

П. Г. Полищук, В. Е. Кузьмин, А. Г. Артеменко, С. Ю. Макан,  
академик НАН Украины С. А. Андронати

## Использование метода классификационных деревьев для анализа связи структура—аффинитет лигандов 5-HT<sub>1A</sub> рецепторов

*The analysis of the structure-affinity relationship for 347 ligands of 5-HT<sub>1A</sub> receptors is made by the method of classification trees. The procedure which allows one to estimate the relative influence of structural fragments on the investigated activity in the context of this algorithm is developed. Structural fragments and their influence on the affinity to 5-HT<sub>1A</sub> receptors are determined.*

Несмотря на успехи в лечении различных психических расстройств и заболеваний, поиск и создание новых анксиолитиков и антидепрессантов по-прежнему остается важным направлением исследований медицинской химии. Одним из путей решения данной проблемы является поиск новых эффективных лигандов 5-HT<sub>1A</sub> рецепторов (5-HT<sub>1A</sub>R), участвующих в регуляции состояний тревоги, страха, депрессий и др. [1]. Для установления связи структура-аффинитет синтезировано и изучено большое число соединений — лигандов 5-HT<sub>1A</sub>R. Обобщение и анализ этих данных определяет перспективное направление в создании новых соединений.

С этой целью на основании данных экспериментальных работ [1, 2] и публикаций [3–12] была создана выборка из 347 лигандов 5-HT<sub>1A</sub>R. Структуры соединений обучающей выборки приведены на рис. 1, где наиболее широко представлены соединения, содержащие адамантильные и диазабициклические фрагменты Y, а также соединения, содержащие в качестве фрагмента Ar *o*-метоксифенил.

Поскольку в работе использовались величины аффинитета исследуемых соединений к 5-HT<sub>1A</sub>R, полученные из различных литературных источников, мы сочли более корректным применить ранговую шкалу значений аффинитета. Для чего все соединения обучающей выборки были разделены на четыре ранга (табл. 1) в соответствии со значениями аффинитета ( $pK_i = -\log K_i$ ).

Для анализа свойств, выраженных в ранговой шкале, применение регрессионных методов анализа (метода частичных наименьших квадратов, множественной линейной регрессии и т. п.) становится неэффективным. В связи с этим наиболее целесообразно использовать удобный, простой и интуитивно понятный метод классификационных деревьев [13].

Таблица 1. Принцип деления соединений на ранги активности

Ранг активности	Диапазон значений $pK_i$	Число соединений данного ранга
1	$5,0 \leq pK_i \leq 6,3$	92
2	$6,3 < pK_i \leq 7,2$	82
3	$7,2 < pK_i \leq 8,0$	89
4	$8,0 < pK_i \leq 10,1$	84

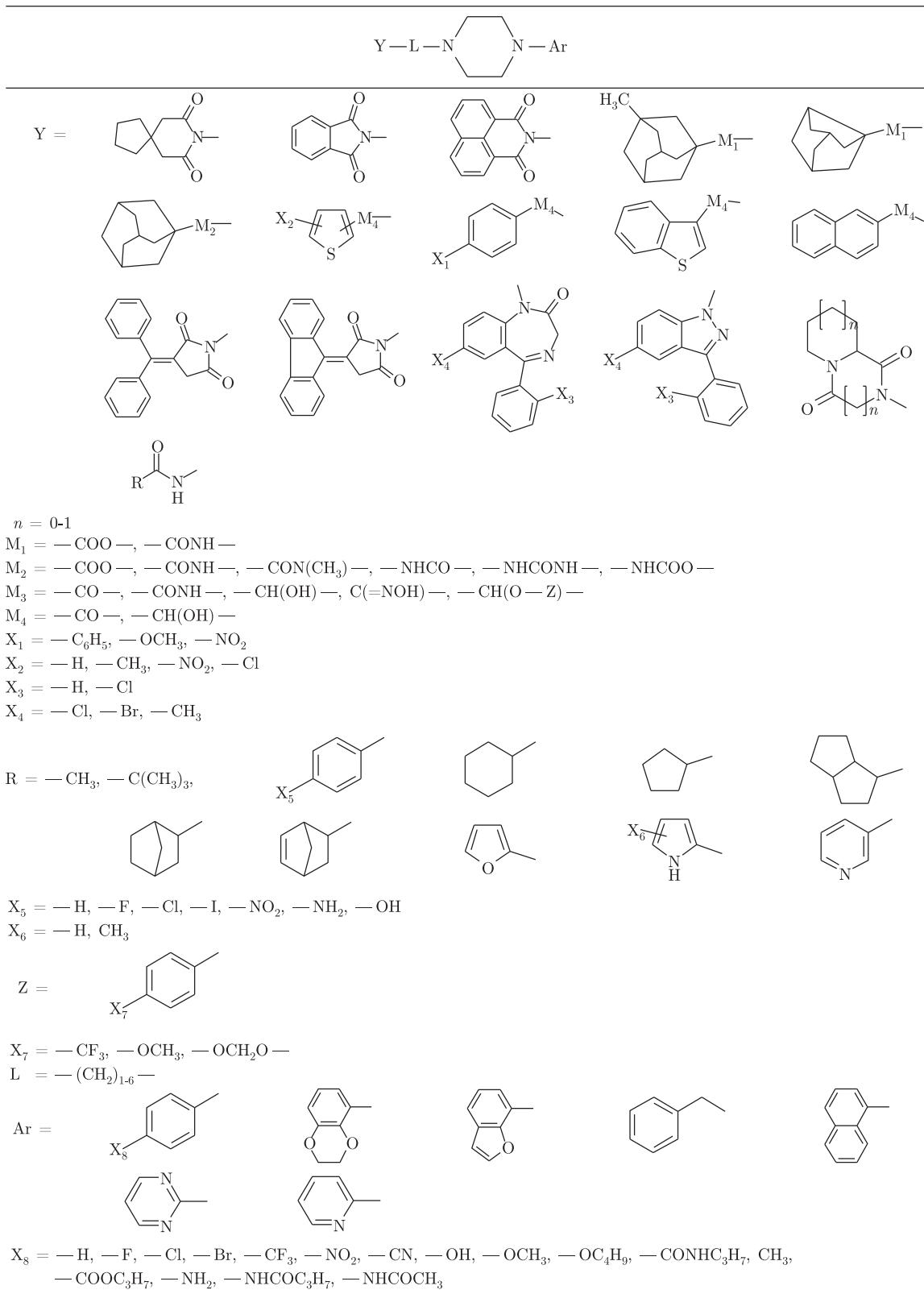


Рис. 1

Построение классификационного дерева происходит следующим образом. В корневую вершину дерева подаются все соединения обучающей выборки, которые характеризуются набором дескрипторов. Далее алгоритм делит все соединения на группы (вершины дерева) по значению одного из дескрипторов, который выбирается с учетом оценки максимального коэффициента эффективности для данного дескриптора. Такое деление соединений происходит до тех пор, пока в соответствующей вершине не останутся соединения только одного ранга аффинности. По сути, такой алгоритм — это способ представления правил классификации в иерархической последовательной структуре.

Главным недостатком моделей, получаемых этим методом, является трудность оценки относительного влияния структурных характеристик молекул. Для преодоления этого недостатка мы использовали новый подход, разработанный на основе метода тренд-вектора [14]

$$T_j = \frac{1}{m} \sum_{i=1}^m [(A_i - A_{\text{mean}})], \quad (1)$$

где  $T_j$  — относительный вклад  $j$ -го дескриптора в исследуемую активность;  $m$  — число соединений в данной вершине;  $A_i$  — ранг аффинитета  $i$ -го соединения;  $A_{\text{mean}}$  — среднее значение ранга аффинитета по всей выборке.

Используя формулу (1), в каждой вершине полученного классификационного дерева мы определяем относительный вклад каждого дескриптора в активность. При этом каждому значению вклада  $T_j$  ставится в соответствие диапазон значений дескриптора, для которого этот вклад применим.

В настоящем сообщении для описания соединений нами использовано симплексное представление молекулярной структуры [2]. Симплексы — четырехатомные молекулярные фрагменты фиксированной структуры, хиральности и симметрии. Дескриптором в данном случае служит число симплексов конкретного типа. Вершины в симплексах можно дифференцировать по типу атомов, липофильности, частичному заряду на атоме, рефракции и способности образовывать водородные связи. Для этого проводят расчет конкретной характеристики для каждого атома и затем разделяют все атомы на конечное число групп (как правило 3–7)<sup>1</sup> [2].

После генерации более 10000 симплексных дескрипторов для данной выборки нами был проведен отсев взаимокоррелирующих параметров. С помощью метода PLS<sup>2</sup> были отобраны 423 наиболее значимых дескриптора, которые были использованы для построения классификационной модели. Формирование модели осуществлялось с помощью программы SPSS AnswerTree [15] (алгоритм CART<sup>3</sup> [13]).

Итоговая модель содержала 35 вершин, из которых 18 — концевые. Для построения модели было использовано 17 симплексных дескрипторов. Модель классификационного дерева содержала 30% ошибок классификации, что считается приемлемым для такого рода задач.

Применяя к каждой вершине формулу (1), были вычислены вклады каждого задействованного в модели дескриптора и определены соответствующие диапазоны значений де-

<sup>1</sup>Параметры деления атомов на группы (римские цифры обозначают тип вершины симплекса):

для частичных зарядов — I ≤ −0,1 < II ≤ −0,05 < III ≤ −0,01 < IV ≤ 0,01 < V ≤ 0,05 < VI ≤ 0,1 < VII;

для липофильности — I ≤ −1,0 < II ≤ −0,5 < III ≤ −0,1 < IV ≤ 0,1 < V ≤ 0,5 < VI ≤ 1,0 < VII;

для рефракции — I ≤ 1,5 < II ≤ 3 < III ≤ 8 < IV;

донорно-акцепторного взаимодействия: I — донор, II — акцептор, III — индифферентный центр.

<sup>2</sup>PLS — partial least squares.

<sup>3</sup>CART — Classification and Regression Trees.

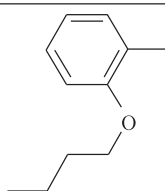
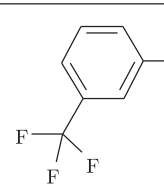
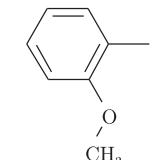
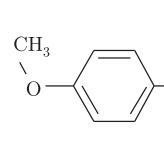
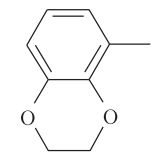
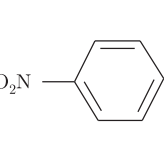
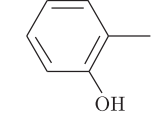
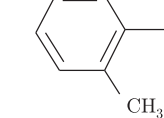
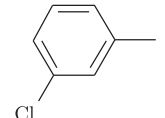
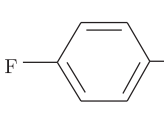
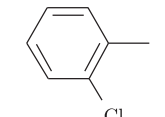
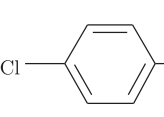
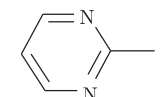
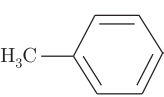
скрипторов. На основании оценок вкладов симплексных дескрипторов выделены структурные фрагменты молекул исследованных соединений, устойчиво влияющие на аффинитет к 5-НТ<sub>1A</sub>Р.

Как следует из данных, представленных в табл. 2, молекулярные фрагменты Аг, содержащие *p*-электронодонорные заместители в *o*-положении фенильного кольца (*o*-ОС<sub>4</sub>Н<sub>9</sub>, *o*-ОСН<sub>3</sub>, *o*-ОН), а также остаток 2,3-дигидробензо[1,4]диоксин-5-ила наиболее благоприятны для проявления высокого аффинитета. В свою очередь, заместители в *n*-положении оказывают отрицательное влияние на аффинитет к 5-НТ<sub>1A</sub>Р.

Установлено, что для проявления высокого аффинитета к 5-НТ<sub>1A</sub> рецепторам цепочками оптимальной длины является  $-(\text{CH}_2)_4-$  фрагмент L.

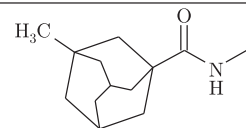
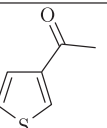
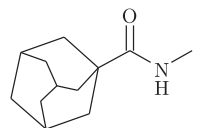
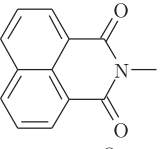
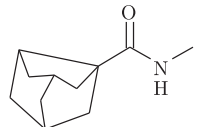
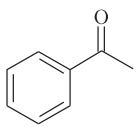
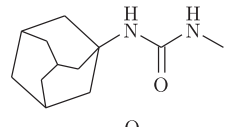
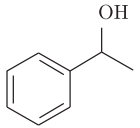
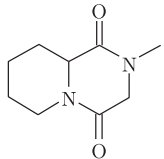
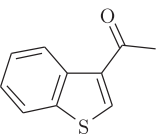
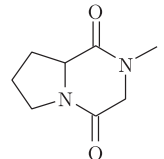
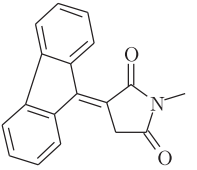
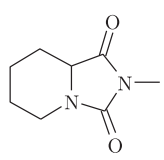
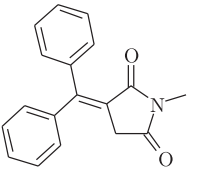
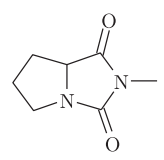
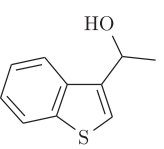
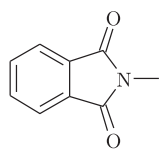
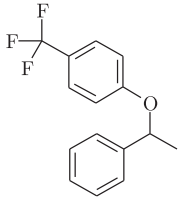
Анализ относительного влияния фрагментов Y исследованных соединений на аффинитет к 5-НТ<sub>1A</sub>Р (табл. 3) показал, что объемные насыщенные фрагменты (адамантил,

Таблица 2. Средние значения относительных вкладов фрагментов Аг исследованных соединений в аффинитет к 5-НТ<sub>1A</sub>Р

Фрагмент Аг	Вклад (T)	Фрагмент Аг	Вклад (T)
	2,8		-0,4
	1,5		-1,0
	1,2		-1,4
	0,9		-1,5
	0,7		-1,5
	0,6		-1,6
	-0,3		-1,7

3-метиладамантил и норадамантил), а также насыщенные диазабициклические фрагменты и остаток фталимида способствуют проявлению высокого аффинитета. Объемные аро-

Таблица 3. Средние значения относительных вкладов фрагментов Y исследованных соединений в аффинитет к 5-HT<sub>1A</sub>R

Фрагмент Y	Вклад (T)	Фрагмент Y	Вклад (T)
	1,6		0
	1,4		-0,2
	1,3		-0,5
	1,1		-0,7
	0,5		-1,0
	0,4		-1,4
	0,4		-1,6
	0,3		-2,1
	0,1		-2,5

матические и гетероароматические фрагменты, напротив, можно отнести к фрагментам, отрицательно влияющим на аффинитет к 5-HT<sub>1A</sub>R.

Нам представлялось интересным рассмотреть влияние молекулярного окружения на относительный вклад фталимида в аффинитет к 5-HT<sub>1A</sub>R. В случае, когда Y = фталимид и Ar = фенил, влияние цепочки на вклад фталимидного фрагмента было следующим: для L = -(CH<sub>2</sub>)<sub>2</sub>- это -0,162, для -(CH<sub>2</sub>)<sub>3</sub>- и -(CH<sub>2</sub>)<sub>4</sub>- это 0,067 и 0,075 соответственно. Если взять случай, когда Y = фталимид и L = -(CH<sub>2</sub>)<sub>4</sub>-, то при Ar = 2,5-диметоксифенил и 2-метоксифенил относительные вклады фталимида в аффинитет к 5-HT<sub>1A</sub>R составили 0,305 и 0,280 соответственно, что можно считать вполне высокими значениями, а при Ar = 1-нафтил вклад составил -0,111. Следовательно, вклад фрагмента не является постоянным, а зависит от его молекулярного окружения. Низкое среднее значение относительного вклада фталимида можно объяснить малой представительностью в обучающей выборке высокоаффинных соединений, содержащих этот фрагмент.

Нам представляется, что метод классификационных деревьев, предусматривающий использование данных по аффинитету для разнородных групп соединений, полученных из различных литературных источников, может быть эффективным инструментом предварительного скрининга биологически активных веществ.

Таким образом, полученные результаты в рамках описанного алгоритма предлагаем использовать как для предварительного скрининга, так и для дизайна новых высокоаффинных лигандов 5-HT<sub>1A</sub> рецепторов.

1. *Andronati S. A., Makan S. Yu.* Nitrogen-Containing Heterocyclic Compounds as Ligands for Serotonin Receptors // *The Chemistry and Biological Activity of Nitrogen-Containing Heterocyclic and Alkaloids* / Ed. V. G. Kartsev, G. A. Tolstikov. – Moscow: Iridium Press, 2001. – 1. – P. 33–50.
2. *Kuz'min V. E., Artemenko A. G., Polischuk P. G. et al.* Hierarchic system of QSAR models (1D – 4D) on the base of simplex representation of molecular structure // *J. Mol. Model.* – 2005. – 11. – P. 457–467.
3. *Richard A. Glennon, Noreen A. Naiman, Robert A. Lyon, Milt Titelert.* Arylpiperazine Derivatives as High-Affinity 5-HT<sub>1A</sub> Serotonin Ligands // *J. Med. Chem.* – 1988. – 31. – P. 1968. – 1971.
4. *Revathi K. Raghupathi, Laura Rydelek-Fitzgerald, Milt Teitler, Richard A. Glennon* Analogues of the 5-HT<sub>1A</sub> Serotonin Antagonist 1-(2-Methoxyphenyl)-4-[4-(2-phthalimido)butyl]piperazine with Reduced al-Adrenergic Affinity // *Ibid.* – 1991. – 34. – P. 2633–2638.
5. *Bart J. van Steen, Ineke van Wijngaarden, Martin Th. M. Tulp Willem Soudijn* Structure-Affinity Relationship Studies on 5-HT<sub>1A</sub> Receptor Ligands. 2. Heterobicyclic Phenylpiperazines with No 4. – Aralkyl Substituents // *Ibid.* – 1994. – 37. – P. 2761–2773.
6. *Aurelio Orjales, Luisa Alonso-Cires, Luis Labeaga, Reyes Corcostegui.* New (2-Methoxyphenyl)piperazine Derivatives as 5-HT<sub>1A</sub> Receptor Ligands with Reduced α<sub>1</sub> – Adrenergic Activity. Synthesis and Structure-Affinity Relationships // *Ibid.* – 1995. – 38. – P. 1273–1277.
7. *Bart J. van Steen, Ineke van Wijngaarden, Martin Th. M. Tulp, Willen Soudijn.* A Series of No 4. – Imidoethyl Derivatives of 1-(2,3-Dihydro-1,4-benzodioxin-5-yl)piperazine as 5-HT<sub>1A</sub> Receptor Ligands: Synthesis and Structure-Affinity Relationships // *Ibid.* – 1995. – 38. – P. 4303–4308.
8. *Maria L. Lopez-Rodriguez, M Luisa Rosado, Bellinda Benhamu et al.* Synthesis and Structure-Activity Relationships of a New Model of Arylpiperazines. 1.2-[[4-(o-Methoxyphenyl)piperazin-1-yl]methyl]-1,3-dioxoperhydroimidazo[1,5-a]pyridine: A Selective 5-HT<sub>1A</sub> Receptor Agonist // *Ibid.* – 1996. – 39. – P. 4439–4450.
9. *Maria L. Lopez-Rodriguez, M. Jose Morcillo, Tandu K. Rovat et al.* Synthesis and Structure Activity Relationships of a New Model of Arylpiperazines. 4.1-[ω-(4-Arylpiperazin-1-yl)alkyl]-3-(diphenylmethylene)-2,5-pyrrolidinediones and -3-(9H-fluoren-9-ylidene)-2,5-pyrrolidinediones: Study of the Steric Requirements of the Terminal Amide Fragment on 5-HT<sub>1A</sub> Affinity/Selectivity // *Ibid.* – 1999. – 42. – P. 36–49.
10. *Magid A. Abou-Gharbia, Wayne E. Childers, Jr., Horace Fletcher et al.* Synthesis and SAR of Adatanserin: Novel Adamantyl Aryl- and Heteroarylpiperazines with Dual Serotonin 5-HT<sub>1A</sub> and 5-HT<sub>2</sub> Activity as Potential Anxiolytic and Antidepressant Agents // *Ibid.* – 1999. – 42. – P. 5077–5094.

11. *Javier Martinez-Esparza, Ana-M. Oficialdegui, Silvia Perez-Silanes et al.* New 1-Aryl-3-(4-arylpiperazin-1-yl)propane Derivatives, with Dual Action at 5-HT<sub>1A</sub> Serotonin Receptors and Serotonin Transporter, as a New Class of Antidepressants // *Ibid.* – 2001. – **44**. – P. 418–428.
12. *Maria L. Lopez-Rodriguez, M. Jose Morcillo, Esther Fernandez et al.* Synthesis and Structure-Activity Relationships of a New Model of Arylpiperazines. 6. Study of the 5-HT<sub>1A</sub>/α<sub>1</sub> – Adrenergic Receptor Affinity by Classical Hansch Analysis, Artificial Neural Networks, and Computational Simulation of Ligand Recognition // *Ibid.* – 2001. – **44**. – P. 198–207.
13. *Breiman L., Friedman J. H., Olshen R. A., Stone C. T.* Classification and Regression Trees. – Belmont, CA: Wadsworth, 1984. – 368 p.
14. *Carhart R. E., Smith D. H., Venkataraghavan R.* Atom pairs as molecular features in structure – activity studies. Definition and application // *J. Chem. Inf. Comput.* – 1985. – **25**, No 2. – P. 64.
15. *Trial* версия программы с <http://www.spss.com>.

Фізико-хімічний інститут  
ім. А. В. Богатського НАН України, Одеса

Поступило в редакцію 23.07.2007