Journal of Enzyme Inhibition and Medicinal Chemistry, 2011; 26(2): 288–294

ISSN 1475-6366 print/ISSN 1475-6374 online
DOI: 10.3109/14756366.2010.496364

informa
healthcare

SHORT COMMUNICATION

Prediction of HIV-1 protease inhibitory activity of 4-hydroxy-5,6-dihydropyran-2-ones: QSAR study

V. Ravichandran[1,2], V. K. Mourya[3], and R. K. Agrawal[1]

[1]Pharmaceutical Chemistry Research Laboratory, Department of Pharmaceutical Sciences, Dr. Hari Singh Gour University, Sagar, Madhya Pradesh, India, [2]Department of Pharmacy, AIMST University, Semeling, Malaysia, and [3]Government College of Pharmacy, Osmanpura, Aurangabad, Maharashtra, India

Abstract

Inhibition of human immunodeficiency virus 1 (HIV-1) protease is an important strategy for the treatment of HIV and acquired immune deficiency syndrome (AIDS). Therefore, HIV-1 protease inhibitory activity of dihydropyranone derivatives has been analyzed with different physico-chemical parameters. In the present work, QSAR studies were performed on a series of 4-hydroxy-5,6-dihydropyran-2-ones to explore the physico-chemical parameters responsible for their HIV-1 protease inhibitory activity. Physico-chemical parameters were calculated using WIN CAChe 6.1. Stepwise multiple linear regression analysis was performed to derive QSAR models which were further evaluated for statistical significance and predictive power by internal and external validation. The selected best QSAR model was having correlation coefficient (R) = 0.875 and cross-validated squared correlation coefficient (Q^2) = 0.707. The developed significant QSAR model indicates that hydrophobicity of whole molecule and the substituent present at sixth position of dihydropyranones play an important role in the HIV-1 protease inhibitory activities of 4-hydroxy-5,6-dihydropyran-2-ones.

Keywords: QSAR, HIV-1 protease inhibitory activity, multiple linear regressions, 4-hydroxy-5, 6-dihydropyran-2-ones

Introduction

Acquired immune deficiency syndrome (AIDS) is a formidable pandemic that is still wreaking havoc world wide. The catastrophic potential of this virally caused disease may not have been fully realized. The causative moiety of the disease is human immunodeficiency virus (HIV), which is a retrovirus of the lentivirus family[1]. The three viral enzymes, reverse transcriptase, protease, and integrase, encoded by the gag and gag-pol genes of HIV play an important role in the virus replication cycle. Among them, viral protease catalyzes the formation of viral functional enzymes and proteins necessary for its survival. The viral particles at this stage are called virions. The virus particles after the protease action have all the necessary constituents of mature virus and are capable of invading other T4 cells and repeating the life cycle of proviral DNA from viral RNA, the key stage in viral replication. Its central role in virus maturation makes protease is a prime target for anti-HIV-therapy[2].

QSAR analyses of HIV-1 reverse transcriptase inhibitors[3], HIV-1 protease inhibitors[4–7], and HIV-1 integrase inhibitors[8,9] were reported. The present group of authors has developed a few quantitative structure-activity relationship models to predict biological activity of different group of compounds[10–16]. In continuation of such efforts, in this article, we have performed QSAR analysis for HIV-1 protease inhibitory activity of 4-hydroxy-5, 6-dihydropyran-2-one derivatives[17] using modelling software WIN CAChe 6.1 (molecular modelling software) and statistical software STATISTICA 6.

The purpose of the present study is to investigate the physico-chemical parameters responsible for the HIV-1

Address for Correspondence: Professor R. K. Agrawal, Pharmaceutical Chemistry Research Laboratory, Department of Pharmaceutical Sciences, Dr. H. S. Gour Vishwavidyalaya, Sagar, Madhya Pradesh, 470 003, India. Tel: +91 7582-265457. Fax: +91 7582-264236. E-mail: dragrawal2001@yahoo.co.in or phravi75@rediffmail.com

(Received 08 May 2010; revised 00 00 0000; accepted 24 May 2010)

protease inhibitory activity of 4-hydroxy-5,6-dihydropyran-2-ones, explore the correlation between them, and expected to get more information for designing novel dihydropyranone derivatives with potent HIV-1 protease inhibitory activity.

There is high structural diversity and a sufficient range of the biological activity in the selected series of 4-hydroxy-5,6-dihydropyran-2-one derivatives. It insists as to select these series of compounds for our QSAR studies. We carried out QSAR analysis and established a QSAR model to guide further structural optimization and predict the potency and physiochemical properties of clinical drug candidates.

Materials and methods

Materials

All of the molecular modelling studies reported herein were performed using WIN CAChe 6.1 (product of Fujitsu Private Limited, Nagano, Japan, http://www.cachesoftware.com /contacts/japan.shtml) and Chem Draw Ultra version 8 (Cambridge Soft, Swann Road, CB5 8LA, UK.) modelling software and the QSAR models were executed with STATISTICA version 6 (Softstat, Inc., Tulsa, OK) software.

Biological data

In the present work we have taken 59 4-hydroxy-5,6-dihydropyran-2-one compounds (Table 1) and their HIV-1 protease inhibitory activity from the reported work[17]. All the HIV-1 protease inhibitory activities used in the present study were expressed as $pIC_{50} = -\log_{10} IC_{50}$; where IC_{50} is the micro molar concentration of the compounds producing 50% reduction in the HIV-1 protease activity is stated as the means of at least two experiments. The compounds which did not show confirmed HIV-1 protease inhibitory activity and the compounds having particular functional groups present only once at a particular position in the above cited literature have not been taken for our study.

Optimization of molecules structure

All the 59 compounds (50 compounds in training set and 9 in test set—test and training set compounds were chosen manually such that low, moderate, and high activity compounds were present approximately equal proportions in both sets[18]) were built on workspace of molecular modelling software WIN CAChe 6.1. The energy minimization was done by geometry optimization of molecules using molecular mechanics followed by semi-empirical PM3 method available in molecular orbital package module until the root mean square gradient value becomes smaller than 0.001 kcal/mol Å. Most stable structure for each compound was generated and used for calculating various physico-chemical descriptors like thermodynamic, steric, and electronic values of descriptors.

Descriptors calculation

The physico-chemical properties were calculated on project leader file of the modelling software WIN CAChe 6.1. In the present study, the calculated descriptors were conformational minimum energies, zero-order connectivity index, first-order connectivity index, second-order connectivity index, dipole moment, total energy at its current geometry after optimization of structure, heat of formation at its current geometry after optimization of structure, highly occupied molecular orbital, low unoccupied molecular orbital, octanol-water partition coefficient (logP), squared partition coefficient ($logP^2$), molar refractivity, shape index order 1, shape index order 2, zero-order valance connectivity index, first-order valance connectivity index, second-order valance connectivity index, and solvent accessible surface area (physico-chemical parameters data will be provided on request). We have also considered some indicator variables for QSAR model development. Those are IR (if the substituent at R is $-(CH_2)_2C_6H_5$ then the value of IR is 1 otherwise 0) and IR′ (if any substituent is present at R′ then the value of IR′ is 1 otherwise 0).

QSAR models development and validation

The above properties were fed manually into statistical software, STATISTICA version 6, and a correlation matrix was made to select the parameters having very less intercorrelation and maximum correlation with activity. This was followed by a stepwise multiple linear regression analysis to achieve best model. Statistical measures used were number of compounds in regression (n), correlation coefficient (R), squared correlation coefficient (R^2), F-test (Fischer's value) for statistical significance, standard error of estimation (SEE), cross validated correlation coefficient (Q^2), and correlation matrix to show intercorrelation among the parameters.

Internal validation was carried out by leave one out (LOO) method using statistical software STATISTICA. Q^2 was calculated using the following equation:

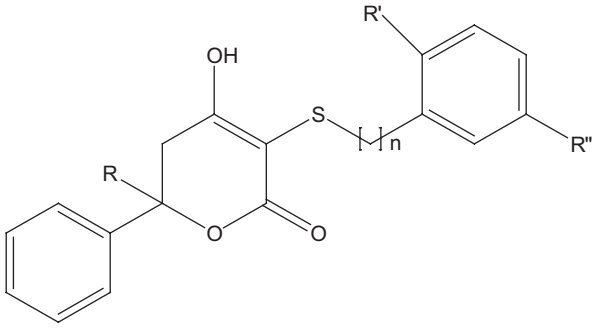
$$Q^2 = 1 - \frac{\text{PRESS}}{\sum_{i=1}^{N} (y_i - y_m)^2},$$

$$\text{PRESS} = \sum_{i=1}^{N} (y_{\text{pred,i}} - y_i)^2,$$

where y_i is the activity for training set compounds, y_m is the mean observed value, corresponding to the mean of the values for each cross-validation group, and $y_{\text{pred,i}}$ is the predicted activity for y_i. The predictive ability of the selected model was also confirmed by external R^2 and R^2CV_{ext}.

$$R^2\,CV_{ext} = 1 - \frac{\sum_{i=1}^{test} (y_{exp} - y_{pred})^2}{\sum_{i=1}^{test} (y_{exp} - y_{tr})^2}$$

Table 1. Structure, selected parameters, and their HIV-1 protease inhibitory activity of 4-hydroxy-5,6-dihydropyran-2-one analogues.



Compound No.	*N*	*R*	*R'*	*R"*	logP	IR	IR′
1[a]	1	H	H	H	2.356	0.000	0.000
2	1	C_3H_7	H	H	3.299	0.000	0.000
3	1	C_4H_9	H	H	3.695	0.000	0.000
4	1	C_5H_{11}	H	H	4.091	0.000	0.000
5	1	C_6H_{13}	H	H	4.488	0.000	0.000
6	1	CH_2CHMe_2	H	H	3.629	0.000	0.000
7	1	$CH_2CH_2CHMe_2$	H	H	4.026	0.000	0.000
8	1	$CH_2CH_2CH_2CHMe_2$	H	H	4.422	0.000	0.000
9	1	CH_2-c-pentyl	H	H	3.918	0.000	0.000
10	1	CH_2-c-hexyl	H	H	4.314	0.000	0.000
11	1	C_6H_5	H	H	3.868	0.000	0.000
12	1	$(CH_2)_2C_6H_5$	H	H	4.516	1.000	0.000
13[a]	0	$CH_2CH_2CH_2CHMe_2$	H	H	4.327	0.000	0.000
14	0	C_6H_5	H	H	3.773	0.000	0.000
15	0	$(CH_2)_2C_6H_5$	H	H	4.421	1.000	0.000
16	2	H	H	H	2.607	0.000	0.000
17[a]	2	C_3H_7	H	H	3.550	0.000	0.000
18	2	C_4H_9	H	H	3.947	0.000	0.000
19	2	C_5H_{11}	H	H	4.343	0.000	0.000
20	2	C_6H_{13}	H	H	4.739	0.000	0.000
21[a]	2	CH_2CHMe_2	H	H	3.881	0.000	0.000
22	2	$CH_2CH_2CHMe_2$	H	H	4.277	0.000	0.000
23	2	$CH_2CH_2CH_2CHMe_2$	H	H	4.673	0.000	0.000
24	2	CH_2-c-pentyl	H	H	4.169	0.000	0.000
25	2	C_6H_5	H	H	4.120	0.000	0.000
26	2	$(CH_2)_2C_6H_5$	H	H	4.768	1.000	0.000
27	0	C_6H_5	H	H	3.773	0.000	0.000
28	0	C_6H_5	Me	H	4.240	0.000	1.000
29	0	C_6H_5	$CHMe_2$	H	4.967	0.000	1.000
30	0	C_6H_5	CHMeEt	H	5.363	0.000	1.000
31	0	C_6H_5	c-Pentyl	H	5.256	0.000	1.000
32	0	C_6H_5	c-Hexyl	H	5.652	0.000	1.000
33	0	C_6H_5	CMe_3	H	5.400	0.000	1.000
34	0	C_6H_5	$CHMe_2$	Me	5.434	0.000	1.000
35	0	C_6H_5	$CHMe_2$	$CHMe_2$	5.830	0.000	1.000
36[a]	0	C_6H_5	CMe_3	Me	5.867	0.000	1.000
37	0	$(CH_2)_2$ C_6H_5	H	H	4.025	1.000	0.000
38	0	$(CH_2)_2$ C_6H_5	Me	H	4.492	1.000	1.000
39	0	$(CH_2)_2$ C_6H_5	$CHMe_2$	H	5.219	1.000	1.000
40	0	$(CH_2)_2$ C_6H_5	CHMeEt	H	5.615	1.000	1.000
41	0	$(CH_2)_2$ C_6H_5	c-Pentyl	H	5.904	1.000	1.000
42[a]	0	$(CH_2)_2$ C_6H_5	c-Hexyl	H	5.952	1.000	1.000
43	0	$(CH_2)_2$ C_6H_5	CMe_3	Me	5.686	1.000	1.000
44	0	$(CH_2)_2$ C_6H_5	$CHMe_2$	$CHMe_2$	6.413	1.000	1.000

Table 1. continued on next page

Table 1. Continued.

Compound No.	*N*	*R*	*R'*	*R''*	logP	IR	IR′
45	0	$(CH_2)_2C_6H_5$	$CHMe_2$	Me	6.119	1.000	1.000
46	0	$(CH_2)_2C_6H_5$	CMe_3	$CHMe_2$	6.846	1.000	1.000
47	0	$(CH_2)_2C_6H_5$	$CHMe_2$	Me	5.686	1.000	1.000
48	2	$(CH_2)_2COOH$	H	H	2.351	0.000	0.000
49	2	$(CH_2)_3COOH$	H	H	2.768	0.000	0.000
50	2	$(CH_2)_4COOH$	H	H	3.185	0.000	0.000
51	2	$(CH_2)_3CONH_2$	H	H	2.117	0.000	0.000
52	2	$(CH_2)_4CONH_2$	H	H	2.535	0.000	0.000
53[a]	2	4-Pyridyl	H	H	2.908	0.000	0.000
54	2	$CH_2N(Me)C_6H_5$	H	H	3.587	0.000	0.000
55	1	$CH_2OC_6H_5$	H	H	3.896	0.000	0.000
56	1	$(CH_2)_4OH$	H	H	2.275	0.000	0.000
57	1	$(CH_2)_{2-}$	H	H	2.597	0.000	0.000
58[a]	1	$(CH_2)_{3-}$	H	H	3.654	0.000	0.000
59[a]	1	$(CH_2)_4CH_{3-}$	H	H	2.374	0.000	0.000

where y_i is the averaged value for the dependent variable for the training set.

Another term to check the external predictability of the selected model is $r^2_{m,}$ which was proposed by Roy and Roy (2007)[19] and it was calculated by the following formula:

$$r^2_m = r^2\left(1-\left|\sqrt{r^2-r_0^2}\right|\right),$$

where r^2 is squared correlation coefficient between observed and predicted values and r_0^2 is squared correlation coefficient between observed and predicted values with intercept value set to zero. A value of r^2_m is greater than 0.5 may be taken as an indicator of good external predictability.

The robustness of a QSAR model was checked by Y–randomization test. In this technique, new QSAR models were developed by shuffling the dependent variable vector randomly and keeping the original independent variable as such. The new QSAR models are expected to have low R^2 and Q^2 values. If the opposite happens, then an acceptable QSAR model cannot be obtained for the specific modelling method and data.

Results

The QSAR studies of 4-hydroxy-5,6-dihydropyran-2-one series resulted in several QSAR equations. Equation 1 is the best equation when one parameter was considered.

$$\mathrm{pIC_{50}} = 4.963(\pm0.262)\pm0.459(\pm0.058)\mathrm{logP} \quad (1)$$

$n=50$, $R=0.750$, $R^2=0.562$, $R^2_{adj}=0.553$, SEE = 0.466, $F=18.95$, $p<0.001$, $Q^2=0.482$.

The above equation is statistically not a significant one. The R^2 and internal predictivity of the model is poor. Equation 2 is the best equation when two parameters were considered.

$$\mathrm{pIC_{50}} = 5.184(\pm0.299)+0.396(\pm0.072)\mathrm{logP} + 0.312(\pm0.212)\mathrm{IR} \quad (2)$$

$n=50$, $R=0.762$, $R^2=0.582$, $R^2_{adj}=0.564$, SEE = 0.460, $F=32.69$, $p<0.001$, $Q^2=0.483$.

The above equation is also statistically not a significant one. The R^2 and internal predictivity of the model is poor. Equation 3 is the best equation when three parameters were considered.

$$\mathrm{pIC_{50}} = 3.531(\pm0.785)+1.254(\pm0.385)\mathrm{logP} - 0.105(\pm0.046)\mathrm{logP}^2+0.765(\pm0.232)\mathrm{IR} \quad (3)$$

$n=50$, $R=0.790$, $R^2=0.624$, $R^2_{adj}=0.599$, SEE = 0.441, $F=25.42$, $p<0.001$, $Q^2=0.521$.

The above equation is statistically significant. The R^2 and internal predictivity of the model is good. Equation 4 is the best equation when four parameters were considered.

$$\mathrm{pIC_{50}} = 3.200(\pm0.719)+1.523(\pm0.390)\mathrm{logP} - 0.155(\pm0.050)\mathrm{logP}^2+0.578(\pm0.223)\mathrm{IR}+0.503 \quad (4)$$

$n=50$, $R=0.812$, $R^2=0.660$, $R^2_{adj}=0.629$, SEE = 0.424, $F=21.84$, $p<0.001$, $Q^2=0.599$.

When we considered five parameters for developing model, there was no significant improvement in R^2 and Q^2. Equations 3 and 4 were selected as the best, significant model on the basis of high Q^2 values and R^2 values. The values given in the parentheses are 95% confidence intervals of the regression coefficients. The residual of observed and calculated activity of compound number 50 and 57 was found two times larger than the standard deviation of Equations 3 and 4. Hence they were removed as outliers; we rebuilt the models with 48 compounds and got the following equations (Equations 5 and 6):

Table 2. Calculated and predicted (LOO) activities of training and test set using equation 5 and 6.

		pIC_{50} (M)			
		Equation 5		Equation 6	
Compound No.	Observed activity (M)[a]	Cal. Act.	Pred. Act.	Cal. Act.	Pred. Act.
1[b]	5.05	5.924	—	5.961	—
2	6.54	6.5	6.498	6.518	6.517
3	6.57	6.695	6.699	6.68	6.685
4	6.91	6.863	6.861	6.801	6.796
5	6.74	7.004	7.014	6.878	6.886
6	6.36	6.664	6.675	6.656	6.667
7	7.01	6.837	6.831	6.784	6.775
8	6.74	6.983	6.992	6.868	6.875
9	7.06	6.793	6.783	6.753	6.742
10	6.86	6.946	6.949	6.85	6.849
11	6.58	6.772	6.778	6.738	6.744
12	7.22	7.013	7.006	6.882	6.862
13[b]	5.85	6.940	—	6.852	—
14	6.96	6.73	6.722	6.708	6.698
15	6.89	6.983	6.986	6.868	6.867
16	5.68	6.093	6.141	6.131	6.184
17[b]	5.29	6.618	—	6.626	—
18	6.40	6.805	6.819	6.762	6.776
19	7.08	6.956	6.951	6.855	6.844
20	6.95	7.08	7.085	6.905	6.902
21[b]	5.92	6.767	—	6.741	—
22	7.02	6.933	6.930	6.842	6.834
23	6.81	7.061	7.071	6.9	6.907
24	6.29	6.893	6.914	6.819	6.842
25	6.82	6.874	6.876	6.808	6.807
26	7.29	7.088	7.079	6.907	6.875
27	6.96	6.73	6.722	6.708	6.698
28	6.41	6.919	6.938	6.835	6.854
29	7.64	7.139	7.114	7.418	7.382
30	7.77	7.22	7.175	7.401	7.354
31	7.85	7.201	7.155	7.41	7.352
32	6.90	7.262	7.310	7.362	7.431
33	7.31	7.226	7.219	7.397	7.408
34	7.42	7.231	7.214	7.393	7.390
35	6.92	7.281	7.348	7.326	7.404
36[b]	7.92	8.118	—	7.875	—
37	6.89	6.837	6.835	7.286	7.432
38	7.14	7.547	7.691	7.433	7.546
39	7.85	7.735	7.712	7.966	8.000
40	8.00	7.798	7.771	7.923	7.910
41	7.49	7.828	7.870	7.864	7.911
42[b]	8.52	8.124	—	7.347	—
43	8.15	7.807	7.762	7.91	7.875
44	7.85	7.845	7.843	7.705	7.669
45	8.00	7.84	7.818	7.805	7.777
46	7.33	7.823	8.094	7.514	7.674
47	8.22	7.807	7.752	7.91	7.864
48	5.92	5.922	5.922	5.955	5.961
49	6.57	6.195	6.162	6.232	6.202
50	8.30	*	*	*	*
51	5.89	5.755	5.714	5.778	5.744
52	6.47	6.046	5.990	6.083	6.031
53[b]	6.10	6.637	—	6.640	—
54	6.76	6.784	6.293	6.747	6.330

Table 2. continued on next page

Table 2. Continued.

Compound No.	Observed activity (M)[a]	pIC_{50} (M) Equation 5 Cal. Act.	Equation 5 Pred. Act.	Equation 6 Cal. Act.	Equation 6 Pred. Act.
55	5.64	5.928	6.784	5.532	6.746
56	5.90	5.869	5.862	5.899	5.899
57	4.93	*	*	*	*
58[b]	6.49	6.666	—	6.663	—
59[b]	5.64	5.931	—	5.968	—

Cal. Act., calculated activity; LOO, leave one out; Pred. Act., predicted activity (LOO).
[a]The experimental IC_{50} values (in nano molar) were converted into $-\log IC_{50}$ (pIC_{50}, in molar).
[b]Test set compound.
*Outlier.

Table 3. Correlation matrix of physico-chemical descriptors and HIV protease inhibitory activity.

	pIC_{50}	logP	$logP^2$	IR	IR′	VIF Equation 5	VIF Equation 6
pIC_{50}	1	1	1	1	1	—	—
logP	0.750	0.987	0.652	0.571		—	—
$logP^2$	0.723	0.600				—	—
IR	0.561	0.714	0.725			2.019	2.021
IR′	0.641					—	2.133

IR, if the substituent at R is $-(CH_2)_2C_6H_5$ then the value of IR is 1 otherwise 0); IR′, if any substituent is present at R′ then the value of IR′ is 1 otherwise 0; VIF, variance inflation factor.

$$pIC_{50} = 3.817(\pm0.564) + 1.098(\pm0.274)\log P - 0.087(\pm0.033)\log P^2 + 0.854(\pm0.162)IR \quad (5)$$

$n=48$, $R=0.875$, $R^2=0.765$, $R^2_{adj}=0.749$, SEE = 0.309, $F=47.86$, $p<0.001$, $Q^2=0.707$, Standard deviation of Predicted residual sum of square (S_{PRESS}) = 0.337, standard deviation of error of predictions (SDEP) = 0.335, $r^2_m=0.245$, optimum logP ($logP_0$) = 4.723.

Equation 5 could explain 76.5% and predict 74.9% of the variance of the HIV-1 protease inhibitory activity data. The calculated HIV-1 protease inhibitory activity values by Equation 5 are given in Table 2. There was no intercorrelation between the descriptors (Table 3). This model showed good correlation coefficient (R) of 0.875 between descriptors (logP, $logP^2$, and IR) and HIV-1 protease inhibitory activity. This model also indicates statistical significance > 99.9% with F-value $F_{(3,44)}=47.86$.

$$pIC_{50} = 3.504(\pm0.518) + 1.362(\pm0.259)\log P - 0.136(\pm0.033)\log P^2 + 0.554(\pm0.149)IR + 0.503 \quad (6)$$

$n=48$, $R=0.902$, $R^2=0.814$, $R^2_{adj}=0.796$, SEE = 0.278, $F=46.90$, $p<0.001$, $Q^2=0.763$, $S_{PRESS}=0.303$, SDEP = 0.301, $r^2_m=0.196$.

Equation 6 could explain 81.4% and predict 79.6% of the variance of the HIV-1 protease inhibitory activity data. The calculated HIV-1 protease inhibitory activity values by Equation 6 are given in Table 2. This model showed good correlation coefficient (R) of 0.902 between descriptors (logP, $logP^2$, IR, and IR′) and HIV-1 protease inhibitory activity. This model also indicates statistical significance > 99.9% with F-value $F_{(4,43)}=46.90$.

Both the models showed good internal predictivity. The predictive ability of the selected models was also confirmed by external r^2CV_{ext} method. The proposed QSAR model is predictive as it satisfies the conditions $r^2CV_{ext} > 0.5$ and $R^2_{Pred} > 0.6$[20]. Both the models satisfied these conditions (Equation 5: $r^2CV_{ext}=0.624$ and $R^2_{Pred}=0.843$, Equation 6: $r^2CV_{ext}=0.555$ and $R^2_{Pred}=0.694$). The robustness of this model was checked by Y–randomization test (maximum R^2 value is 0.210 and maximum Q^2 is 0.080). The low R^2 and Q^2 values indicate that the good results in our original model are not due to a chance correlation or structural dependency of the training set. But both the models failed to satisfy the condition that r^2_m should be > 0.5.

Discussion

Both Equations 5 and 6 showed statistical significance and good internal predictivity but Equation 5 showed better external predictivity than Equation 6. Hence the Equation 5 was selected as the best QSAR model to predict the HIV-1 protease inhibitory activity of 4-hydroxy-5,6-dihydropyran-2-one derivatives.

The QSAR shows a parabolic relationship of HIV-1 protease inhibitory activity with the logP. It is interesting to note that the optimum logP value in this case is 4.723. This value lies in the (4.49–6.96) range determined by Garg and Patel[5]. Negative sign of $logP^2$ indicates that highly hydrophobic groups are not good for improving the activity of the series. The positive coefficient of IR

showed that the presence of $-(CH_2)_2C_6H_5$ group at sixth position of dihydropyranone nucleus is essential for better activity.

The proposed model, due to the good internal and external predictive ability, can therefore act as a useful aid to the costly and time consuming experiments for determining the molar concentration of compounds required to achieve 50% inhibition in HIV-1 protease activity. Our results lead to the conclusion that the HIV-1 protease inhibitory activity of 4-hydroxy-5,6-dihydropyran-2-ones can be increased if substituents that bring about changes in the molecules as mentioned above are attached to it.

Acknowledgements

One of the authors V. Ravichandran is thankful to AICTE, New Delhi for providing (QIP) Senior Research Fellowship. The authors are thankful to the reviewers for their valuable suggestions to improve the quality of this paper.

Declaration of interest

The authors report no conflict of interest.

References

1. Clercq ED. Toward improved anti-HIV chemotherapy: therapeutic strategies for intervention with HIV infections. J Med Chem 1995;38:2491–17.
2. Milton J, Slater MJ, Bird AJ, Spinks D, Scott G, Price CE, Downing S, Green DV, Madar S, Bethell R, Stammers DK. Biaryl acids: novel non-nucleoside inhibitors of HIV reverse transcriptase types 1 and 2. Bioorg Med Chem Lett 1998;8:2623–28.
3. Pungpo P, Hannongbua S. Three-dimensional quantitative structure-activity relationships study on HIV-1 reverse transcriptase inhibitors in the class of dipyridodiazepinone derivatives, using comparative molecular field analysis. J Mol Graph Model 2000;18:581–90, 601.
4. Nair AC, Jayatilleke P, Wang X, Miertus S, Welsh WJ. Computational studies on tetrahydropyrimidine-2-one HIV-1 protease inhibitors: improving three-dimensional quantitative structure-activity relationship comparative molecular field analysis models by inclusion of calculated inhibitor-and receptor-based properties. J Med Chem 2002;45:973–83.
5. Kumar S, Jacob RR, Tiwari M. 3D-QSAR study of some 5,6-dihydropyran-2-ones as protease inhibitors. Indian J Pharm Sci 2005;67:30–36.
6. Garg R, Patel D. Hydrophobicity in the design of P2/P2′ tetrahydropyrimidinone HIV protease inhibitors. Bioorg Med Chem Lett 2005;15:3767–70.
7. Bhhatarai B, Garg R. From SAR to comparative QSAR: role of hydrophobicity in the design of 4-hydroxy-5,6-dihydropyran-2-ones HIV-1 protease inhibitors. Bioorg Med Chem 2005;13:4078–84.
8. Buolamwini JK, Assefa H. CoMFA and CoMSIA 3D QSAR and docking studies on conformationally-restrained cinnamoyl HIV-1 integrase inhibitors: exploration of a binding mode at the active site. J Med Chem 2002;45:841–52.
9. Raghavan K, Buolamwini JK, Fesen MR, Pommier Y, Kohn KW, Weinstein JN. Three-dimensional quantitative structure-activity relationship (QSAR) of HIV integrase inhibitors: a comparative molecular field analysis (CoMFA) study. J Med Chem 1995;38:890–7.
10. Ravichandran V, Agrawal RK. Predicting anti-HIV activity of PETT derivatives: CoMFA approach. Bioorg Med Chem Lett 2007;17:2197–202.
11. Ravichandran V, Jain PK, Mourya VK, Agrawal RK. QSAR study on some arylsulfonamides as anti-HIV agents. Med Chem Res 2007;16:7–9.
12. Ravichandran V, Mourya VK, Agrawal RK. QSAR study of novel 1, 1, 3-trioxo [1, 2, 4]-thiadiazine (TTDs) analogues as potent anti-HIV agents. Arkivoc 2007;XIV:204–12.
13. Ravichandran V, Mourya VK, Agrawal RK. QSAR prediction of HIV-1 reverse transcriptase inhibitory activity of benzoxazinone derivatives. Internet Electron J Mol Des 2007;6:363–74.
14. Ravichandran V, Prashanthakumar BR, Sankar S, Agrawal RK. Comparative molecular similarity indices analysis for predicting anti-HIV activity of phenyl ethyl thiourea (PET) derivatives. Med Chem Res 2008;17:1–11.
15. Sahu KK, Ravichandran V, Jain PK, Sharma S, Mourya VK, Agrawal RK. QSAR analysis of chicoric acid derivatives as HIV-1 integrase inhibitors. Acta Chimi Slov 2008;55:138–45.
16. Sahu KK, Ravichandran V, Mourya VK, Agrawal RK. QSAR analysis of caffeoyl naphthalene sulphonamide derivatives as HIV-1 Integrase inhibitors. Med Chem Res 2007;15:418–30.
17. Tait BD, Hagen S, Domagala J, Ellsworth EL, Gajda C, Hamilton HW, Prasad JV, Ferguson D, Graham N, Hupe D, Nouhan C, Tummino PJ, Humblet C, Lunney EA, Pavlovsky A, Rubin J, Gracheck SJ, Baldwin ET, Bhat TN, Erickson JW, Gulnik SV, Liu B. 4-hydroxy-5,6-dihydropyrones. 2. Potent non-peptide inhibitors of HIV protease. J Med Chem 1997;40:3781–3792.
18. Roy K. On some aspects of validations of predictive QSAR models. Expert Opin Drug Discov 2007;2:1567–77.
19. Roy PP, Roy K. On some aspects of variable selection for partial least squares regression models. QSAR Comb Sci 2007;27:302–13.
20. Tropsha A, Gramatica P, Gombar VK. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. Quant Struct Act Rel 2003;22:1–9.