Journal of Enzyme Inhibition and Medicinal Chemistry, 2002 Vol. 17 (4), pp. 219–225

Taylor & Francis healthsciences

Quantitative Structure–Activity Relationship Study of 2-arylsulfonyl-6-substituted Benzonitriles as Non-nucleoside Reverse Transcriptase Inhibitors of HIV-1

B.K. SHARMA, RAJENDRA KUMAR and P. SINGH*

Department of Chemistry, S.K. Government College, Sikar 332 001, India

(Received 8 December 2001)

The reverse transcriptase inhibition of HIV-1, the most common form of HIV, by non-nucleoside 2-arylsulfonyl-6-substituted benzonitriles is analysed through Fujita-Ban and Hansch approaches. The analyses have helped to ascertain the role of different substituents in explaining the observed inhibitory actions of these compounds. From both approaches it appeared that SO_2 instead of SO or S at *X*; and NH_2 instead of F at *Y* (see Figure 1) are advantageous to improving the activity of a compound against HIV-1. This in turn leads to the suggestion that the 2-arylsulfonyl-6-aminobenzonitrile scaffold is the only appropriate structural entity that may further result into potential compounds. Further, the compounds having a OMe substituent at the *ortho*-position, the bulkier substituents at *meta*-positions and "no" substituent at *para*-position of 2-arylsulfonyl moiety are beneficial in raising the activity. The two quantitative structure–activity relationship (QSAR) analyses, differing in parametric approach, therefore, provided the grounds for rationalizing the substituent selection in designing more potent compounds of the series.

Keywords: 2-Arylsulfonyl-6-substituted benzonitriles; Non-nucleoside reverse transcriptase inhibitors of HIV-1; QSAR analysis; Fujita-Ban and Hansch approaches; Physicochemical properties

INTRODUCTION

The human immunodeficiency virus (HIV) is a pathogenic retrovirus and causative agent of Acquired Immunodeficiency Syndrome (AIDS) and its related disorders. The first successful combination therapy for the treatment of HIV-1, the most common form of HIV, infections with the non-nucleoside reverse transcriptase inhibitor (NNRTI) nevirapine[1] as a component in the treatment regimen has stimulated interest in the search for novel and potent NNRTIs. Instead of acting at the catalytic site of HIV reverse transcriptase (RT) by terminating DNA synthesis,[2] NNRTIs bind in a region of the enzyme at a certain distance (10 Å) away from the catalytic site. This results in the distortion of the catalytic site because of changes in the position of the key aspartic acid residues, which affects the ability of the enzyme to carry out its catalytic functions. Further, NNRTIs do not require anabolism for activation, as they are not analogues of natural compounds and do not utilize the biochemical processes involved within the host cells. Milder side effect, such as rash, at the clinical level[3] are sometimes observed on treatment with NNRTIs. This therapy was therefore restricted initially. Their potential for the treatment of HIV infections was, however, broadly realized when nevirapine in combination with ddI and AZT was found to lead to a sustained viral load reduction.[1] Thus the approval by the FDA of nevirapine for combination therapy was immediately followed by that for delavirdine[4] and efavirenz.[5] Such multiple-drug treatment approach has greatly reduced morbidity and mortality among HIV-infected patients.[6,7] HIV chemotherapy is still in its primitive stage because of a high pill burden, the high cost of treatment and toxicity profile. These limitations may, however, be overcome by a treatment regimen containing a potent NNRTI with a limited side effect profile. NNRTIs

*Corresponding author. E-mail: imran_khan_sonu@yahoo.com

ISSN 1475-6366 print/ISSN 1475-6374 online © 2002 Taylor & Francis Ltd
DOI: 10.1080/1475636021000010932

are generally noncompetitive potent inhibitors of HIV-1 RT with low cost compounds possessing a permissible toxicity profile. In view of this Chan *et al.*[8] have recently reported the synthesis and structure–activity relationship (SAR) studies that led to the identification of a group of analogues with the 2-arylsulfonyl-6-aminobenzonitrile scaffold having potent antiviral activity against HIV-1 and good therapeutic indices. Their SAR studies on these analogues were aimed only at the alteration of substituents at different positions of the parent moiety and provided no rationale to reduce the trial-and-error factors. Hence, a quantitative SAR (QSAR) study on these analogues is conducted so as to provide the rationale for drug-design and to explore the possible mechanism of their action.

MATERIALS AND METHODS

The reported series of compounds[8] consists of 2-arylsulfonyl-6-substituted benzonitriles, 2-arylsulfinyl-6-substituted benzonitriles and 2-arylthio-6-substituted benzonitriles. The numbering indicated therein are as shown in the general structure given by Figure 1 and as compiled in Table I. In order to establish important QSAR, both the Fujita-Ban[9] and the Hansch type analyses were carried out on these congeners. The biological inhibitory effects against HIV-1-infected cells and relevant physicochemical parameters are also included in Table I. The biological assay assessed the reduction of the cytopathic effect of HIV-1 on MT4 cells in the presence of compounds, and the results were reported as IC_{50} values, which are the concentrations of compounds that would produce a 50% decrease in the cytopathic effect. In the present work the same are, however, expressed as $-\log IC_{50}$ on a molar basis. The most suitable parameters were found to be the hydrophobic

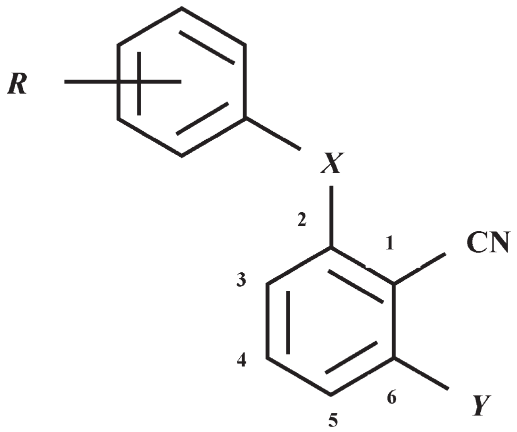

FIGURE 1 Structure of 2-Arylsulfonyl-6-substituted benzonitrile derivatives.

constant, π and the molar refraction parameter, *MR* that were taken directly from the compilation of Hansch *et al.*[10] Additionally, some indicator variables were also employed to account for the binary effects due to certain substituents at the varying positions of the parent structure.

In the Fujita-Ban approach, which is based on an additivity principle, the biological activity is expressed as,

$$-\log IC_{50} = \sum a_i X_i + \mu \quad (1)$$

where a_i is the group contribution of the ith substituent, X_i has a value of 1 if the ith substituent is present otherwise the value is 0 and μ is the theoretical biological activity of the (unsubstituted) reference compound of the series. The linear equations generated using Equation (1) were solved by the multiple regression analysis[11] (MRA) employing the method of least squares[12] for the unknowns a_i and μ.

The Hansch approach involves finding the "best" fit of a dependent variable, $-\log IC_{50}$ to a linear combination of the independent variables, usually the physicochemical parameters, by the method of least squares. This is more formally expressed as in Equation (2)

$$-\log IC_{50} = a_0 + a_1X_1 + a_2X_2 + \cdots\cdots\cdots + a_nX_n \quad (2)$$

where the descriptors $X_1, X_2\ldots\ldots\ldots X_n$ are the physicochemical parameters. Step-wise regression was used to develop the best QSAR from the available descriptors. Once the best equation is established, it may be used to increase the understanding of the mechanisms of actions of sets of congeners and to direct drug design in congeneric series, as well as to attempt to predict biological activities of untested compounds quantitatively. This methodology has also been called the extrathermodynamic, linear free energy, multiple parameter and physicochemical structure–activity relationship (PSAR) approach. In addition to physicochemical parameters the indicator variables, representing the presence or absence of certain structural characteristics, are sometimes also used in this approach.

RESULTS AND DISCUSSION

Initially, fifty nine compounds in Table I were retained in the construction of the Fujita-Ban matrix with compound **1** as the reference congener. The frequency of occurrence of certain group in eight compounds (**4**, **10**, **14**, **27–30** and **63**; Table I) at a specified position is only one and the same were not included in the training data set. Tabulation of the resulting matrix of 59 linear equations in

TABLE I QSAR parameters and inhibition potency of 2-arylsulfonyl-6-substituted benzonitrile derivatives against HIV-1 (see Figure 1 for structure)

											$-\log IC_{50}$(M)			
Compound No.	*R*	*X*	*Y*	π_2	MR_4	MR_5	I_{Br}	I_{Me}	I_X	I_Y	Obsd[a]	Calc. F.B.	Calc. Eq. (3)	Calc. Eq. (4)
1	H	SO_2	NH_2	0.00	0.103	0.103	0	0	1	0	5.70	5.70	5.86	5.84
2	2-OMe	SO_2	NH_2	−0.02	0.103	0.103	0	0	1	0	6.22	6.26	5.87	5.85
3	3-OMe	SO_2	NH_2	0.00	0.103	0.103	0	0	1	0	6.05	5.77	5.86	5.84
4	4-OMe	SO_2	NH_2	0.00	0.787	0.103	0	0	1	0	4.60	–[c]	4.88	4.88
5	2-Me	SO_2	NH_2	0.56	0.103	0.103	0	0	1	0	5.64	5.45	5.57	5.55
6	3-Me	SO_2	NH_2	0.00	0.103	0.103	0	0	1	0	6.40	6.27	5.86	5.84
7	4-Me	SO_2	NH_2	0.00	0.565	0.103	0	0	1	0	5.02	5.05	5.20	5.19
8	2-Cl	SO_2	NH_2	0.71	0.103	0.103	0	0	1	0	5.39	5.70	5.49	5.48
9	3-Cl	SO_2	NH_2	0.00	0.103	0.103	0	0	1	0	6.23	6.08	5.86	5.84
10	4-Cl	SO_2	NH_2	0.00	0.603	0.103	0	0	1	0	5.52	–[c]	5.14	5.14
11	2-Br	SO_2	NH_2	0.86	0.103	0.103	0	0	1	0	5.30	5.37	5.41	5.40
12	3-Br	SO_2	NH_2	0.00	0.103	0.103	1	0	1	0	6.27	6.29	6.38	6.39
13	4-Br	SO_2	NH_2	0.00	0.888	0.103	0	0	1	0	4.70	5.21	4.73	4.74
14	2-F	SO_2	NH_2	0.14	0.103	0.103	0	0	1	0	5.52	–[c]	5.79	5.76
15	3-F	SO_2	NH_2	0.00	0.103	0.103	0	0	1	0	5.52	5.70	5.86	5.84
16	2-CN	SO_2	NH_2	−0.57	0.103	0.103	0	0	1	0	5.27	5.85	6.16	6.12
17	2-CN	SO_2	NH_2	0.00	0.103	0.103	0	0	1	0	5.62	6.04	5.86	5.84
18	2-CN	SO_2	NH_2	0.00	0.633	0.103	0	0	1	0	4.10	4.66	5.10	5.09
19	3-CF_3	SO_2	NH_2	0.00	0.103	0.103	0	0	1	0	5.46	5.54	5.86	5.84
20	2,5-Cl_2	SO_2	NH_2	0.71	0.103	0.603	0	0	1	0	6.52	6.40	6.59	6.57
21	3,5-Cl_2	SO_2	NH_2	0.00	0.103	0.603	0	0	1	0	7.15	6.78	6.97	6.93
22	3,5-Me_2	SO_2	NH_2	0.00	0.103	0.565	0	1	1	0	8.00	7.51	7.71	7.70
23	3-Br,5-Me	SO_2	NH_2	0.00	0.103	0.565	1	0	1	0	7.70	7.54	7.41	7.40
24	3-Cl,5-Me	SO_2	NH_2	0.00	0.103	0.565	0	0	1	0	7.52	7.32	6.88	6.85
25	3-OMe,5-Me	SO_2	NH_2	0.00	0.103	0.565	0	0	1	0	7.30	7.02	6.88	6.85
26	3-OMe,5-CF_3	SO_2	NH_2	0.00	0.103	0.502	0	0	1	0	7.05	6.63	6.74	6.71
27	3-OH,5-Me	SO_2	NH_2	0.00	0.103	0.565	0	0	1	0	6.37	–[c]	6.88	6.85
28	3-OEt,5-Me	SO_2	NH_2	0.00	0.103	0.565	0	0	1	0	7.22	–[c]	6.88	6.85
29	3-OPr,5-Me	SO_2	NH_2	0.00	0.103	0.565	0	0	1	0	7.22	–[c]	6.88	6.85
30	3-OBu, 5-Me	SO_2	NH_2	0.00	0.103	0.565	0	0	1	0	6.22	–[c]	6.88	6.85
31	3,5-Cl_2	SO_2	F	0.00	0.103	0.603	0	0	1	1	4.38	4.51	4.61	4.60
32	3,5-Me_2	SO_2	F	0.00	0.103	0.565	0	1	1	1	5.82	5.24	5.36	5.37
33	3-OMe,5-Me	SO_2	F	0.00	0.103	0.565	0	0	1	1	4.29	4.75	4.53	4.52
34	2-OMe	SO	NH_2	−0.02	0.103	0.103	0	0	0	0	5.32	5.24	4.95	4.94
35	3-OMe	SO	NH_2	0.00	0.103	0.103	0	0	0	0	4.80	4.75	4.94	4.93
36	2-Me	SO	NH_2	0.56	0.103	0.103	0	0	0	0	4.03	4.43	4.65	4.64
37	3-Me	SO	NH_2	0.00	0.103	0.103	0	0	0	0	4.53	5.25	4.94	4.93
38	4-Me	SO	NH_2	0.00	0.565	0.103	0	0	0	0	4.31	4.03	4.28	4.28
39	2-Br	SO	NH_2	0.86	0.103	0.103	0	0	0	0	4.41	4.35	4.49	4.49
40	3-Br	SO	NH_2	0.00	0.103	0.103	1	0	0	0	7.10[b]	–	–	–
41	4-Br	SO	NH_2	0.00	0.888	0.103	0	0	0	0	4.69	4.19	3.82	3.83
42	2-CN	SO	NH_2	−0.57	0.103	0.103	0	0	0	0	5.41	4.83	5.24	5.21
43	3-CN	SO	NH_2	0.00	0.103	0.103	0	0	0	0	4.85	5.02	4.94	4.93
44	3-CF_3	SO	NH_2	0.00	0.103	0.103	0	0	0	0	4.40	4.52	4.94	4.93
45	3,5-Me_2	SO	NH_2	0.00	0.103	0.565	0	1	0	0	6.47	6.49	6.80	6.79
46	2,5-Cl_2	SO	NH_2	0.71	0.103	0.603	0	0	0	0	5.01	5.38	5.68	5.66
47	3-Cl,5-Me	SO	NH_2	0.00	0.103	0.565	0	0	0	0	6.49	6.30	5.97	5.94
48	3-OMe,5-CF_3	SO	NH_2	0.00	0.103	0.502	0	0	0	0	5.68	5.62	5.83	5.80
49	H	S	NH_2	0.00	0.103	0.103	0	0	0	0	4.84	4.84	4.94	4.93
50	2-OMe	S	NH_2	−0.02	0.103	0.103	0	0	0	0	5.37	5.40	4.95	4.94
51	3-OMe	S	NH_2	0.00	0.103	0.103	0	0	0	0	5.22	4.91	4.94	4.93
52	2-Me	S	NH_2	0.56	0.103	0.103	0	0	0	0	4.80	4.59	4.65	4.64
53	3-Me	S	NH_2	0.00	0.103	0.103	0	0	0	0	5.21	5.40	4.94	4.93
54	4-Me	S	NH_2	0.00	0.565	0.103	0	0	0	0	3.94	4.19	4.28	4.28
55	2-Cl	S	NH_2	0.71	0.103	0.103	0	0	0	0	5.39	4.83	4.57	4.57
56	3-Cl	S	NH_2	0.00	0.103	0.103	0	0	0	0	5.13	5.22	4.94	4.93
57	2-Br	S	NH_2	0.86	0.103	0.103	0	0	0	0	4.52	4.51	4.49	4.49
58	3-Br	S	NH_2	0.00	0.103	0.103	1	0	0	0	5.29	5.43	5.47	5.48
59	3-F	S	NH_2	0.00	0.103	0.103	0	0	0	0	5.01	4.83	4.94	4.93
60	3-CN	S	NH_2	0.00	0.103	0.103	0	0	0	0	5.76	5.18	4.94	4.93
61	4-CN	S	NH_2	0.00	0.633	0.103	0	0	0	0	4.36	3.80	4.18	4.18
62	3-CF_3	S	NH_2	0.00	0.103	0.103	0	0	0	0	4.89	4.68	4.94	4.93
63	3-NH_2	S	NH_2	0.00	0.103	0.103	0	0	0	0	4.50	–[c]	4.94	4.93
64	3,5-Me_2	S	NH_2	0.00	0.103	0.565	0	1	0	0	6.37	6.65	6.80	6.79
65	3-Cl,5-Me	S	NH_2	0.00	0.103	0.565	0	0	0	0	5.75	6.46	5.97	5.94
66	3-OMe,5-Me	S	NH_2	0.00	0.103	0.565	0	0	0	0	5.70	6.15	5.97	5.94
67	3-OMe,5-CF_3	S	NH_2	0.00	0.103	0.502	0	0	0	0	5.29	5.77	5.83	5.80

[a]Inhibition concentration, on molar scale, that produces a 50% decrease in the HIV-induced cytopathic effect; taken from Ref. 8. [b]'Outlier' compound in the present study. [c]Compounds not considered in formulating Fujita-Ban matrix.

TABLE II Fujita-Ban contributions of substituents and parent moiety to the anti HIV-1 activities of title compounds

Position	Substitution	Contribution to $-\log IC_{50}$	
		n = 59	n = 58
R	2-Br	−0.373	−0.328
	2-CN	0.084	0.150
	2-Cl	−0.005	−0.002
	2-Me	−0.294	−0.249
	2-OMe	0.520	0.565
	3-Br	1.039	0.594
	3-CF_3	−0.201	−0.156
	3-CN	0.294	0.339
	3-Cl	0.406	0.380
	3-F	−0.001	−0.001
	3-Me	0.570	0.568
	3-OMe	0.085	0.075
	4-Br	−0.558	−0.492
	4-CN	−1.040	−1.040
	4-Me	−0.692	−0.647
	5-CF_3	0.806	0.861
	5-Cl	0.640	0.701
	5-Me	1.156	1.243
X	SO	−0.881	−1.019
	S	−0.855	−0.862
Y	F	−2.202	−2.271
	Contribution due to parent moiety, μ	5.695	5.698

22 unknowns including the contribution of the parent compound is avoided here for the sake of brevity. These equations were solved by the method of least squares for the unknowns, μ and a_i. The contributions of various substituents obtained thereby are summarized in the first column of Table II and the resulting statistical parameters of the study are:

$$n = 59,\ r = 0.915,\ s = 0.495,\ F(22, 36) = 8.392$$

where n, r, s and F are respectively the number of data points, multiple regression coefficient, standard error of estimate and F-ratio between the variances of calculated and observed activities. The datum point **40** is the only compound whose calculated $-\log IC_{50}$ value (5.85) was found to be much lower than the observed value. Perhaps this compound entails an error in the observed activity value. This data point was, therefore, ignored to improve the results further. In doing so, the corresponding row was removed from the Fujita-Ban matrix and the MRA of resulting matrix lead to the results summarized in the last column of Table II. The improved statistical parameters of the study are:

$$n = 58,\ r = 0.934,\ s = 0.433,\ F(22, 35) = 10.911$$

The r^2-value now accounts for 87% of the variance and the F-value obtained is significant at 99% level [$F_{22,35}(0.01) = 2.41$]. The calculated values of $-\log IC_{50}$, listed in Table I, are also in close agreement with the observed ones. The substituents, to be incorporated at various positions of the parent moiety, that make higher positive contributions to activity may only be used to design more active compounds of the series in future. From Table II, the possible combinations of various substituents have the following pattern:

R (Positions)			*X*	*Y*
2	3	5		
OMe	Br	Me	SO_2	NH_2
	Me	CF_3		
		Cl		

The negative contributions obtained for SO and S relative to SO_2 at *X*; and F relative to NH_2 at *Y* are favoring only the 2-arylsulfonyl-6-aminobenzonitriles scaffold having potential for antiviral activity against HIV-1 over that of either 2-arylsulfinyl-6-aminobenzonitriles or 2-arylthio-6-aminobenzonitriles. The optimal activities seem to be manifested by compounds where 2-, 3- and 5-positions in the arylsulfonyl moiety are substituted respectively by OMe, Br (or Me) and Me (or CF_3, Cl) groups. Such a prediction may help in synthesizing the analogues, having trisubstitution in the arylsulfonyl moiety, which have not yet been reported.

It is important to note that the Fujita-Ban approach cannot extrapolate beyond the substituents used in the training set whereas the Hansch approach, used next for the same data set of 58 compounds, can do so. In a preliminary study of this approach, a large number of descriptors pertaining to hydrophobic, electronic and steric interactions were examined for varying sites of the molecules in various possible ways. For this purpose, a set of descriptors such as π (hydrophobicity), *HD* (hydrogen-bond donor), *HA* (hydrogen-bond acceptor), σ (electronic, *meta* and *para*), *F* (field), *R* (resonance), μ (dipole moment), E_S (Taft's steric), *MR* (molar refraction), *MW* (molecular weight), V_W (van der Waals volume) were initially choosen for *R* substituents of each of the 2-, 3-, 4-, and 5-positions of the aryl sulfonyl moiety. However, no descriptor was found suitable for 3-*R* substituents of this moiety, rather two indicator variables I_{Br} and I_{Me} (selected arbitrarily for 3-Br and 3,5-Me_2) were able to explain the variation in observed activity values. Two more indicator variables, I_X and I_Y (accounting for binary variations of *X* and *Y* substituents) were also incorporated. Thus, a large number of regression equations in seven independent variables were obtained through the MRA. These equations were then subjected to various statistical tests and the highest significant correlation as shown by Equation (3) is finally obtained. The steps of its

TABLE III Stepwise development of regression Equation (3) $-\log IC_{50} = a_0 + a_1\pi_2 + a_2MR_4 + a_3MR_5 + a_4I_{Br} + a_5I_X + a_6I_Y$

a_0	a_1	a_2	a_3	a_4	a_5	a_6	r	s	$F^a_{k,n-k-1}$	
5.158	−0.496(±0.72)				0.983(±0.45)	−1.310(±1.00)	0.553	0.813	7.935	(i)
5.546	−0.672(±0.62)	−2.125(±0.97)			0.995(±0.39)	−1.491(±0.87)	0.700	0.703	12.765	(ii)
4.876	−0.611(±0.43)	−1.508(±0.68)	2.575(±0.65)		0.919(±0.26)	−2.296(±0.63)	0.877	0.478	34.688	(iii)
4.849	−0.579(±0.42)	−1.453(±0.67)	2.574(±0.64)	0.464(±0.57)	0.893(±0.26)	−2.248(±0.62)	0.884	0.470	30.279	(iv)

[a]The F-statistics obtained for n (= 58) data points and k (= 3, 4, 5, 6) independent variable(s).

development are described through the correlation Equations (i)–(iv) in Table III.

$$-\log IC_{50} = -0.523(\pm 0.39)\pi_2 - 1.437(\pm 0.61)MR_4 + 2.210(\pm 0.63)MR_5 + 0.524(\pm 0.52)I_{Br} + 0.830(\pm 0.50)I_{Me} + 0.917(\pm 0.24)I_X - 2.355(\pm 0.57)I_Y + 4.864 \quad n = 58, \quad r = 0.906, s = 0.429, F(7,50) = 32.816 \tag{3}$$

The binary variations, either SO_2 or other insertion such as SO and S at *X*; and NH_2 or F at *Y* are described respectively by the indicator variables I_X and I_Y. A value of 1 for I_X is taken for SO_2 and 0 otherwise. Likewise, a value of 1 or 0 for I_Y, in that order, indicates the presence or absence of NH_2. The variables π_2, MR_4 and MR_5 are, respectively the hydrophobicity of the 2-substituent and the molar refraction constants of 4- and 5-substituents in the arylsulfonyl moiety. The molar refractivity, generally, accounts for polarizability or steric bulk of the substituents. The substitutions 3-Br and 3,5-Me_2 in this moiety are further highlighted through the indicator variables, I_{Br} and I_{Me}. Thus, $I_{Br} = 1$ for 3-Br and $= 0$ otherwise. Similarly the presence or absence of 3,5-Me_2 in the arylsulfonyl moiety are assigned respectively the value 1 or 0 to I_{Me}. The parameters, *r*, *s*, and *F* obtained above for Equation (3) denote statistically sound results and the equation as such reflects the parameteric requirement of various substitutions at different positions in the 2-arylsulfonyl-6-substituted benzonitrile derivatives that are having potent antiviral activity against HIV-1. The r^2-value accounts for 82% of variance in the observed activity values and the *F*-value is significant at 99% level [$F_{7,50}(0.01) = 3.02$]. The calculated activity values, using Equation (3) and listed in Table I, are in close agreement with the observed ones. In addition, the activities of eight compounds that were outside the training set are also calculated using this equation. Their predicted activities, in this way, are also listed in Table I and are closely related to the observed ones. That the variables used in deriving Equation (3) have no mutual correlation is shown in Table IV. The validity of Equation (3) is strengthened further through consideration of all the compounds of Table I. The test set that was outside the training set earlier was also combined to give a relatively larger data set. The MRA on complete data have resulted in the correlation Equation (4) with the earlier compound (**40**) as the "outlier".

$$-\log IC_{50} = -0.505(\pm 0.38)\pi_2 - 1.400(\pm 0.55)MR_4 + 2.186(\pm 0.57)MR_5 + 0.551(\pm 0.51)I_{Br} + 0.856(\pm 0.48)I_{Me} + 0.909(\pm 0.22)I_X - 2.327(\pm 0.55)I_Y + 4.845 \quad n = 66, \quad r = 0.908, s = 0.427, F(7,58) = 38.999 \tag{4}$$

The statistical significance of the above equation is similar to that of Equation (3) and reflects slightly improved results, as in this equation, derived for a larger data set, the *r*-value is improved. Also the *F*-value, significant at 99% level [$F_{7,58}(0.01) = 2.96$] is increased. The mutual orthogonality requirements among the independent variables of Equation (4) are shown in Table V and the calculated activity values that closely resemble the observed ones are listed in Table I. For the compounds in the test set,

TABLE IV The intercorrelation matrix[a] amongst the independent variables of Equation (3)

	π_2	MR_4	MR_5	I_X	I_{Br}	I_{Me}	I_Y
π_2	1.000	0.015	0.001	0.002	0.006	0.008	0.006
MR_4		1.000	0.058	0.000	0.007	0.010	0.007
MR_5			1.000	0.024	0.000	0.166	0.132
I_X				1.000	0.011	0.001	0.067
I_{Br}					1.000	0.004	0.003
I_{Me}						1.000	0.060
I_Y							1.000

[a]Matrix elements are the r^2-values.

TABLE V The intercorrelation matrix[a] amongst the independent variables of Equation (4)

	π_2	MR_4	MR_5	I_X	I_{Br}	I_{Me}	I_Y
π_2	1.000	0.016	0.003	0.004	0.005	0.007	0.005
MR_4		1.000	0.075	0.003	0.007	0.010	0.007
MR_5			1.000	0.041	0.000	0.134	0.104
I_X				1.000	0.005	0.000	0.048
I_{Br}					1.000	0.003	0.002
I_{Me}						1.000	0.062
I_Y							1.000

[a]see footnote under Table IV.

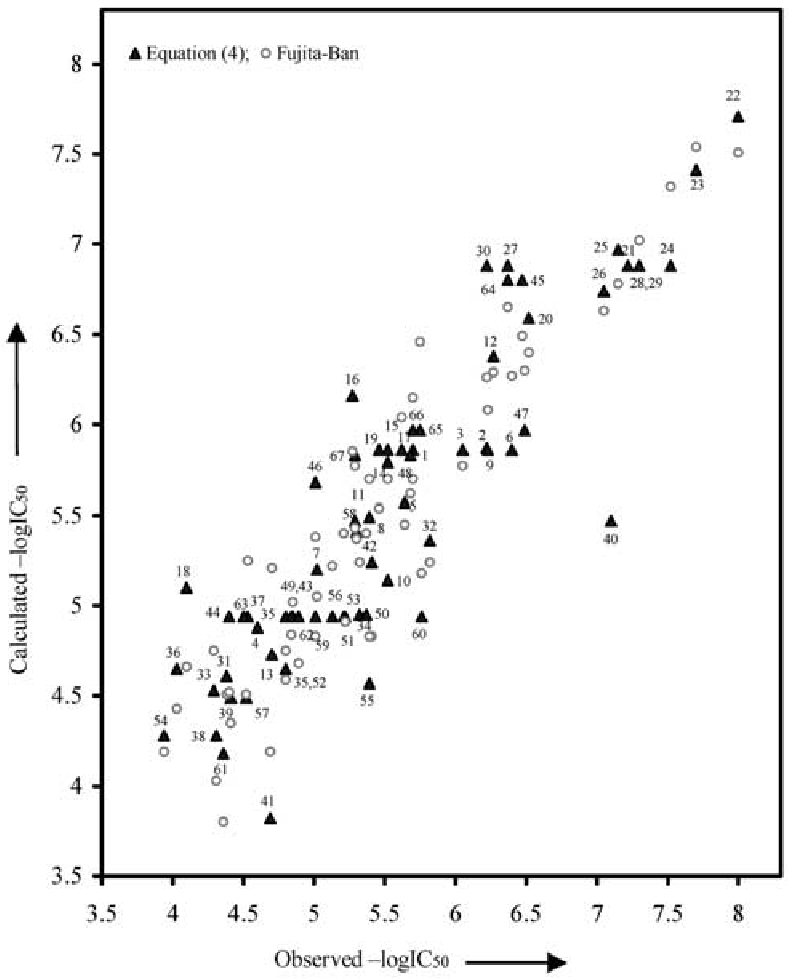

FIGURE 2 Plot of observed versus calculated $-\log IC_{50}$ values.

the observed and the calculated activities using Equation (3) and Equation (4) respectively are quite similar. From both of these equations, it appears that a less hydrophobic 2-substituent, less bulky/polarizable 4-substituent, more bulky/polarizable 5-substituent are helpful in improving the activity of a compound. Likewise, 3-Br and/or 3,5-Me_2 at *R*, SO_2 at *X*; and NH_2 at *Y* are also beneficial. An attempt to incorporate the electronic parameter instead of the molar refraction parameter, however, did not improve the statistical results rationally. The plot of observed versus predicted $-\log IC_{50}$s using Equation (4) and the Fujita-Ban study is also given in Figure 2 to demonstrate the goodness of fit and to show systematic variations of observed versus predicted activities in the present congeneric series.

Based on the QSAR results, obtained from both the Fujita-Ban and the Hansch approaches, the following conclusions may be drawn:

1. From Equations (3) or (4), it appears that the substituent such as SO_2 instead of SO or S at *X*; and NH_2 instead of F at *Y* are advantageous to improve the activity of a compound against HIV-1. The Fujita-Ban study, in conformity with this, assigned negative contributions to SO, S and F substituents. Thus, it may be suggested that the scaffold of 2-arylsulfonyl-6-aminobenzonitrile is only the prerequisite to have potent compounds and may be further explored in the near future.
2. The *para*-position of 2-arylsulfonyl moiety seems to be the least optimal position to place a substituent. The contributions obtained in the Fujita-Ban study for various 4-substituents are all negative. The Hansch study, in addition, predicts that the substituents having smaller values of molar refraction are appropriate. The *para*-substitution seems to have unfavorable interaction in the region of HIV-1 RT into which the substituents are projected. It is therefore, better to have this position unsubstituted.
3. The *meta*-positions of 2-arylsulfonyl moiety are highly sensitive to various substitutions that may lead to improved activity profiles. The Hansch study urged the presence of either 3-Br or 3,5-Me_2. The same was supported by the Fujita-Ban analysis in which higher positive contributions were obtained for such substituents.

4. Only the OMe substituent at the *ortho*-position of the 2-arylsulfonyl moiety is found, in the Fujita-Ban analysis, to have a positive substituent contribution that may additively improve the activity of a compound. The Hansch study, on the other hand, favors a less hydrophobic substituent at this position.

In conclusion, the compounds having a OMe at 2-position, bulkier (than Me) groups at 3,5-positions and "no" substitution at 4-position may further be synthesized to give more active analogues of the 2-arylsulfonyl-6-aminobenzonitrile scaffold. Also, the two analyses in the present study provide the ground for rationalizing substituent selection in designing more potent compounds of the series.

Acknowledgements

We are thankful to the Department of Chemistry and the Institution for providing the facilities to complete this work.

References

[1] Montaner, J.S.G., Reiss, P., Cooper, D., Vella, S., Harris, M., Conway, B., Wainberg, M.A., Smith, D., Robinson, P., Hall, D., Myers, M. and Lange, J.M.A. (1998) *J. Am. Med. Assoc.* **279**, 930–937.

[2] To date six nucleosides have been approved by the FDA: AZT, 3TC, d4T, ddI, ddC and abacavir.

[3] Moeremans, M., Raeymaeker, M.D., Broeck, R.V.D., Stoffels, P., Brabander, M.E., Cree, J.D., Hertogs, K., Pauwels, R., Staszewski, S. and Andries, K. (1995) *Abstract of the 4th International Workshop on HIV Drug Resistance*, Sardinia, Italy, July 6–9, p. 33.

[4] Freimuth, W.W. (1996) *Adv. Exp. Med. Bio.* **394**, 279–289.

[5] Young, S.D., Britcher, S.F., Tran, L.O., Payne, L.S., Lumma, W.C., Lyle, T.A., Huff, J.R., Anderson, P.S., Olsen, D.B., Carroll, S.S., Pettibone, D.J., O'Brien, J.A., Ball, R.G., Balani, S.K., Lin, J.H., Chen, I.-W., Schleif, W.A., Sardana, V.V., Long, W.J., Byrnes, V.W. and Emini, E.A. (1995) *Antimicrob. Agents Chemother.* **39**, 2602–2605.

[6] Hogg, R.S., Rhone, S.A., Yip, B., Sherlock, C., Conway, B., Schechter, M.T., O'Shaughnessy, M.V. and Montaner, J.S.G. (1998) *AIDS* **12**, 279–284.

[7] Palella, Jr, F.J., Delaney, K.M., Moorman, A.C., Loveless, M.O., Fuhrer, J., Satten, G.A., Aschman, D.J. and Holmberg, S.D. (1998) *N. Engl. J. Med.* **338**, 853–860.

[8] Chan, J.H., Hong, J.S., Hunter, III, R.N., Orr, F., Cowan, J.R., Sherman, D.B., Sparks, S.M., Reitter, B.E., Andrews, III, C.W., Hazen, R.J., Clair, M.S., Boone, L.R., Ferris, R.G., Creech, K.L., Roberts, G.B., Short, S.A., Weaver, K., Ott, R.J., Ren, J., Hopkins, A., Stuart, D.I. and Stammers, D.K. (2001) *J. Med. Chem.* **44**, 1866–1882.

[9] Fujita, T. and Ban, T. (1971) *J. Med. Chem.* **14**, 148–152.

[10] Hansch, C. and Leo, A.J. (1979) *Substituents Constants for Correlation Analysis in Chemistry and Biology* (John Wiley, New York).

[11] Snedecor, G.W. and Cochran, W.G. (1967) *Statistical Methods.* (Iowa State University Press, Ames).

[12] Draper, N.R. and Smith, H. (1966) *Applied Regression Analysis* (Wiley, New York).