G&I Genomics & Informatics

pISSN 1598-866X eISSN 2234-0742
Genomics Inform 2013;11(4):282-288
http://dx.doi.org/10.5808/GI.2013.11.4.282

ORIGINAL ARTICLE

Molecular Vibration-Activity Relationship in the Agonism of Adenosine Receptors

Hyun Keun Chee[1], S. June Oh[2]*

[1]Department of Thoracic and Cardiovascular Surgery, Konkuk University Medical Center, Konkuk University School of Medicine, Seoul 143-729, Korea,
[2]Department of Pharmacology, Inje University College of Medicine, Busan 614-735, Korea

The molecular vibration-activity relationship in the receptor-ligand interaction of adenosine receptors was investigated by structure similarity, molecular vibration, and hierarchical clustering in a dataset of 46 ligands of adenosine receptors. The resulting dendrogram was compared with those of another kind of fingerprint or descriptor. The dendrogram result produced by corralled intensity of molecular vibrational frequency outperformed four other analyses in the current study of adenosine receptor agonism and antagonism. The tree that was produced by clustering analysis of molecular vibration patterns showed its potential for the functional classification of adenosine receptor ligands.

Keywords: corralled intensity of molecular vibrational frequency, G-protein-coupled receptors, molecular descriptor, molecular vibration-activity relationship, purinergic P1 receptors

Introduction

Membrane proteins play crucial roles in organismal transduction of information. Molecular recognition in biological membrane systems is the initial process triggering external factors to be passed into the inner part of a cell. G protein-coupled receptors (GPCRs), which consist of 7 transmembrane segments, play pivotal roles for signal transduction in higher eukaryotes and organize the largest families of proteins in the human genome [1]. The fact that drugs targeting GPCRs constitute over one-third of pharmaceuticals shows the importance of GPCRs as drug targets [2].

Adenosine receptors (AdoRs) belong to a family of rhodopsin-like class A GPCRs, which constitute the largest portion of GPCRs in humans, including olfactory receptors, and there are four known subtypes of AdoRs—namely, AdoR1, AdoRA2A, AdoRA2B, and AdoR3 [3, 4]. Although these AdoR subtypes have distinct amino acid sequences and tissue-specific distribution, their endogenous agonist and ligand-binding pockets are highly conserved.

The biogenic nucleoside adenosine (Fig. 1) exerts its effects by interacting with AdoRs that are involved in various diseases, such as cardiac ischemia, arrhythmia, neurodegeneration, diabetes, glaucoma, and inflammation [5]. The physiological effects of AdoRs can be used for diagnoses and planning of the surgical strategies, such as in coronary artery diseases [6].

Olfaction, the sense of smell, is also mediated by class A GPCRs—namely, olfactory receptors. The fundamental mechanism of olfaction is now under some controversy [7-9]. Various attempts were introduced to describe the molecular mechanism of receptor-ligand recognition in olfaction, such as the binding theory and vibration theory

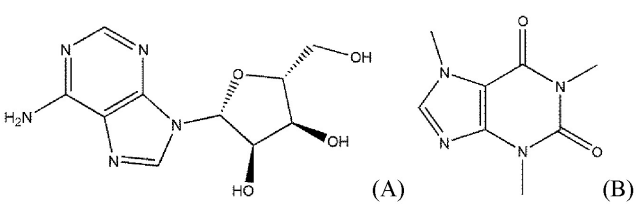

Fig. 1. The chemical structures of adenosine and caffeine as an agonist and an antagonist of adenosine receptors, respectively. (A) Adenosine, an endogenous agonist of adenosine receptors. (B) Caffeine, a typical antagonist of adenosine receptors.

Received October 29, 2013; Revised November 12, 2013; Accepted November 14, 2013
***Corresponding author:** Tel: +82-51-893-3569, Fax: +82-51-797-8569, E-mail: o@biophilos.org
Copyright © 2013 by the Korea Genome Organization
It is identical to the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/).

[10-12]. The former, where the specificity of ligand is explained by molecular shape, has been developed into a concept of a pharmacophore and is generally accepted by researchers. Nevertheless, this is not sufficient to give the account of ligand variety and agonism complexity of GPCRs.

In recent decades, there have been many efforts to find an efficient way to seek or make appropriate ligands working on the target receptors through a chemogenomic approach. Application of molecular descriptors can make an efficient way to discriminate ligand binding modes in receptor-ligand relationships in such cases as in silico chemogenomic screening [13]. Molecular descriptors encoding information about the molecular structure can be classified by the dimensionality of their molecular representation, and they are almost on the structural, topological, and geometrical bases [14]. They are also including information on dipole moment, electric polarizability, and electrostatic potential in their data types.

Lately, a computational approach was carried out to search for a molecular vibration-activity relationship in the agonism of histamine receptors, and the author suggested that the molecular vibrational frequency pattern may play a role as a possible molecular descriptor for histamine receptor ligands [15]. The EigenVAlue (EVA) descriptor is also based on infrared range molecular vibrations among the various molecular descriptors and is a unique one, based on eigenvalues corresponding to individual calculated normal modes among various molecular descriptors [16, 17].

In the present paper, we tested the potential of the corralled intensity of molecular vibrational frequency (CIMVF) [15] as a molecular descriptor, compared to the EVA descriptor and other descriptors for the classification of AdoR ligands.

Methods

Dataset

All 46 ligand molecules in the dataset, comprising 23 AdoR agonists and 23 antagonists, are shown in Table 1, with the Tanimoto distance of each ligand from adenosine.

The Tanimoto coefficient is a widely used measure of molecular structural similarity. The coefficient is defined as $Tc = N_{ab}/(N_a + N_b - N_{ab})$, with N_{ab} being the number of common bits, N_a the unique bits in molecule a, and N_b the unique bits in molecule b, using a molecular fingerprint [18]. In this study, the molecular similarity to adenosine was calculated as the Tanimoto coefficient using the 38-bit set. The simplified molecular-input line-entry system (SMILES)

Table 1. Dataset of adenosine receptor ligands used in the present study and their molecular similarity in the Tanimoto coefficient to adenosine

PID	Compound	Tanimoto	Function	PID	Compound	Tanimoto	Function
60961	Adenosine	1	Agonist	2519	Caffeine	0.343284	Antagonist
8974	2-CADO	0.887097	Agonist	1329	DPCPX	0.339394	Antagonist
4402	CPA	0.814815	Agonist	64627	KW-3902	0.333333	Antagonist
1585	CV1808	0.797101	Agonist	9953065	SLV320	0.315476	Antagonist
5311431	ENBA	0.797101	Agonist	9882625	PSB-11	0.295964	Antagonist
164305	SDZWAG884	0.791367	Agonist	6918429	AMP-579	0.29148	Agonist
104795	NECA	0.75	Agonist	5697	XAC	0.288557	Antagonist
93205	R-PIA	0.733333	Agonist	176408	SCH58261	0.2723	Antagonist
123807	CCPA	0.728477	Agonist	5311037	KW-6002	0.267327	Antagonist
56972171	2-MeCCPA	0.714286	Agonist	5353365	CSC	0.259259	Antagonist
10981286	HEMADO	0.691824	Agonist	10196114	CVT-5440	0.246809	Antagonist
5310992	AB-MECA	0.635294	Agonist	9996279	FA385	0.243243	Antagonist
219024	CVT-3146	0.617978	Agonist	5311041	OSIP339391	0.229358	Antagonist
164437	HENECA	0.606742	Agonist	2690	CGS15943	0.22549	Antagonist
16218845	ADAC	0.597826	Agonist	9821511	FR194921	0.201794	Antagonist
3086599	CGS21680	0.593407	Agonist	176407	ZM241385	0.17734	Antagonist
9828356	CP-608039	0.431535	Agonist	9975036	LUF5835	0.173516	Agonist
9833519	UK-432097	0.426877	Agonist	11716665	LAS-38096	0.169643	Antagonist
11248240	MRS3558	0.409574	Agonist	3661570	MRS1523	0.169492	Antagonist
9902054	WRC-0571	0.402597	Antagonist	11717831	BAY60-6583	0.165919	Agonist
1676	Enprofylline	0.386364	Antagonist	6439091	FK453	0.159292	Antagonist
158540	BG9719	0.352601	Antagonist	9936489	Capadenoson	0.12766	Agonist
2153	Theophylline	0.348485	Antagonist	5282339	CGH2466	0.0972222	Antagonist

PID, PubChem ID.

and 3-dimensional structure data format (SDF) files of the dataset were downloaded from the PubChem Compound Database in National Center for Biotechnology Information (NCBI) and used in further analyses.

Molecular descriptors

Molecular fingerprints: MACCS keys, PubChem, and Klekota-Roth fingerprint

Molecular fingerprints are binary bit string representations that capture diverse aspects of molecular structure and properties and are popular tools for virtual screening [19]. In this study, three fingerprints—MACCS keys [20], PubChem fingerprint [21], and Klekota-Roth fingerprint [22]—were tested to be compared with the Tanimoto coefficient and other descriptors. The descriptor numbers of the MACCS, PubChem, and Klekota-Roth fingerprints are 320, 881, and 4860, respectively. The MACCS keys of each ligand were generated by a MACCS key generator, and the other two fingerprints of the ligands were calculated by PaDEL-Descriptor [23].

EVA descriptor

An ideal molecular descriptor should encode all the features of a molecule in numerical form. EVA, based on the normal coordinate eigenvalues, is derived from calculated infrared range vibrational frequencies and consequently has the characteristic feature of molecular-specific vibration. The resulting vibrational frequencies were then convolved using a sum of Gaussian functions to generate a pseudospectrum of 3N-6 overlapping kernels:

$$EVA_x = \sum_{i=1}^{3N-6} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-f_i)^2}{2^{\sigma 2}}}$$

, where N is the number of atoms, σ is a fixed standard deviation of the Gaussian function, and f_i is the i–th frequency of the molecule. We fixed the value of σ to 5 cm^{-1} in this study.

CIMVF descriptor

For a simplified comparison of the molecular vibration patterns, the calculated molecular vibrations of a ligand were sorted in increasing order and taken into each corral with a fixed step (e.g., 5 cm^{-1}) size. The intensities of each molecular vibrational frequency in the same corral were summed up as the representative of the corral. As a molecular descriptor of a ligand, the intensity of each corral is displayed in a 1-dimensional vector containing 800 elements from the vibrational frequency range of 0–4,000 cm^{-1}. The calculations of CIMVF were performed by in-house scripts written in Python.

EVA and CIMVF descriptors were produced from PubChem SDF data of ligands through geometry optimization and calculations of molecular vibrations with the GAMESS program package [24]. The similarity matrices obtained from these five cases of molecular representations were subjected to hierarchical clustering using complete linkage method.

Calculation of molecular vibration

Since the EVA and CIMVF descriptors require geometry optimization of a given molecule, each provided theoretical 3-D conformer SDF underwent single low-energy conformation using the GAMESS program package [24]. Restricted Hartree-Fock calculations using the BLYP DFT method with the 6-31G basis set were performed to optimize the geometries of the molecules. Each result was taken as the representative conformation of the molecule, although the calculation of molecular vibrational frequency has some dependence on conformation. The results of geometry optimization were subjected to the calculation step for the molecular vibrational frequency with RUNTYP of HESSIAN in the GAMESS program.

Hierarchical clustering of molecular descriptors and dendrogram structure comparison

To test the availability of CIMVF as a molecular descriptor for the classification of AdoR ligands, a kind of molecular calculation using agglomerative hierarchical clustering was adopted in this work.

Each fingerprint or descriptor of ligands was gathered into a matrix of cognate fingerprint or descriptor. Finally, the similarity matrix, comprising descriptors of 46 ligands of AdoRs, was then subjected to hierarchical clustering in the agglomerative manner. In this study, each similarity matrix was finally clustered to make a dendrogram of 46 vertices. To compare the structures of the resulting dendrogram, we used a program for pairwise comparison of phylogenies, which shows the topological difference between two dendrograms [25]. The multiple comparison of the 5 dendrograms was also carried out using a meta-tree-generating program for comparing multiple alternative dendrograms [26].

Results and Discussion

As shown in Fig. 1, adenosine and caffeine, a typical agonist and antagonist of AdoRs, respectively, share the structure of purine. However, the Tanimoto coefficient between the two ligands is not high. The Tanimoto coefficient of some agonists to adenosine was lower than 0.3, though they did not take a large share. Moreover, the lowest Tanimoto

coefficient of the agonists was 0.12766 (capadenoson) and is a lower value as an antagonist.

The results of hierarchical clustering of the similarity matrices from the MACCS keys, PubChem, and Klekota-Roth fingerprints; EVA; and CIMVF are shown in Fig. 2.

The dendrograms obtained from MACCS keys and

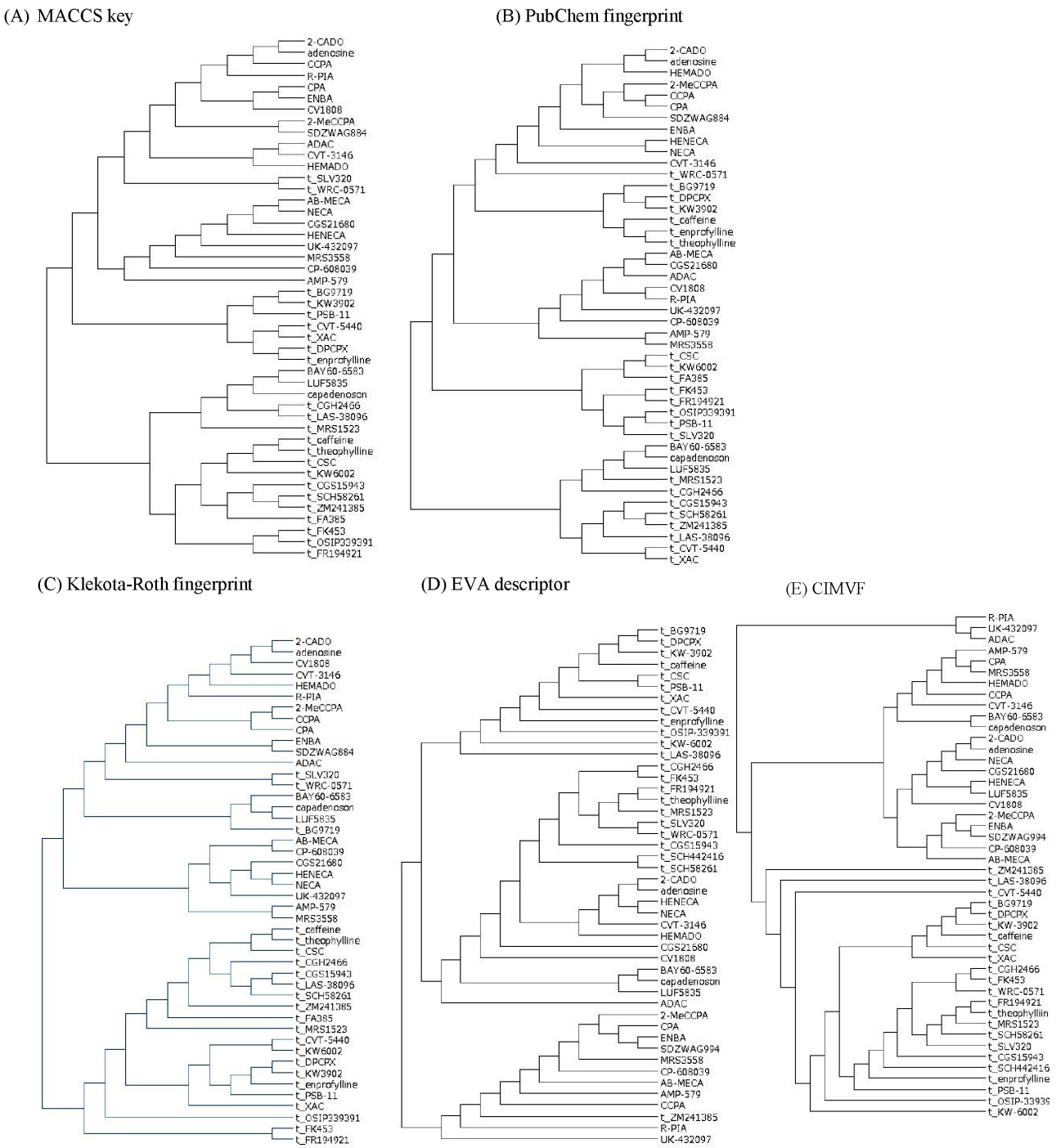

Fig. 2. Dendrograms of clustering using fingerprints, EigenVAlue (EVA), and corralled intensity of molecular vibrational frequency (CIMVF) by complete linkage method. Antagonists were tagged with "t_" to their chemical names as a prefix. (A) MACCS keys. Agonists with low Tanimoto coefficient to adenosine were clustered to a subtree of antagonists. (B) PubChem fingerprint. The clustering result of the PubChem fingerprint was similar to that of MACCS clustering. (C) Klekota-Roth fingerprint. Some antagonists were located in a subtree of agonists. (D) EVA. About half of the antagonists were located in the cluster of agonists. An antagonist, ZM241385, was clustered in the agonist subtree. (E) CIMVF. All agonists and antagonists were categorized to each cluster.

PubChem fingerprint were shown to be similar to the Tanimoto coefficient pattern. The result of the Klekota-Roth fingerprint analysis showed a somewhat different pattern to the fingerprint-based one, as only two antagonists were located in the cluster of agonists. However, the three fingerprint-based results did not match the known facts of AdoR agonism.

In the case of EVA, Takane and Mitchell [27] reported a structure-odor relationship using the EVA descriptor and hierarchical clustering. The authors reported that the dendrograms that were produced by the EVA method outperformed those from UNITY 2D in a classification of odorant molecules. However, the EVA analysis in the current study was positive but not perfect for the molecular property-activity relationship in AdoR agonism. The larger σ got, the more the clusters in the dendrogram got entangled (data not shown). It might come from the Gaussian function that smeared out the vibrational frequencies, such that vibrations of similar frequency overlapped together. The dendrogram obtained from the EVA descriptor showed an

Table 2. Topological difference between dendrograms in Fig. 2

Descriptor	CIMVF	EVA	Klekota-Roth	PubChem	MACCS
CIMVF	100	59.6	35.8	37.2	33.8
EVA	-	100	38.4	39.9	37.7
Klekota-Roth	-	-	100	48.1	55.1
PubChem	-	-	-	100	49.5
MACCS	-	-	-	-	100

CIMVF, corralled intensity of molecular vibrational frequency; EVA, EigenVAlue.

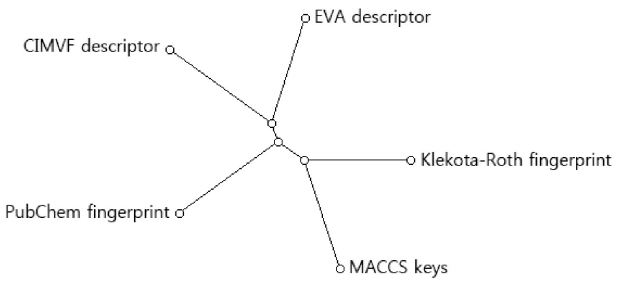

Fig. 3. Meta-tree build from multiple comparison of 5 dendrograms. CIMVF, corralled intensity of molecular vibrational frequency; EVA, EigenVAlue.

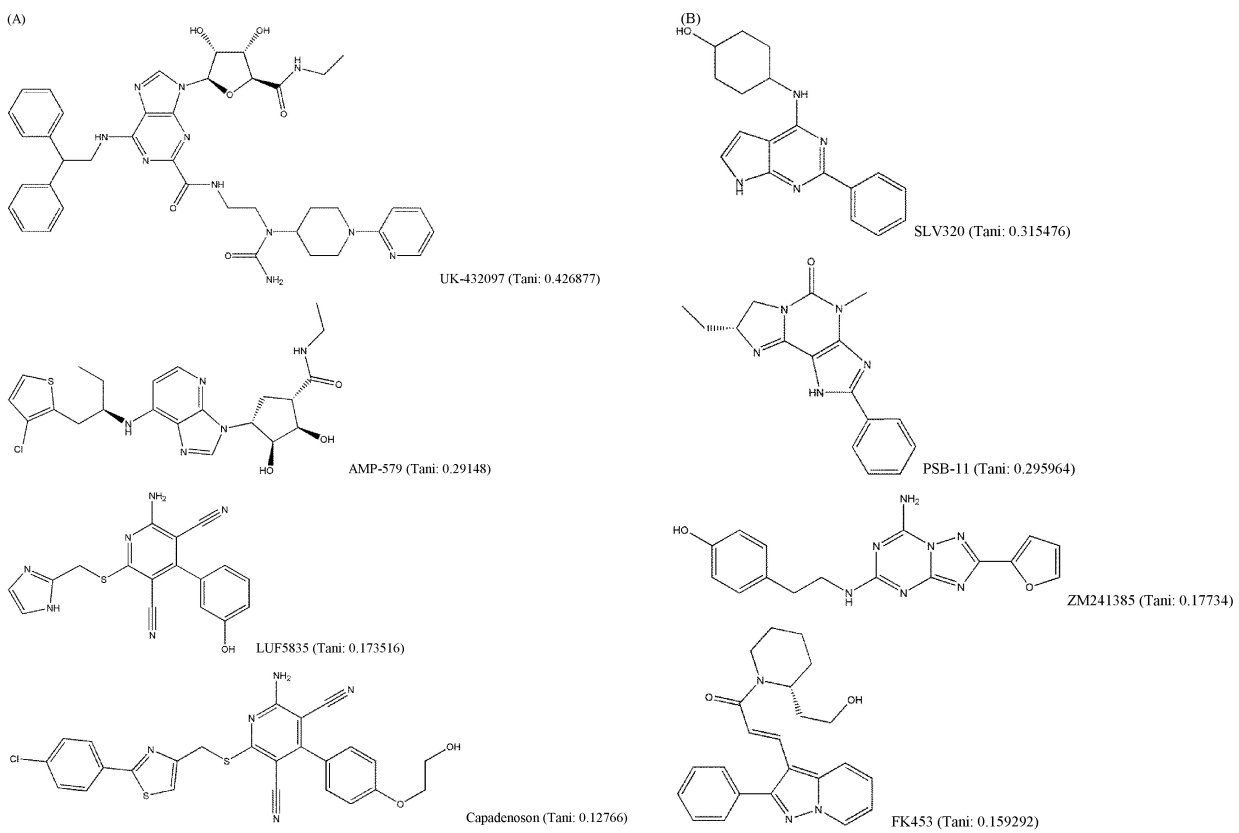

Fig. 4. Structural variety of adenosine receptors (AdoR) ligands (Tani: Tanimoto coefficient to adenosine). (A) AdoR agonists. (B) AdoR antagonists.

eligible tree for the agonism of AdoRs on the whole (Fig. 2D). One exception in which an antagonist (ZM241385) was clustered in a subtree of agonists might have resulted from smoothing by the Gaussian convolution of molecular vibration data.

All antagonists were clustered into a subtree in the dendrogram obtained from CIMVF analysis, as shown in Fig. 2E. We can tell the regional difference between agonists and antagonists in the dendrogram and also find that the information from the CIMVF analysis can play a role in the discrimination of agonists from antagonists in AdoR agonism. The discreteness of CIMVF data seems to facilitate the performance of the binary classification of AdoR ligands.

To find the topological difference between dendrograms, similarity scores between every two dendrograms were shown in Table 2. The similarity between two dendrograms was represented as a percentage. For example, the score of topological similarity between the two dendrograms of CIMVF and EVA analyses is 59.6%. The value has the highest similarity among all comparisons. A distance meta-tree build from five dendrograms of 46 AdoR ligands was illustrated in Fig. 3. As shown in the tree, the resulting dendrogram produced from the CIMVF data analysis has the nearest position to that produced from the EVA data analysis and is located the farthest from the dendrogram produced from the analysis of MACCS keys.

It is generally known that structurally similar molecules have similar properties or functions; however, a small change in molecular structure can fairly affect its vibrational frequencies. The structural variety of AdoR ligands was shown in Fig. 4. The Tanimoto coefficients of LUF5835, BAY60-6583, and capadenoson to adenosine were smaller than 0.18, and the ligands share a hydroxyphenyl pyridine dicarbonitrile backbone (Fig. 4A). They seem to be dissimilar to other agonists that share the structure of purine. Several antagonists (Fig. 4B) do not have a similar structure to caffeine; however, they act as antagonists of AdoRs.

As mentioned above, the current experiment does provide circumstantial evidence for the molecular vibrational information to AdoR agonism, at least. With a more concentrated study on the relationship between the molecular vibrational frequency and pharmacological function of a ligand, the vibrational spectrum of a ligand molecule may explicitly propose a novel path to the field of receptor-ligand interaction mechanisms.

Acknowledgments

The authors would like to express their sincere appreciation to Prof. C. H. Choi for the use of his cluster computer and his valuable advice. This work was supported by the 2011 Inje University research grant.

References

1. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, *et al*. The sequence of the human genome. *Science* 2001; 291:1304-1351.
2. Rask-Andersen M, Almén MS, Schiöth HB. Trends in the exploitation of novel drug targets. *Nat Rev Drug Discov* 2011; 10:579-590.
3. Fredholm BB, IJzerman AP, Jacobson KA, Klotz KN, Linden J. International Union of Pharmacology. XXV. Nomenclature and classification of adenosine receptors. *Pharmacol Rev* 2001;53:527-552.
4. Fredholm BB, IJzerman AP, Jacobson KA, Linden J, Müller CE. International Union of Basic and Clinical Pharmacology. LXXXI. Nomenclature and classification of adenosine receptors: an update. *Pharmacol Rev* 2011;63:1-34.
5. Jacobson KA, Gao ZG. Adenosine receptors as therapeutic targets. *Nat Rev Drug Discov* 2006;5:247-264.
6. Ko SM, Choi JW, Hwang HK, Song MG, Shin JK, Chee HK. Diagnostic performance of combined noninvasive anatomic and functional assessment with dual-source CT and adenosine-induced stress dual-energy CT for detection of significant coronary stenosis. *AJR Am J Roentgenol* 2012;198: 512-520.
7. Haffenden LJ, Yaylayan VA, Fortin J. Investigation of vibrational theory of olfaction with variously labelled benzaldehydes. *Food Chem* 2001;73:67-72.
8. Keller A, Vosshall LB. A psychophysical test of the vibration theory of olfaction. *Nat Neurosci* 2004;7:337-338.
9. Franco MI, Turin L, Mershin A, Skoulakis EM. Molecular vibration-sensing component in *Drosophila melanogaster* olfaction. *Proc Natl Acad Sci U S A* 2011;108:3797-3802.
10. Malcolm Dyson G. The scientific basis of odour. *J Soc Chem Ind* 1938;57:647-651.
11. Wright RH. Odour and chemical constitution. *Nature* 1954; 173:831.
12. Wright RH. Odor and molecular vibration: neural coding of olfactory information. *J Theor Biol* 1977;64:473-502.
13. Harris CJ, Stevens AP. Chemogenomics: structuring the drug discovery process to gene families. *Drug Discov Today* 2006; 11:880-888.
14. Scior T, Bender A, Tresadern G, Medina-Franco JL, Martínez-Mayorga K, Langer T, *et al*. Recognizing pitfalls in virtual screening: a critical review. *J Chem Inf Model* 2012;52:867-881.
15. Oh SJ. Characteristics in molecular vibrational frequency patterns between agonists and antagonists of histamine receptors. *Genomics Inform* 2012;10:128-132.
16. Ferguson AM, Heritage T, Jonathon P, Pack SE, Phillips L, Rogan J, *et al*. EVA: a new theoretically based molecular descriptor for use in QSAR/QSPR analysis. *J Comput Aided Mol Des* 1997;11:143-152.
17. Turner DB, Willett P, Ferguson AM, Heritage T. Evaluation of a novel infrared range vibration-based descriptor (EVA) for QSAR studies. 1. General application. *J Comput Aided Mol Des*

1997;11:409-422.
18. Godden JW, Xue L, Bajorath J. Combinatorial preferences affect molecular similarity/diversity calculations using binary fingerprints and Tanimoto coefficients. *J Chem Inf Comput Sci* 2000;40:163-166.
19. Bajorath J. Integration of virtual and high-throughput screening. *Nat Rev Drug Discov* 2002;1:882-894.
20. Durant JL, Leland BA, Henry DR, Nourse JG. Reoptimization of MDL keys for use in drug discovery. *J Chem Inf Comput Sci* 2002;42:1273-1280.
21. PubChem. Bethesda: National Center for Biotechnology Information, U.S. National Library of Medicine, 2013. Accessed 2013 Nov 1. Available from: http://pubchem.ncbi.nlm.nih.gov.
22. Klekota J, Roth FP. Chemical substructures that enrich for biological activity. *Bioinformatics* 2008;24:2518-2525.
23. Yap CW. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. *J Comput Chem* 2011;32:1466-1474.
24. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, *et al*. General atomic and molecular electronic structure system. *J Comput Chem* 1993;14:1347-1363.
25. Nye TM, Liò P, Gilks WR. A novel algorithm and web-based tool for comparing two alternative phylogenetic trees. *Bioinformatics* 2006;22:117-119.
26. Nye TM. Trees of trees: an approach to comparing multiple alternative phylogenies. *Syst Biol* 2008;57:785-794.
27. Takane SY, Mitchell JB. A structure-odour relationship study using EVA descriptors and hierarchical clustering. *Org Biomol Chem* 2004;2:3250-3255.