approved HIV-PI[16] and the optimum log P (log P_0) observed in our QSAR models[5–8] as shown below:

FDA Approved Protease Inhibitors[a]	Clog P[b]
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®)	6.09

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

[a] All FDA approved HIVPI are peptidic in nature.

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

Furthermore, research on P2/P2′ tetrahydropyrimidinone analogues that are now obsolete may provide valuable insight into the binding pattern regarding its substituents' interaction with enzyme binding pockets. We hope that our results will provide useful clues for further developing these inhibitors and current drugs in clinical trails for improved antiviral activity, enzyme inhibition, and pharmacokinetic profile.

Acknowledgments

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

References and notes

1. Chrusciel, R. A.; Strohbach, J. W. *Curr. Top. Med. Chem* **2004**, *4*, 1097.
2. Hansch, C.; Leo, A. *Exploring QSAR, Fundamentals and Applications in Chemistry and Biology*; American Chemical Society: Washington, DC, 1995.
3. Gupta, S. P.; Babu, M. S.; Garg, R.; Sowmya, S. *J. Enzyme Inhib.* **1998**, *13*, 399.
4. Gayathri, P.; Pande, V.; Sivakumar, R.; Gupta, S. P. *Bioorg. Med. Chem.* **2001**, *9*, 3059.
5. Garg, R.; Gupta, S. P.; Gao, H.; Mekapati, S. B.; Debnath, A. K.; Hansch, C. *Chem. Rev.* **1999**, *99*, 3525.
6. Kurup, A.; Mekapati, S. B.; Garg, R.; Hansch, C. *Curr. Med. Chem* **2003**, *10*, 1819.
7. Garg, R.; Bhhatarai, B. *Bioorg. Med. Chem.* **2004**, *12*, 5819.
8. Bhhatarai, B.; Garg, R. *Bioorg. Med. Chem.* **2005**, *13*, 4078.
9. De Lucca, G. V.; Liang, J.; Aldrich, P. E.; Calabrese, J.; Cordova, B.; Klabe, R. M.; Rayner, M. M.; Chang, C.-H. *J. Med. Chem.* **1997**, *40*, 1707.
10. Cramer, R. D., III; Bunce, J. D.; Patterson, D. E.; Frank, I. E. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18.
11. CQSAR program, Biobyte Corp., 201 W. 4th street, Claremont, CA 91711.
12. Hansch, C.; Hoekman, D.; Leo, A.; Weininger, D.; Selassie, C. D. *Chem. Rev.* **2002**, *102*, 783.
13. Selassie, C. D.; Garg, R.; Kapur, S.; Kurup, A.; Verma, R. P.; Mekapati, S. B.; Hansch, C. *Chem. Rev.* **2002**, *102*, 2585.
14. Leo, A. *Chem. Rev.* **1993**, *93*, 1281.
15. Babine, R. E.; Bender, S. L. *Chem. Rev.* **1997**, *97*, 1359.
16. http://www.niadi.nih.gov/daids/dtpdb/FDADRUG.HTM (accessed 04/28/2005).