Journal of Molecular Graphics and Modelling 28 (2010) 883–890

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

In silico classification of adenosine receptor antagonists using Laplacian-modified naïve Bayesian, support vector machine, and recursive partitioning

Jin Hee Lee, Sunkyung Lee, Sun Choi *

College of Pharmacy, Division of Life & Pharmaceutical Sciences, and National Core Research Center for Cell Signaling & Drug Discovery, Ewha Womans University, Seoul 120-750, Republic of Korea

ARTICLE INFO

Article history:
Received 7 November 2009
Received in revised form 18 March 2010
Accepted 21 March 2010
Available online 27 March 2010

Keywords:
Adenosine receptor
Antagonist
Classification
Laplacian-modified naïve Bayesian
Recursive partitioning
Support vector machine

ABSTRACT

Adenosine receptors (ARs) belong to the G-protein-coupled receptor (GPCR) superfamily and consist of four subtypes referred to as A_1, A_{2A}, A_{2B}, and A_3. It is important to develop potent and selective modulators of ARs for therapeutic applications. In order to develop reliable in silico models that can effectively classify antagonists of each AR, we carried out three machine learning methods: Laplacian-modified naïve Bayesian, recursive partitioning, and support vector machine. The results for each classification model showed values high in accuracy, sensitivity, specificity, area under the receiver operating characteristic curve and Matthews correlation coefficient. By highlighting representative antagonists, the models demonstrated their power and usefulness, and these models could be utilized to predict potential AR antagonists in drug discovery.

1. Introduction

Adenosine receptors (ARs) belong to the superfamily of G-protein-coupled receptors (GPCRs), and consist of four subtypes: A_1, A_{2A}, A_{2B}, and A_3. A_1 and A_3 ARs are coupled to G_i proteins, which inhibit adenylyl cyclase, leading to a decrease in intracellular levels of cAMP [1]. Whereas A_{2A} and A_{2B} ARs are coupled to G_s proteins and their simulation increases adenylyl cyclase activity [1]. Having over 60% of sequence similarity among subtypes, ARs mediate many biological effects [2]. These interactions between each AR and its ligands offer very broad therapeutic potential, thus a number of selective AR agonists and antagonists have been developed.

Most AR agonists are modified within the adenosine part of the molecules, while AR antagonists have greater structure diversity with a variety of scaffolds such as xanthine, adenosine, or polyheterocycle. In addition, the search for selective antagonists held greater appeal than selective agonists, not only for their potential therapeutic applications but also considering the fact that antagonists are preferred molecular probes for pharmacological characterization of receptors [2]. Antagonists of A_1 AR are of potential use for cardiac anti-arrhythmic agents, cognitive disorders and anxiety disorders [3,4]. A_{2A} AR antagonists are currently considered promising agents in dopamine replacement therapy for Parkinson's disease [5], and A_{2B} AR antagonists have indicated a variety of potential therapeutic applications for cancer, gastrointestinal disorders, and neurological disorders [6]. A_3 AR antagonists are being investigated as anti-asthma, anti-glaucoma, and anti-inflammatory agents [7,8]. Since the development of potent and selective antagonists of each AR is challenging and of great importance, and effective classification models will have certain potential for practical use, this study focused on the in silico classification of these antagonists.

In developing in silico models, we classified antagonists for each AR using machine learning methods. Machine learning methods are considered as one of the best ways to efficiently classify a large number of molecules due to their robustness, speed and reproducibility, and these methods are becoming more and more frequently used in the early stage of drug discovery. In this study, we utilized three machine learning classification methods: Laplacian-modified naïve Bayesian, recursive partitioning (RP), and support vector machine (SVM). Laplacian-modified naïve Bayesian can handle large amounts of data and requires less parameter configuration and fewer actives to build a model. RP provides a graphical tree as well as statistics, so it is easy to interpret which descriptors are more important than others. SVM can build an effective classification model from any complex data set because of its scalability and automated model tuning. Using these three methods, classification models for the AR antagonists were constructed and their performances were evaluated through analysis.

* Corresponding author. Tel.: +82 2 3277 4503; fax: +82 2 3277 2851.
E-mail address: sunchoi@ewha.ac.kr (S. Choi).

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.03.008

Table 1
Summary of the data sets used in this study.

Subtype	Set	No. of actives	No. of inactives	Total no.
A_1	Training/test	260/27	791/70	1051/97
A_{2A}	Training/test	462/43	589/54	1051/97
A_{2B}	Training/test	154/12	897/85	1051/97
A_3	Training/test	175/15	876/82	1051/97

2. Methods

2.1. Computational methodologies

All computation calculations were undertaken on a Linux (Cent OS release 4.6) workstation. Preparation of data sets, calculation of descriptors, building and evaluation of the classification models were carried out using SciTegic Pipeline Pilot v.7.0.1 [9].

2.2. Data sets and descriptors

All the molecular structures used in this study were collected from Prous Science Integrity [10] and taken from literature [11,12]. Counter ions and solvent molecules were removed, and hydrogens were added before calculating descriptors. To build classification models, we defined actives and inactives in data sets, so that if a molecule is defined by the literature as active for one subtype, it is defined as inactive for the other subtypes. Duplicate compounds in more than one subtype were removed from the data sets, and all the data sets were divided randomly into training and test sets in the ratio of 9:1 using random percent filter component [13] in Pipeline Pilot. Table 1 shows the number of compounds selected for the training and test sets for each AR subtype. The training sets were used to construct models, and the test sets were used to evaluate the classification accuracy of the models.

The descriptors used in this study are: AlogP, molecular weight, number of hydrogen bond donors, number of hydrogen bond acceptors, number of rotatable bonds, and the functional-class fingerprints (FCFP_6). FCFPs are 2D and extended-connectivity fingerprints based on the functional roles of an atom and its neighbors (i.e., hydrogen bond donor/acceptor, positively/negatively ionizable, aromatic, and halogen). The number displayed in the fingerprint name is the maximum diameter of the features generated, thus FCFP_6 generates features around each atom up to six bonds [14]. FCFPs are widely used in machine learning and clustering along with other physicochemical descriptors, and they have shown good performance in many literature [15,16].

2.3. Laplacian-modified naïve Bayesian

Based on statistical probability estimation, Laplacian-modified naïve Bayesian assumes that the descriptors in the training set are independent and of equal importance [17]. This method produces a high-dimensional representation of the molecule by using extended-connectivity fingerprints. In this study, the model building and analysis were performed by Pipeline Pilot v.7.0.1 with the default parameters. This program automatically carried out an extremely fast, leave-one-out cross-validation as the model was built, providing a variety of statistics to allow assessment of the model's quality.

2.4. Recursive partitioning

As a decision tree method, recursive partitioning (RP) divides data by using a hierarchical set of yes/no questions per node to

Table 2
Summary of the statistics for the AR antagonist models by three classification methods: Laplacian-modified naïve Bayesian, support vector machine (SVM), and recursive partitioning (RP).

Subtype	Method	Actives		Inactives		Accuracy (%)	Sensitivity (%)	Specificity (%)
		TP	FN	FP	TN			
(a) Training sets								
A_1	Bayesian	246	14	43	748	94.58	94.62	94.56
	SVM	234	26	11	780	96.48	90.00	98.61
	RP	246	14	59	732	93.05	94.62	92.54
A_{2A}	Bayesian	451	11	19	570	97.15	97.62	96.77
	SVM	441	21	13	576	96.76	95.45	97.79
	RP	446	16	41	548	94.58	96.54	93.04
A_{2B}	Bayesian	147	7	42	855	95.34	95.45	95.32
	SVM	129	25	1	896	97.53	83.77	99.89
	RP	150	4	83	814	91.72	97.40	90.75
A_3	Bayesian	160	15	41	835	94.67	91.43	95.32
	SVM	137	38	4	872	96.00	78.29	99.54
	RP	162	13	108	768	88.49	92.57	87.67
(b) Test sets								
A_1	Bayesian	22	5	2	68	92.78	81.48	97.14
	SVM	24	3	0	70	96.91	88.89	100.00
	RP	19	8	5	65	86.60	70.37	92.86
A_{2A}	Bayesian	42	1	3	51	95.88	97.67	94.44
	SVM	41	2	1	53	96.91	95.35	98.15
	RP	37	6	6	48	87.63	86.05	88.89
A_{2B}	Bayesian	11	1	4	81	94.85	91.67	95.29
	SVM	10	2	1	84	96.91	83.33	98.82
	RP	12	0	8	77	91.75	100.00	90.59
A_3	Bayesian	14	1	2	80	96.91	93.33	97.56
	SVM	14	1	1	81	97.94	93.33	98.78
	RP	15	0	13	69	86.60	100.00	84.15

TP: true positive; FN: false negative; FP: false positive; TN: true negative.

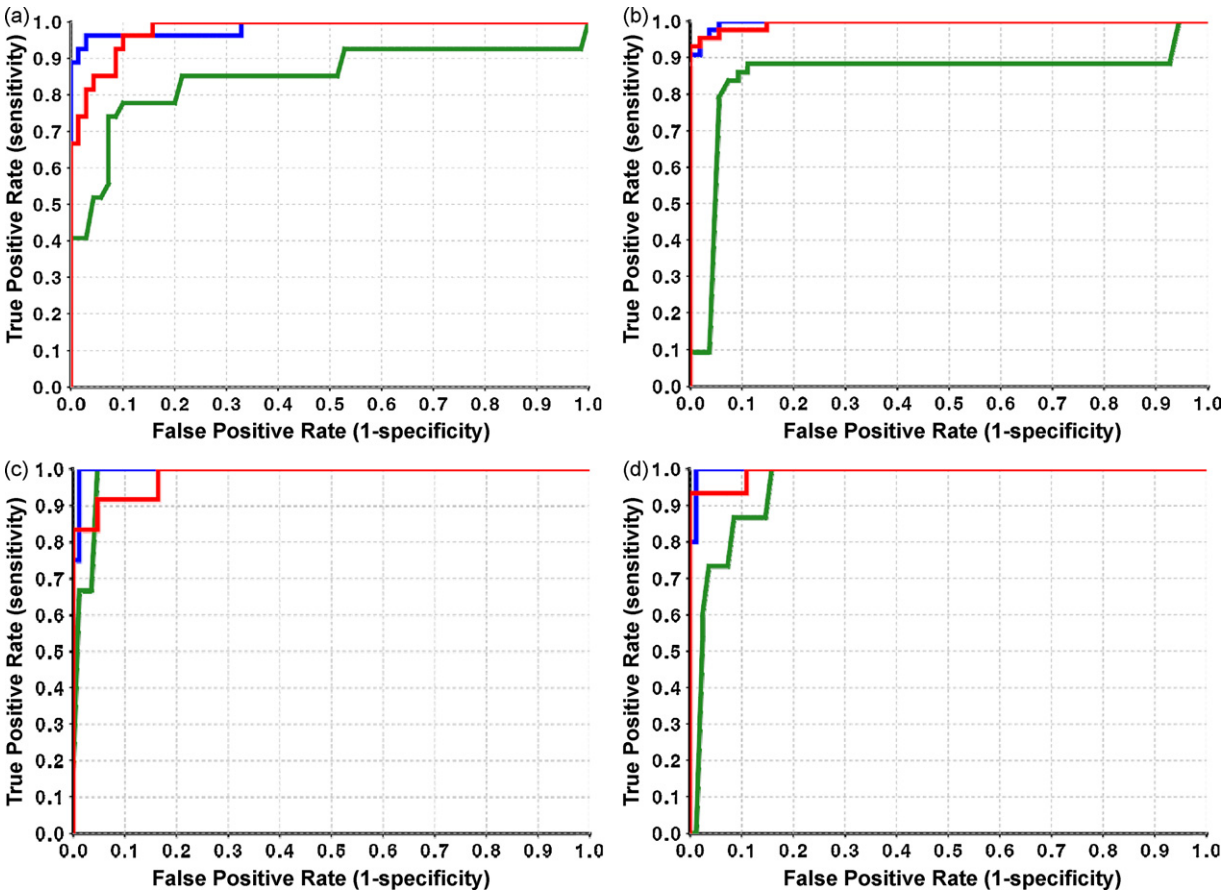


Fig. 1. Comparison of the receiver operating characteristic (ROC) plots of the AR antagonist test set models generated by three machine learning methods: Laplacian-modified naïve Bayesian (red), SVM (blue), RP (green). (a) A_1 AR antagonist models, (b) A_{2A} AR antagonist models, (c) A_{2B} AR antagonist models, and (d) A_3 AR antagonist models.

split a data set into smaller subsets. The splitting process continues until no more significant nodes are obtained or when a minimum number of samples per node is reached. RP was performed using the CART algorithm implemented in Pipeline Pilot v.7.0.1 [9]. In our models, 10-fold cross-validation was used, the classes were weighted equally, and splits were scored using Gini index. To avoid excessive partitioning, the minimum samples per node and the maximum tree depth were limited to 10.

2.5. Support vector machine

Support vector machine (SVM) is a supervised machine learning method, developed by Vapnik [18]. This method has become more frequently used in drug discovery due to its good performance and fast speed. SVM constructs a hyperplane in a high-dimensional molecular descriptor space not only to separate the actives from the inactives but also to minimize errors. In model building, we used 10-fold cross-validation to avoid overfitting of the training set, and SVM calculations of the data sets were performed using e1071 library in R package [19].

2.6. Classification model validation and evaluation

The performance of each model was evaluated by calculating the accuracy, sensitivity, and specificity, along with other statistics such as receiver operating characteristic (ROC) curve and Matthews correlation coefficient (MCC). Accuracy of a model is defined to correctly classify the actives from the inactives. Sensitivity measures the proportion of actual positives which are correctly identified as such. Specificity measures the proportion of negatives which are correctly identified. In other words, sensitivity is the ability of the model to avoid false negatives, and specificity is the ability to avoid false positives. The accuracy, sensitivity, and specificity were calculated in the following way:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}},$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

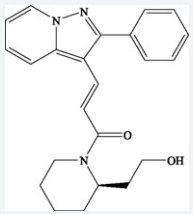
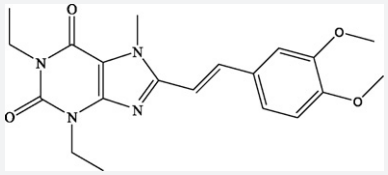
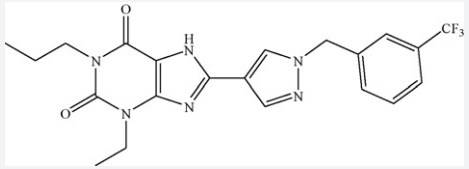
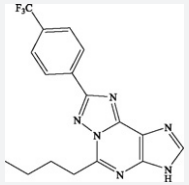
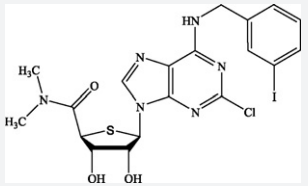
$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

Table 3
AUC and MCC values for the AR antagonist models by three classification methods: Laplacian-modified naïve Bayesian, SVM, and RP.

Subtype	Method	Training set		Test set	
		AUC	MCC	AUC	MCC
A_1	Bayesian	0.988	0.86	0.979	0.82
	SVM	0.984	0.90	0.986	0.92
	RP	0.957	0.83	0.883	0.66
A_{2A}	Bayesian	0.995	0.94	0.995	0.92
	SVM	0.994	0.93	0.997	0.94
	RP	0.956	0.89	0.910	0.75
A_{2B}	Bayesian	0.979	0.84	0.982	0.79
	SVM	0.986	0.90	0.997	0.85
	RP	0.938	0.75	0.987	0.74
A_3	Bayesian	0.978	0.82	0.993	0.89
	SVM	0.985	0.85	0.998	0.92
	RP	0.893	0.68	0.961	0.67

AUC: the area under of receiver operating characteristic (ROC) curve; MCC: Matthews correlation coefficient.

Table 4
Affinity of AR antagonists at the A_1, A_{2A}, A_{2B}, and A_3 ARs.

Subtype	Name	Structure	Phase	K_i[a] values for ARs (nM)				Reference
				A_1	A_{2A}	A_{2B}	A_3	
A_1	FK-453		Phase II	18	1300	ND	2800	[22]
A_{2A}	KW-6002		Phase III	2830	36	1800	>3000	[22,23]
A_{2B}	CVT-6883		Phase I	>10,000	>10,000	8.3	>10,000	[22]
A_3	OT-7999		Preclinical	>10,000[b]	>10,000[b]	>10,000[b]	0.95	[22]
A_3	LJ-1256		Biological testing	6220	>10,000	>10,000	15.5	[24]

[a] Binding experiments at recombinant human A_1, A_{2A}, A_{2B}, and A_3 ARs.
[b] IC_{50} values.

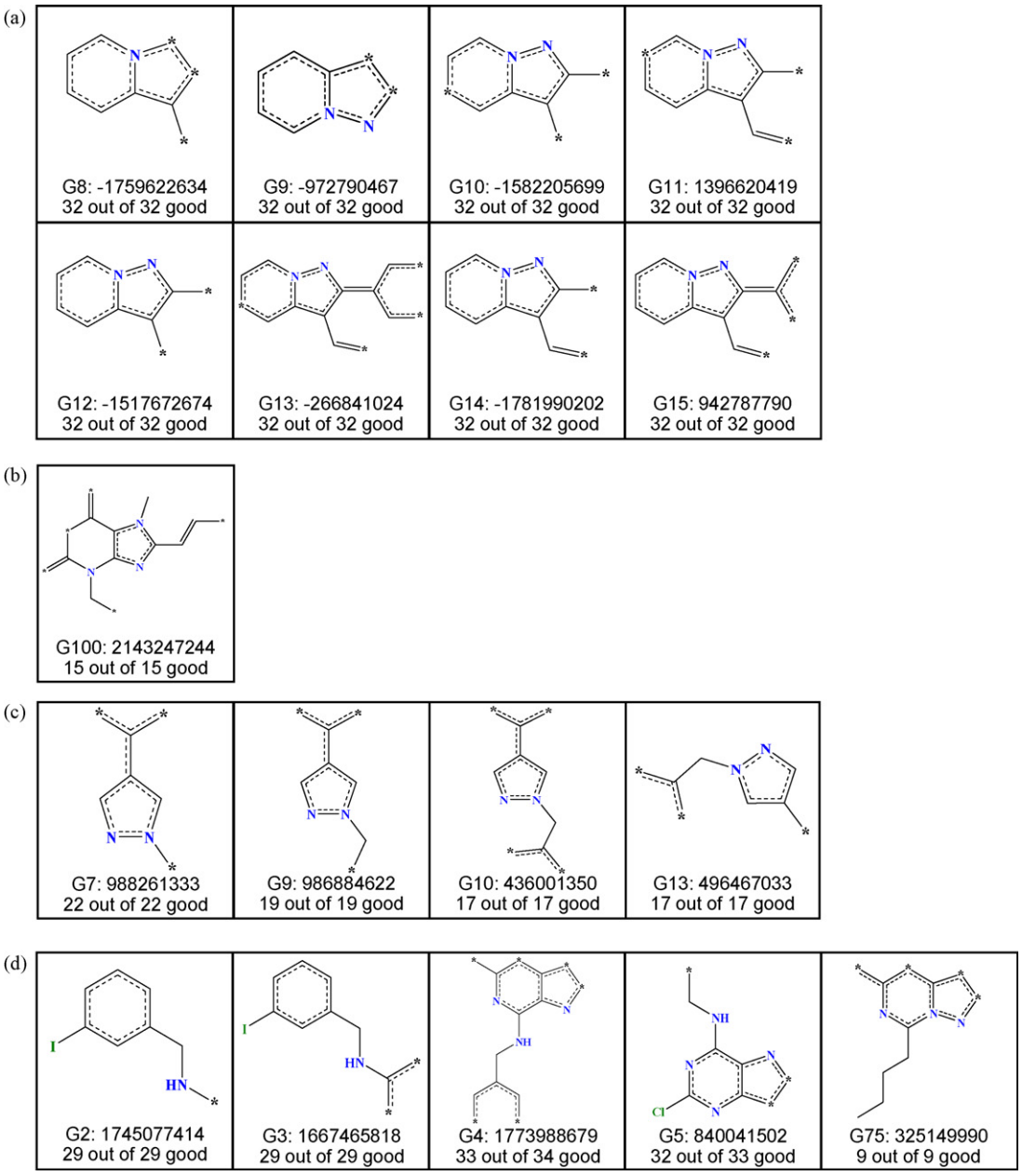


Fig. 2. The good features identified by Laplacian-modified naïve Bayesian method based on FCFP_6 from the models of (a) A_1 AR antagonists, (b) A_{2A} AR antagonists, (c) A_{2B} AR antagonists, and (d) A_3 AR antagonists. The identifiers and bin values of the good features are shown along with the frequencies of their occurrence in actives.

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives (i.e., inactives predicted as actives), and FN is the number of false negatives (i.e., actives predicted as inactives).

Receiver operating characteristic (ROC) curve and Matthews correlation coefficient (MCC) were applied as additional criteria to evaluate the performance of each model. The ROC curve is a plot of a fraction of true positives vs. that of false positives for different cutoff values of a score predicted by a binary classification model [20]. The area under the ROC curve (AUC) is a measure of the accuracy of a model and provides a diagnostic tool to select optimal classification models. An area of 1 represents a perfect model and that of 0.5 represents that the model has no predictive ability. As the other criterion to evaluate the quality of the models, Matthews correlation coefficient (MCC) measures how the normalized variables tend to have the same sign and magnitude [21]. It is expressed by a value between −1 and +1, where a value of +1 represents a perfect prediction, 0 an average random prediction and −1 an inverse prediction. MCC was calculated from the confusion matrix in the following way:

$$\mathrm{MCC} = \frac{(\mathrm{TP} \times \mathrm{TN}) - (\mathrm{FP} \times \mathrm{FN})}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}}$$

3. Results and discussion

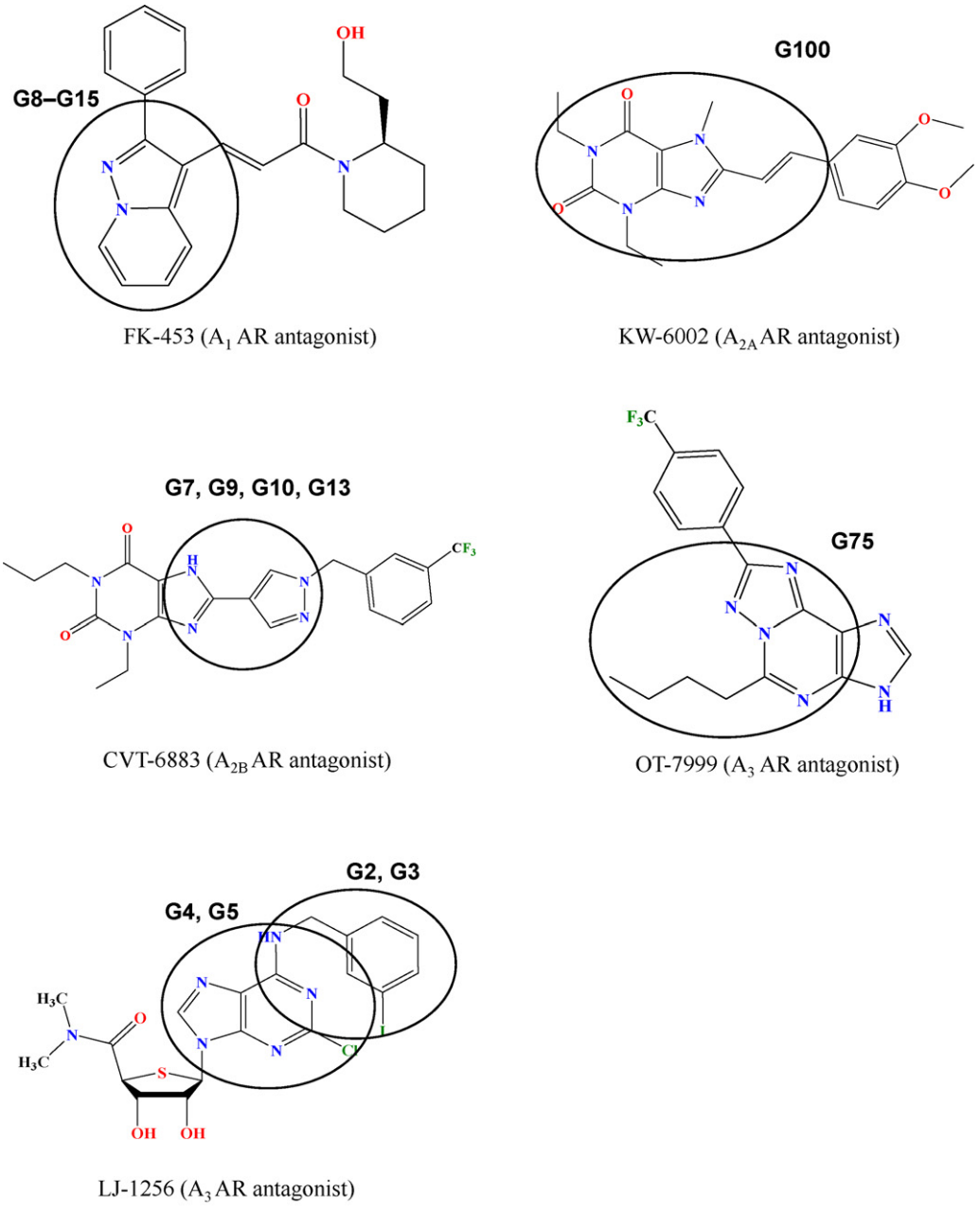
In order to obtain reliable classification models for AR antagonists, three machine learning methods were applied: Laplacian-modified naïve Bayesian, recursive partitioning (RP), and support vector machine (SVM). The performance of each model was then evaluated by accuracy, sensitivity, specificity, AUC and MCC values as listed in Tables 2 and 3.

In A_1 AR antagonist models, all three classification methods performed well. Accuracy, sensitivity, and specificity values for all three models were higher than 90% in the training set. For the three models in the test set, although their sensitivity values were generally lower, that of SVM was relatively high, thus SVM would be more useful than other methods to avoid false negatives. All the classification models for A_1 AR antagonists yielded satisfactory results in AUC and MCC values (Table 3). In addition, their ROC plots of Laplacian-modified naïve Bayesian and SVM were almost perfect in prediction, and that of RP also performed well (Fig. 1(a)).

In A_{2A} AR antagonist models, the accuracy, sensitivity, and specificity for classification under Laplacian-modified naïve Bayesian, SVM, and RP methods showed excellent results with near or above 95% both in the training and test sets. Though the performance of RP in the test set was not as good as in the other models, it was still acceptable. Also, AUC and MCC values of the three classification models were near or above 0.90 in most cases. Specifically, AUC values of Laplacian-modified naïve Bayesian and SVM were above 0.95, and ROC plots from both methods demonstrated strong results (Fig. 1(b)).

In A_{2B} AR antagonist models, Laplacian-modified naïve Bayesian and RP methods produced successful results in which the three evaluated values were higher than 90% both in the training and test sets. SVM method demonstrated outstanding results in accuracy and specificity, but in terms of sensitivity, it showed a slightly lower value compare to other methods. As shown in Fig. 1(c), ROC plots of A_{2B} AR antagonists demonstrated that the three classification methods had excellent results with all their AUC values being above 0.95.

In A_3 AR antagonist models, Laplacian-modified naïve Bayesian yielded strong results both in the training and test sets. The accuracy and specificity of SVM showed a performance similar to Laplacian-modified naïve Bayesian, but the sensitivity of SVM was relatively lower in the training set. In the RP model, the accuracy



FK-453 (A_1 AR antagonist)

KW-6002 (A_{2A} AR antagonist)

CVT-6883 (A_{2B} AR antagonist)

OT-7999 (A_3 AR antagonist)

LJ-1256 (A_3 AR antagonist)

Fig. 3. Representative AR antagonists and their good features identified by this study. The good features are highlighted in circles and their identifiers are marked.

and specificity values were slightly below 90%. All three classification methods generated AUC values of 0.9 or above and MCC values above 0.65. Laplacian-modified naïve Bayesian and SVM showed almost perfect ROC plots, and RP displayed very good ROC plots as well (Fig. 1(d)).

Laplacian-modified naïve Bayesian method consistently showed superior performance in all of the AR antagonist classification models based on accuracy, sensitivity, specificity, AUC and MCC values. SVM produced similar results as Laplacian-modified naïve Bayesian in most of the models, but its sensitivity was not as good for A_{2B} and A_3 AR antagonist classification models. Although the performance of RP was not better than the others, it had acceptable results for all AR subtypes.

To be used as examples, several representative compounds were identified (highlighted in Table 4), which are selective and potent AR antagonists mostly in clinical trials. Among these compounds, FK-453 is an A_1 AR antagonist and currently in phase II clinical trial for the treatment of heart failure [22]. KW-6002 is an A_{2A} AR antagonist in the last stage of clinical trials for the treatment of Parkinson's disease [22,23]. CVT-6883, an A_{2B} AR antagonist, is in phase I for the management of chronic pulmonary diseases [22]. As of now, there is no A_3 AR antagonist in clinical trial, but OT-7999 and LJ-1256 are very promising and notable compounds. As a triazolopurine derivative, OT-7999 displayed a high affinity for A_3 AR and had a 10,000-fold selectivity relative to other AR subtypes [22]. As a nucleoside analogue, LJ-1256 could overcome species-dependence and increase aqueous compatibility [24].

In Fig. 2, the good features, which appear mostly in actives of the training set for each AR, are shown. These good features were identified by Laplacian-modified naïve Bayesian method based on FCFP_6. The representative active antagonists tested in this study and their good features are shown in Fig. 3. As an active A_1 AR antagonist, FK-453 was correctly classified by Laplacian-modified naïve Bayesian. Under analysis the good features, G8–G15 (Figs. 2(a) and 3), were identified and shown to contain pyrazolopyridine parts found in 32 actives in the training set. In the RP model, FK-453 was classified as an active A_1 AR antagonist following the nodes of 0, 2, 6, 9, and 16 in the RP tree (Fig. 4 and Supplementary Information S1). In the case of SVM, this method only classified each compound as active or inactive, and FK-453 was correctly classified as an active.

As a known active A_{2A} AR antagonist, KW-6002, was correctly classified by SVM, Laplacian-modified naïve Bayesian, and RP. It was reported that the introduction of the (*E*)-styryl group in the 8 position of xanthine contributes to the selectivity of A_{2A} AR antagonists [25–28]. In the Laplacian-modified naïve Bayesian model, most part of the styrylxanthine moiety was detected as the good feature G100 (Figs. 2(b) and 3), which was shown in 15 actives in the training set and also found in KW-6002. In the RP model, that moiety was the criteria to split compounds at the node 12, and KW-6002 was correctly classified following the nodes of 0, 2, 6, 12, and 15 (Supplementary Information S2). Also, as an active A_{2B} AR antagonist, CVT-6883 was correctly classified by SVM as well as Laplacian-modified naïve Bayesian. Under analysis the good features, G7, G9, G10, and G13 (Figs. 2(c) and 3), were identified in CVT-6883. The G7 feature with the pyrazole moiety was seen in 22 actives in the training set. In the RP model, it was correctly predicted following the nodes of 0, 1, 4, and 7 (Supplementary Information S3).

As representative examples of A_3 AR antagonists, OT-7999 and LJ-1256 were correctly classified as active by SVM and Laplacian-modified naïve Bayesian. Under analysis the good features, G75 and G2–G5 features (Figs. 2(d) and 3), were identified in OT-7999 and LJ-1256, respectively by the Laplacian-modified naïve Bayesian. Furthermore, they were correctly spilt by the RP model, so that OT-7999 followed the nodes of 0 and 1, and LJ-1256 followed the nodes of 0, 2, 6, 11, and 18 (Supplementary Information S4). Being different from other subtypes, antagonists of A_3 AR include nucleoside analogues. We assumed that adenine may play an important role in the classification of A_3 AR antagonists, and it was confirmed

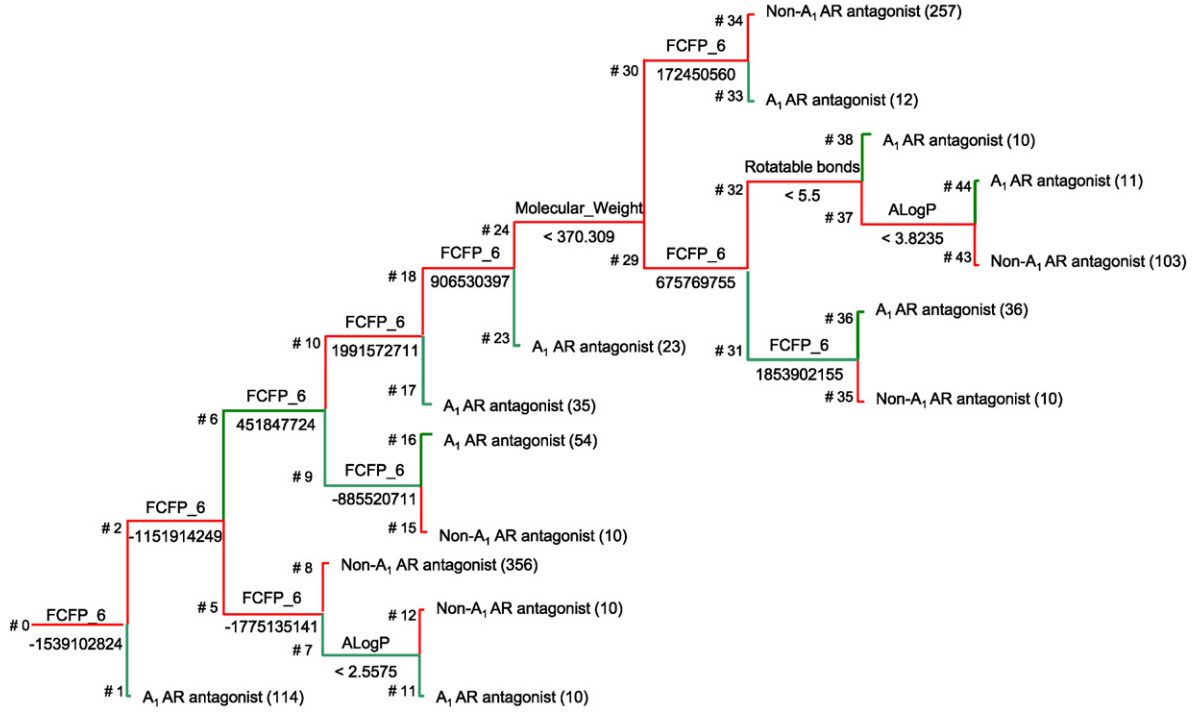


Fig. 4. The recursive partitioning (RP) tree for A_1 AR antagonist model. Molecules were split by proper descriptors at each node, and the node numbers are marked after # sign. The green line means antagonists, and the red line means non-antagonists. The number of molecules classified from the training set is shown in parenthesis. The bin value, which is the fingerprint bit that defines the bin, is marked under FCFP_6.

by analysis using FCFP_6 fingerprints in Laplacian-modified naïve Bayesian and the RP tree.

4. Conclusions

In order to build reliable in silico classification models of AR antagonists, we adopted six descriptors and performed three machine learning methods: Laplacian-modified naïve Bayesian, RP, and SVM. The resulting classification models of the antagonists for each AR had excellent results that yielded high accuracy, sensitivity, specificity, area under the receiver operating characteristic (ROC) curve and Matthews correlation coefficient (MCC) values. The application of representative antagonists to these classification models for each AR demonstrated the power and utility of these models. These models could be further utilized in the prediction of potential AR antagonists in drug discovery.

Acknowledgments

This work was supported by the second stage of the Brain Korea 21 Project and the grant no. R15-2006-020 from the National Core Research Center (NCRC) program of the Ministry of Education, Science and Technology (MEST) and the Korea Science and Engineering Foundation (KOSEF) through the Center for Cell Signaling & Drug Discovery Research at Ewha Womans University.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2010.03.008.

References

[1] K.A. Jacobson, Z.G. Gao, Adenosine receptors as therapeutic targets, Nat. Rev. Drug Discov. 5 (2006) 247–264.
[2] S. Moro, Z.G. Gao, K.A. Jacobson, G. Spalluto, Progress in the pursuit of therapeutic adenosine receptor antagonists, Med. Res. Rev. 26 (2006) 131–159.
[3] A.K. Dhalla, J.C. Shryock, R. Shreeniwas, L. Belardinelli, Pharmacology and therapeutic applications of A1 adenosine receptor ligands, Curr. Top. Med. Chem. 3 (2003) 369–385.
[4] W. Soudijn, I. van Wijngaarden, A.P. Ijzerman, Medicinal chemistry of adenosine A1 receptor ligands, Curr. Top. Med. Chem. 3 (2003) 355–367.
[5] K. Fuxe, S. Ferre, M. Canals, M. Torvinen, A. Terasmaa, D. Marcellino, S.R. Goldberg, W. Staines, K.X. Jacobsen, C. Lluis, A.S. Woods, L.F. Agnati, R. Franco, Adenosine A2A and dopamine D-2 heteromeric receptor complexes and their function, J. Mol. Neurosci. 26 (2005) 209–219.
[6] M.L. Peyot, A.P. Gadeau, F. Dandre, I. Belloc, F. Dupuch, C. Desgranges, Extracellular adenosine induces apoptosis of human arterial smooth muscle cells via A2b-purinoceptor, Circ. Res. 86 (2000) 76–85.
[7] B.B. Fredholm, I.J. AP, K.A. Jacobson, K.N. Klotz, J. Linden, International union of pharmacology. XXV. Nomenclature and classification of adenosine receptors, Pharmacol. Rev. 53 (2001) 527–552.
[8] M.Y. Avila, R.A. Stone, M.M. Civan, Knockout of A3 adenosine receptors reduces mouse intraocular pressure, Invest. Ophthalmol. Vis. Sci. 43 (2002) 3021–3026.
[9] Pipeline Pilot, Version 7.0.1, Accelrys Software Inc., San Diego, CA, 2008.
[10] Prous Science Integrity, http://integrity.prous.com.
[11] S. Pal, W.J. Choi, S.A. Choe, C.L. Heller, Z.G. Gao, M.S. Chinn, K.A. Jacobson, X.Y. Hou, S.K. Lee, H.O. Kim, L.S. Jeong, Structure–activity relationships of truncated adenosine derivatives as highly potent and selective human A3 adenosine receptor antagonists, Bioorg. Med. Chem. 17 (2009) 3733–3738.
[12] L.S. Jeong, S. Pal, S.A. Choe, W.J. Choi, K.A. Jacobson, Z.G. Gao, A.M. Klutz, X.Y. Hou, H.O. Kim, H.W. Lee, S.K. Lee, D.K. Tosh, H.R. Moon, Structure–activity relationships of truncated D- and L-4′-thioadenosine derivatives as species-independent A3 adenosine receptor antagonists, J. Med. Chem. 51 (2008) 6609–6613.
[13] The random percent filter component in Pipeline Pilot applies the probability to data sets, so the total percent of data set may slightly differ from the actual percent value.
[14] Accelrys Software Inc., Chemistry Collection: Basic Chemistry User Guide, SciTegic Pipeline Pilot Release 7.0, Accelrys Software Inc., San Diego, 2008.
[15] H.M. Sun, An accurate and interpretable Bayesian classification model for prediction of hERG liability, Chem. Med. Chem. 1 (2006) 315–322.
[16] D. Rogers, R.D. Brown, M. Hahn, Using extended-connectivity fingerprints with Laplacian-modified Bayesian analysis in high-throughput screening follow-up, J. Biomol. Screen. 10 (2005) 682–686.
[17] X.Y. Xia, E.G. Maliski, P. Gallant, D. Rogers, Classification of kinase inhibitors using a Bayesian model, J. Med. Chem. 47 (2004) 4463–4470.
[18] V. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 1995.
[19] R-package, http://www.R-project.org.
[20] A. Nicholls, What do we know and when do we know it? J. Comp. Aid. Mol. Des. 22 (2008) 239–255.
[21] P. Baldi, S. Brunak, Y. Chauvin, C.A.F. Andersen, H. Nielsen, Assessing the accuracy of prediction algorithms for classification: an overview, Bioinformatics 16 (2000) 412–424.
[22] P.G. Baraldi, M.A. Tabrizi, S. Gessi, P.A. Borea, Adenosine receptor antagonists: translating medicinal chemistry and pharmacology into clinical utility, Chem. Rev. 108 (2008) 238–263.
[23] J. Shimada, N. Koike, H. Nonaka, S. Shiozaki, K. Yanagawa, T. Kanda, H. Kobayashi, M. Ichimura, J. Nakamura, H. Kase, F. Suzuki, Adenosine A2A antagonists with potent anti-cataleptic activity, Bioorg. Med. Chem. Lett. 7 (1997) 2349–2352.
[24] Z.G. Gao, B.V. Joshi, A.M. Klutz, S.K. Kim, H.W. Lee, H.O. Kim, L.S. Jeong, K.A. Jacobson, Conversion of A3 adenosine receptor agonists into selective antagonists by modification of the 5′-ribofuran-uronamide moiety, Bioorg. Med. Chem. Lett. 16 (2006) 596–601.
[25] J. Shimada, F. Suzuki, H. Nonaka, A. Ishii, S. Ichikawa, (E)-1,3-Dialkyl-7-methyl-8-(3,4,5-trimethoxystyryl)xanthines: potent and selective A2 antagonists, J. Med. Chem. 35 (1992) 2342–2345.
[26] C.E. Müller, B. Stein, Adenosine receptor antagonists: structures and potential therapeutic applications, Curr. Pharm. Des. 2 (1996) 501–530.
[27] P.G. Baraldi, B. Cacciari, G. Spalluto, A. Borioni, M. Viziano, S. Dionisotti, E. Ongini, Current developments of A2A adenosine receptor antagonists, Curr. Med. Chem. 2 (1995) 707–722.
[28] E. Ongini, A. Monopoli, B. Cacciari, P.G. Baraldi, Selective adenosine A2A receptor antagonists, Farmaco 56 (2001) 87–90.