Med Chem Res (2007) 15:418–430
DOI: 10.1007/s00044-006-0020-2

MEDICINAL CHEMISTRY RESEARCH

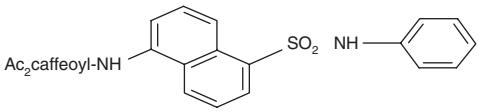
ORIGINAL RESEARCH

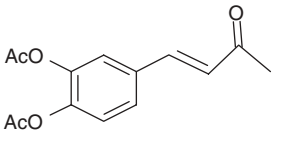
QSAR analysis of caffeoyl naphthalene sulfonamide derivatives as HIV-1 integrase inhibitors

Kamlesh Kumar Sahu · Veerasamy Ravichandran · Vishnu Kant Mourya · Ram Kishore Agrawal

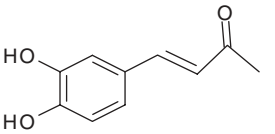
Received: 20 July 2006 / Accepted: 2 October 2006 / Published online: 26 May 2007

Abstract Human immunodeficiency virus type 1 (HIV-1) integrase is a potential target for anti-HIV therapy. It is an essential enzyme required for replication of the acquired immunodeficiency syndrome (AIDS) virus. Caffeoyl naphthalene sulfonamide derivatives act against HIV integrase and thus have the potential to become a part of an anti-HIV drug regimen. Although caffeoyl naphthalene sulfonamide derivatives have all the features required of good anti-HIV agents such as the presence of bis-catechol moieties, polyaromatic rings, and a central linker, they do





A



c_2-caffeoyl caffeoyl

Ac_2caffeoyl = acetylated caffeoyl group

not perform well as anti-HIV agents in cell-based assays, that is, they do not stop viral replication at nontoxic concentration. We carried out a quantitative structure–activity relationship (QSAR) study of caffeoyl naphthalene sul-

Kamlesh K. Sahu (✉) · V. Ravichandran · R. K. Agrawal
Department of Pharmaceutical Sciences, Dr. H. S. Gour University, Sagar (M.P.), India
e-mail: kamleshsemail@yahoo.co.in

V. K. Mourya
Government College of Pharmacy, Osmanpura, Aurangabad, Maharashtra, India

fonamide derivatives via the software WIN CAChe 6.1 and STATISTICA to improve its activity. QSAR reveals that if partition coefficient, connectivity index, and shape index of these molecules are altered, the activity is likely to increase. On the basis of the QSAR model, we designed a new series of compounds, calculated the activities, and found that they were more potent than the existing compounds.

Keywords Caffeoyl · Human immunodeficiency virus · Integrase · Inhibition · Partition coefficient · Quantitative structure–activity relationship · Sulfonamide · Total energy

Introduction

General

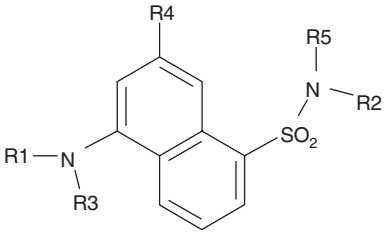
Among the innumerable viruses that cause infection and disease in humans and animals is a large family called retroviruses, which cause cancers and leukemias. Some retroviruses, called immunodeficiency viruses, cause immune deficiency in cattle, monkeys, and humans. Retroviruses have some very interesting properties. Once a person is infected, the infection remains life-long. In most people the infection causes no disease for a long period. Yet the infected person can be the source of infection for others.

Human immunodeficiency virus type 1 (HIV-1) integrase is an enzyme required for viral replication (De Clercq, 1995). HIV integrase catalyzes integration of viral DNA into host genome in two separate but chemically similar reactions known as 3′ processing and DNA strand transfer (Craigie *et al*., 1990; Katz *et al*., 1990). In 3′ processing, integrase (IN) removes a dinucleotide next to a conserved cytosine-adenine sequence from each 3′- end of the viral DNA. IN then attaches the processed 3′-end of the viral DNA to the host cell DNA in the strand transfer reaction. As there is no known human counterpart of HIV IN, it is an attractive target for antiretroviral drug design (Chen *et al*., 2000). A large number of HIV IN inhibitors have been discovered (Neamati, 2002). However, the mechanism of action is incompletely understood (Parril, 2003).

Several families of IN inhibitors have been identified. Most of them can be classified into three groups: DNA ligands, C-terminal domain ligands, and compounds that interfere with the catalytic domain of the protein. The first family contains nonspecific intercalating agents (Carteau *et al*., 1993; Fesen *et al*., 1993) as well as more specific oligonucleotide targeting IN binding sites on both long terminal repeats (LTRs; Mouscadet *et al*., 1994).

While many IN inhibitors have now been developed, only a handful displayed antiviral activity in cell culture. This group comprises lignanolides (Eich *et al*., 1996), curcumin (Majumder *et al*., 1995), aurintricarboxylic acids (Cushman and Sherman, 1992), dicaffeoyl quinic acids and analogues

Table 1 Physicochemical parameters and the antiviral activity data used to derive QSAR



Compound no.	R	R_1	CI0	TE	LOGP	BKO1	Log ($1/IC_{50}$) Obs.	Calc.
1	Ac_2caffeoyl	C_6H_5	27.99	–289.94	3.664	31.93	–2	–1.846
2	Ac_2caffeoyl	$CH_2C_6H_5$	23.42	–240.69	4.098	26.07	–1.0864	–1.136
3	Ac_2caffeoyl	$CH_2CH_2C_6H_5$	28.7	–297.07	3.759	32.9	–2	–1.675
4	Ac_2caffeoyl	4-Fluorobenzene	24.13	–247.8	4.193	27.05	–0.9345	–0.982
5	Ac_2caffeoyl	4-Chlorobenzene	29.41	–304.24	4.011	33.88	–1.0755	–1.411
6	Ac_2caffeoyl	4-Bromobenzene	24.84	–254.97	4.444	28.02	–0.6532	–0.737
7	Ac_2caffeoyl	4-Methylbenzene	28.86	–305.86	3.804	32.9	–1.3909	–1.403
8	Ac_2caffeoyl	3-Nitrobenzene	24.29	–256.6	4.237	27.05	–0.8976	–1.514
9	Ac_2caffeoyl	2-Methylbenzene	28.86	–301.71	4.182	32.9	–2	–1.984
10	Ac_2caffeoyl	2,3-Dimethylbenzene	24.29	–252.44	4.616	27.05	–1.3096	–1.292
11	Caffeoyl	C_6H_5	28.86	–299.84	4.456	32.9	–2	–1.825
12	Caffeoyl	$CH_2C_6H_5$	24.29	–250.57	4.889	27.05	–1.29	–1.133
13	Caffeoyl	$CH_2CH_2C_6H_5$	28.86	–297.13	4.131	32.9	–2	–2.007
14	Caffeoyl	4-Fluorobenzene	24.29	–247.87	4.565	27.05	–2	–1.315
15	Caffeoyl	4-Chlorobenzene	30.28	–321.64	3.712	34.87	–2	–2.02
16	Caffeoyl	4-Bromobenzene	25.71	–272.36	4.145	28.99	–1.2856	–1.364
17	Caffeoyl	4-Methylbenzene	28.86	–297.12	4.132	32.9	–2	–2.007
18	Caffeoyl	3-Nitrobenzene	24.29	–247.85	4.565	27.05	–1.3096	–1.315
19	Caffeoyl	2-Methylbenzene	29.73	–304.31	4.599	33.88	–2	–2.166
20	Caffeoyl	2,3-Dimethylbenzene	25.17	–255.05	5.032	28.02	–1.2967	–1.495

(Robinson *et al*., 1996a,b), diarylsulfones (Mazumder *et al*., 1996), and finally G-rich oligonucleotides (Hansch, 1969).

Computational chemistry has developed into an important contributor to rational drug design. Quantitative structure–activity relationship (QSAR) modeling results in a quantitative correlation between chemical structure and biological activity (Parril, 2003). In this study, we performed QSAR analysis of caffeoyl naphthalene sulfonamide derivatives using WIN CAChe 6.1 and STATISTICA.

The software

The software WIN CAChe 6.1 was used to draw the molecules, to minimize the energy of the molecules, and to calculate the physicochemical properties (Table 1) on a project leader provided in the software. This was followed by

regression analysis, which was performed via STATISTICA (Softstat). WIN CAChe 6.1 is a product of Fujitsu private limited (http://www.cachesoftware.com/contacts/japan.shtml and http://www.cachesoftware.com/techsupport/download.shtml).

Objective of the study

Some of the features of anti- HIV agents include polyaromatic rings separated by central linker and presence of catechol moieties. It has been reported that the majority of natural products endowed with anti-IN activity were characterized by one or two 3,4- dihydroxycinnamoyl moieties (Santo *et al*., 2003). Despite these features, caffeoyl naphthalene sulfonamides do not prevent the replication of HIV at nontoxic concentrations. Hence the present study was conducted to design molecules with improved potency.

Molecular descriptors

A property calculated from mathematical and physical entities of a compound's molecular structure is called as descriptor. Examples of descriptor types are topological, thermodynamic, spatial, and electronic.

Spatial descriptors describe the molecule's ''solvent-accessible'' surface areas and their charges. Electronic descriptors describe the electron orientation and charge.

Topological descriptors are based on graph/structure concepts and geometric features such as shape, size, and branching. Thermodynamic descriptors describe energy of molecules and their conversions. Quantum mechanical descriptors are calculated using semiempirical methods that are likely to be more accurate.

The following list includes some of the experiments that are available with Project Leader of CAChe 6.1. These experiments can be performed on one or more chemical samples simultaneously to calculate the properties:

- Optimization to find a low energy structure for steric energy, heat of formation, or total energy
- Net positive and negative charge for a molecule
- Molecular formula, weight, and refractivity
- Ring count and size
- Investigation of molecular orbital energies such as HOMO and LUMO energies
- Calculation of the dielectric, steric, total energy, and heat of formation of a structure at its current geometry
- Investigation of visible and UV-visible spectra data
- Zero-order, first-order, and second-order molecular or valence connectivity indices
- Dipole moment and dipole vectors x, y, and z
- Electron affinity

- Shape index order 1, 2, and 3
- Octanol–water partition coefficient.

In the present study, the descriptors calculated were zero-order molecular or valence connectivity index (CI0), first-order molecular or valence connectivity index (CI1), dipole moment (DM), electron affinity (EA), total energy at its current geometry after optimization of structure (TE), heat of formation at its current geometry after optimization of structure (HF), highest occupied molecular orbital energies (HOMO), lowest unoccupied molecular orbital energies (LUMO), UV-visible spectra data (LMAX), octanol–water partition coefficient (LOGP), conformational minimum energies (ME), molar refractivity (MR), ionization potential (IP), shape index order 1 or basic kappa order 1 (BKO1), and solvent-accessible surface area (SAS).

Results and Discussion

We searched for a molecule having the same nucleus but better biological activity as the existing caffeoyl naphthalene sulfonamide derivatives. After regression analysis via STATISTICA, the best equation (Eq. 1) obtained was:

$$\begin{aligned}\text{Log } (1/\text{IC50}) = {} & (-3.3466 \pm 0.976)\text{CIO} + (0.00131 \pm 0.00054)\text{TE} \\ & + (0.5726 \pm 0.295)\text{LOGP} + (2.5455 \pm 0.775)\text{BKO1} \\ & + (8.8528 \pm 2.360)\end{aligned} \quad (1)$$

n = 20, r = 0.841, s = 0.2868, calculated F-ratio = 9.03, t-test value = 2.131 (95%), r^2 = 0.7073, Y-variance = 0.470, Y-mean = –1.526, variance in Y explained via the regression = 70.7%. IC_{50} is the molar concentration of the drug leading to 50% inhibition of integrase, CI0 = connectivity index, TE = total energy, LOGP = partition coefficient, BKO1 = basic kappa order 1 (shape index), n = number of data points, r = correlation coefficient, s = standard error of regression, F-ratio = F-ratio between variances of calculated and observed value, t-test = Student's t-test for statistical significance.

Regression analysis of caffeoyl naphthalene sulfonamides, calculated coefficients, and estimates of error are given in Table 2. This equation reveals that biological activity can be increased if (1) the partition coefficient (LOG P) of the molecule is increased by attaching groups that impart good partition coefficient (alkyl groups, aromatic rings, trifluromethane -CF_3 etc.); (2) the connectivity index is decreased; (3) the shape index is increased.

Equation (1) is obtained after QSAR of training set of compounds.

$$\begin{aligned}\text{Log } (1/\text{IC50}) = {} & (-3.7721 \pm 1.161)\text{CIO} + (0.00136 \pm 0.00061)\text{TE} \\ & + (0.4101 \pm 0.355)\text{LOGP} + (2.8629 \pm 0.915)\text{BKO1} \\ & + (11.2711 \pm 3.387)\end{aligned} \quad (2)$$

Table 2 Regression analysis of caffeoyl naphthalene sulfonamides: calculated coefficients and Estimates of error

Variable	Variable	Coefficient	Std. error	95% CI	*t*-Value	Mean	Variance
0	Intercept	8.8528	2.36	5.03			
1	CI0	–3.3466	0.976	2.081	–3.428	26.758	2.421
2	TE	0.00131	0.00054	1.20E–03	2.42	–246.667	133
3	LOGP	0.5726	0.295	0.629	1.942	4.262	0.381
4	BKO1	2.55	0.775	1.651	3.286	30.268	3.106

No. of compounds (n) = 20; t-test value = 2.131 (95%); std. error of regression (s) = 0.2868; Y-variance = 0.470; Y-mean = –1.526; Calculated F-ratio = 9.03; multiple correlation coefficient® = 0.841; variance in Y explained via the regression = 70.7%

N = 16, r = 0.841, s = 0.3162, calculated F-ratio = 6.35, t-test value = 2.201 (95%), r^2 = 0.698, Y-variance = 0.493, Y-mean = –1.495, variance in Y explained via the regression = 69.8%.

The values of variables present in the Eq. (2) and observed and predicted values for the test set of compounds (Table 3) show that the prediction by the equation obtained via QSAR is very close to the observed values.

The predicted activities of a newly designed series (Table 4) of compounds show that they all have predicted activities ranging from IC_{50} = 1.0 μg/ml to 1.11 μg/ml whereas in the reported series of caffeoyl naphthalene sulfonamide derivatives (Xu *et al*., 2003) the most potent compound has an activity of 4.5 μg/ml.

Conclusion

Equation 1 predicts that increase in partition coefficient, total energy, and shape index would increase the biological activity of the compound. The effect of an increase of total energy on increase in biological activity is less compared to that of the other three variables. The biological activity is increased when the connectivity index is decreased. Thus we conclude that if the groups that bring about the aforementioned changes in the molecule are attached to it, the biological activity will be increased.

Table 3 Physicochemical properties, observed and predicted activities of test set of compounds

					Activity (Log1/IC_{50})	
Compound	CI0	TE	LOGP	BKO1	Observed	Predicted
17	28.86	–297.12	4.132	32.9	–2	–2.112
18	24.29	–247.85	4.565	27.05	–1.3096	–1.377
19	29.73	–304.31	4.599	33.88	–2	–2.407
20	25.17	–255.05	5.032	28.02	–1.2967	–1.73

Table 4 Physicochemical properties and predicted activities of a designed series of compounds

Comp.no.	CIO	TE	Log P	BKO1	Constt.	Predicted activity (Log1/IC_{50})	Predicted IC_{50} (μg/ml)
1	31.976	–326.09	5.592	36.285	8.8528	–3.019	1.001
2	34.424	–347.542	6.649	39.213	8.8528	–3.181	1.000
3	32.227	–340.516	6.182	34.106	8.8528	–9.087	1.000
4	29.57	–304.311	4.528	33.884	8.8528	–1.660	1.022
5	29.57	–304.303	4.528	33.884	8.8528	–1.660	1.022
6	29.57	–304.302	4.528	33.884	8.8528	–1.660	1.022
7	30.277	–311.466	4.924	34.865	8.8528	–1.312	1.050
8	30.277	–311.484	4.924	34.865	8.8528	–1.312	1.050
9	30.277	–311.457	4.924	34.865	8.8528	–1.312	1.050
10	30.441	–311.444	4.858	34.865	8.8528	–1.898	1.013
11	30.441	–311.468	4.858	34.865	8.8528	–1.899	1.013
12	30.441	–311.46	4.858	34.865	8.8528	–1.899	1.013
13	30.985	–318.603	5.32	35.847	8.8528	–0.964	1.115
14	30.985	–318.527	5.118	35.847	8.8528	–1.080	1.087
15	31.363	–344.881	4.547	35.847	8.8528	–2.706	1.002
16	31.363	–344.852	4.547	35.847	8.8528	–2.706	1.001
17	31.363	–344.881	4.547	35.847	8.8528	–2.706	1.001
18	31.363	–344.796	5.656	35.847	8.8528	–2.071	1.009
19	36.924	–395.316	7.339	42.149	8.8528	–3.742	1.000
20	34.733	–399.715	6.539	39.781	8.8528	–2.902	1.001
21	34.733	–399.724	6.539	39.781	8.8528	–2.902	1.001
22	32.07	–354.023	5.5	36.829	8.8528	–2.039	1.009
23	33.863	–378.659	4.555	38.796	8.8528	–3.606	1.000
24	37.941	–442.59	7.161	43.724	8.8528	–3.301	1.001
25	37.941	–442.605	7.161	43.724	8.8528	–3.301	1.001
26	37.941	–426.67	4.813	43.724	8.8528	–4.624	1.000
27	32.07	–338.076	3.152	36.829	8.8528	–3.362	1.000
28	32.07	–338.062	3.152	36.829	8.8528	–3.362	1.000
29	32.07	–338.025	3.922	36.829	8.8528	–2.922	1.001
30	41.311	–481.607	6.034	47.675	8.8528	–5.218	1.000

Experimental

Method

The biological activities of all 20 compounds were collected from the literature (Xu *et al*., 2003). All 20 compounds were built on workspace of WIN CAChe 6.1, followed by minimization of energy by geometry optimization of molecules using MM3 (Molecular Mechanics version 3). The physicochemical properties were calculated via the project leader file of the software (Table 1). This was followed by regression analysis on STATISTICA. Regression analysis included the correlation matrix (Table 5), observed and estimated values with residuals (Table 6), and calculated coefficients and estimates of error (Table 2). Observed and estimated values have been plotted as a graph (Fig. 1).

Table 5 Correlation matrix

S.No	VARIABLES	Log(1/IC_{50}) (μg/ml)	CI0	CI1	DM	EA	TE	HF	HOMO	LUMO	LMAX	LOGP	ME	MR	IP	BKO1	SAS
1	IC50	1															
2	CIO	0.677	1														
3	CI1	0.656	0.997	1													
4	DM	0.605	0.854	0.849	1												
5	EA	0.024	0.122	0.128	0.516	1											
6	TE	–0.29	0.013	0	0.085	–0.05	1										
7	HF	–0.457	–0.67	–0.64	–0.37	0.551	–0.27	1									
8	HOMO	–0.101	–0.43	–0.44	–0.6	–0.58	–0.14	–0.04	1								
9	LUMO	–0.014	–0.12	–0.13	–0.51	–1	0.034	–0.57	0.578	1							
10	LMAX	–0.401	–0.37	–0.36	–0.43	–0.01	0.094	0.336	0.303	–0.02	1						
11	LOGP	–0.296	–0.52	–0.57	–0.59	–0.17	–0.18	0.324	0.295	0.163	0.047	1					
12	ME	0.368	0.551	0.555	0.795	0.867	–0.07	0.178	–0.62	–0.87	–0.12	–0.4	1				
13	MR	0.679	0.972	0.969	0.748	–0.01	–0.09	–0.68	–0.38	0.006	–0.41	–0.39	0.43	1			
14	IP	0.167	0.13	0.119	0.428	0.629	–0.05	0.232	–0.61	–0.61	–0.19	–0.09	0.61	0.06	1		
15	BKO1	0.668	0.999	0.999	0.852	0.122	0.008	–0.66	–0.43	–0.12	–0.37	–0.55	0.55	0.97	0.12	1	
16	SAS	0.689	0.988	0.986	0.808	0.041	–0.04	–0.68	–0.41	–0.04	–0.41	–0.47	0.47	0.99	0.08	0.99	1

CI0 = connectivity index 0, CI1 = connectivity index 1, DM = dipole moment, EA = electron affinity, TE = total energy, HF = heat of formation, HOMO = highest occupied molecular orbital, LUMO = lowest unoccupied molecular orbital, LMAX = lambda max for UV transition, LOGP = partition coefficient, ME = conformation minimum energy, MR = molar refractivity, IP = ionization potential energy, BKO1 = basic kappa order 1, SAS = solvent accessible surface

BIRKHÄUSER

Table 6 Observed and estimated values with residuals for a series of caffeoyl naphthalene sulfonamide derivatives

Sorting residuals			Tables of residuals				
Original order				Sorted on *Y*'s			
Cpd	*Y* (obs)	*Y* (est)	Residual	Cpd	*Y* (obs)	*Y* (est)	Residual
1	–2	–1.846	–0.154	6	–0.653	–0.737	0.084
2	–1.086	–1.136	0.05	8	–0.898	–1.514	0.617
3	–2	–1.675	–0.325	4	–0.934	–0.982	0.048
4	–0.934	–0.982	0.048	5	–1.076	–1.411	0.336
5	–1.076	–1.411	0.336	2	–1.086	–1.136	0.05
6	–0.653	–0.737	0.084	16	–1.286	–1.364	0.079
7	–1.391	–1.403	0.012	12	–1.29	–1.133	–0.157
8	–0.898	–1.514	0.617	20	–1.297	–1.495	0.199
9	–2	–1.984	–0.016	10	–1.31	–1.292	–0.018
10	–1.31	–1.292	–0.018	18	–1.31	–1.315	0.005
11	–2	–1.825	–0.175	7	–1.391	–1.403	0.012
12	–1.29	–1.133	–0.157	9	–2	–1.984	–0.016
13	–2	–2.007	0.007	1	–2	–1.846	–0.154
14	–2	–1.315	–0.685	14	–2	–1.315	–0.685
15	–2	–2.02	0.02	3	–2	–1.675	–0.325
16	–1.286	–1.364	0.079	11	–2	–1.825	–0.175
17	–2	–2.007	0.007	17	–2	–2.007	0.007
18	–1.31	–1.315	0.005	13	–2	–2.007	0.007
19	–2	–2.166	0.166	19	–2	–2.166	0.166
20	–1.297	–1.495	0.199	15	–2	–2.02	0.02

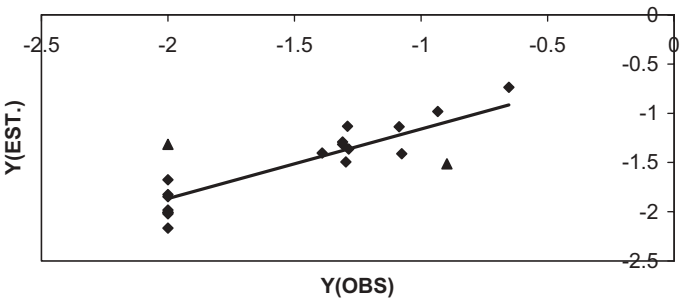


Fig. 1 Plot of observed versus estimated anti-HIV values for a series of caffeoyl naphthalene sulfonamide derivatives. The symbol ▲ indicates the outliers

Model validation

The model was validated by taking the first 16 compounds of the series as training set and the last four as the test set. The QSAR was done for the training set and the equation thus obtained (Eq. 2) was used to predict the biological activities of the remaining four compounds of the series. The

Table 7 Observed and estimated values with residuals for a series of the first 16 caffeoyl naphthalene sulfonamide derivatives

Sorting residuals			Tables of residuals				
Original order				Sorted on *Y*'s			
Cpd	*Y* (obs)	*Y* (est)	Residual	Cpd	*Y* (obs)	*Y* (est)	Residual
1	–2	–1.815	–0.185	6	–0.653	–0.729	0.076
2	–1.086	–1.087	0.001	8	–0.898	–1.55	0.652
3	–2	–1.65	–0.35	4	–0.934	–0.941	0.007
4	–0.934	–0.941	0.007	5	–1.076	–1.418	0.342
5	–1.076	–1.418	0.342	2	–1.086	–1.087	0.001
6	–0.653	–0.729	0.076	16	–1.286	–1.37	0.084
7	–1.391	–1.425	0.034	12	–1.29	–1.275	–0.015
8	–0.898	–1.55	0.652	10	–1.31	–1.389	0.079
9	–2	–2.098	0.098	7	–1.391	–1.425	0.034
10	–1.31	–1.389	0.079	1	–2	–1.815	–0.185
11	–2	–1.983	–0.017	3	–2	–1.65	–0.35
12	–1.29	–1.275	–0.015	9	–2	–2.098	0.098
13	–2	–2.113	0.113	11	–2	–1.983	–0.017
14	–2	–1.404	–0.596	13	–2	–2.113	0.113
15	–2	–2.037	0.037	15	–2	–2.037	0.037
16	–1.286	–1.37	0.084	14	–2	–1.404	–0.596

Table 8 Regression analysis: calculated Coefficients and estimates of error for data of first 16 caffeoyl naphthalene sulfonamide derivatives

Variable	Variable	Coefficient	Std. error	95% CI	*t*-Value	Mean	Variance
0	Intercept	11.2711	3.387	7.454			
1	CI0	–3.7721	1.161	2.556	–3.248	26.694	2.442
2	TE	0.00136	0.00061	0.0013	2.23	–239.313	148
3	LOGP	0.4101	0.355	0.782	1.154	4.182	0.35
4	BKO1	2.86	0.915	2.015	3.128	30.219	3.139

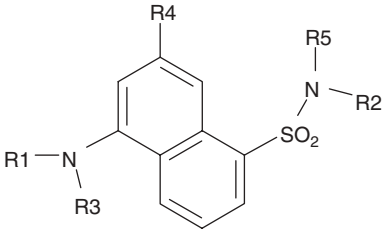
No. of compounds (n) = 16; t-test value = 2.201 (95%); std. error of regression (s) = 0.3162; Y-variance = 0.493; Y-mean = –1.495; calculated F-ratio = 6.35; multiple correlation coefficient® = 0.835; variance in Y explained via the regression = 69.8%

observed and estimated values with residuals, calculated coefficients, and estimates of error are given in Tables 7 and 8, respectively.

Designed molecules

On the basis of Eq. (1), a series of 30 compounds (Table 9) was designed to find molecules with higher potencies than existing caffeoyl naphthalene sulfonamide derivatives. The independent variables were calculated and used in Eq. (1) to obtain predicted biological activities of all 30 compounds of the

Table 9 Series designed on the basis of Eq. (1)



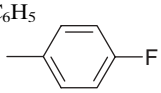
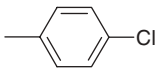
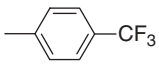
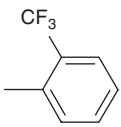
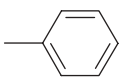
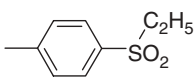
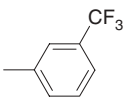
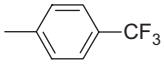
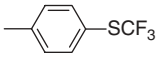
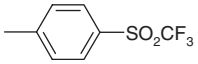
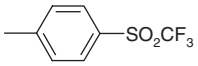
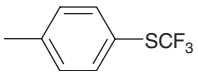
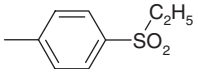
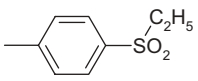
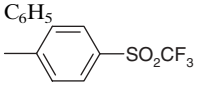
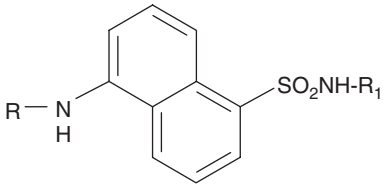
Comp. no.	R_1	R_2	R_3	R_4	R_5
1	Ac_2caffeoyl	C_6H_5	C_6H_5	H	H
2	Ac_2caffeoyl	C_6H_5	C_6H_5	CH_3	C_2H_5
3	Ac_2caffeoyl	C_6H_5	C_6H_5	H	C_2H_5
4	Ac_2caffeoyl	2-Ethyl benzene	H	H	H
5	Ac_2caffeoyl	3-Ethyl benzene	H	H	H
6	Ac_2caffeoyl	4-Ethyl benzene	H	H	H
7	Ac_2caffeoyl	2-Propyl benzene	H	H	H
8	Ac_2caffeoyl	3-Propyl benzene	H	H	H
9	Ac_2caffeoyl	4-Propyl benzene	H	H	H
10	Ac_2caffeoyl	2-Isopropyl benzene	H	H	H
11	Ac_2caffeoyl	3-Isopropyl benzene	H	H	H
12	Ac_2caffeoyl	4-Isopropyl benzene	H	H	H
13	Ac_2caffeoyl	C_6H_5	H	*n*-Butyl	H
14	Ac_2caffeoyl			H	*n*-Butyl
15	Ac_2caffeoyl		H	H	H
16	Ac_2caffeoyl		H	H	H
17	Ac_2caffeoyl		H	H	H
18	Ac_2caffeoyl		CF_3	H	H
19	Ac_2caffeoyl		4-Ethyl benzene	CF_3	
20	Ac_2caffeoyl		CF_3	H	H

Table 9 Continued

Comp. no.	R_1	R_2	R_3	R_4	R_5
21	Ac_2caffeoyl		CF_3	H	H
22	Ac_2caffeoyl		H	H	H
23	Ac_2caffeoyl		H	H	H
24	Ac_2caffeoyl		SCF_3	H	H
25	Ac_2caffeoyl		SO_2CF_3	H	H
26	Ac_2caffeoyl		SO_2CF_3	H	H
27	Ac_2caffeoyl		H	H	H
28	Ac_2caffeoyl	C_6H_5	H	$SO_2C_2H_5$	H
29	Ac_2caffeoyl	C_6H_5	$SO_2C_2H_5$	H	H
30	Ac_2caffeoyl		CF_3	H	H

designed series. The structures of these 30 molecules show some relationship with the activities. The general structure of the designed molecule is



The presence of aromatic groups at the R2 position increases the predicted activity of caffeoyl naphthalene sulfonamide derivatives. The presence of alkyl groups such as ethyl, propyl, and isopropyl on this aromatic ring improves the partition coefficient of the molecule without decreasing the predicted activity of the molecule. Similarly, the presence of groups such as CF_3, SO_2CF_3, and SCF_3 increases the partition coefficient and predicted activity of the compound. The presence of a caffeoyl group is essential for good predicted activity. Lower alkyl groups can be attached at positions R4 and R5 to improve the partition coefficient without any decrease in predicted activity.

Acknowledgment Kamlesh K. Sahu thanks the All India Council for Technical Education (AICTE), Delhi, India for providing a fellowship to fund this work.

References

Carteau S, Mouscadet JF, Goulaouic H, Subra F, Auclair C (1993) Effect of topoisomerase inhibitors on the in vitro HIV DNA integration reaction. Biochem Biophys Res Commun 192:1409

Chen IJ, Neamati N, Nicklaus MC, Orr A, Anderson L, Barchi JJ Jr, Kelly JA, Pommier Y, MacKerell AD Jr (2000) Identification of HIV-1 integrase inhibitors via three-dimensional database searching using ASV and HIV-1 integrases as targets. Bioorg Med Chem 8:2385

Craigie R, Fujiwara T, Bushman F (1990) The IN protein of Moloney murine leukemia virus processes the DNA viral ends and accomplishes their integration in vitro. Cell 62:829

Cushman M, Sherman P (1992) Inhibition of HIV-1 integration protein by aurintricarboxylic acid monomers, monomer analogs, and polymer fractions. Biochem Biophys Res Commun 185:58

De Clercq E (1995) Toward improved anti-HIV chemotheraphy: therapeutic strategies for intervention with HIV infections. J Med Chem 38:2491

Eich E, Pertz H, Kaloga M, Schulz J, Fesen MR, Mazumder A, Pommier Y (1996) (–)-Arctigenin as a lead structure for inhibitors of human immunodeficiency virus type-1 integrase. J Med Chem 39:86

Fesen MR, Kohn KW, Leteurtre F, Pommier Y (1993) Inhibitors of human immunodeficiency virus integrase. Proc Natl Acad Sci USA 90:2399

Hansch C (1969) A quantitative approach to biochemical structure-activity relationships. Acc Chem Res 2:232

Katz RA, Merkel G, Kulkosk T, Leis J, Salka AM (1990) The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. Cell 63:87

Mazumder A, Raghawan K, Weinstein J, Kohn KW, Pommier Y (1995) Inhibition of human immunodeficiency virus type-1 integrase by curcumin. Biochem Pharmacol 49:1165

Mazumder A, Neamati N, Sommadossi JP, Gosseli G, Schinazi RF, Pommier Y (1996) Effects of nucleotide analogues on human immunodeficiency virus type 1 intrgrase. Mol Pharmacol 49:621

Mouscadet JF, Carteau S, Goulaouic H, Subra F, Auclair C (1994) Triplex-mediated inhibition of HIV DNA integration in vitro. J Biol Chem 269:21635

Neamati N (2002) Patented small molecule inhibitors of HIV-1 integrase: a 10-year saga. Expert Opin Ther Patents 12:709

Parril AL (2003) HIV-1 integrase inhibition: binding sites, structure activity relationships and future perspectives. Curr Med Chem 10:1811

Robinson WE Jr, Cordeiro M, Abdel-Malek S, Jia Q, Chow SA, Reineck MG, Mitchell WM (1996a) Dicaffeoylguinic acid inhibitors of human immunodeficiency virus integrase: inhibition of the core catalytic domain of human immunodeficiency virus integrase. Mol Pharmacol 50:846

Robinson WE Jr, Reineck MG, Abdel-Malek S, Jia Q, Chow SA (1996b) Inhibitors of HIV-1 replication inhibit HIV integrase. Proc Natl Acad Sci USA 93:6326

Santo RD, Costi R, Artico M, Tramontano E, Colla PL, Pani A (2003) HIV-1 integrase inhibitors that block HIV-1 replication in infected cells. Planning synthetic derivatives from natural products. Pure Appl Chem 75:195

Xu YW, Zhao GS, Shin CG, Zang HC, Lee CK, Lee YS (2003) Caffeoyl naphtalenesulfonamide derivatives as HIV integrase inhibitors. Bioorg Med Chem 11:3589