Available online at www.sciencedirect.com
SCIENCE DIRECT
Bioorganic & Medicinal Chemistry
Bioorganic & Medicinal Chemistry 14 (2006) 1454–1461

Classification of dopamine antagonists using functional feature hypothesis and topological descriptors

Hye-Jung Kim,[a,b] Yong Seo Cho,[a] Hun Yeong Koh,[a] Jae Yang Kong,[c] Kyoung Tai No[b] and Ae Nim Pae[a,*]

[a]*Biochemicals Research Center, Korea Institute of Science and Technology, PO Box 131, Cheongryang, Seoul 130-650, Republic of Korea*
[b]*Department of Biotechnology, Yonsei University 220, Sinchon-dong, Seodaemun-gu, Seoul 120-749, Republic of Korea*
[c]*Korea Research Institute of Chemical Technology, PO Box 107, Yusung-gu, Taejeon 305-343, Republic of Korea*

Received 18 August 2005; revised 28 September 2005; accepted 29 September 2005
Available online 25 October 2005

Abstract—The designing of selective dopamine antagonists for their own subreceptors can be useful in individual therapy of various neuropsychiatric disorders. Three-dimensional pharmacophore hypothesis and two-dimensional topological descriptors were used to investigate and compare different classes of dopamine antagonists. The structurally diverse D_3 and D_4 antagonists above preclinical trials were selected to map common structural features of highly selective and efficacious antagonists. The generated pharmacophore hypotheses were successfully employed as discriminative probe for database screening. To filter out the false positive from screening hits, the classification models by two-dimensional topological descriptors were built. Molconn-Z and BCUT topological descriptors were employed to develop a classification model for 1328 dopamine antagonists from MDDR database. The soft independent modeling of class analogy and artificial neural network, two supervised classification techniques, successfully classified D_1, D_3, and D_4 antagonists at the average of 80% rates into their own active classes. The mean classification rates for D_2 antagonists were obtained to 60% due to insufficient selective D_2 antagonists. In this paper, we report the validity of our models generated using functional feature hypotheses and topological descriptors. The combining both of classification using functional feature hypotheses and topological descriptors would be a useful tool to predict selective antagonists.

1. Introduction

Two types of dopamine receptors have been classified as D1-like and D2-like receptors. The D1-like receptors comprise of D_1 and D_5 receptor subtypes, while the D2-like receptors encompass D_2, D_3, and D_4 receptors. Traditionally, antipsychotics have been considered to act via the blockade of the classical 'D_2 receptor'.[1,2] In early 1990, the discovery of dopamine D_3 and D_4 receptor and their distribution in brain allowed us to consider a new target for antipsychotics devoid of extrapyramidal side-effects.[3–6] Many classical D_2 selective antipsychotics reported earlier were recently shown to possess high affinity for D_2, D_3, and D_4 receptors.[7,8] Unfortunately, clozapine, marketed as an atypical antipsychotic (approximately 10-fold D_4/D_2 selectivity), exhibits agranulocytosis in some patients and is currently of limited therapeutic use.[9] Therefore, the identification of highly selective antagonists via discriminating ligands on one or several dopamine subreceptors can be useful for designing highly subtype selective antagonists.[10,11]

For that reason, the goal of this study was to create robust pharmacophore model and classification model that can categorize and predict selective dopamine antagonists (DAs) for their relevant classes.

The classification of DAs by topological fragment spectral method has been reported.[12] The pharmacophore model for D4 antagonists has been identified by superimposition[13] and QSAR model for some D4As has been reported.[14] In this paper, the common features for D3A and D4A are identified using structurally diverse antagonists above preclinical trial to generate the phar-

Keywords: Dopamine antagonist; Classification; SIMCA; ANN; Molconn-Z; BCUT; Pharmacophore hypothesis.
* Corresponding author. Tel.: +82 02 958 5185; fax: +82 02 958 5189; e-mail: anpae@kist.re.kr

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.09.072

macophore based classification model. The HipHop program in Catalyst[15] was used to generate the pharmacophore for distinguishing the common features of active antagonists for each subtype. In addition to this, the 2D structure based classification models by topological descriptors were established. All available DAs from MDDR (MDL drug data report)[16] were collected and each DA was assigned a subtype classification (D1A = 1 to D4A = 4). SIMCA[17] (soft independent modeling by class analog) and ANN[18] (artificial neural network) methods were applied for discriminating DAs using Molconn-Z[19] and BCUT[20] topological descriptors. Our models that combine 2D and 3D structural descriptors could be used for virtual screening and designing of selective and potent antagonists.

2. Methods

2.1. Dataset

The 1475 dopamine antagonists of four different types of receptors from MDDR database (D1A: 137 compounds, D2A: 440 compounds, D3A: 249 compounds, and D4A: 649 compounds) were collected for data set. Counter ions and solvent molecules were eliminated and charged groups were neutralized by adding or removing protons. Duplicates within and between the classes were removed. To generate a 2D structure based classification model, all data set was divided into two subgroups: the training and test sets. The training set consists of 1328 compounds (90% of the total data) and the test set consists of 147 compounds (10% of the total data). To generate pharmacophore hypothesis, structurally diverse six D3As[21–24] above preclinical trial and six D4As[25–30] above phase I trial were selected based on the potency and selectivity. The structures and the pharmacological information of selected DAs are shown in Figure 1 and Table 1.

2.2. Pharmacophore hypothesis generation

For the pharmacophore generation, the conformational models of selected six antagonists having up to 250 conformers were built using the 'best conformer generation' method with a 20 kcal/mol energy cutoff.[31] Pharmacophore analyses for D3As and D4As were carried out using the HipHop module in Catalyst,[15] which consisted only of identification and overlay of common feature shared by all antagonists without taking their biological data into account. The parameter values of

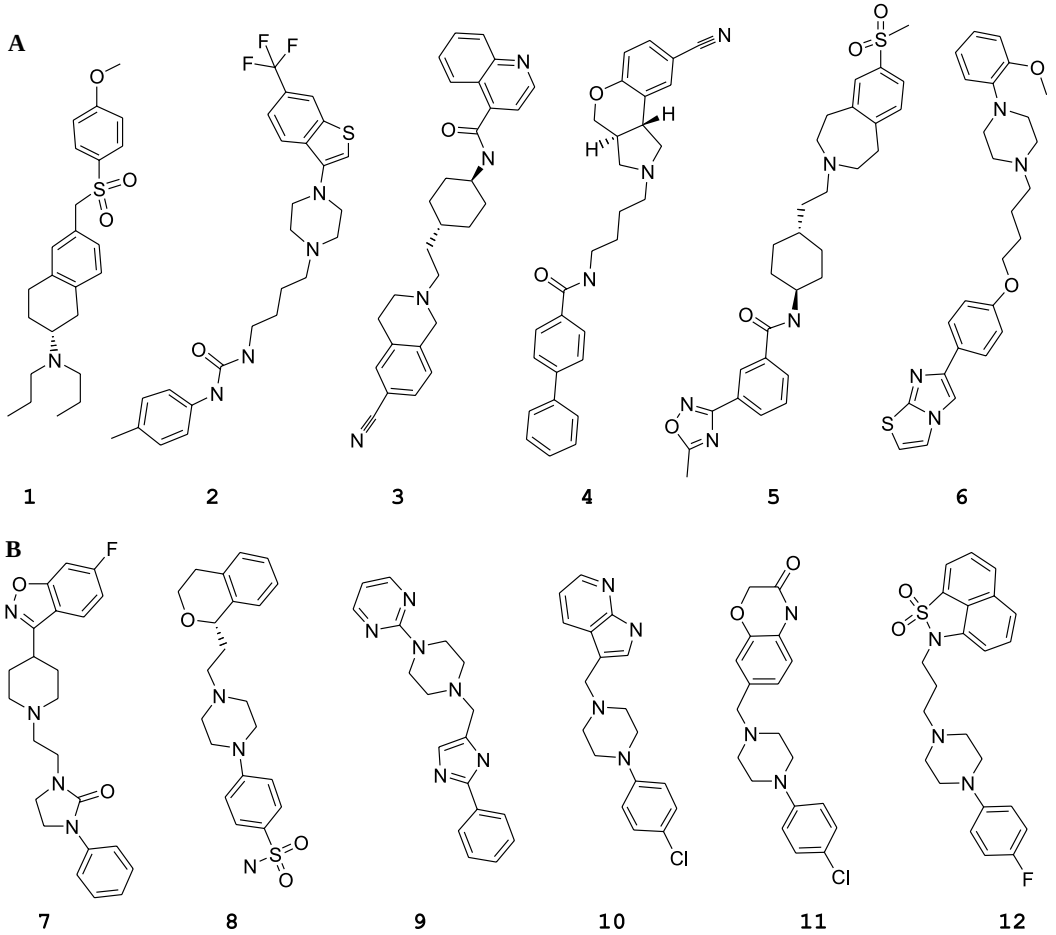

Figure 1. Chemical structures used for hypothesis generation. (A) The selective D3 antagonists above preclinical trial. (B) The selective D4 antagonists above phase I trial.

Table 1. Functional data for D3 and D4 antagonists selected for hypothesis generation

Compound	Candidate	IC_{50} (nM)	Selectivity (to D2)
D3 antagonist			
1	GR-218231 (preclinical)	6.3	~10000
2	AVE-5997 (phase I)	3.0	~500
3	SB-277011 (preclinical)	10.0	~100
4	S-33084 (preclinical)	0.3	~100
5	SB-414796 (preclinical)	4.0	~100
6	RGH-1756 (preclinical)	0.2	~40
D4 antagonist			
7	Fananserin (phase II)	2.9	~20000
8	L-745870 (phase II)	0.4	~2200
9	NGD-94-1 (phase I)	3.8	~2000
10	Sonepiprazole Mesilate (phase II)	6.8	~400
11	PD-172760 (phase I)	4.3	~100
12	S-18327 (phase I)	7.9	~10

(1, 1) and (2, 2) for miss-feature and miss-compound were set. All 20 hypotheses were obtained for D3 and D4 antagonists. To generate pharmacophore features, four chemical functions (hydrogen bonding acceptor, aromatic ring, positive ionizable, and hydrophobic group) were used. After deleting the redundant hypotheses, the diverse hypotheses were selected according to ranking scores and fitting scores. To evaluate discriminative powers of selected pharmacophore hypotheses, database screening was performed on all the dopamine antagonists.

2.3. Descriptor calculation

All data set collected from MDDR database was imported to molecular spreadsheet of SYBYL program.[32] 332 Molconn-Z descriptors were calculated on the basis of a SD file. The Molconn-Z carries out the computation of a wide range of topological indices of molecular structure. These indices reflect the intrinsic features of molecules; atoms, their electronic state, and their relationship to one another.[33] Molconn-Z's families of descriptors are based on the pioneering work of Lowell Hall and Lemont Kier.[34] BCUT descriptors were calculated using DiverseSolutions version 6.2 module[35] in SYBYL. BCUT descriptor is a symmetrical matrix representing molecular properties and bonding information in its diagonal and off-diagonal.[20] Atomic charge, polarizabilities, and H-bond acceptor–donor abilities were considered as diagonal elements. These two types of information were combined through a process called matrix diagonalization using auto-scaling option. The resulting 45 highest and lowest eigenvalues were used to form the BCUT descriptors. A significant advantage of these topological descriptors is that geometry optimization of the structure is not needed. The most relevant Molconn-Z and BCUT descriptors in differentiating the classes were selected by comparing their correlations with each other. The correlation matrices of descriptors were built, and 27 Molconn-Z and 14 BCUT descriptors were selected on the basis of the correlation threshold of $R = 0.9$ (Tables S1 and S2, Supporting information). These narrowed 27 Molconn-Z and 14 BCUT descriptors were used for a 2D structure based classification model. The mixed descriptors by combining these two descriptors also applied to the SIMCA and ANN models.

2.4. Classification model development

Each descriptor set was used independently in separate SIMCA and ANN models, and the classification model was also generated using both Molconn-Z and BCUT descriptors. The SYBYL implementation was used to generate SIMCA model.[36] The SIMCA method applies principal component analysis (PCA) separately to each class of objects and uses the principal components (PCs) to define hyper-volumes in the descriptor space. Prediction consists of calculating the distance to the hyper-volume for each category. The number of significant components was calculated with pre-scaled descriptors in a 4-fold cross-validation procedure. The optimal number of components for each category was selected based on the sum of squared residuals for each compound.

A fully connected, three layer-backpropagation neural network (BPNN)[37] was trained using Cerius2 program.[38] The network tries to minimize errors between the estimated classes and the true classes. Each descriptor set was submitted to the ANN input signals for the input neurons. The number of hidden layer neurons was varied to find the best classification model. The best model was determined by trial and error. Two kinds of tests were performed to validate the classification models. First, the prediction test by test set, 10% compounds of all data set was checked. Then, cross-validation process was tested to avoid overfitting and to improve generalization of the classification models. One-tenth of all compounds as a test set and the remainder as a training set were taken. Ten different test sets were generated and each antagonist in total set was predicted once in a test set.

3. Results and discussion

3.1. Pharmacophore hypothesis generation and validation by database screening

The HipHop module of Catalyst was used to map common features for our training set, which has potent receptor binding affinity and good antipsychotic efficacy (Fig. 1, Table 1). All 20 hypotheses were generated by setting the parameter options, miss-feature and miss-compounds, to (1, 1) and (2, 2). The redundant pharmacophores with same feature configuration were deleted by examining the ranking scores of each hypothesis and fitting scores to training compounds. Therefore, three representative hypotheses with best scores are reported in Table 2. Although the scores of top three models are almost same, the composition of each pharmacophore model and disposition of features were completely different from each other.

Table 2. Summary of the common feature hypotheses run

No.[a]	Composition[b]	Ranking score[c]	Direct hit mask[d]	Partial hit mask[e]
D3 antagonists				
1	RPHHA	91.91	110111	001000
2	RPHHA	91.68	111111	000000
3	PHHAA	90.83	111011	001000
D4 antagonists				
1	RPHH	58.99	111101	000010
2	RRPH	57.02	001111	110000
3	PHHA	56.70	111100	000011

[a] Numbers for the hypothesis are consistent with the order obtained by ranking score and fitting score.
[b] R, aromatic ring; H, hydrophobic group; A, hydrogen bond acceptor; P, positive ionizable.
[c] The higher the ranking score, the less likely it is that the molecules in training set fit the hypothesis by chance correlation. Best hypotheses have highest ranks.
[d] A training set molecule mapped every feature in the hypothesis. 1 means yes and 0 means no.
[e] A training molecule mapped all but one feature in the hypothesis.

Figure 2 displays the pharmacophore model in order of scores.

To determine the hypothesis that can be used as a rational query, 1475 dopamine antagonists were used as a test set. Catalyst database for all dopamine antagonists was built containing 300 conformations. The 'best flexible search' option was selected to screen compounds that shared the same 3D representations of the functional group in some conformations. The recognition rates from screening are reported in Table 3. The hits used as the training compounds were excluded the counting. The results of pharmacological screening placed in evidence that our mappings onto the pharmacophore models are quite reasonable. The best hypotheses with better sensitivity and specificity were determined comparing the correct prediction rate. Particularly, the results of the validation tests showed that the second best models in ranking score for each hypothesis are efficient in recognizing selective antagonists. In this work, we realize that the validation test is important because small difference in model configuration can give large difference in the number of hits. Moreover, the pharmacophore models with second ranking score presented better hit number than the first ranked one. The best D3A pharmacophore model reported great true positive recognition rate of 81% and gave reasonable false positive rate of about 30% for D2As and D4As. Specially, the hit rate of 0% in D1As confirms the validity of our D3A pharmacophore modeling approach. The best D4A pharmacophore model recorded high true positives of 90%, while it exhibited poor false positives of 57% and 67% for D2As and D3As, respectively. It is possibly explained that the D3As contain some interaction sites for D4 receptor due to the larger size than D4As. Many known D2 antagonists are actually dual antagonists for D2 and D4 receptors. It is difficult to distinguish the D2 and D4 antagonists with only pharmacophore model without receptor structure. To remove these false positives, a 2D structure based classification model was developed. Following this simultaneous application of D3A and D4A pharmacophore hypothesis, classification can be a good approach to develop efficient antipsychotics.

3.2. 2D structure based classification model

For the entire set, 1328 DAs were set to the training (D1: 124, D2: 396, D3: 224, and D4: 584) and the remaining 147 DAs were used as a test set (D1: 13, D2: 44, D3: 25, and D4: 65). All 2D structure based classification models by SIMCA and ANN were generated using Molconn-Z, BCUT, and mixed descriptors of Molconn-Z and BCUT.

For the training set composed of 1328 dopamine antagonists, SIMCA analyses exhibited good recognition rates on the three component model. Table 4 shows the success rate for training and test sets. These models account for the success recognition rates of 70.9%, 70.3%, and 75.0% for Molconn-Z, BCUT, and mixed descriptors, respectively. The combining of two kinds of descriptors improved the results than that separatly employed. In the case of D2A, it showed poor recognition rate due to the lack of selective D2As. The results detailed in Table 4 show that nearly half of D2As were misclassified as antagonists of another class. We analyzed the distribution of the false positives of D2As misclassified to other classes from all test sets. Figure 3 shows that misclassification of D2As as another class depends upon algorithm not descriptor. The false positives from SIMCA analyses were distributed in D3 and D4 classes. Most of the false positives from ANN analyses were misclassified into the D3 class. Early D2As are actually D2-like antagonists. Most of post D2As developed after defining D3 and D4 subtypes are dual antagonists, which display the main activity for serotonin receptor with low contribution for the dopamine receptor. These antagonists also have low selectivity for D2-like receptors. These insufficient selective D2As provide understandable limitation to our classification model. Classification rates of D1As showed the best recognition rate using Molconn-Z (93.5%) than using other descriptor sets. The classification rates for D3A and D4A recorded more than 80% using mixed descriptors. SIMCA models by Molconn-Z and BCUT descriptor sets yielded an almost similar classification rate and the SIMCA model by mixed descriptors resulted in better result than than other two models. The prediction tests on 147 DAs by three kinds of descriptors led to good predictability. The model was able to predict D1As, D3As, and D4As with a success rate greater than 70% in all cases for the descriptor set. For the misclassified D2As, most of them were assigned to D3As or D4As.

For training set, the optimum three hidden layer was selected for the Molconn-Z ANN model and the best BCUT ANN model was generated by a four hidden layer. Figure 4 displays the relationship between the success rate and the number of hidden layers. ANN models for training set achieved a better recognition rate than SIMCA models (Table 4). Specially, the mixed descriptor has categorized the training set 80.0% correctly. Although the classification rate for D2As is poorer

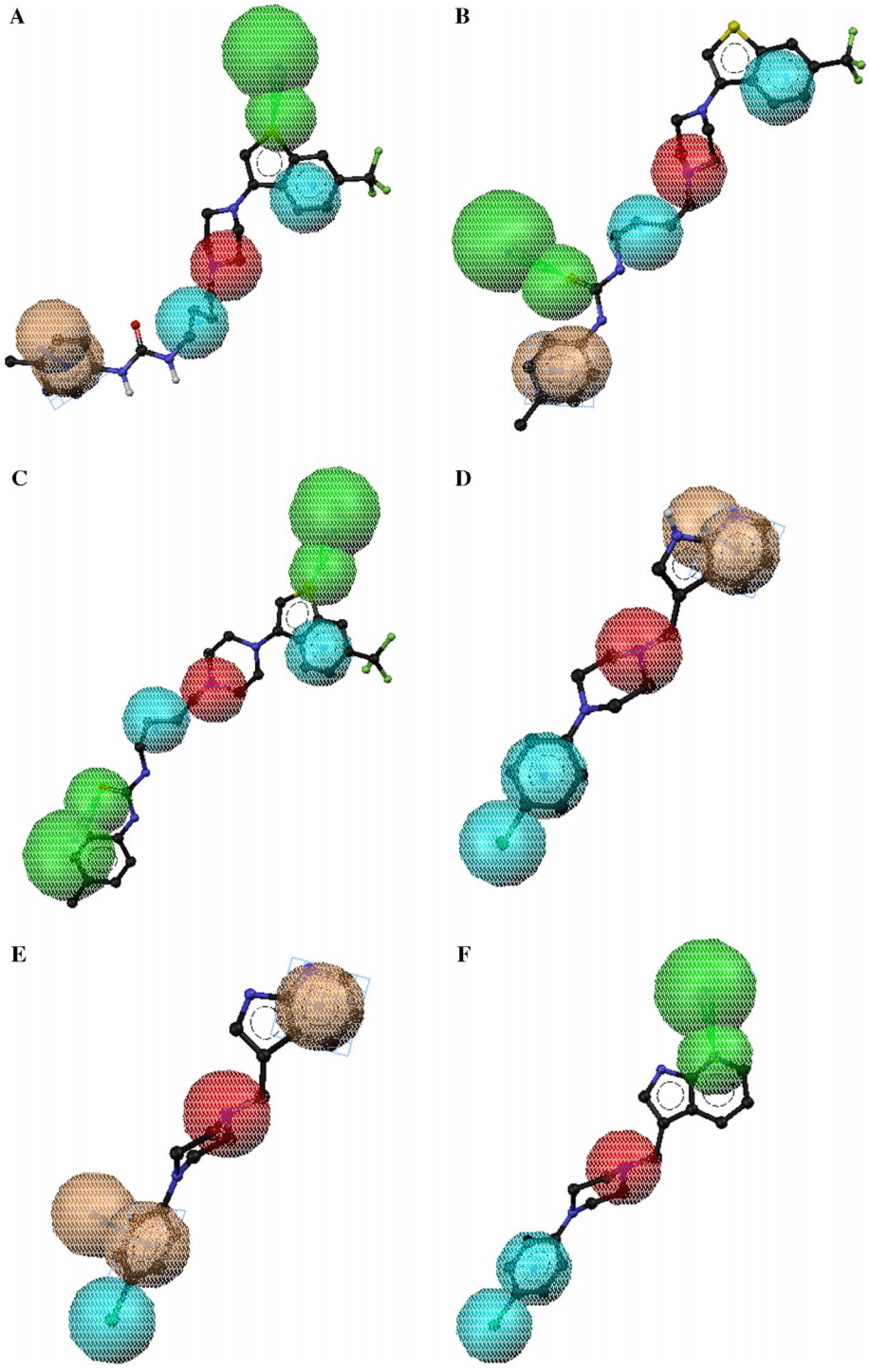

Figure 2. Diverse pharmacophore hypotheses used as query. They display in order of ranking scores and fitting scores. (A–C) D3A Pharmacophore models aligned with AVE-5997. (D–F) D4A pharmacophore models aligned with L-745870. The second best model for each antagonist was determined as the best hypothesis by database screening. Pharmacophore features are color coded (cyan, hydrophobic center; green, H-bond acceptor; orange, aromatic ring; red, positive charge or positive ionizable).

than those of other classes, the prediction rate at 68.2% of the ANN model using mixed descriptors can be useful to filter the false positives for D2 antagonists. For D1As, D3As, and D4As, ANN models gave a classification percentage much greater than 70% in each descriptor set. The prediction test by external test set recorded good performance. The ANN model by separate Molconn-Z and BCUT predicted the test set of about 70% correctly. The mixed descriptor correctly classified 80.3% of the test set.

These are important results as they indicate that the Molconn-Z and BCUT descriptors are powerful tools to predict the pharmacological active class of dopamine

Table 3. The screening results for all dopamine antagonists using pharmacophore models selected in order of ranking score and matching score

D3A hypothesis	D1: false positive (%)	D2: false positive (%)	D3: true positive (%)	D4: false positive (%)
1	1	39	49	30
2	0	26	81	29
3	1	22	35	29
D4A hypothesis	**D1: false positive (%)**	**D2: false positive (%)**	**D3: false positive (%)**	**D4: true positive (%)**
1	4	57	70	84
2	1	57	67	90
3	3	56	53	51

D1A: 137 compounds; D2A: 440 compounds; D3A: 249 compounds; D4A: 649 compounds.

Table 4. Classification results of SIMCA and ANN for the training and the prediction sets using Molconn-Z, BCUT, and mixed descriptor

SIMCA class	Molconn-Z	BCUT	Mix	ANN class	Molconn-Z	BCUT	Mix
Training (*all*: *1328*, *D1*: *124*, *D2*: *396*, *D3*: *224*, *D4*: *584*)							
All	942 (70.9)	933 (70.3)	996 (75.0)	all	1016 (76.5)	982 (73.9)	1063 (80.0)
D1	116 (93.5)	88 (71.0)	107 (86.3)	D1	106 (85.5)	91 (73.4)	95 (76.6)
D2	202 (51.0)	223 (56.3)	225 (56.8)	D2	248 (62.6)	258 (65.2)	269 (67.9)
D3	161 (71.9)	171 (76.3)	189 (84.4)	D3	198 (88.4)	190 (84.8)	198 (88.4)
D4	463 (79.3)	451 (77.2)	475 (81.3)	D4	464 (79.5)	443 (75.9)	501 (85.6)
Prediction (*all: 147, D1: 13, D2: 44, D3: 25, D4: 65*)							
All	103 (71.4)	104 (70.7)	111(75.5)	all	105 (71.4)	104 (70.7)	118 (80.3)
D1	11 (84.6)	10 (76.9)	11 (84.6)	D1	9 (69.2)	10 (76.9)	10 (76.9)
D2	22 (50.0)	24 (54.5)	26 (59.0)	D2	28 (63.6)	26 (59.0)	30 (68.2)
D3	19 (76.0)	20 (80.0)	22 (88.0)	D3	19 (76.0)	20 (80.0)	22 (88.0)
D4	53 (81.5)	50 (76.9)	52 (80.0)	D4	49 (75.3)	48 (73.8)	56 (86.2)

antagonists. The combination of these two descriptors is particularly suited to such an approach to calculate the class of dopamine antagonists. Therefore, our classification model using mixed descriptors can filter out the false positives from virtual screening hits by pharmacophore only.

3.3. Model validation by cross-validation test

In the previous section, our classification models were validated by an external test set. We performed a 10-fold cross-validation process as the second validation. For the entire data set of 1475 DAs containing training and test, 10 different trial data sets were built. The average and standard deviation results of SIMCA classification for 10 different trials are summarized in Table 5. Different ten-test sets generated reasonable prediction accuracy as well as training set. All descriptor sets pre-

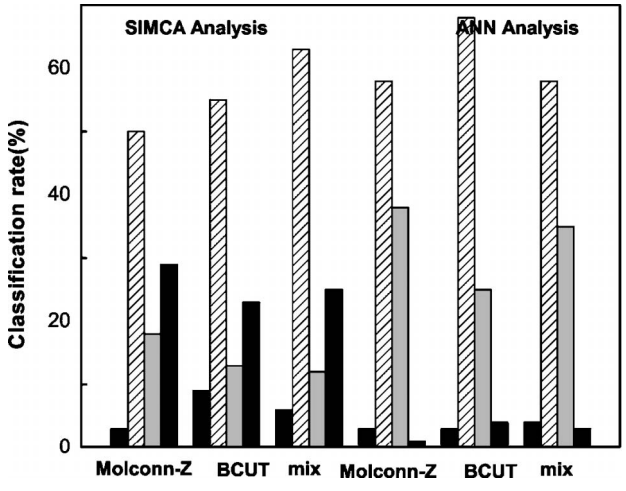

Figure 3. The distribution for D2 antagonists of test sets classified by six different classification models (black, D1As; hashed white, D2As; light gray, D3As; dark gray, D4As).

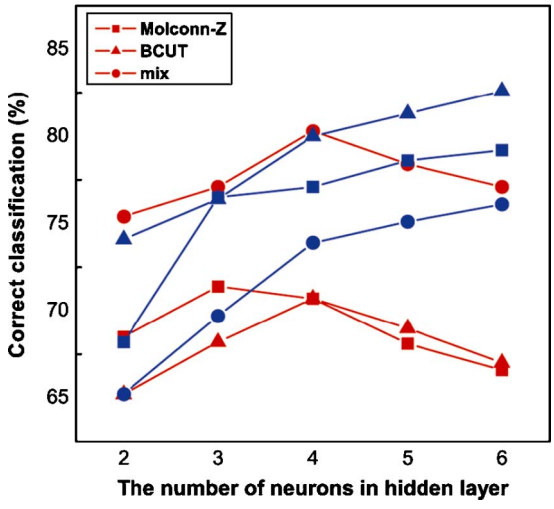

Figure 4. Classification rates of the training and the prediction test by three kinds of descriptor based artificial neural network with different number of neurons in the hidden layer. For the training set, 1328 antagonists were used and 147 antagonists were predicted external test set. The plots are results for the selected best model. The training and test sets are colored by blue and red, respectively.

Table 5. Mean values and standard deviation for the training and the prediction obtained in a 10-fold cross-validation test

SIMCA class	Molconn-Z	BCUT	Mix	ANN class	Molconn-Z	BCUT	Mix
Training (*recognition rate %*)							
All	70.6 ± 1.8	72.3 ± 1.9	74.8 ± 1.7	All	76.9 ± 2.0	73.0 ± 1.9	80.0 ± 1.7
D1	92.4 ± 2.4	72.0 ± 1.8	85.7 ± 2.8	D1	82.3 ± 2.1	75.6 ± 2.0	76.6 ± 2.3
D2	50.9 ± 1.3	57.6 ± 1.4	55.7 ± 1.2	D2	65.9 ± 1.8	58.5 ± 1.9	67.6 ± 2.2
D3	70.4 ± 1.8	75.3 ± 1.9	83.9 ± 1.8	D3	87.1 ± 2.2	79.1 ± 2.0	87.7 ± 2.0
D4	78.6 ± 2.0	76.7 ± 2.0	81.1 ± 1.4	D4	80.1 ± 2.1	77.8 ± 2.4	86.2 ± 1.5
Prediction (*prediction rate %*)							
All	71.6 ± 1.8	70.8 ± 1.8	75.1 ± 2.8	All	72.4 ± 1.1	71.1 ± 1.8	80.3 ± 3.6
D1	89.6 ± 2.2	76.3 ± 1.9	83.8 ± 5.2	D1	74.9 ± 1.9	74.8 ± 1.9	76.9 ± 4.9
D2	48.6 ± 1.2	54.8 ± 1.6	54.4 ± 3.4	D2	56.9 ± 1.3	58.1 ± 1.4	65.1 ± 2.1
D3	73.2 ± 2.1	79.1 ± 2.0	84.4 ± 4.9	D3	77.1 ± 1.7	81.5 ± 2.2	88.0 ± 2.8
D4	80.0 ± 2.1	77.0 ± 2.0	80.6 ± 2.9	D4	76.9 ± 1.9	73.1 ± 1.9	85.8 ± 1.2

dicted successfully test sets at an average rate of more than 70%. From the prediction rates, we can conclude that our models obtained by the SIMCA method are reliable and stable.

We also tested the validity of ANN models using the same method as the SIMCA models. The total average recognition rates and standard deviation of training and test sets are reported in Table 5. The average classification rates of test sets are consistent with results of training sets. We achieved quite similar and good results using SIMCA and ANN algorithm, and using Molconn-Z and BCUT descriptors, which can differentiate each subtype with its chemical features. Therefore, it can be considered that our models are valid to classify and predict dopamine antagonists strongly.

4. Conclusion

In this paper, we suggested the merge of pharmacophore based virtual screening and 2D structure based classification to increase the hit rates for selective dopamine antagonists. We generated the pharmacophoric models for D3As and D4As, which can be used directly for high-throughput virtual screening. Despite the small training set, the best models showed highly quantitative predictabilities. The D3A pharmacophoric model achieved good discriminative power between D4As and D3As. The D4A model has more false positives that might be due to the smaller number of features than the D3A model, but the false positives predicted as D3As and D2As could be eliminated by a 2D structure based classification model. The classification methods are likely to be useful in the filtering of false positives from virtual screening by the pharmacophore model for efficient antipsychotics. In this paper, we have compared SIMCA and ANN in their classification performances for all the currently available dopamine antagonists. We also investigate the effectiveness of the three descriptor set, Molconn-Z, BCUT and mixed descriptor, which represent different structural information and some physicochemical properties. The cross-application of two different algorithms and three different descriptors achieved a reasonable degree of classification and good prediction value for test set. Furthermore, we obtained a consensus predictive potential for 10 different test sets and training sets using all classification models. Overall good classification accuracies suggest that the Molconn-Z and BCUTs can be combined with SIMCA and ANN to recognize selective antagonists for antipsychotics. Specially, the hit rates for D3 and D4 antagonists would be increased when the virtual screening hits were put into the classification model as test compounds. We can conclude that the combining of database screening by pharmacophoric model and classification can make it possible to complement the limitation from insufficient selective D2As.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2005.09.072.

References and notes

1. Creese, I.; Burt, D. R.; Snyder, S. H. *Science* **1976**, *192*, 481.
2. Seeman, P.; Lee, T.; Chau-Wong, M.; Wong, K. *Nature* **1976**, *261*, 717.
3. Sokoloff, P.; Giros, B.; Martres, M. P.; Bouthenet, M. L.; Schwartz, J. C. *Nature* **1990**, *347*, 146.
4. Van Tol, H. H.; Bunzow, J. R.; Guan, H. C.; Sunahara, R. K.; Seeman, P., et al. *Nature* **1991**, *350*, 610.
5. Bouthenet, M. L.; Souil, E.; Martres, M. P.; Sokoloff, P.; Giros, B., et al. *Brain Res.* **1991**, *564*, 203.
6. Seeman, P. *Neuropsychopharmacology* **1992**, *7*, 261.
7. Seeman, P.; Guan, H. C.; Van Tol, H. H. *Nature* **1993**, *365*, 441.
8. Ogren, S. O.; Hall, H.; Kohler, C.; Magnusson, O.; Sjostrand, S. E. *Psychopharmacology* **1986**, *90*, 287.
9. Lahti, R. A.; Evans, D. L.; Stratman, N. C.; Figur, L. M. *Eur. J. Pharmacol.* **1993**, *236*, 483.
10. Audinot, V.; Newman-Tancredi, A.; Gobert, A.; Rivet, J. M.; Brocco, M. *J. Pharmacol. Exp. Ther.* **1998**, *287*, 187.
11. Strange, P. G. *Pharmacol. Rev.* **2001**, *53*, 119.
12. Fujishima, S.; Takahashi, Y. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1006.
13. Bostrom, J.; Gundertofte, K.; Liljeforsa, T. *J. Comput. Aided Mol. Des.* **2000**, *14*, 769.

14. Bostrom, J.; Bohm, M.; Gundertofte, K.; Klebe, G. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1020.
15. Accelrys Inc. Catalyst, Version 4.9; San Diego, CA, USA, 2003; http://www.accelrys.com.
16. MDL Drug Data Report. http//:www.discoverygate.com.
17. Wold, S. *Pattern Recognit.* **1976**, *8*, 127.
18. Zupan, J.; Gasteiger, J. *Neural Networks in Chemistry and Drug Design*, 2nd ed.; Weinheim: Wiley-VCH, 1999.
19. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure–Activity Analysis*; Research Studies Press Ltd/ Wiley: Letchworth, Hertfordshire, England, 1986.
20. Pearlman, R. S.; Smith, K. M. *Persp. Drug Discov. Des.* **1998**, *9*, 339.
21. Millan, M. J.; Dekeyne, A.; Rivet, J. M.; Dubuffet, T.; Lavielle, G., et al. *J. Pharmacol. Exp. Ther.* **2000**, *293*, 1063.
22. Stemp, G.; Ashmeade, T.; Branch, C. L.; Hadley, M. S.; Hunter, A. J., et al. *J. Med. Chem.* **2000**, *43*, 1878.
23. Dubuffet, T.; Newman-Tancredi, A.; Cussac, D.; Audinot, V.; Loutz, A., et al. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2059.
24. Macdonald, G. J.; Branch, C. L.; Hadley, M. S.; Johnson, C. N.; Nash, D. J., et al. *J. Med. Chem.* **2003**, *46*, 4952.
25. Comte, M. I.; Gueremy, C.; Malleron, J. L.; Mignani, S.; Peyronel, J. F., et al. EP 0350403, **1989**.
26. Leeson, P. D.; Smith, A. L.; Ridgill, M. P.; Baker, R.; Curtis, N. R., et al. EP 0623618, **1995**.
27. Boyfield, I.; Brown, T. H.; Coldwell, M. C.; Cooper, D. G.; Hadley, M. S., et al. *J. Med. Chem.* **1996**, *39*, 1946.
28. TenBrink, R. E.; Ennis, M. D.; Lin, C.-H.; Lahti, R. A.; Romero, A. G., et al. WO 95/018118, **1999**.
29. Belliotti, T. R.; Wise, L. D.; Wustrow, D. J. Benzoxazinone dopamine D4 receptor antagonists. WO 97/045419. **1998**.
30. Millan, M. J.; Brocco, M.; Rivet, J.-M.; Audinot, V.; Newman-Tancredi, A., et al. *J. Pharmacol. Exp. Ther.* **2000**, *292*, 54.
31. Smellie, A.; Teig, S. L.; Towbin, P. *J. Comput. Chem.* **1995**, *16*, 171.
32. TRIPOS Inc. SYBYL, Version 7.0; Hanley Rd, St. Louis, Missouri, USA, 2004; http://www.tripos.com.
33. Hall, L. H. Computational Aspects of Molecular Connectivity and its Role in Structure-Property Modeling. In *Computational Chemical Graph Theory*, Rouvray, D. H., Ed.; Nova Press, New York. 1990; Chapter 8, pp 203–233.
34. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.
35. Menard, P. R.; Mason, J. S.; Morize, I.; Bauerschmidt, S. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1204.
36. Dunn, W. J., III; Wold, S. *Simca Pattern Recognition and Classification. Chemometric Methods in Molecular Design*; VCH: New York, 1995, pp 179–193.
37. Hornik, K.; Stinchombe, M.; White, H. *Neural Networks* **1989**, *2*, 359.
38. Accelrys, Inc. Cerius 2, Verson 4.9; San Diego, CA, USA, 2003; http://www.accelrys.com.