

УДК 577.29

# СРАВНИТЕЛЬНЫЙ АНАЛИЗ ТОЧНОСТИ KNN-QSAR И ОСЧЕМ МОДЕЛЕЙ В ПРОГНОЗИРОВАНИИ АКТИВНОСТИ ХИМИЧЕСКИХ СОЕДИНЕНИЙ

*Аддылова Ф. Т., Давронов Р. Р.*

fatadilova@gmail.com; rifqat@gmail.com

Институт математики им. В. И. Романовского АН РУз, г. Ташкент, ул. Мирзо Улугбека, 81

Модели, разработанные с помощью метода kNN-QSAR, показали самую высокую прогностическую силу по сравнению с 9 моделями ОСЧЕМ для протестированных данных на обучающей, тестовой и внешней выборке. Дополнительно включенная проверка на областях применимости моделей также продемонстрировала высокие качества моделей прогноза активности, построенных методом kNN-QSAR.

Наши исследования позволяют сделать вывод о том, что при наличии достаточного количества данных и строгой проверки, модели kNN-QSAR должны быть выбраны в качестве предпочтительного инструмента виртуального скрининга.

**Ключевые слова:** модели структура-активность, точность прогнозирования, виртуальный скрининг

**Цитирование:** Аддылова Ф. Т., Давронов Р. Р. Сравнительный анализ точности kNN-QSAR и ОСЧЕМ моделей в прогнозировании активности химических соединений // Проблемы вычислительной и прикладной математики. — 2018. — № 2(14). — С. 114–121.

## 1 Введение

В настоящее время разработка новых лекарств в значительной степени зависит от применения вычислительных методов. Предсказание на этой базе биологических и физико-химических свойств соединений существенно снижает количество экспериментов.

Виртуальный скрининг (ВС) является распространенным и эффективным подходом к решению этой проблемы [1, 2]. Методы виртуального скрининга делятся на методы на основе лиганда (LBVs), и на основе структуры (SBVS). Любой инструмент LBVs основан на принципе сходства, т.е. соединения со сходными химическими структурами, как ожидается, имеют сходные биологические свойства. Тогда можно прогнозировать биологическую активность химически подобных соединений, для которых уже известны активности.

Сегодня известны два основных подхода LBVs, реализованные в виде пакета SEA (Similarity Ensemble Approach) [3] и пакета [4] PASS (Prediction of Activity Spectra for Substances), работающего на основе предсказания QSAR [2]. В работе [5] сравнили предсказательную силу этих двух методов, используя наборы данных соединений против нескольких мишеней – G-белок спаренных рецепторов (GPCRs). На этом же материале была сопоставлена точность предсказания активных/неактивных соединений в обучающих и тестовых наборах с помощью kNN-QSAR моделей [5] с прогнозами PASS и SEA, доступными онлайн. При этом прогностическая эффективность нескольких методов рассматривалась с учетом точности прогнозирования как на обучающих, так и на внешних наборах данных, называемых областями применимости моделей или AD (Applicable Domain). Исследования показали, что kNN-QSAR модели, строго построенные и проверенные, показали самую высокую прогностическую

силу для почти всех протестированных наборов данных GPCR. Программное обеспечение PASS, которое опирается на простой автоматизированный QSAR подход, продемонстрировало умеренную прогностическую силу. Подход SEA, базирующийся на химическом сходстве, показал низкую точность предсказания для всех испытанных семи случаев GPCR.

Известная интернет-среда химического моделирования (OCHEM) является открытой веб-платформой для хранения данных, разработки моделей и публикации химической информации [6]. Поскольку метод kNN-QSAR превосходит пакеты PASS и SEA, то представляется интересным сравнить его с более мощным по функциональности пакетом OCHEM. Ранее [7] мы убедились, что алгоритм kNN-QSAR с включенными процедурами  $k$  — ближайших соседей, кросс-валидации и отжига при выборе дескрипторов, и принятыми критериями оценки качества моделей, генерирует надежные прогнозные модели.

Поэтому целью настоящей работы является сопоставление двух подходов к построению точных моделей прогноза активности kNN-QSAR и OCHEM по принятым статистическим критериям, включая точность прогнозирования на допустимой области их применения (Applicable Domain, AD).

## 2 Материалы и методы

Онлайн-платформа для исследований QSAR/QSPR – Online Chemical Modeling Environment (OCHEM, <http://www.ochem.eu>), позволяет выполнять полный цикл QSAR/QSPR моделирования в полуавтоматическом режиме. Платформа имеет две основные подсистемы: базу данных экспериментально измеренных свойств химических соединений и модули моделирования. Подсистема базы данных включает в себя хранение экспериментальных конечных точек и инструменты для эффективного ввода, поиска и обработки данных. Структура моделирования предоставляет возможности для использования этих данных в процессе моделирования и выполнения всех этапов типового рабочего процесса: подготовка данных, расчет и фильтрация молекулярных дескрипторов, применение методов машинного обучения и анализ эффективности моделей.

OCHEM позволяет выполнять полный цикл разработки модели QSAR, который включает в себя:

- управление наборами экспериментальных данных. Пользователи могут создавать и использовать повторно наборы данных;
- расчет молекулярных дескрипторов: OCHEM поддерживает более 20 типов современных молекулярных дескрипторов от разных поставщиков;
- запуск выбранного метода машинного обучения;
- протокол проверки модели;
- расчет статистик модели;
- применение модели к новым соединениям.

В настоящее время OCHEM поддерживает следующие методы машинного обучения: ASNN-Associative Neural Networks, Bagging, FSMLR, J48 decision tree (Weka implementation), KNN (k nearest neighbors), LIBRARY meta-learning, Partial Least Squares (PLS), XGBoost (Scalable and Flexible Gradient Boosting), Deep Neural Networks.

Когда модель будет построена и просмотрена прогнозная статистика принимается решение – сохранить или отбросить модель. Обычно используемые показатели прогностичности построенной модели – это среднеквадратичная ошибка (RMSE), сред-

няя абсолютная ошибка (MAE) и квадрат коэффициента корреляции:  $q^2$  (обучающая выборка) и  $R^2$  (тестовая выборка) [8].

Моделирование области применимости (AD) является активной областью современного исследования QSAR. Универсальной методики определения AD нет [9–11] – каждое определение AD обычно основывается на некотором произвольно определенном расстоянии (или сходстве) анализируемой молекулы из соединений обучающей выборки и/или модели для данного свойства.

В подходе kNN-QSAR для построенных моделей AD [2] рассчитывали из распределения сходства между каждым соединением и его ближайшими соседями в наборах обучения. Распределение расстояний (попарное сходство) между каждым соединением и его ближайшими соседями в обучающем наборе вычисляется для получения порога  $D_T$  (Domain Threshold), рассчитанного для каждой модели kNN следующим образом:

$$D_T = \bar{y} + Z\sigma. \quad (1)$$

Здесь  $\bar{y}$  – среднее евклидово расстояние  $k$  ближайших соседей каждого соединения в обучающем наборе,  $\sigma$  – стандартное отклонение этих евклидовых расстояний, а  $Z$  – произвольный параметр для контроля уровня значимости. Как правило, значение для этого параметра по умолчанию устанавливается равным 0,5, что формально помещает границу, для которой соединения будут предсказаны при половине стандартного отклонения (если предположить больцмановское распределение расстояний между каждым соединением и его  $k$  ближайшими соседями в обучающей выборке). Если расстояние внешнего соединения от всех его ближайших соседей в обучающей выборке превышает этот порог, предсказание считается ненадежным.

Известно несколько методов определения AD [12], но в ОСНЕМ мы взяли метод LEVERAGE для модели, построенной методом машинного обучения ASNN.

Критерий LEVERAGE разработан в Университете Insubria, Италия [13], где модель строилась в виде обычной регрессии, а генетический алгоритм использовали для расчета наилучшей линейной модели с использованием дескрипторов, доступных в программном обеспечении Dragon [14]. Значения из матрицы, представляющие «расстояние» молекулы до модельного структурного пространства, вычислялись как

$$\text{LEVERAGE} = x^T(X^T X)^{-1}x, \quad (2)$$

где  $x$  – вектор дескрипторов соединения запроса, а  $X$  – матрица, образованная строками, соответствующими дескрипторам молекул из обучающего набора. В линейном моделировании плечо, которое часто обозначается как  $h$ , колеблется между  $1/N$  и 1, и средними  $(K + 1)/N$  для  $N$  соединений в наборе данных обучения, где  $K$  – число дескрипторов модели. Остаток соединения в наборе данных имеет дисперсию  $\sigma^2(1 - h)$ , но если внешнее соединение имеет плечо  $h$ , его ошибка прогнозирования имеет дисперсию  $\sigma^2(1 + h)$ . Когда рычаг внешнего соединения становится большим, начинается экстраполяция вне диапазона данных обучения и больше не стоит считать, что сама модель достоверна.

Молекулы с

$$\text{LEVERAGE} = h > 3(K + 1)/N \quad (3)$$

были идентифицированы как структурно расположенные вне пределов исходной модели. Таким образом, мы сравниваем 2 метода построения QSAR (kNN-QSAR  $\approx$  ОСНЕМ) и 2 метода определения AD ( $D_T \approx$  LEVERAGE) на одном полигоне данных.

По современной химической номенклатуре алкалоиды гармалы обыкновенной хиназолинового строения и их производные, исследованные в [15], разделяются на две группы:

- 1) соединения собственно хиназолинового строения или дезоксипеганинового ряда (соединения от 1 до 43);
- 2) соединения хиназолонового строения или дезоксивазиционного ряда (соединения от 44 до 65).

Химические структуры исследуемых соединений записаны в формате SMILES с использованием программы MedChemDesing 3.0. Пакетом *rcdk* вычислили 91 исходный дескриптор. После стандартной фильтрации соединений и дескрипторов (пропуски, выбросы, порог коэффициента корреляции) осталось 55 соединений и 22 дескриптора, выбранных методом отжига (SA).

Исходное множество соединений трижды методом простой рандомизации было разделено на три подмножества (выборки) в следующем соотношении: 60% – обучение, 20% – тестирование (контрольная выборка), 20% – внешняя выборка (по отношению к обучающей и контрольной), на которой определялось AD.

Проводились два вычислительных эксперимента (ВЭ): в первом ВЭ получили прогностические модели, удовлетворяющую статистическим критериям точности на обучающей и тестовой выборке методом kNN-QSAR и 9 методами из OCHEM. Второй ВЭ показал, насколько точно определяются AD в методе kNN-QSAR и пакете OCHEM, для чего рассчитывали AD на внешней выборке вышеописанными способами, одновременно с точностью прогноза моделей на полученном AD по значению RMSE.

### 3 Обсуждение результатов

#### 3.1 Вычислительный эксперимент 1

Вычисления проводились на веб-платформе OCHEM в соответствии с инструкцией <http://docs.ochem.eu/display/MAN/OCHEM+Introduction>. Все 9 вышеупомянутых методов машинного обучения, включая Deep Neural Networks были использованы для построения прогностических моделей под контролем статистических показателей RMSE, MAE,  $q^2$  и  $R^2$ . Для сравнения была построена прогностическая модель методом kNN-QSAR, процедура построения которой включает в себя отбор переменных, т.е. сходство вычисляется, используя только подмножество дескрипторов, отобранных с помощью имитации отжига (SA), что оптимизирует модель.

В таблице 1 приведены построенные модели с оценками их статистических характеристик. Из таблицы следует, что модель, полученная методом kNN-QSAR, является лучшей.

#### 3.2 Вычислительный эксперимент 2

Из внешней выборки при помощи двух процедур определения AD – LEVERAGE и  $D_T$  выделили  $AD_1$  и  $AD_2$  соответственно.  $AD_1$  содержало 66%, а  $AD_2$  – 58% соединений из исходной внешней выборки. Точность прогнозирования вычислялась дважды: на «родных» и «чужих» AD, тем самым удалось показать, что метод определения AD для одной модели не всегда подходит для другой. В таблице 2 приведены результаты вычислений.

Из таблицы 2 следует, что на исходной внешней выборке точность прогнозирования модели kNN-QSAR на порядок выше, чем модели ASNN. На «родной»  $AD_2$

**Таблица 1** Модели QSAR, полученные на платформе OCHEM и методом kNN-QSAR

№	Метод	Статистические критерии качества моделей			
		$R^2$	$q^2$	RMSE	MAE
1	Ассоциативные нейронные сети (Associative Neural Networks)	$0,5 \pm 0,1$	$0,5 \pm 0,2$	$0,29 \pm 0,06$	$0,2 \pm 0,04$
2	Быстрая многоступенчатая линейная регрессия (Fast Stagewise Multiple Linear Regression)	$0,5 \pm 0,1$	$0,3 \pm 0,3$	$0,33 \pm 0,06$	$0,25 \pm 0,04$
3	k-ближайшие соседи (k-nearest neighbors)	$0,75 \pm 0,08$	$0,75 \pm 0,08$	$0,2 \pm 0,04$	$0,14 \pm 0,03$
4	Модель библиотеки (модель коррекции локального смещения, основанная на другой модели ASNN) (Library model (A local bias correction model based on another ASNN model))	$0,81 \pm 0,07$	$0,79 \pm 0,07$	$0,18 \pm 0,04$	$0,12 \pm 0,02$
5	LibSVM с оптимизацией параметров сетки-поиска (LibSVM with grid-search parameter optimisation)	$0,09 \pm 0,07$	$0,08 \pm 0,07$	$0,38 \pm 0,04$	$0,32 \pm 0,04$
6	Множественная линейная регрессия (Multiple Linear Regression)	$0,08 \pm 0,1$	$0 \pm 0$	$0,8 \pm 0,2$	$0,5 \pm 0,1$
7	Частичные наименьшие квадраты (Partial Least Square)	$0,4 \pm 0,1$	$0,3 \pm 0,2$	$0,33 \pm 0,06$	$0,23 \pm 0,04$
8	Масштабируемое и гибкое усиление градиента (XGBoost (Scalable and Flexible Gradient Boosting))	$0,6 \pm 0,1$	$0,6 \pm 0,1$	$0,25 \pm 0,04$	$0,2 \pm 0,03$
9	Глубокие нейронные сети (Deep Neural Networks)	$0,7 \pm 0,1$	$0,5 \pm 0,2$	$0,29 \pm 0,04$	$0,24 \pm 0,03$
10	kNN-QSAR	0,961	0,73	0,0012	0,09

**Таблица 2** Сравнение методов определения AD для двух QSAR моделей

Модели	$R^2$	$q^2$	RMSE		
			Исходная внешняя выборка	AD1	AD2
ASNN	0,5	0,5	0,35	0,3	0,3
kNN-QSAR	0,961	0,73	0,04	0,195	0,02

точность модели kNN-QSAR – 98%, а на «чужой» AD<sub>1</sub> – 80,5%. Однако, модель ASNN дает одинаково более низкую точность прогнозирования на обеих AD – 70%.

## 4 Выводы

Таким образом, наши исследования показывают, что метод kNN-QSAR лучше некоторых методов, представленных на веб-платформе OCHEM. Строго построенные и проверенные, модели kNN-QSAR показали высокую прогностическую силу для протестированных наборов данных. Возможно, дополнительные исследования на базе OCHEM с использованием большего количества целей помогут устранению недостатков, выявленных нашими исследованиями.

Тем не менее, данная работа ещё раз подтверждает, что при достаточно больших наборах данных можно строить модели QSAR, которые служат наиболее надежными инструментами для мишень-специфичных LBVs.

## Литература

- [1] *Kapetanovic I. M.* Computer-aided drug discovery and development (CADD): in silico-chemico-biological approach // *Chem Biol Interact*, 2008. №171. С. 165–176. doi: <http://dx.doi.org/10.1016/j.cbi.2006.12.006>.
- [2] *Tropsha A., Golbraikh A.* Predictive QSAR modeling workflow, model applicability domains, and virtual screening // *Curr Pharm Des*, 2007. №13(34). С. 3494-504.
- [3] *Keiser M. J. et.all.* Relating protein pharmacology by ligand chemistry // *Nat Biotechnol*, 2000. Т. 25. №2. С. 197–206. doi: <http://dx.doi.org/10.1038/nbt1284>.
- [4] *Lagunin A., Stepanchikova A., Filimonov D., Poroikov V.* PASS: prediction of activity spectra for biologically active substances // *Bioinformatics*, 2000. Т. 16. №8. С. 747–748. doi: <http://dx.doi.org/10.1093/bioinformatics/16.8.747>.
- [5] *Luo M., Wang X. S., Tropsha A.* Comparative Analysis of QSAR-based vs. Chemical Similarity Based Predictors of GPCRs Binding Affinity // *Molecular informatics*, 2016. Т. 35. №1. С. 36–41. doi: <http://dx.doi.org/10.1002/minf.201500038>.
- [6] *Sushko I. et.all.* Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information // *J Comput Aided Mol Des*, 2011. Т. 25. №6. С. 533–554. doi: <http://dx.doi.org/10.1007/s10822-011-9440-2>.
- [7] *Адълова Ф. Т., Давронов Р. Р., Жамилов У. У., Муродов Ш. Н.* Оценка эффективности подхода kNN-QSAR моделирования в хемоинформатике // *Доклады Академии наук Республики Узбекистан*, 2017. №4. С. 3–6.
- [8] *Golbraikh A., Tropsha A.* Beware of  $q^2$ ! // *Journal of Molecular Graphics and Modelling*, 2002. Т. 20. №4. С. 269–276. doi: [http://dx.doi.org/10.1016/S1093-3263\(01\)00123-1](http://dx.doi.org/10.1016/S1093-3263(01)00123-1).
- [9] *Netzeva T. I. et.all.* Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships // *Altern Lab Anim*, 2005. Т. 33. №3. С. 155–173.
- [10] *Zhu H. et.all.* Combinatorial QSAR Modeling of Chemical Toxicants Tested against *Tetrahymena pyriformis* // *J Chem Inf Model*, 2008. Т. 48. №4. С. 766–784. doi: <http://dx.doi.org/10.1021/ci700443v>.
- [11] *Tetko I. V. et.all.* Can we estimate the accuracy of ADME-Tox predictions // *Drug DiscoVery Today*, 2006. №11. С. 700–707. doi: <http://dx.doi.org/10.1016/j.drudis.2006.06.013>.
- [12] *Tetko I. V. et.all.* Critical Assessment of QSAR Models of Environmental Toxicity against *Tetrahymena pyriformis*: Focusing on Applicability Domain and Overfitting by Variable Selection // *J Chem Inf Model*, 2008. Т. 48. №9. С. 1733–1746. doi: <http://dx.doi.org/10.1021/ci800151m>.
- [13] *Tropsha A., Gramatica P., Gombar V. K.* The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models //

- Molecular informatics, 2003. T. 22. № 1. С. 69–77. doi: <http://dx.doi.org/10.1002/qsar.200390007>.
- [14] Todeschini R., Consonni V. Handbook of Molecular Descriptors. —Weinheim: WILEY-VCH, 2000. 667 с.
- [15] Адылова Ф. Т., Жамилов У. У., Давронов Р. Р., Муродов Ш. Н., Азаматов А. А. Прогноз ЛД<sub>50</sub> активности алкалоидов гармалы обыкновенной хиназолинового, хиназолонного строения и их производных на основе QSAR моделей // Кимёвий технология назорат ва бошқарув, 2015. № 5. С. 21–27.

Поступила в редакцию 02.02.2018

UDC 577.29

## COMPARATIVE ANALYSIS OF THE ACCURACY OF KNN-QSAR AND OCHEM MODELS FOR PREDICTING THE ACTIVITY OF CHEMICAL COMPOUNDS

*Adilova F. T., Davronov R. R.*

fatadilova@gmail.com; rifqat@gmail.com

Institute of Mathematics named after V. I. Romanovskiy NAS RUz, 81 Mirzo Ulugbek st.,  
Tashkent, Uzbekistan

Models developed using the kNN-QSAR method showed the highest predictive power compared to the 9 OCHEM models for the tested data on the training, test and external samples. Additionally included testing on model applicability areas also demonstrated the high quality of the activity forecast models constructed using the kNN-QSAR. Our studies allow us to conclude that if there is enough data and strict verification, the kNN-QSAR models should be chosen as the preferred virtual screening tool.

**Keywords:** structure-activity models, predicting accuracy, virtual screening

**Citation:** Adilova F. T., Davronov R. R. 2018. Comparative analysis of the accuracy of kNN-QSAR and OCHEM models for predicting the activity of chemical compounds. *Problems of Computational and Applied Mathematics*. 2(14):114–121.

## References

- [1] Kapetanovic, I. M. 2008. Computer-aided drug discovery and development (CADDD): in silico-chemico-biological approach. *Chem Biol Interact* 171(2):165–76. doi: <http://dx.doi.org/10.1016/j.cbi.2006.12.006>.
- [2] Tropsha, A., A. Golbraikh. 2007. Predictive QSAR modeling workflow, model applicability domains, and virtual screening. *Curr Pharm Des* 13(34):3494–504.
- [3] Keiser, M. J. et. all. 2000. Relating protein pharmacology by ligand chemistry. *Nat Biotechnol* 25(2):197–206. doi: <http://dx.doi.org/10.1038/nbt1284>.
- [4] Lagunin, A., A. Stepanchikova, D. Filimonov, V. Poroikov. 2000. PASS: prediction of activity spectra for biologically active substances. *Bioinformatics* 16(8):747–748. doi: <http://dx.doi.org/10.1093/bioinformatics/16.8.747>.
- [5] Luo, M., X. S. Wang, A. Tropsha. 2016. Comparative Analysis of QSAR-based vs. Chemical Similarity Based Predictors of GPCRs Binding Affinity. *Molecular informatics* 35(1):36–41. doi: <http://dx.doi.org/10.1002/minf.201500038>.

- [6] Sushko, I. et.al. 2011. Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information. *J Comput Aided Mol Des* 25(6):533–554. doi: <http://dx.doi.org/10.1007/s10822-011-9440-2>.
- [7] Adylova, F. T., R. R. Davronov, U. U. Zhamilov, Sh. N. Murodov. 2017. Otsenka effektivnosti podkhoda kNN-QSAR modelirovaniya v khemoinformatike [Estimation of the efficiency of the kNN-QSAR modeling approach in chemoinformatics]. *Doklady Akademii nauk Respubliki Uzbekistan* [Reports of the Academy of Sciences of the Republic of Uzbekistan] 4:3–6. (In Russian)
- [8] Golbraikh, A., A. Tropsha. 2002. Beware of  $q^2$ !. *Journal of Molecular Graphics and Modelling* 20(4):269–276. doi: [http://dx.doi.org/10.1016/S1093-3263\(01\)00123-1](http://dx.doi.org/10.1016/S1093-3263(01)00123-1).
- [9] Netzeva, T. I. et.al. 2005. Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships. *Altern Lab Anim* 33(3):155–173.
- [10] Zhu, H. et.al. 2008. Combinatorial QSAR Modeling of Chemical Toxicants Tested against *Tetrahymena pyriformis*. *J Chem Inf Model* 48(4):766–784. doi: <http://dx.doi.org/10.1021/ci700443v>.
- [11] Tetko, I. V. et.al. 2006 Can we estimate the accuracy of ADME-Tox predictions. *Drug DiscoVery Today* 11(15-16):700–7 doi: <http://dx.doi.org/10.1016/j.drudis.2006.06.013>.
- [12] Tetko, I. V. et.al. 2008. Critical Assessment of QSAR Models of Environmental Toxicity against *Tetrahymena pyriformis*: Focusing on Applicability Domain and Overfitting by Variable Selection. *J Chem Inf Model* 48(9):1733–1746. doi: <http://dx.doi.org/10.1021/ci800151m>.
- [13] Tropsha, A., P. Gramatica, V. K. Gombar. 2003. The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. *Molecular informatics* 22(1):69–77. doi: <http://dx.doi.org/10.1002/qsar.200390007>.
- [14] Todeschini, R., V. Consonni. 2000. *Handbook of Molecular Descriptors*. Weinheim: WILEY-VCH. 667 p.
- [15] Adylova, F. T., U. U. Zhamilov, R. R. Davronov, Sh. N. Murodov, A. A. Azamatov. 2015. Prognoz LD<sub>50</sub> aktivnosti alkaloidov garmaly obyknovennoy khinazolinogo, khinazolonogo stroeniya i ikh proizvodnykh na osnove QSAR modeley [Forecast *text*LD<sub>50</sub> activity of normalum quinazoline normal, quinazolone structure alkaloids and their derivatives based on QSAR models]. *Kimeviy tekhnologiya nazorat va boshkaruv* [Chemical technology control and management] 5:21–27. (In Russian)

Received February 02, 2018