European Journal of Medicinal Chemistry 44 (2009) 1377–1382

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

Topological descriptors in modeling the agonistic activity of human A_3 adenosine receptor ligands: The derivatives of 2-chloro-N^6-substituted-4′-thioadenosine-5′-uronamide

Susheela Sharma [a], Brij K. Sharma [b,*], Sanjeev K. Sharma [b], Prithvi Singh [b], Yenamandra S. Prabhakar [c]

[a] Department of Engineering Chemistry, Sobhasaria Engineering College, Sikar 332 021, Rajasthan, India
[b] Department of Chemistry, S.K. Government College, Silver Jubilee Road, Sikar 332 001, Rajasthan, India
[c] Medicinal and Process Chemistry Division, Central Drug Research Institute, Lucknow 226 001, India

ARTICLE INFO

Article history:
Received 11 April 2008
Received in revised form 10 September 2008
Accepted 15 September 2008
Available online 30 September 2008

Keywords:
Quantitative structure–activity relationship (QSAR)
2-Chloro-N^6-substituted-4′-thioadenosine-5′-uronamide derivatives as human A3 receptor ligands
Combinatorial Protocol in Multiple Linear Regression (CP-MLR)

ABSTRACT

The human A_3 adenosine receptor agonistic activity of 2-chloro-N^6-substituted-4′-thioadenosine-5′-uronamides has been analyzed through Combinatorial Protocol in Multiple Linear Regression (CP-MLR) using 488 topological descriptors obtained from DRAGON software for the energy minimized 3D-structures of these molecules. Among the various descriptor classes considered in the study, the human A_3 adenosine receptor agonistic activity is correlated with simple topological descriptors (TOPO), Modified Burden eigenvalues (BCUT) and functional group (FUNC) classes of descriptors. The average valence connectivity index of order zero, X0Av, the sum of topological distances between O and Cl, T (O···Cl) from the TOPO class, the lowest eigenvalue n.2 of Burden matrix/weighted by atomic masses, BELm2, from the BCUT class and the number of secondary aliphatic amides, nCONHR, from FUNC class have contributed significantly in the development of a statistical sound models. The models developed and the participating descriptors suggest that the substituent groups at N^6-position and/or 5′-uronamide of 2-chloro-N^6-substituted-4′-thioadenosine-5′-uronamide derivatives hold scope for structural modification in the optimization of activity.

1. Introduction

Adenosine is an endogenous signaling molecule which regulates many physiological functions through specific subtypes (A_1, A_{2A}, A_{2B}, and A_3) of adenosine receptors (ARs) and at least one of which is expressed on mast cells [1]. The recently identified A_3 subtype is reported to be involved in many a diseases such as cardiac [2] and cerebral ischemia [3], cancer [4], asthma [5] glaucoma [6] and inflammation [7,8]. Hence, it is regarded as a good target for the development of clinically useful agents for these disease states.

A number of nucleoside and nonnucleoside derivatives (e.g. IB-MECA, Cl-IB-MECA) have been tested for binding affinities at the human adenosine receptor subtypes [9–16]. IB-MECA was found to show potent in vivo antitumor activity [4] and is now undergoing phase II clinical trials whereas Cl-IB-MECA is being used extensively as a pharmacological tool for studying A_3 AR [17]. The structure–activity relationship study on the cyclopropyl-fused carbocyclic nucleoside (MRS 3558) [18], adopting a fixed conformation of the pseudoribose ring, indicated that the binding site of the human A_3 AR preferred the North sugar ring puckering. Recently, on the basis of a bioisosteric rationale and the high binding affinity and selectivity of above-mentioned compounds, Jeong et al. [19] have reported a new series of 2-chloro-N^6-substituted-4′-thioadenosine-5′-uronamide derivatives as highly potent and selective A_3 AR agonists. However, the structure–activity relationship (SAR) in these analogues is limited to substituent's variation at different positions of (N^6-position and/or 5′-uronamide) the moiety. Several of the recent QSAR studies on adenosine receptor agonists are focused on A_1, A_{2A}, A_{2B} subtypes [20–27]. Also a QSAR study of rat A_3 adenosine affinity of Cl-IB-MECA analogues with BCUT descriptors is reported [28]. In the present communication, a quantitative SAR (QSAR) study on 2-chloro-N^6-substituted-4′-thioadenosine-5′-uronamide analogues (Fig. 1, Table 1) was conducted, in order to provide the rationale for drug-design and to explore the possible mechanism of their action at molecular levels.

2. Methodology

2.1. Dataset

For the QSAR study, the human A_3 AR agonistic activity of 2-chloro-N^6-substituted-4′-thioadenosine-5′-uronamide derivatives

* Corresponding author. Tel.: +91 1572 251039.
E-mail address: bksharma_sikar@rediffmail.com (B.K. Sharma).

0223-5234/$ – see front matter © 2008 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2008.09.022

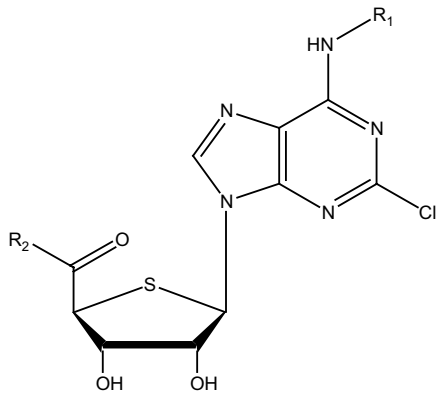

Fig. 1. Structures of 2-chloro-4′-thioadenosine-5′-uronamide derivatives.

(Table 1) expressed in CHO cells has been considered from the literature report [19]. In QSAR study the binding affinities are expressed as pK_i on a molar basis. The structures of the compounds have been drawn in ChemDraw [29] using the standard procedure. All these structures have been ported to DRAGON software [30] for computing the parameters corresponding to 0D-, 1D-, and 2D-descriptor classes. This has resulted in 488 descriptors corresponding to all the descriptor classes. These descriptor classes along with their definitions and scope in addressing the structural features have been presented in Table 2. As the total number of descriptors involved in this study is very large, only the names of descriptor classes and the actual descriptor involved in the models have been addressed in the discussion.

2.2. CP-MLR

The Combinatorial Protocol in Multiple Linear Regression (CP-MLR) has been used for the QSAR model development. It is a 'filter'-based variable selection procedure for the model identification and development in QSAR studies [31–36]. It involves a combinatorial strategy with appropriately placed 'filters' interfaced with MLR and extracts diverse models having unique combination of descriptors from the dataset. The filters set the thresholds for the descriptors in terms of inter-parameter correlation cutoff limits in subset regressions (filter-1), *t*-values of the regression coefficients (filter-2), internal explanatory power (filter-3; square root of adjusted multiple correlation coefficient of regression equation, *r*-bar), and the external consistency (filter-4; Q^2 i.e. cross-validated R^2 from the leave-one-out procedure). Throughout this study, for the filters-1, 2 and 4 the thresholds were assigned as 0.3, 2.0, and $0.3 \leq Q^2 \leq 1.0$, respectively. The filter-3 was assigned an initial value of 0.71. In

Table 1
Observed and modeled binding affinities of 2-chloro-4′-thioadenosine-5′-uronamide derivatives at the human A_3 receptors (see Fig. 1 for structures).

Compound number	R_1	R_2	pK_i (M)				
			Obsd.[a]	Eq. (13)		Eq. (14)	
				Calc.	LOO	Calc.	LOO
1	H	$NHCH_3$	9.40	8.54	8.46	8.54	8.46
2	CH_3	$NHCH_3$	9.55	8.75	8.69	8.76	8.71
3	CH_3	$NH(CH_3)_2$	5.82	6.80	6.99	6.94	7.18
4	CH_3	Cyclopropyl-NH	8.55	8.70	8.71	8.71	8.72
5	CH_3	Cyclopropylmethyl-NH	8.96	8.58	8.56	8.60	8.57
6	CH_3	Isoamyl-NH	8.79	8.38	8.36	8.43	8.41
7	CH_3	Morpholine	5.41	6.02	6.08	5.71	5.76
8	CH_3	4-Benzyl-piperidine	5.46	5.10	5.06	5.24	5.21
9	CH_3	4-(4-F-Phenyl)-piperazine	5.57	5.61	5.62	5.73	5.75
10	CH_3	3-F-Benzyl-NH	7.76	7.61	7.59	7.61	7.59
11	CH_3	2-(3-F-Phenyl)ethyl-NH	7.07	7.69	7.75	7.69	7.75
12	CH_3	3,3-Diphenylpropyl-NH	6.38	6.73	6.79	6.75	6.82
13	Cyclopropyl	$NHCH_3$	8.65	8.62	8.62	8.63	8.63
14	Cyclopropyl	$NHCH_2CH_3$	9.17	8.29	8.25	8.31	8.26
15	Cyclopropyl	Cyclopropyl-NH	8.25	8.60	8.63	8.60	8.63
16	Cyclopropyl	4-Benzylpiperidine	5.40	5.00	4.96	5.14	5.11
17	Cyclopropyl	Morpholine	5.35	6.04	6.13	5.73	5.82
18	Cyclopentyl	$NHCH_3$	8.37	8.36	8.36	8.38	8.38
19	Cyclopentyl	$NHCH_2CH_3$	8.55	8.38	8.37	8.40	8.39
20	3-Iodobenzyl	$NHCH_3$	9.42	8.99	8.92	9.05	8.99
21	3-Iodobenzyl	$NHCH_2CH_3$	9.05	8.99	8.97	9.05	9.05
22	3-Iodobenzyl	Isoamyl-NH	7.38	7.90	8.01	7.99	8.13
23	3-Iodobenzyl	Cyclopropyl-NH	8.53	8.85	8.88	8.90	8.95
24	3-Iodobenzyl	Cyclopropylmethyl-NH	8.44	8.88	8.93	8.94	9.00
25	3-Iodobenzyl	Cyclobutyl-NH	7.74	7.22	7.14	7.30	7.23
26	3-Iodobenzyl	4-Benzylpiperidine	6.06	5.17	5.03	5.35	5.22
27	3-Iodobenzyl	4-(4-F-Phenyl)-piperazine	5.84	4.97	4.85	5.14	5.02
28	3-Iodobenzyl	Morpholine	6.29	6.19	6.17	5.94	5.85
29	3-Iodobenzyl	3-Cl-Benzyl-NH	6.51	7.17	7.27	6.33	5.94
30	3-Iodobenzyl	3-(Trifluoromethyl)benzyl-NH	6.45	6.62	6.65	6.67	6.70
31	3-Iodobenzyl	2-Phenylethyl-NH	6.36	7.03	7.12	7.10	7.21
32	3-Iodobenzyl	3,3-Diphenylpropyl-NH	6.46	6.87	6.93	6.93	7.00
33	2-Methylbenzyl	$NHCH_3$	8.66	8.12	8.10	8.14	8.11
34	2-Methylbenzyl	$NHCH_2CH_3$	8.60	8.17	8.15	8.19	8.17
35	2-Methylbenzyl	$N(CH_3)_2$	6.20	6.38	6.41	6.52	6.56
36	2-Methylbenzyl	Cyclopropyl-NH	7.56	8.03	8.06	8.04	8.07
37	2-Methylbenzyl	Cyclopropylmethyl-NH	7.53	8.08	8.11	8.09	8.12
38	2-Methylbenzyl	Morpholine	5.12	5.48	5.52	5.18	5.19
39	2-Methylbenzyl	4-Benzylpiperidine	4.31	4.06	4.00	4.22	4.19

[a] The affinities, determined in terms of binding constant, K_i, for human A_3 adenosine receptor and derived from the inhibition of [^{125}I]I-AB-MECA as a radioligand on Chinese hamster ovary (CHO) cells stably transfected with cDNA encoding the human ARs; taken from Ref. [19].

Table 2
Descriptor classes used for the analysis of binding affinities of 2-chloro-4′-thioadenosine-5′-uronamide derivatives at the human A_3 receptors and identified categories in modeling the activity.

Descriptor class (acronyms)[a]	Definition and scope	Descriptor category[b]
Constitutional (CONS)	Dimensionless or 0D descriptors; independent from molecular connectivity and conformations	I
Topological (TOPO)	2D-descriptor from molecular graphs and independent conformations	I
Molecular walk counts (MWC)	2D-descriptors representing self-returning walk counts of different lengths	III
Modified Burden eigenvalues (BCUT)	2D-descriptors representing positive and negative eigenvalues of the adjacency matrix, weighs the diagonal elements and atoms	III
Galvez topological charge indices (GVZ)	2D-descriptors representing the first 10 eigenvalues of corrected adjacency matrix	I
2D-autocorrelations (2D-AUTO)	Molecular descriptors calculated from the molecular graphs by summing the products of atom weights of the terminal atoms of all the paths of the considered path length (the *lag*)	I
Functional groups (FUNC)	Molecular descriptors based on the counting of the chemical functional groups	I
Atom centered fragments (ACF)	Molecular descriptors based on the counting of 120 atom centered fragments, as defined by Ghose-Crippen	I
Empirical (EMP)	1D-descriptors represent the counts of nonsingle bonds, hydrophilic groups and ratio of the number of aromatic bonds and total bonds in an H-depleted molecule	III
Properties (PROP)	1D-descriptors representing molecular properties of a molecule	I

[a] Ref. [30].
[b] Descriptor categories identified at the end of second stage; in this the filter values are as follows: filter-1 as 0.3, filter-2 as 2.0, filter-3 as 0.71, and filter-4 as $0.3 \leq Q^2 \leq 1.0$, the number of compounds in each dataset is 39.

order to collect the descriptors with higher information content, the threshold of filter-3 was successively incremented with increasing number of descriptors (per equation) by considering the *r*-bar value of the preceding optimum model as the new threshold for the next generation.

2.3. Descriptor classification protocol

This modeling study involves nine descriptor classes belonging to 0D-, 1D-, and 2D-descriptor classes. A three-stage descriptor classification protocol [32] is implemented with the two-descriptor combinations (baseline models), as they are the simplest to understand and explain the activity. In the first stage of the classification protocol, the correlations of the activity with two-descriptor combinations from the individual descriptor classes (DCs) of the dataset were used to sort the DCs into four categories. They are primary contributors (category I: a DC forms a model with its constituent descriptors), collective contributors (category II: a DC unable to form a model with its constituent descriptors, but forms model(s) in combination with a descriptor from another such DC), secondary contributors (category III: a DC forms a model(s) only in combination with category I) and non-contributors (category IV: a DC unable to form a model(s) in any manner like that of categories I, II, and III). The sorted DCs were collated in the second stage to identify all the three-descriptor models across the categories. In the last stage, the individual descriptors of all three-descriptor models were pooled to discover the higher models for the activity.

All the models identified in the last stage have further been put to a randomization test [33,37] by repeated randomization of

Table 3
The selected QSAR models that emerged in one- and two-descriptors from different descriptor classes belonging to category I.

Descriptor class	Constant	Descriptor-1	Descriptor-2	r^a	s	F	Q^2	Eq.
CONS	9.703	−1.112nR06	–	0.726	1.026	41.237	0.479	(1)
	5.642	0.453AMW	−1.224nR06	0.801	0.906	32.141	0.590	(2)
TOPO	21.280	−2.487MAXDP	–	0.833	0.825	84.139	0.665	(3)
	35.939	−2.699MAXDP	−38.147PW3	0.874	0.735	58.127	0.715	(4)
GVZ	7.954	−5.334GGI5	28.403JGI1	0.793	0.921	30.517	0.575	(5)
2D-AUTO	12.228	−0.035ATS5e	–	0.722	1.032	40.279	0.472	(6)
	10.730	−0.029ATS5e	26.518MATS1p	0.862	0.766	52.125	0.688	(7)
FUNC	−1.313	2.318nHDon	–	0.804	0.888	67.498	0.612	(8)
	6.491	−0.173nCaH	2.382nCONHR	0.896	0.670	73.701	0.768	(9)
ACF	−1.313	2.318H-050	–	0.804	0.888	67.498	0.612	(10)
	−0.115	−0.158C-024	2.159H-050	0.886	0.701	65.739	0.741	(11)
PROP	27.494	−0.267PSA	−0.650MLOGP	0.826	0.852	30.729	0.625	(12)

[a] In all, the number of compounds (n) is 39, r is the correlation coefficient, Q^2 is cross-validated index from leave-one-out (LOO) procedure, s is the standard error of the estimate and F is the F-ratio between the variances of calculated and observed activities.

Table 4
Descriptors identified for modeling the binding affinities of 2-chloro-4′-thioadenosine-5′-uronamide derivatives at the human A_3 receptors along with the average regression coefficients and the total incidences.

Descriptor class	Descriptor[a]	Avg reg coeff (incidence)[b]	Descriptor[a]	Avg reg coeff (incidence)[b]	Descriptor[a]	Avg reg coeff (incidence)[b]
CONS	AMW	0.197(1)	Sv	−0.110(2)	Ss	−0.060(1)
	nSK	−0.132(2)	nBO	−0.116(1)	SCBO	−0.084(1)
TOPO	IAC	−0.050(1)	ZM1V	−0.010(1)	Dz	−0.065(3)
	LPRS	−0.018(1)	VDA	−0.011(1)	Xu	−0.159(3)
	WA	−0.681(3)	RDSUM	−0.024(1)	QW	−0.001(1)
	TI2	−0.653(4)	Rww	−0.129(3)	D/D	−0.007(1)
	MAXDP	−2.281(2)	X0	−0.199(1)	X1	−0.269(3)
	X2	−0.297(1)	X0Av	13.772(42)	X1Av	25.026(27)
	RDCHI	−1.398(3)	BLI	8.346 (25)	UNIP	−0.015(1)
	IDM	−1.501(3)	IDDM	−2.948(3)	IVDE	10.628(1)
	IVDM	−1.789(1)	HDcpx	−17.461(3)	TIC0	−0.050(1)
	Eig1e	−0.012(1)	SEigv	0.260(2)	SEige	−1.417(3)
	SEigp	0.167(2)	VEA2	32.329(3)	VRA1	−0.001(4)
	VRA2	−0.035(4)	VED1	−1.534(3)	VED2	47.461(3)
	VEZ1	−1.517(3)	VEZ2	47.776(3)	VEv1	−1.502(3)
	VEv2	49.366(3)	VEe1	−1.523(3)	VEe2	47.958(1)
	VEp1	−1.499(3)	VEp2	49.153(3)	CID	−0.063(1)
	BID	−0.133(2)	D/Dr09	−0.025(3)	T (O···Cl)	−0.046(4)
MWC	SRW02	−0.058(1)	–	–	–	–
BCUT	BELm2	−8.863(5)	BELe3	−6.780(3)	–	–
2D-AUTO	MATS2m	59.841(2)	MATS1v	17.446(2)	MATS1p	29.746(1)
FUNC	nCONHR	2.056(104)	nHDon	1.823(1)	–	–

[a] The descriptors are identified from the three-parameter model that emerged from CP-MLR protocol with filter-1 as 0.30; filter-2 as 2.0; filter-3 as 0.91; filter-4 as $0.3 \leq Q^2 \leq 1.0$; number of compounds in the study are 39; **CONS**: AMW, average molecular weight; Sv, sum of atomic van der Waals volumes; Ss, sum of Kier–Hall electro-topological states; nSK, number of non-H atoms; nBO, number of non-H-bonds; SCBO, sum of conventional bond orders (H-depleted); **TOPO**: IAC, total information index of atomic composition; ZM1V, first Zagreb index by valence vertex degrees; Dz, Pogliani index; LPRS, log of product of row sums (PRS); VDA, average vertex distance degree; Xu, Xu index; WA, mean Wiener index; RDSUM, reciprocal distance Wiener-type index; QW, quasi-Wiener index (Kirchhoff number); TI2, second Mohar index; Rww, reciprocal hyper-detour index; D/D, distance/detour index; MAXDP, maximal electrotopological positive variation; X0, X1 and X2, connectivity indices of order 0, 1 and 2, respectively; X0Av and X1Av, average valence connectivity indices of 0 and 1 order; RDCHI, reciprocal distance Randic-type index; BLI, Kier benzene-likeliness index; UNIP, unipolarity; IDM and IDDM, mean information contents on the distance magnitude and distance degree magnitude, respectively; IVDE and IVDM, mean information content on the vertex degree equality and magnitude, respectively; HDcpx, graph distance complexity index (log); TIC0, total information content index (neighborhood symmetry of 0-order); Eig1e, leading eigenvalue from electronegativity weighted distance matrix; SEigv, SEige and SEigp, eigenvalue sum from van der Waals volume, electronegativity and polarizability weighted distance matrix, respectively; VEA2, average eigenvector coefficient sum from adjacency matrix; VRA1 and VRA2, Randic and average Randic-type eigenvector-based index from adjacency matrix; VED1 and VED2; VEZ1 and VEZ2; VEv1 and VEv2; VEe1 and VEe2; VEp1 and VEp2, eigenvector and average eigenvector coefficient sum from distance matrix, Z weighted distance matrix (Barysz matrix), van der Waals weighted distance matrix, electronegativity weighted distance matrix, polarizability weighted distance matrix, respectively; CID and BID, Randic and Balaban ID number, respectively; D/Dr09, distance/detour ring index of order 9; T (O···Cl), sum of topological distances between O···Cl; **MWC**: SRW02, self-returning walk count of order 02; **BCUT**: BELm2, lowest eigenvalue n.2 of Burden matrix/weighted by atomic masses; BELe3, lowest eigenvalue n.3 of Burden matrix/weighted by atomic Sanderson electronegativities; **2DAUTO**: MATS2m, Moran autocorrelation −*lag* 2/weighted by atomic masses; MATS1v and MATS1p, Moran autocorrelation −*lag* 1/weighted by atomic van der Waals volumes and polarizabilities, respectively; **FUNC**: nCONHR, number of secondary aliphatic amides; nHDon is the number of donor atoms for H-bonds (with N and O).

[b] The average regression coefficient of the descriptor corresponding to all models and the total number of its incidences; the arithmetic sign of the coefficient represents the actual sign of the regression coefficient in the models.

the activity to discover the chance correlations, if any, associated with them. For this every model has been subjected to 100 simulation runs with scrambled activity. The scrambled activity models with regression statistics better than or equal to that of the original activity model have been counted to express the percent chance correlation of the model under scrutiny. The model development procedure has been further verified by creating divergent training and test sets from the compounds of the study.

3. Results and discussion

Initially, the human A_3 receptor binding affinity of 39 analogues of 2-chloro-N^6-substituted-4′-thioadenosine-5′-uronamide is investigated with a variety of 0D-, 1D- and 2D-descriptors obtained from DRAGON software. Several of these descriptors recognized from seven different classes have shown significant correlations and are identified as the primary contributors (category I) in modeling the inhibitory activity of these compounds. Table 3 lists various models (Eqs.(1)–(12)) derived using such descriptors from each class.

Among the one- and two-descriptor models, the binding affinity of the compounds has been explained well by category I descriptor classes. The CONS descriptors that appeared in models favor the bulkiness of the compound (AMW, average molecular weight) in addition to the least preference of six-membered rings (nR06, number of six-membered rings) in a structure for optimum activity. The TOPO class descriptors, MAXDP (maximal electrotopological positive variation) and PW3 (path/walk 3 – Randic shape index) are negatively correlated to the activity. These descriptors recommend minimum positive field effects and lower path/walk for better activity. In model 5, the Galvez topological charge index of order 5 (GGI5), which corresponds to the fifth eigenvalue of the corrected adjacency matrix of the molecule, advocates a lower value of this index for better activity. On the other hand, a higher value of mean topological charge index of order 1 (JGI1) supplements the activity. The 2D-AUTO descriptors, ATS5e (Broto-Moreau autocorrelation of *lag* 5/weighted by atomic Sandersons electronegativities) and MATS1p (Moran autocorrelation −*lag* 1/weighted by atomic polarizabilities) suggest the importance of *lags* 5 and 2 weighted by the respective properties. Among the FUNC class of descriptors, nHDon (number of donor atoms for H-bonds with N and O) and

Table 5
Predicted residual activity of different test sets (10 compounds each) of compounds of Table 1 and training and test sets statistics.

Compound number	Residual[a]					
	Des Clus[b]		Act Clus[c]		Random[d]	
	D_{13}[e]	D_{14}[f]	A_{13}[e]	A_{14}[f]	Ra_{13}[e]	Ra_{14}[f]
1					0.96	0.93
5			0.32	0.32	0.47	0.44
7	−0.57	−0.08	−0.61	−0.39		
9					−0.09	−0.26
11			−0.74	−0.74		
12			−0.39	−0.44		
13	−0.08	−0.11			0.12	0.09
15	−0.46	−0.50				
17					−0.73	−0.40
20			0.61	0.55		
21					0.13	0.11
24	−0.50	−0.49	−0.29	−0.35		
25			0.75	0.63	0.53	0.58
26	0.96	0.86	1.12	0.96		
28	0.17	0.71				
29	−0.73	0.75			−0.66	0.52
32	−0.50	−0.45				
33	0.44	0.42			0.61	0.60
34			0.40	0.38		
36	−0.57	−0.60				
37					−0.47	−0.48
39			0.34	0.19		
r	0.945	0.950	0.936	0.941	0.937	0.942
s	0.543	0.520	0.530	0.521	0.541	0.532
F	69.664	55.533	59.197	46.499	60.169	47.231
Q^2_{LOO}	0.852	0.856	0.834	0.836	0.835	0.829
Q^2_{LGO}	0.839	0.864	0.827	0.800	0.857	0.851
Test set r^2	0.775	0.774	0.854	0.885	0.869	0.892

[a] The difference of observed pK_i and predicted pK_i; the training models have 29 compounds each.
[b] Test set from the cluster analysis of full descriptor dataset of the compounds.
[c] Test set from the cluster analysis of the activity of the compounds.
[d] Test set from random selection of the compounds.
[e] D, A and Ra, respectively, represent Descriptor Cluster, Activity Cluster, and Random selections; the training set equations are formed with the descriptors of Eq. (13).
[f] The training set equations are formed with the descriptors of Eq. (14).

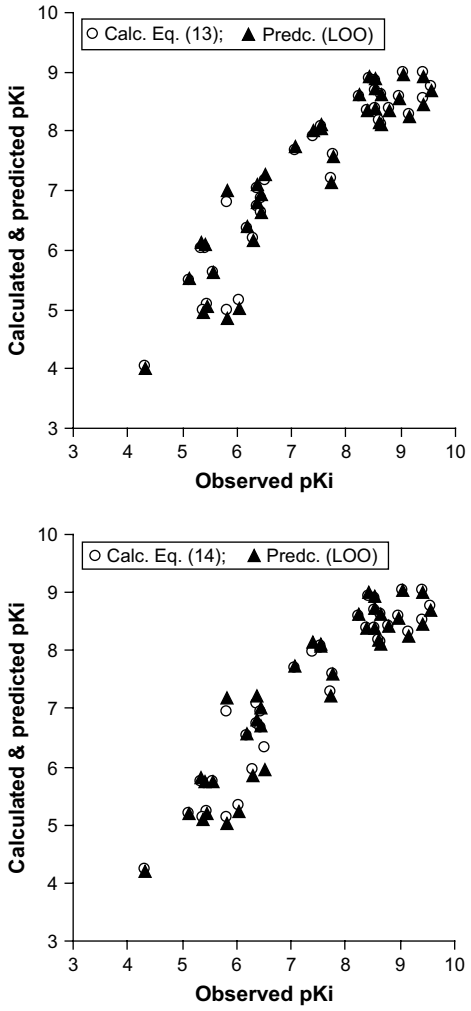

Fig. 2. Plot of observed versus calculated and predicted pK_i.

nCONHR (number of secondary amides, aliphatic) correlate positively to the activity and favor such functional groups in the molecule. The descriptor, nCaH {number of unsubstituted aromatic C(sp2)} favors the least substituted aromatic carbon atoms in the aromatic ring. The ACF class descriptors, C-024 (structure fragment type: R–CH–R) and H-050 (H attached to hetero atom) recommend certain structural fragments in a molecule for better activity. Fragment-based polar surface area (PSA), a molecular property suggests lower polar area whereas, Moriguchi octanol–water partition coefficient (MLOGP) recommends higher hydrophobicity of a molecule to improve the activity.

The significance of various emerged models, listed in Table 3, may be ascertained through the statistical parameters, the correlation coefficient r, the standard error of the estimate s, and the Fisher's ratio F. Additionally, the cross-validated index Q^2, obtained from leave-one-out (LOO) procedure, may assist in identifying the robustness of these models.

Even though each individual descriptor class is enriched with information corresponding to the activity, different descriptor classes together have led to the models with optimum explained variance. These models (with collective descriptor) were identified in CP-MLR by successively incrementing the filter-3 with increasing number of descriptors (per equation). For this the optimum r-bar value of the preceding level model has been used as the new threshold of filter-3 for the next generation. At the end of a search, 108 three-parameter models sharing 63 descriptors among themselves were identified. All these 63 descriptors are shown in Table 4 along with their brief meaning, average regression coefficients and total incidences corresponding to three-parameter models which will serve as a measure of their estimate across the models. These descriptors hold scope for even higher models. The following models are selected three- and four-parameter models evolved from these descriptors.

$$\mathrm{p}K_i = 11.875 + 15.797(4.222)\mathrm{X0Av} - 8.598(1.047)\mathrm{BELm2} + 1.904(0.201)\mathrm{nCONHR}$$

$$n = 39,\ r = 0.936,\ s = 0.538,\ F = 83.099,\ Q^2_{LOO} = 0.846,\ Q^2_{LGO} = 0.844 \quad (13)$$

$$\mathrm{p}K_i = 12.025 + 16.777(4.030)\mathrm{X0Av} - 0.035(0.016)\mathrm{T(O\cdots Cl)} - 8.460(0.995)\mathrm{BELm2} + 1.798(0.197)\mathrm{nCONHR}$$

$$n = 39,\ r = 0.944,\ s = 0.511,\ F = 70.456,\ Q^2_{LOO} = 0.858,\ Q^2_{LGO} = 0.862 \quad (14)$$

Now the r-values are increased to account, respectively, for 88% and 89% of variance in the observed activities and the s-

values are decreased. In addition, the data within parentheses are lowered and the Q^2 index (both from LOO and LGO) is increased. The latter index hints at a reasonable robust QSAR model.

Above model equations were further subjected to randomization process, where 100 simulations per model were carried out but none of the identified models have shown any chance correlation. Additionally, the above model equation was also validated through three test sets, each containing 10 compounds out of the 39 active ones listed in Table 1. Of the three test sets, two were generated in the SYSTAT [38] using the single linkage hierarchical cluster procedure involving the Euclidean distances of the respective descriptors or the activity as the case may be. The selection of the test set from the cluster tree was done in such a way to keep the test compounds at a maximum possible distance from each other. The third test of the compounds corresponds to the random selection procedure. These test sets represent different cross-sections of compounds. The predictions of the test sets have been done with the models developed using the 29 compounds remaining in the training sets. The residuals of the predictions and the corresponding predictive r^2-, s-, Q^2- and F-values of each test set have been given in Table 5. The predictions corresponding to all the test set compounds are within the reasonable limits of their actual values (Fig. 2, Table 5).

As far as the physical meaning of the newly appeared descriptors is concerned, in above equations X0Av (average valence connectivity index, chi-0), and T (O···Cl) belong to TOPO class. The descriptor X0Av encodes information about size, branching, cyclization, unsaturation and heteroatom content in the molecule whereas, T (O···Cl) is the sum of topological distances between O and Cl. The coefficients of these descriptors suggest that a larger value of X0Av and a smaller value of T (O···Cl) would be beneficiary to the activity. The BCUT class descriptor, BELm2, is the lowest eigenvalue n.2 of Burden matrix/weighted by atomic masses. This descriptor indicates the role of atomic masses of the constituent elements of the molecule in the activity. A higher value of this parameter, BELm2, is detrimental to the activity. Moreover, the presence of secondary aliphatic amide (nCONHR) is beneficiary to augment the activity of 2-chloro-N^6-substituted-4′-thioadenosine-5′-uronamide derivatives.

4. Conclusions

The study has provided detailed structure–activity relationship of the binding affinity of 2-chloro-N^6-substituted-4′-thioadenosine-5′-uronamide derivatives (Table 1) in terms of structural requirements. The activity of the compounds is a cumulative influence of different structural features which are identified in terms of individual descriptors. The positive regression coefficient of parameter nCONHR clearly favors the presence of secondary aliphatic amide group for better activity. The regression coefficient of descriptor X0Av emphasizes the importance of the branching of the molecules for improved activity. The descriptor T (O···Cl) hints that a closer topological path between O and Cl in the structural frame would be better for the activity. The associated regression coefficients of the Burden matrix corresponding to atomic masses suggest that heavy massed molecule poses detrimental effect to the activity. All the test sets (from the descriptors of the study, the activity, and the random selection procedures) have been predicted well by their corresponding training set equations. The models obtained and the participating descriptors suggest that the substituent groups at N^6-position and/or 5′-uronamide of 2-chloro-N^6-substituted-4′-thioadenosine-5′-uronamides derivatives hold scope for further modification. These results will be helpful to advance the understanding of modeling aspects of 2-chloro-N^6-substituted-4′-thioadenosine-5′-uronamide derivatives as human A_3 receptor ligands.

Acknowledgement

We express our sincere gratitude to our Institutions for providing necessary facilities to carry out this work. CDRI communication no. 7489.

References

[1] M.E. Olah, G.L. Stiles, Pharmacol. Ther. 85 (2000) 55–75.
[2] B.T. Liang, K.A. Jacobson, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 6995–6999.
[3] D.K.J.E. von Lubitz, W. Ye, J. McClellan, R.C.-S. Lin, Ann. N.Y. Acad. Sci. 90 (1999) 93–106.
[4] P. Fishman, L. Madi, S. Bar-Yehuda, F. Barer, L. Del Valle, K. Khalili, Oncogene 21 (2002) 4060–4064.
[5] C.J. Meade, J. Mierau, I. Leon, H.A. Ensinger, J. Pharmacol. Exp. Ther. 279 (1996) 1148–1156.
[6] M.Y. Avila, R.A. Stone, M.M. Civan, Invest. Opthalmol. Vis. Sci. 43 (2002) 3021–3026.
[7] F.G. Sajjadi, K. Takabayashi, A.C. Foster, R.C. Domingo, G.S. Firestein, J. Immunol. 156 (1996) 3435–3442.
[8] V. Ramkumar, G.L. Stiles, M.A. Beaven, H. Ali, J. Biol. Chem. 268 (1993) 16887–16890.
[9] K.-N. Klotz, Naunyn Schmiedebergs Arch. Pharmacol. 362 (2000) 382–391.
[10] P.G. Baraldi, B. Cacciari, R. Romagnoli, S. Merighi, K. Varani, P.A. Borea, G. Spalluto, Med. Res. Rev. 20 (2000) 103–128.
[11] E.W. van Tilburg, J. von Frijtag Drabbe Kunzel, M. de Groote, R.C. Vollinga, A. Lorenzen, A.P. Ijzerman, J. Med. Chem. 42 (1999) 1393–1400.
[12] E.W. van Tilburg, P.A.M. van der Klein, J. von Frijtag Drabbe Kunzel, M. de Groote, C. Stannek, A. Lorenzen, A.P. Ijzerman, J. Med. Chem. 44 (2001) 2966–2975.
[13] Z.-G. Gao, S.K. Kim, T. Biadatti, W. Chen, K. Lee, D. Barak, S.G. Kim, C.R. Johnson, K.A. Jacobson, J. Med. Chem. 45 (2002) 4471–4484.
[14] Z.-G. Gao, J.B. Blaustein, A.S. Gross, N. Melman, K.A. Jacobson, Biochem. Pharmacol. 65 (2003) 1675–1684.
[15] Z.-G. Gao, L.S. Jeong, H.R. Moon, H.O. Kim, W.J. Choi, D.H. Shin, E. Elhalem, M.J. Comin, N. Melaman, L. Mamedova, A.S. Gross, J.B. Rodriguez, K.A. Jacobson, Biochem. Pharmacol. 67 (2004) 893–901.
[16] H.O. Kim, X.-D. Ji, S.M. Siddiqi, M.E. Olah, G.L. Stiles, K.A. Jacobson, J. Med. Chem. 37 (1994) 3614–3621.
[17] K.A. Jacobson, L.J.S. Knutsen, in: M.P. Abbracchio, M. Williams (Eds.), Handbook of Experimental Pharmacology, vol. 151/I, Springer-Verlag, Berlin, Germany, 2001, pp. 129–175.
[18] K.A. Jacobson, X.-D. Xi, A.-H. Li, N. Melman, M.A. Siddiqui, K.-J. Shin, V.E. Marquez, R.G. Ravi, J. Med. Chem. 43 (2000) 2196–2203.
[19] L.S. Jeong, H.W. Lee, K.A. Jacobson, H.O. Kim, D.H. Shin, J.A. Lee, Z.-G. Gao, C. Lu, H.T. Duong, P. Gunaga, S.K. Lee, D.Z. Jin, M.W. Chun, H.R. Moon, J. Med. Chem. 49 (2006) 273–281.
[20] M.P. González, C. Teran, Bull. Math. Biol. 66 (2004) 907–920.
[21] M.P. González, C. Teran, Bioorg. Med. Chem. 12 (2004) 2985–2993.
[22] M.P. González, C. Teran, Y. Fall, M. Teijeira, P. Besada, Bioorg. Med. Chem. 13 (2005) 601–618.
[23] M.P. González, C. Teran, M. Teijeira, P. Besada, Bioorg. Med. Chem. Lett. 15 (2005) 2641–2645.
[24] M.P. González, C. Teran, M. Teijeira, P. Besada, M.J. Gonzalez-Moa, Eur. J. Med. Chem. 40 (2005) 1080–1086.
[25] M.P. González, C. Teran, M. Teijeira, P. Besada, A.M. Helguera, Bull. Math. Biol. 69 (2007) 347–359.
[26] M.P. González, C. Teran, M. Teijeira, A.H. Morales, Curr. Med. Chem. 13 (2006) 2253–2266.
[27] M.P. González, C. Teran, M. Teijeira, A.H. Morales, Eur. J. Med. Chem. 41 (2006) 56–62.
[28] M.P. González, C. Teran, M. Teijeira, P. Besada, M.J. Gonzalez-Moa, Bioorg. Med. Chem. Lett. 15 (2005) 3491–3495.
[29] ChemDraw Ultra 6.0 and Chem3D Ultra, Cambridge Soft Corporation, Cambridge, USA, 2001.
[30] R. Todeschini, V. Consonni, DRAGON Software (version 1.11-2001), Milano, Italy, 2003.
[31] Y.S. Prabhakar, QSAR Comb. Sci. 22 (2003) 583–595.
[32] M.K. Gupta, Y.S. Prabhakar, J. Chem. Inf. Model. 46 (2006) 93–102.
[33] Y.S. Prabhakar, V.R. Solomon, R.K. Rawal, M.K. Gupta, S.B. Katti, QSAR Comb. Sci. 23 (2004) 234–244.
[34] Y.S. Prabhakar, Internet Electron J. Mol. Des. 3 (2004) 150–162.<http://www.biochempress.com>.
[35] M.K. Gupta, R. Sagar, A.K. Shaw, Y.S. Prabhakar, Bioorg. Med. Chem. 13 (2005) 343–351.
[36] M. Saquib, M.K. Gupta, R. Sagar, Y.S. Prabhakar, A.K. Shaw, R. Kumar, P.R. Maulik, A.N. Gaikwad, S. Sinha, A.K. Srivastava, V. Chaturvedi, R. Srivastava, B.S. Srivastava, J. Med. Chem. 50 (2007) 2942–2950.
[37] S.S. So, M. Karplus, J. Med. Chem. 40 (1997) 4347–4359.
[38] SYSTAT, Version 7.0, SPSS Inc., 444 North Michigan Avenue, Chicago, IL, 60611.