DECISION TREE FOR MECHANISM
OF ANTITUMOR DRUGS ACTION PREDICTION

N.N. Palko, *inn_74@list.ru*
V.A. Potemkin, *potemkinva@susu.ru*
M.A. Grishina, *marialgri29@gmail.com*
South Ural State University, Chelyabinsk, Russian Federation

Classification by the mechanism of action of antitumor drugs using quantitative descriptors has been carried out in this study. The dataset of drugs includes 115 compounds with known activity and mechanism of action. The structures have been taken from the National Cancer Institute database. The dataset includes: 30 structures with the mechanism of alkylation action, 23 compounds are topoisomerase I inhibitors, 16 structures are topoisomerase II inhibitors, 17 compounds are DNA/RNA antimetabolites (dihydrofolate reductase inhibitors), 16 molecules are DNA antimetabolites, and 13 compounds are antimitotic drugs. A decision tree has been constructed for determination of each class of the compounds using 3D descriptors that have been computed by MERA software. Analysis of the results has shown that each mechanism of drug action is characterized by a set of descriptors. Important quantum-chemical descriptors of structures with mechanism for alkylation action, DNA/RNA antimetabolites, and DNA antimetabolites have been determined. Quantum-chemical and geometric descriptors for structures of topoisomerase II inhibitors and geometric descriptors for structures of topoisomerase I inhibitors have been established. An important energy descriptor for antimitotic drugs has been determined. The typical descriptor values for active and inactive structures for each mechanism of action have been determined. The quality recognition of active structures for each mechanism of drugs action has been determined. The highest recognition quality value of active and inactive compounds is observed in the decision tree of topoisomerase II inhibitors. The minimal recognition quality value of active and inactive compounds is observed in the decision tree of DNA/RNA-antimetabolites. The suggested decision trees can be used for determination of the action mechanism of antitumor drugs.
Keywords: antitumor drugs, mechanism of action, structural 3D descriptors, decision tree.

Introduction

Creation of antitumor agents is an actual problem of modern science. The development of antitumor drugs with a minimal side effect and maximal activity is of a particular interest. The development of new drugs, the synthesis and determination of their activity does not always lead to expected positive result, since requires a large amount of resources. The usage of virtual screening techniques and computer modeling can minimize a time and cost for the development of new drugs. The creation of decision tree [1–7] is one of virtual screening method. The method allows determining rules for classification of compounds. The determined rules and descriptors can be further used for virtual design and screening of molecules with desirable properties [8–10].

Research Methodology

In this paper, a theoretical study of the relationship between structural descriptors and mechanism of action of antitumor drugs has been carried out. The dataset of drugs includes 115 compounds with known activity and mechanism of action. The structures have been taken from the National Cancer Institute database [11]. The dataset includes:

– 30 structures acting by alkylation mechanism;
– 23 compounds are topoisomerase I inhibitors;
– 16 structures are topoisomerase II inhibitors;
– 17 compounds are DNA/RNA antimetabolites (dihydrofolate reductase inhibitor);
– 16 molecules are DNA antimetabolites;
– 13 compounds are antimitotic agents.

All the studied structures possess an antitumor activity against some cell lines elucidated at National Cancer Institute.

The energy minimization along with conformational analysis and the calculation of energies,

18

Bulletin of the South Ural State University. Ser. Chemistry.
2019, vol. 11, no. 1, pp. 18–24

charges, geometric characteristics and other descriptors for drugs have been carried out by using the MERA model for all molecules. More than 500 geometric and quantum-chemical descriptors have been calculated. Calculations have been performed using the online laboratory Chemosophia [12–20]. The decision trees creation has been performed for all studied mechanisms and molecules and the descriptors determining the mechanism of drugs action have been found.

Discussion

Analysis of the results showed that each mechanism of drugs action was characterized by their own set of descriptors. Analysis of the decision tree for antitumor drugs possessing the alkylating mechanism of action showed, that drugs were characterized by the portion of the overlapped surface of carbon atoms in molecules. The recognition quality of for this decision tree is 76 %.

The highest descriptor values (0.48 ÷ 0.54) are typical for antitumor drugs with alkylation action, while inactive structures have lower descriptor values (0.45 ÷ 0.48) (Fig. 1).

Carbon-including substitution increasing the strengthening tension and increasing the values of the descriptor that in turn leads to strengthening the carbon as an alkylating agent. The electronegative atoms in structure provide an interaction between a target and a drug, necessary for alkylation.
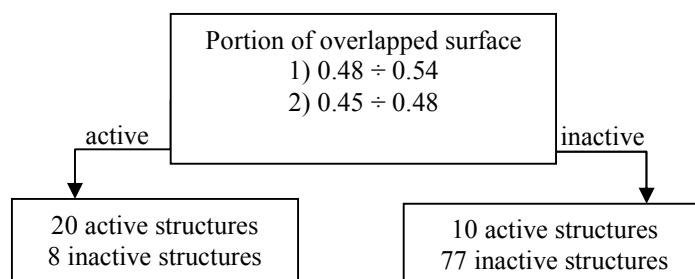

Fig. 1. Decision tree for anticancer drugs of alkylating action

Analysis of decision tree for topoisomerase I inhibitors showed, that the important characteristic feature for these drugs is the standard deviation of all coordinates of atoms respectively to the center of mass. The highest descriptor values (5.72 ÷ 13.30) are characteristic features of topoisomerase I inhibitors, inactive structures have lower values (0.45 ÷ 6.77). The recognition quality of the decision tree is 71 % (Fig. 2).

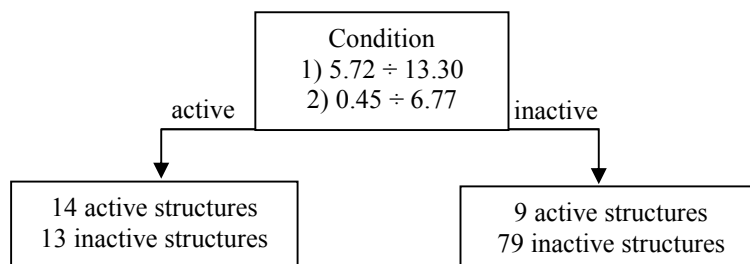**Fig. 2. Decision tree for anticancer drugs of the topoisomerase I inhibitors**
(Condition – the standard deviation of all coordinates of atoms respectively to the center of mass)

Research of the decision tree for topoisomerase II inhibitors showed, that the drugs are characterized by number of electrons in overlaps of oxygen atoms. The highest descriptor values (1.17 ÷ 1.81) are typical for the inhibitors of topoisomerase II, while inactive structures have lower values (0.60 ÷ 1.12).

Also, the important characteristic is the dimensions of topoisomerase II inhibitors, namely the maximal distance between the center of mass and the most remote atom along the second rotational invariant. The active structures should have the distance, which lies in the range 6.91 ÷ 8.97, while the inactive structures have lower values of the descriptor ranging from 4.53 ÷ 6.95. The recognition quality for the decision tree is 80 % (Fig. 3).
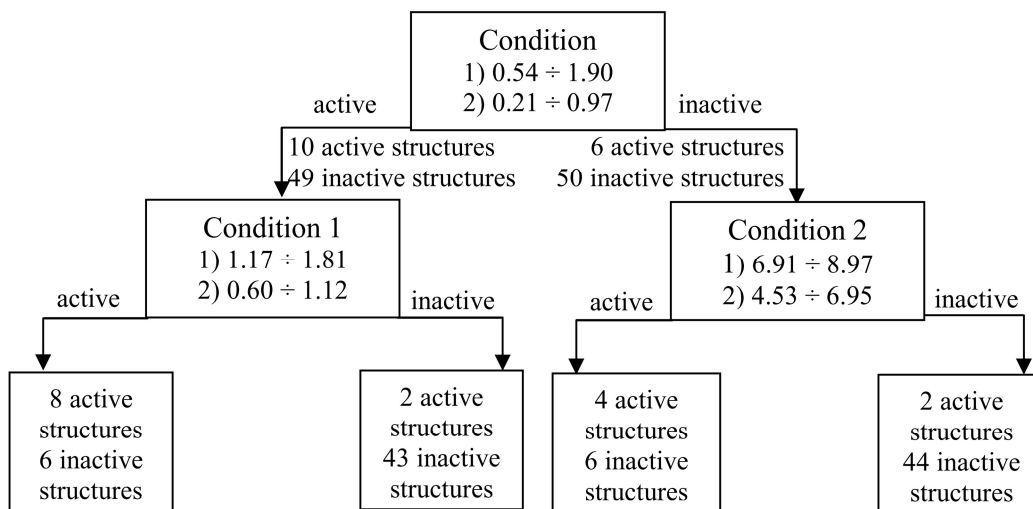
Fig. 3. Decision tree for antitumor topoisomerase II inhibitors
(Condition, Condition 1 – the number of electrons in overlaps of oxygen atoms,
Condition 2 – the maximal distance between the center of mass
and the most remote atom along the second rotational invariant)

The dataset of the DNA/RNA-antimetabolites includes methotrexate derivatives and the pyrimidine antagonists, inhibitors of a purine synthesis and nucleotides. Analysis of the results showed, that the action was determined by the portion of the overlapped surface and by the number of electrons of nitrogen atoms in the antitumor drugs molecules. The active structures should have highest values of the latter descriptor ($0.90 \div 3.11$), while the inactive structures have lower values of the descriptor ($0.27 \div 1.67$). The recognition quality of active structures is 62 % (Fig. 4).
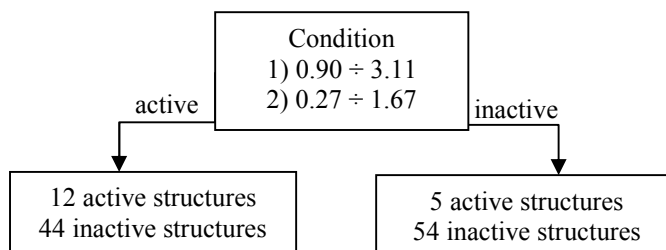

Fig. 4. Decision tree for antitumor DNA/RNA-antimetabolites
(Condition – portion of the overlapped surface
and by the number of electrons of nitrogen atoms in the antitumor drugs)

The important descriptor for these structures are a portion of the non-overlapped outer-shell electrons and a portion of non-overlapped volumes of sulfur atoms. The active structures should have the non-overlapped outer-shell electrons in the range $0.22 \div 0.30$, while for inactive structures the descriptor have lower values ($0.20 \div 0.22$). The values of non-overlapped volumes of sulfur atoms for active structures ranging from 0.007 to 0.087 and for inactive structures, it is in the interval $0.0025 \div 0.0055$. The recognition quality is 76 % (Fig. 5).

Increase of value of non-overlapped electrons and of non-overlapped volume of sulfur atoms leads to increase of nucleophilic properties of the molecules that defines selective effect of drugs. At the same time, it provides the formation of intermolecular interactions in the "receptor – ligand" complexes and the further reaction between a drug and the DNA of tumor cells.
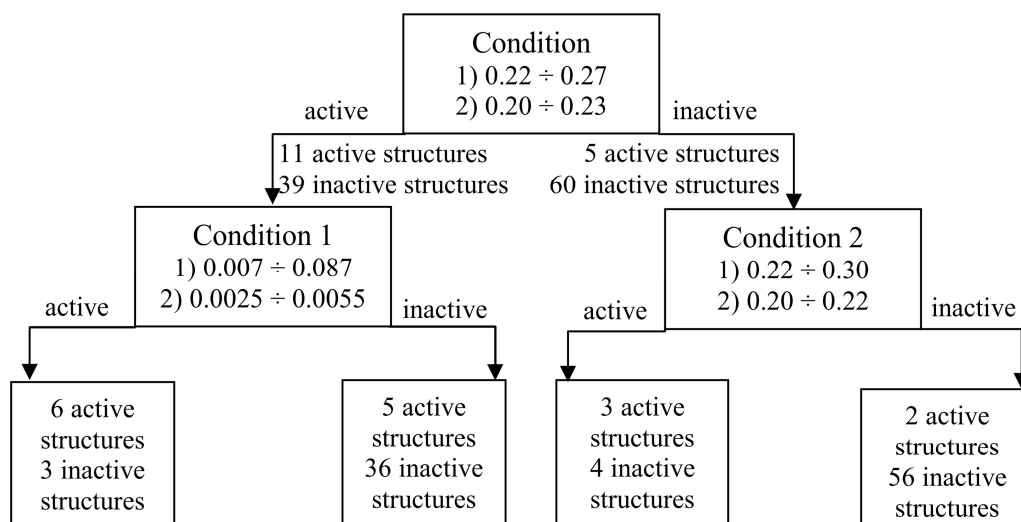
20

Bulletin of the South Ural State University. Ser. Chemistry.
2018, vol. 11, no. 1, pp. 18–24

Fig. 5. Decision tree for anticancer drugs of the DNA-antimetabolites
(Condition, Condition 2 – portion of the non-overlapped outer-shell electrons,
Condition 1 – portion of non-overlapped volumes of sulfur atoms)

Increase of value of non-overlapped electrons and of non-overlapped volume of sulfur atoms leads to increase of nucleophilic properties of the molecules that defines selective effect of drugs. At the same time, it provides the formation of intermolecular interactions in the "receptor – ligand" complexes and the further reaction between a drug and the DNA of tumor cells.

Analysis of results for the antimitotic drugs showed, that a major factor determining the mechanism of action is the energy of intramolecular van der Waals interactions. The active structures should have the lower descriptor values (–279.26 ÷ –87.72), while inactive structures possess highest values of the descriptor (–97.16 ÷ 25.38). The recognition quality for the decision tree is 69 % (Fig. 6). Active structures should have large number of substituents that can form intramolecular interactions, leading to stabilization of the system and decrease its energy.
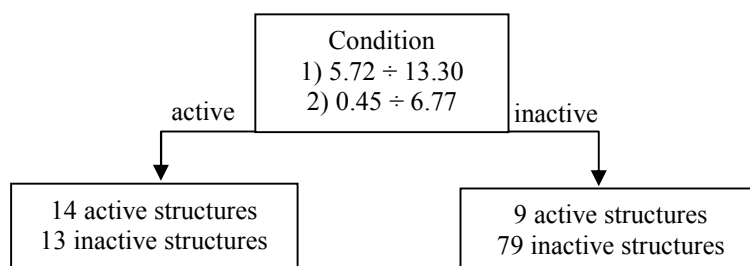

Fig. 6. Decision tree for antimitotic drugs
(Condition – energy of intramolecular van der Waals interactions)

Conclusion

Thus, the decision trees creation has been performed for antitumor drugs possessing the alkylating mechanism of action, topoisomerase I inhibitors, topoisomerase II inhibitors, DNA/RNA-antimetabolites, DNA-antimetabolites and antimitotic drugs. A set of descriptors for each mechanism of drugs action have been determined. Descriptor values for active and inactive structures have been established. The suggested decision trees can be used to for determination of mechanism of antitumor agents action.

References

1. Plewczynski D., Spieser S.A., Koch U. Assessing Different Classification Methods for Virtual Screening. *Journal of Chemical Information and Modeling*, 2006, vol. 46, no. 3, pp. 1098–1106. DOI: 10.1021/ci050519k.

2. Sebban M., Nock R., Chauchat J.H., Rakotomalala R. Impact of Learning Set Quality and Size on Decision Tree Performances. *International Journal of Computer Science and Security*, 2000, vol. 1, no. 1, pp. 85–105. DOI: 10.1.1.62.6365.

3. Quinlan J.R. Induction of Decision Trees. *Machine Learning*, 1986, vol. 1, no. 1, pp. 81–106. DOI: 10.1007/BF00116251.

4. Cruciani G., Mannhold R., Kubinyi H., Folkers G. *Molecular interaction fields: applications in drug discovery and ADME prediction.* Weinheim, Wiley WILEY-VCH Verlang GmbH & Co. KGaA, 2005. 328 p.

5. Breiman L., Friedman J.H., Olshen R.A., Stone C.J. *Classification and regression trees.* Boca Raton, FL, Chapman & Hall/CRC Taylor & Francis Grou, 2017. 368 p.

6. Melville J.L., Burke E.K., Hirst J.D. Machine Learning in Virtual Screening. *Comb. Chem. High Throughput Screen,* 2009, vol. 12, no. 4, pp. 332–343. DOI: 10.2174/138620709788167980.

7. Stouch T.R., Kenyon J.R., Johnson S.R., Chen X.Q., Doweyko A., Li Y. In silico ADME/Tox: why models fail. *J. Comput. Aided Mol. Des.*, 2003, vol. 17, no. 2–4, pp. 83–92. DOI: 10.1023/A:1025358319677.

8. Blower P.E., Kevin P., Cross K.P. Decision Tree Methods in Pharmaceutical Research. *Current Topics in Medicinal Chemistry,* 2006, vol. 6, no. 1, pp. 31–39. DOI: 10.2174/156802606775193301.

9. Grishina M.A., Potemkin V.A., Pogrebnoi A.A., Ivshina N.N. A Study of Conformational States of Substrates of Isoform 3a4 of Cytochrome P450. *Biophysics*, 2008, vol. 53, no. 5, pp. 355–360. DOI: 10.1134/S0006350908050060.

10. Afon'kina E.S., Toreeva N.A., Pal'Ko N.N., Potemkin V.A., Grishina M.A., Matveev G.A. Effect of the Structural Characteristics of Dihydrofolate Reductase Inhibitors on their Metabolic Properties. *Journal of Structural Chemistry*, 2012, vol. 53, no. 2, pp. 365–372. DOI: 10.1134/S0022476612020230.

11. National Cancer Institute. Available at: https://dtp.cancer.gov/databases tools/bulk data.htm (accessed 5 September 2017).

12. Potemkin V.A., Bartashevich E.V., Belik A.V. A Model for Calculating the Atomic Volumetric Characteristics in Molecular Systems. *Russian Journal of Physical Chemistry A,* 1998, vol. 72, no. 4, pp. 561–566.

13. Chemosophia s.r.o. Available at: http://www.chemosophia.com (accessed 5 September 2017).

14. Potemkin V.A., Pogrebnoy A.A., Grishina M.A. Technique for Energy Decomposition in the Study of "Receptor-Ligand" Complexes. *Journal of Chemical Information and Modeling,* 2009, vol. 49, no. 6, pp. 1389–1406. DOI: 10.1021/ci800405n.

15. Bartashevich E.V., Potemkin V.A., Grishina M.A., Belik A.V. A Method for Multiconformational Modeling of the Three Dimensional Shape of a Molecule. *Journal of Structural Chemistry,* 2002, vol. 43, no. 6, pp. 1033–1039. DOI: 10.1023/A:1023611131068.

16. Grishina M.A., Bartashevich E.V., Potemkin V.A., Belik A.V. Genetic Algorithm for Predicting Structures and Properties of Molecular Aggregates in Organic Substances. *Journal of Structural Chemistry,* 2002, vol. 43, no. 6, pp. 1040–1044. DOI: 10.1023/A:1023663115138.

17. Potemkin V.A., Arslambekov R.M., Bartashevich E.V., Grishina M.A., Belik A.V., Perspicace S. and Guccione S. Multiconformational Method for Analyzing the Biological Activity of Molecular Structures. *Journal of Structural Chemistry,* 2002, vol. 43, no. 6, pp. 1045–1049. DOI: 10.1023/a:1023615231976.

18. Potemkin V.A., Grishina M.A. A new paradigm for pattern recognition of drugs. *J. Comput. Aided Mol. Des.*, 2008, vol. 22, pp. 489–505. DOI: 10.1007/s10822-008-9203-x.

19. Potemkin V., Grishina M. Principles for 3D/4D QSAR classification of drugs. *Drug Discovery Today,* 2008, vol. 13, no. 21/22, pp. 952–959. DOI: 10.1016/j.drudis.2008.07.006.

20. Potemkin V., Grishina M. Electron-based descriptors in the study of physicochemical properties of compounds. *Computational and Theoretical Chemistry,* 2018, vol. 1123, pp. 1–10. DOI: org/10.1016/j.comptc.2017.11.010.

22

Bulletin of the South Ural State University. Ser. Chemistry.
2018, vol. 11, no. 1, pp. 18–24

КЛАССИФИЦИРУЮЩЕЕ ПРАВИЛО ДЛЯ ПРОГНОЗА МЕХАНИЗМА ДЕЙСТВИЯ ПРОТИВООПУХОЛЕВЫХ СРЕДСТВ

Н.Н. Палько, В.А Потёмкин, М.А. Гришина
Южно-Уральский государственный университет, г. Челябинск, Россия

Проведена классификация по механизму действия противоопухолевых соединений с использованием количественных дескрипторов. Рассмотрено 115 соединений с известной активностью и механизмом действия: 30 структур алкилирующего действия, 23 соединения – ингибиторы топоизомеразы I, 16 структур – ингибиторы топоизомеразы II, 17 соединений – антиметаболиты ДНК/РНК (ингибиторы дигидрофолатредуктазы), 16 молекул – антиметаболиты ДНК, 13 соединений – антимитотики. Структуры были взяты из базы данных Национального Института Рака (США). С использованием дескрипторов, рассчитанных в рамках модели MERA, построено классифицирующее правило для прогноза механизма действия противоопухолевых лекарственных средств. Определены важные квантово-химические дескрипторы для структур алкилирующего действия, антиметаболитов ДНК/РНК и антиметаболитов ДНК. Установлены квантово-химические и геометрические дескрипторы для ингибиторов топоизомеразы II и геометрические дескрипторы для ингибиторов топоизомеразы I. Определен важный энергетический дескриптор для антимитотиков. Установлены значения дескрипторов для активных и неактивных структур каждого механизма действия. Определено качество распознавания классифицирующего правила для всех рассматриваемых механизмов действия лекарственных средств. Максимальное значение качества распознавания активных и неактивных соединений наблюдаются у ингибиторов топоизомеразы II. Минимальное значение качества распознавания активных и неактивных соединений наблюдаются у антиметаболитов ДНК/РНК. Предлагаемые классифицирующие правила могут использоваться для определения механизма действия противоопухолевых соединений.

Ключевые слова: противоопухолевые средства, механизм действия, структурные 3D дескрипторы, классифицирующее правило.

Литература

1. Plewczynski, D. Assessing Different Classification Methods for Virtual Screening / D. Plewczynski, S.A. Spieser, U. Koch // *Journal of Chemical Information and Modeling*. – 2006. – V. 46, № 3. – P. 1098–1106. DOI: 10.1021/ci050519k.

2. Impact of Learning Set Quality and Size on Decision Tree Performances / M. Sebban, R. Nock, J.H. Chauchat, R. Rakotomalala // *International Journal of Computer Science and Security*. – 2000. – V. 1, № 1. – P. 85–105. DOI: 10.1.1.62.6365.

3. Quinlan, J.R. Induction of Decision Trees / J.R. Quinlan // *Machine Learning*. – 1986. – V. 1, № 1. – P. 81–106. DOI: 10.1007/BF00116251.

4. *Molecular interaction fields: applications in drug discovery and ADME prediction* / G. Cruciani, R. Mannhold, H. Kubinyi, G. Folkers. – Weinheim: Wiley WILEY-VCH Verlang GmbH & Co. KGaA, 2005. – 328 p.

5. *Classification and regression trees* / L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone. – Boca Raton, FL: Chapman & Hall/CRC Taylor & Francis Group, 2017. – 368 p.

6. Melville, J.L. Machine Learning in Virtual Screening / J.L. Melville, E.K. Burke, J.D. Hirst // *Comb. Chem. High Throughput Screen.* – 2009. – V. 12, № 4. – P. 332–343. DOI: 10.2174/138620709788167980.

7. In silico ADME/Tox: why models fail / T.R. Stouch, J.R. Kenyon, S.R. Johnson et al. // *J. Comput. Aided Mol. Des.* – 2003. – V. 17, № 2–4. – P. 83–92. DOI: 10.1023/A:1025358319677.

8. Blower, P.E. Decision Tree Methods in Pharmaceutical Research / P.E. Blower, P. Kevin, K.P. Cross // *Current Topics in Medicinal Chemistry.* – 2006. – V. 6, № 1. – P. 31–39. DOI: 10.2174/156802606775193301.

9. Исследование конформационных состояний субстратов изоформы 3а4 цитохрома р450 / М.А. Гришина, В.А. Потемкин, А.А. Погребной, Н.Н. Ившина // Биофизика. – 2008. – Т. 53, № 5. – С. 758–765.

10. Влияние структурных характеристик ингибиторов дигидрофолатредуктазы на их метаболические свойства / Е.С. Афонькина, Н.Н. Палько, Г.А. Матвеев и др. // Журнал структурной химии. – 2012. – Т. 53, № 2. – С. 364–371.

11. National Cancer Institute. – https://dtp.cancer.gov/databases tools/bulk data.htm.

12. Потемкин, В.А. Модель расчета атомных объемных характеристик в молекулярных системах / В.А. Потемкин, Е.В. Барташевич, А.В. Белик // Журнал физической химии. – 1998. – Т. 72, № 4. – С. 650–656.

13. Chemosophia s.r.o. – http://www.chemosophia.com.

14. Potemkin, V.A. Technique for Energy Decomposition in the Study of "Receptor-Ligand" Complexes / V.A. Potemkin, A.A. Pogrebnoy, M.A. Grishina // *Journal of Chemical Information and Modeling.* – 2009. – V. 49, № 6. – P. 1389–1406. DOI: 10.1021/ci800405n.

15. Метод мультиконформационного моделирования пространственной формы молекулы / Е.В. Барташевич, В.А. Потемкин, М.А. Гришина, А.В. Белик // Журнал структурной химии. – 2002. – Т. 43, № 6. – С. 1112–1119.

16. Генетический алгоритм для прогноза строения и свойств молекулярных агломератов в органических веществах / М.А. Гришина, Е.В. Барташевич, В.А. Потемкин, А.В. Белик // Журнал структурной химии. – 2002. – Т. 43, № 6. – С. 1120–1125.

17. Мультиконформационный метод анализа биологической активности молекулярных структур / В.А. Потемкин, Р.М. Арсламбеков, Е.В. Барташевич и др. // Журнал структурной химии. – 2002. – Т. 43, № 6. – С. 1126–1130.

18. Potemkin, V.A. A new paradigm for pattern recognition of drugs / V.A. Potemkin, M.A. Grishina // *J. Comput. Aided Mol. Des*. – 2008. – V. 22. – P. 489–505. DOI: 10.1007/s10822-008-9203-x.

19. Potemkin, V. Principles for 3D/4D QSAR classification of drugs / V. Potemkin, M. Grishina // *Drug Discovery Today*. – 2008. – V. 13, № 21/22. – P. 952–959. DOI: 10.1016/j.drudis.2008.07.006.

20. Potemkin, V. Electron-based descriptors in the study of physicochemical properties of compounds / V. Potemkin, M. Grishina // *Computational and Theoretical Chemistry*. – 2018. – V. 1123. – P. 1–10. DOI: org/10.1016/j.comptc.2017.11.010.

Палько Надежда Николаевна – младший научный сотрудник Научно-исследовательской лаборатории компьютерного моделирования лекарственных средств Высшей медико-биологической школы, Южно-Уральский государственный университет. 454080, г. Челябинск, проспект Ленина, 76. E-mail: inn_74@list.ru.

Потёмкин Владимир Александрович – кандидат химических наук, начальник Научно-исследовательской лаборатории компьютерного моделирования лекарственных средств Высшей медико-биологической школы, Южно-Уральский государственный университет. 454080, г. Челябинск, проспект Ленина, 76. E-mail: potemkinva@susu.ru.

Гришина Мария Александровна – доктор химических наук, главный научный сотрудник Научно-исследовательской лаборатории компьютерного моделирования лекарственных средств Высшей медико-биологической школы, Южно-Уральский государственный университет. 454080, г. Челябинск, проспект Ленина, 76. E-mail: marialgri29@gmail.com.

24

**Bulletin of the South Ural State University. Ser. Chemistry.
2018, vol. 11, no. 1, pp. 18–24**