Digest Journal of Nanomaterials and Biostructures Vol. 3, No.2, June 2008, p. 63 - 73

QSAR STUDY OF DISUBSTITUTED N^6-CYCLOPENTYLADENINE ANALOGUES AS A ADENOSINE A_1 RECEPTOR ANTAGONIST

Abhishek K. Jain[a], V. Ravichandran[a], Rajesh Singh[a], Simant Sharma[a], V. K. Mourya[b], R. K. Agrawal[a*]

[a]*Pharmaceutical Chemistry Research Laboratory, Department of Pharmaceutical Sciences, Dr. H. S. Gour Vishwavidyalaya, Sagar (M.P) – 470 003, India*
[b]*Government College of Pharmacy, Aurangabad (MH), India*

In pursuit of better adenosine A_1 receptor antagonist agents, QSAR studies were performed on a series of disubstituted N^6-cyclopentyladenine analogues. Stepwise multiple linear regression analysis was performed to derive QSAR models which were further evaluated for statistical significance and predictive power by internal and external validation. The best QSAR model was selected, having correlation coefficient (r) = 0.879, standard error of estimation (SEE) = 0.368 and cross validated squared correlation coefficient (q^2) = 0.664. The predictive ability of the selected model was also confirmed by leave one out cross validation. The QSAR model indicates that the dielectric energy, connectivity index 1, dipole vector Y, dipole vector Z, and HOMO energy play an important role for the A1 receptor antagonist activities. The results of the present study may be useful on the designing of more potent disubstituted N^6-cyclopentyladenine analogues as adenosine A_1 receptor antagonist agents.

(Received April 2, 2008; accepted April 9, 2008)

Keywords: QSAR; disubstituted N^6-cyclopentyladenine analogues, Adenosine A_1 receptor antagonist

1. Introduction

Denosine is a neuromotor which produces many important biological functions by activation of G protein coupled receptors that are classified in to A_1 , A_2B, and A_3 subtypes. Adenosine receptors from different species shows 87-93% amino acid sequence homology ,the only exception being the A_3 subtypes which exhibit 74% primary sequence homology between rat and human [1-3] adenosine receptor are involved in many peripheral and central regulatory mechanism including vasodilation [4], vasoconstriction in the kidney [5], inhibition of lypolysis and insulin release [6] and moderation of cerebral ischemi [7].

The first A_1 receptor antagonists were xantine derivatives , such as theophylline ,since then a variety of different classes of heterocyclic compounds has described to possess antagonist activity at adenosine receptor, xantine, adenines, 7-deazaadenine,7-deaza-8-ajapurine, pyrazolo (3-4-c)quinolines, pyrazolo-(1-5-α) pyridine and 1-8-naphthyridine. E.W Van Tilburg synthesized a series of 4-methyl-(2-phenyl-carboxamido-)-1,3-thiazole derivatives as potential antagonist for the adenosine A_1 receptors [8-14].

* Corresponding author: dragrawal2001@yahoo.co.in

Rianne et al. [15] expressed N_9 and C-8 position for increase adenosine A_1 receptor affinity, small substituents at the 2-position of adenines or adenosines only have limited effects on adenosine A_1 receptor affinity.

Linden and co-workers [16] investigated C-8 position of adenines to some extent. They synthesized 8-substituted N^6-norbornyl-9-methyladenines and found that N-containig group at this position enhances A_1 receptor affinity while introduction of alkyl chain on the C-8 position of adenines led to selective adenosine A_3 receptor antagonist [17].

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. Senior author of the article Dr. R. K. Agrawal and his research team has developed a few quantitative structure-activity relationship models to predict biological activity of different group of compounds [18-26].

2. Results and discussion

A data set of 37 compounds of reported series[15] for adenosine A1 receptor antagonist activity was used for the present QSAR study (Table 1). The QSAR studies of the N^6-cyclopentyladenine analogues series resulted in several QSAR equations. The two best equations are:

pKi = 10.653 (± 1.520) DE – 0.670 (± 0.256) CI1 +0.00311 (± 0.033) MR – 0.279 (± 0.106) DVZ + 0.450 (± 0.102) DVY + 2.782 (± 0.863) HE + 31.380 (± 8.420)……. (1)

n = 33, r = 0.883, r^2 = 0.780, r^2_{adj} = 0.729, q^2 = 0.638, F = 15.35, SEE = 0.3695, S_{PRESS} = 0.430, P < 0.001.

pKi = 11.076 (± 1.451) DE – 0.434 (± 0.066) CI1 – 0.275 (± 0.106) DVZ + 0.508 (± 0.081) DVY + 3.224 (± 0.727) HE + 35.705 (± 7.080)…….. (2)

n = 33, r = 0.879, r^2 = 0.772, r^2_{adj} = 0.730, q^2 = 0.644, F = 18.30, SEE = 0.3689, S_{PRESS} = 0.430, P < 0.001.

In the above equations n is the number of compounds used to derive the model and values in parentheses are the 95% confidence limit of respective coefficient. We extended our study for five-parametric correlations as they are permitted for a data set of 33 compounds in accordance with the lower limit of rule of thumb. Correlation matrix of the parameters in best model is given in table 3.

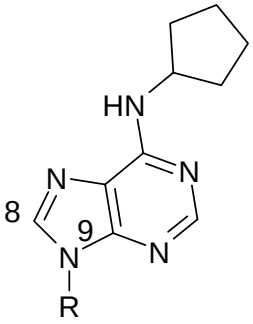
The calculated and predicted (LOO) activities of the compounds by the above models are shown in table 4. Model-1 shows good correlation coefficient (r) of 0.883 between descriptors (DE, CI1, MR, DVY, DVZ, and HE) and A_1 receptor antagonist activity. Squared correlation coefficient (r^2) of 0.780 explains 78.0% variance in biological activity.

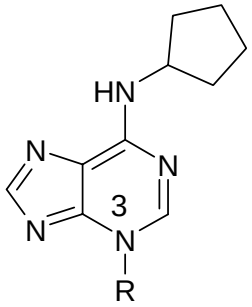
This model also indicates statistical significance > 99.9% with F values F = 15.35. Cross validated squared correlation coefficient of this model was 0.638, which shows the good internal prediction power of this model. Model-2 shows good correlation coefficient (r) of 0.879 between descriptors (DE, CI1, HE, DVY, and DVZ) and A_1 receptor antagonist activity. Squared correlation coefficient (r^2) of 0.772 explains 77.2% variance in biological activity. This model also indicates statistical significance > 99.9% with F values F = 18.30. Cross validated squared correlation coefficient of this model was 0.644, which shows the good internal prediction power of this model.

Consequently equation-2 can be considered as most suitable model with both high statistical significant and excellent predictive ability.

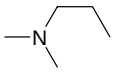
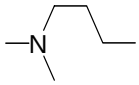
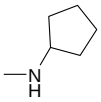
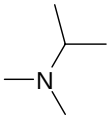
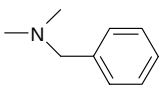
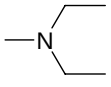
The predictive ability of model-2 was also confirmed by external r^2CVext. The robustness of the selected model was checked by Y – randomization test. The low r^2 and q^2 values indicate (data not shown) that the good results in our original model are not due to a chance correlation or structural dependency of the training set. The predictive ability of this model was also confirmed by external cross validation (equation 3). Consequently equation-2 can be considered as most suitable model with both high statistical significant and excellent predictive ability.

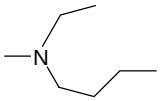
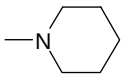
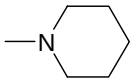
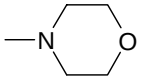
Table 1. Structures, biological activity of the N^6-cyclopentyladenine analogues.





Compd.	**C8**	**N9-R**	**N3-R**	**K_i (nm),A_1 receptor**
1	Br	H	-	2646
2	Br	Methyl	-	43
3c	Br	Methyl	-	467
4	Br	Allyl	-	35
5	Br	Propyl	-	33
6	Br	Benzyl	-	1220
7	H	Benzyl	-	1810
8	Br	-	Methyl	2760
9	Br	-	Propyl	995
10	Br	-	Benzyl	870
11	OCH_3	Allyl	-	208
12	OCH_3	Propyl	-	270
13	OCH_3	Methyl	-	120
14	OC_2H_5	Methyl	-	106
15	$OCH(CH_3)_2$	Methyl	-	40

16	OC_3H_7	Methyl	-	235
17	SC_2H_5	Methyl	-	224
18	=O	Methyl	-	1610
19	$NHCH_3$	-	-	206
20		-	-	169
21		-	-	89
22		-	-	160
23		-	-	1011
24	NHC_2H_5	-	-	344
25		-	-	2040
26		-	-	3560
27		-	-	7.7
28		-	-	5900
29		-	-	75

30		-	-	706
31		-	-	28
32		-	-	68
33		-	-	2840

K_i = Displacement of {^{3}H} DPCPX from CHO- A_1 membrane

Table 2. Selected descriptors involved in developing QSAR models.

Comp.	CI1	DVY	DVZ	DE	HE	MR
1	7.83	7.831	-0.529	-0.61	-8.749	66.162
2	8.254	8.254	-0.406	-0.501	-8.693	71.059
3	5.698	5.698	0.127	-0.707	-8.585	49.153
4	9.292	9.292	-0.795	-0.489	-8.685	80.221
5	9.292	9.292	-0.7	-0.471	-8.685	80.331
6	11.31	11.31	0.021	-0.492	-8.689	95.671
7	10.899	10.899	-0.075	-0.528	-8.614	88.046
8	8.237	8.237	-0.08	-0.914	-8.593	76.053
9	9.275	9.275	-0.025	-0.836	-8.559	85.326
10	11.293	11.293	0.163	-0.833	-8.527	100.666
11	9.83	9.83	-0.358	-0.447	-8.499	78.533
12	9.83	9.83	-0.42	-0.443	-8.507	78.642
13	8.792	8.792	-0.399	-0.468	-8.517	69.37
14	9.292	9.292	-0.485	-0.463	-8.507	74.118
15	9.648	9.648	-1.177	-0.451	-8.487	78.536
16	9.792	9.792	-0.536	-0.457	-8.504	78.642
17	9.292	9.292	-0.439	-0.451	-8.47	80.562

18	8.254	8.254	0.502	-0.574	-8.975	64.748
19	8.792	8.792	-1.451	-0.557	-8.398	73.101
20	9.165	9.165	-0.914	-0.482	-8.563	77.336
21	9.703	9.703	0.616	-0.442	-8.534	82.084
22	10.203	10.203	0.582	-0.466	-8.559	86.608
23	10.703	10.703	0.615	-0.455	-8.482	91.209
24	9.292	9.292	0.263	-0.545	-8.376	77.849
25	10.81	10.81	-1.5	-0.528	-8.394	89.513
26	11.31	11.31	-0.208	-0.569	-8.57	91.393
27	10.075	10.075	-1.977	-0.454	-8.534	86.502
28	11.737	11.737	-0.251	-0.457	-8.497	96.29
29	10.241	10.241	-1.276	-0.437	-8.429	86.832
30	11.241	11.241	-0.625	-0.51	-8.464	95.957
31	10.326	10.326	-1.265	-0.466	-8.494	84.876
32	10.826	10.826	-1.699	-0.465	-8.449	89.477
33	10.826	10.826	-0.688	-0.543	-8.513	86.410

CI1= Connectivity index order 1, DVY= Dipole vector Y, DVZ= Dipole vector Z, DE= Dielectric energy, HE = HOMO energy, MR= Molar refractivity.

Table 3. Correlation matrix between descriptors which are present in model.

	BA	**CI1**	**DVY**	**DVZ**	**DE**	**HE**	**MR**
BA	1						
CI1	-0.211	1					
DVY	-0.285	-0.143	1				
DVZ	-0.415	-0.102	0.300	1			
DE	0.456	0.294	-0.888	-0.277	1		
HE	0.134	0.342	-0.371	-0.344	0.187	1	
MR	-0.191	0.320	0.061	-0.063	0.108	0.339	1

Table 4. Observed, calculated and predicted (LOO) activity of derivatives.

Compd. No.	Obs. Act.[a] (p K_i)	Model-1		Model-2	
		Cal. Act.	Pred. Act.	Cal. Act.	Pred. Act
1	-3.423	-2.924	-2.735	-2.995	-2.865
2	-1.633	-1.682	-1.693	-1.722	-1.739
3	-2.669	-2.794	-2.902	-2.809	-2.929
4	-1.544	-1.733	-1.773	-1.792	-1.835
5	-1.518	-1.633	-1.662	-1.696	-1.732
6	-3.086	-2.916	-2.869	-2.985	-2.963
7	-3.258	-3.330	-3.340	-3.280	-3.282
8	-3.440	-3.337	-3.280	-3.356	-3.3101
9	-2.998	-2.695	-2.544	-2.692	-2.539
10	-2.939	-3.317	-3.628	-3.277	-3.541
11	-2.318	-2.163	-2.133	-2.053	-2.034
12	-2.431	-2.232	-2.205	-2.142	-2.125
13	-2.079	-2.129	-2.140	-2.009	-2.003
14	-2.025	-2.287	-2.322	-2.200	-2.211
15	-1.602	-2.022	-2.062	-1.980	-2.009
16	-2.371	-2.457	-2.464	-2.39	-2.399
17	-2.350	-1.869	-1.785	-1.963	-1.931
18	-3.207	-3.173	-3.140	-3.101	-3.015
19	-2.314	-2.271	-2.261	-2.249	-2.234
20	-2.228	-2.113	-2.103	-2.158	-2.154
21	-1.950	-2.244	-2.321	-2.267	-2.356
22	-2.204	-2.597	-2.691	-2.648	-2.740
23	-3.005	-2.627	-2.517	-2.691	-2.616
24	-2.537	-2.845	-2.925	-2.843	-2.924
25	-3.310	-2.855	-2.749	-2.864	-2.760
26	-3.551	-4.218	-4.599	-4.235	-4.623
27	-0.886	-1.442	-1.623	-1.497	-1.670
28	-3.770	-2.973	-2.857	-2.959	-2.843
29	-1.875	-1.628	-1.592	-1.648	-1.616
30	-2.849	-3.133	-3.230	-3.250	-3.318
31	-2.605	-2.444	-2.427	-2.446	-2.429
32	-1.832	-2.045	-2.086	-2.015	-2.049
33	-3.453	-3.134	-3.068	-3.032	-2.992

[a] All data represent mean values for at least two separate experiments. Obs. Act. - Observed activity, Cal. Act. – Calculated activity, Pred. Act. – Predicted activity by leave one out cross validation.

The predictive ability of model-2 was also confirmed by external r^2CVext. The robustness of the selected model was checked by Y – randomization test. The low r^2 and q^2 values indicate (data not shown) that the good results in our original model are not due to a chance correlation or structural dependency of the training set.The predictive ability of this model was also confirmed by external cross validation (equation 3). The selected model was externally validated by randomly making training set of 27 compounds and test set of 6 compounds (28, 29, 30, 31, 32 and 33) (Table 5). QSAR was performed for training set and a model 3 was developed. This model was used to predict the biological activities of test set of compound.

pKi = 11.289 (± 1.609) DE – 0.391 (± 0.075) CI1 – 0.267 (± 0.116) DVZ + 0.500 (± 0.088) DVY + 3.166 (± 0.793) HE + 34.972 (± 7.850)……. (3)

n = 27, r = 0.884, r^2 = 0.781, r^2_{adj} = 0.729, F = 14.96, SEE = 0.3632, P < 0.001.

Table 5. Predicted activity of test compounds.

Compd.	Observed activity	Predicted activity
28	-3.771	-2.898
29	-1.875	-1.563
30	-2.849	-3.359
31	-2.605	-2.531
32	-1.833	-1.884
33	-3.453	-2.939

The variables used in the selected model have no mutual correlation. This model showed good correlation coefficient (r) of 0.884 between descriptors Dielectric energy, Connectivity index 1, Diploe vector Y, Dipole vector Z and HOMO Energy and A_1 receptor antagonist activity. Squared correlation coefficient (r^2) of 0.781 explains 78.1% variance in biological activity.

The positive contribution of dielectric energy, dipole vector Y and HOMO energy on the biological activity showed that the increase in the values of these parameters lead to better A1-receptor antagonistic properties. The negative coefficient of connectivity index 1 indicated that the increase of CI1 is detrimental to biological activity and the negative coefficient of dipole vector Z is conducive to activity. Based on the developed QSAR model, new A1-receptor antagonist derivatives can be designed with caution.

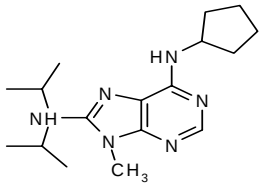
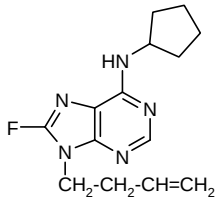
The predicted activities of newly designed series (table 6) of compounds show that they all have predicted activities ranging from **K_i (nm)** = 0.57μM to 6.7 μM whereas the reported series has most active compound with **K_i (nm)** = 7.7 μM.

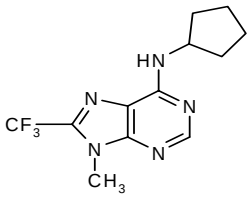
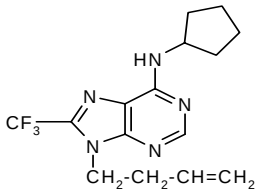
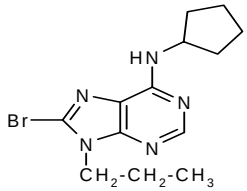
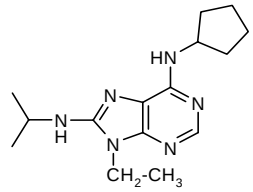
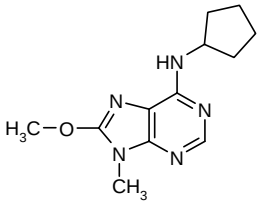
3. Experimental

3.1 General Procedure:

Win CAChe 6.1 (molecular modeling software, a product of Fujitsu private limited, Japan), Molecular modeling pro 6.1.0 (trial version, Cambridge software Corp.), STATISTICA version 6 (StatSoft, Inc., Tulsa, USA).

Table 6. The new designed series of compounds based on model 3

Comp.No.	Compounds Structure	K_i (nm),A_1 receptor
1		6.7
2		4.78

3		1.33
4		0.57
5		4.6
6		0.63
7		4.8

A data set of 33 compounds for A_1-receptor antagonist activity was used for the present QSAR study. The molar concentrations of the compounds required to produce binding at receptor site (in nm) converted to free energy related negative logarithmic values for undertaking the QSAR study.

All 33 compounds' structure were built on workspace of Win CAChe 6.1 (molecular modeling software, a product of Fujitsu private limited, Japan) and energy minimization of the molecules was done using Allinger's MM2 force field followed by semi empirical PM3 method available in MOPAC module until the root mean square gradient value becomes smaller than 0.001 kcal/mol Å. Most stable structure for each compound was generated and used for calculating various physico-chemical descriptors like thermodynamic, steric and electronic values of descriptors.

3.2 Descriptors calculation, QSAR models development and validation

In present study the calculated descriptors were conformational minimum energies (CME), Zero-order connectivity index (CI0), First-order connectivity index (CI1), Second-order connectivity index (CI2), dipole moment (DM), 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), octanol-water partition coefficient(LOGP), molar refractivity(MR), shape index order 1 (SI1), shape index order 2 (SI2), shape index order 3 (SI3), Zero-order valance

connectivity index (VCI0), First-order valance connectivity index (VCI1), Second-order valance connectivity index (VCI2). Some of important descriptor which is present in model is shown in Table 2

All the calculated descriptors (50 descriptors calculated by Win CAChe 6.1 and Molecular modeling pro 6.1.0, the complete descriptors data set of all compounds will be provided on request) were considered as independent variable and biological activity as dependent variable. STATISTICA version 6 (StatSoft, Inc., Tulsa, USA) software was used to generate QSAR models by stepwise multiple linear regression analysis. Statistical measures used were n-number of compounds in regression, r-correlation coefficient, r^2-squared correlation coefficient, F- test (Fischer's value) for statistical significance, SEE- standard error of estimation, q^2- cross validated correlation coefficient and correlation matrix to show correlation among the parameters.

The squared correlation coefficient (or coefficient of multiple determination) r^2 is a relative measure of fit by the regression equation. Correspondingly, it represents the part of the variation in the observed data that is explained by the regression. The correlation coefficient values closer to 1.0 represent the better fit of the regression. The F-test reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High values of the F-test indicate that the model is statistically significant. Standard deviation is measured by the error mean square, which expresses the variation of the residuals or the variation about the regression line. Thus standard deviation is an absolute measure of quality of fit and should have a low value for the regression to be significant.

The predictive ability of the generated correlations was evaluated by cross validation method employing a 'leave-one-out' scheme. Validation parameters considered were cross validated r^2 or q^2, standard deviation based on predicted residual sum of squares (S_{PRESS}) and standard error of prediction (SDEP). The predictive ability of the selected model was also confirmed by external r^2CVext.

$$r^2CVext = 1 - \frac{\sum_{i=1}^{test}(y_{exp} - y_{pred})^2}{\sum_{i=1}^{test}(y_{exp} - \bar{y}_{tr})^2}$$

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

Acknowledgement

One of the authors Mr. Abhishek Jain is grateful to U.G.C for providing fellowship for this work.

References

[1] C. A Salvatore, M. A Jacobson, H. E Taylor, J. Linden, R. Johnson, Proc. Natl. Acad. Sci. U.S.A. **90**, 10365 (1993).
[2] J. Linden, Trends Pharmacol. Sci. **15**, 298 (1994).
[3] J. P. Hannon, H. J. Pfannkuche, J. R. Fozard, Br. J. Pharmacol **115**, 945 (1995).
[4] R. A. Olsson, J. D. Pearson, Pharmacol. Rev. **3**, 761 (1990).
[5] N. F. Rossi, P. C. Churchill, K. A. Jacoson, A. E. Leahy, Pharmacol. Exper. Ther. **240**, 911 (1987).
[6] C. Londson, D. M. Cooper, Woff, J.Prac.Natl.Acad.sci. USA **77**, 2251 (1980).
[7] B. D. Hillaire, G. Bertrand, R. Gross, Eur.J.Pharmacol. **136**, 109 (1987).

[8] L. Betti, G. Biagi, G. Giannaccini, I. Giorgi, O. Livi, A. Lucacchini, C. Manera, V. Scartoni, J. Med. Chem. **41**, 668 (1998).
[9] E. Camaioni, S. Costanzi, S. Vittori, R. Volpini, K. N. Klotz, G. Cristalli, Bioorg. Med. Chem. **6**, 523 (1998).
[10] V. Collota, D. Catarzi, F. Varano, L. Cecchi, G. Filacchioni, C. Martini, L. Trincavelli, A. Lucacchini, J. Med. Chem. **43**, 3118 (2000).
[11] P. L. Ferrarini, C. Mori, C. Manera, A. Martinelli, F. Mori, G. Saccomanni, P. L. Barili, L. Betti, G. Giannaccini, L. Trincavelli, A. Lucacchini, J. Med. Chem. **43**, 2814 (2000).
[12] T. Katsushima, L. Nieves, N. Wells, J. Med. Chem. **33**, 1906 (1990).
[13] S. Kuroda, A. Akahane, H. Itani, S. Nishimura, K. Durkin, T. Kinoshita, Y. Tenda, K. Sakane, Bioorg. Med. Chem. Lett. **9**, 1979 (1999).
[14] C. E. Muller, Expert Opin. Ther. Pat. **5**, 419 (1997).
[15] A. F. Rianne, D. Ligt, A. M. Pieter, V. D. Klein, K. Jacobien, F. Drabbe Kunzel, A. Lorenzen, F. A. Maate, S. Fujikawa, R. V. Westhoven, T. V. D. Hoven, J. Brussee, P. Ijzerman, Bioorg. Med. Chem.**12**, 139 (2004).
[16] P. L. Martin, R. J. Wysocki, R. J. Barrett, J. M. May, J. Linden, J. Pharmacol. Exp. Ther. **276**, 490 (1996).
[17] R. Volpini, S. Costanzi, C. Lambertucci, S. Vittori, A. Lorenzen, K. N. Klotz, G. Cristalli, Bioorg. Med. Chem. Lett. **11**, 1931 (2001).
[18] V. Ravichandran, R. K. Agrawal, Bioorg. Med. Chem. Lett. **17**, 2197 (2007).
[19] V. Ravichandran, V. K. Mourya, R. K. Agrawal, Arkivoc **XIV**, 204 (2007).
[20] K. K. Sahu, V. Ravichandran, V. K. Mourya, R. K. Agrawal, Med. Chem. Res. **15**, 418 (2007).
[21] K. K. Sahu, V. Ravichandran, P. K. Jain, S. Sharma, V. K. Mourya, R. K. Agrawal, Acta Chim. Slov. (Inpress) (2007).
[22] S. Sharma, V. Ravichandran, P. K. Jain, V. K. Mourya, R. K. Agrawal, J. Enzym. Inhib. Med. Chem. (Inpress) (2007).
[23] H. K. Jain, R. K. Agrawal, Inter. Elect. J. Mol. Design **5**, 224 (2006).
[24] H. K. Jain, R. K. Agrawal, Inter. Elect. J. Mol. Design (Inpress) (2007).
[25] H. K. Jain, V. K. Mourya, R. K. Agrawal, Bioorg. Med. Chem. Lett. **16**, 5280 (2006).
[26] S. J. Dighade, H. K. Jain, R. K. Agrawal, Indian J. Pharm. Sci. **65**, 586 (2003).