

**К 40-летию Института физиологически активных веществ РАН****ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ****QSAR МОДЕЛИРОВАНИЕ ОСТРОЙ ОРАЛЬНОЙ ТОКСИЧНОСТИ МЛЕКОПИТАЮЩИХ****О.А. Раевский\*, В.Ю. Григорьев, А.В. Ярков, О.В. Тиньков**Институт физиологически активных веществ Российской академии наук,  
142432, Черноголовка Московской обл., Северный проезд, 1; \*эл. почта: raevsky@ipac.ac.ru

В работе исследован ряд из 7490 органических соединений, обладающих острой оральной токсичностью по отношению к мышам. С использованием оригинального подхода АСС (арифметическое среднее свойство) созданы регрессионные модели, обладающие удовлетворительными статистическими характеристиками. В частности, в лучших моделях с использованием обучающей и тестовой выборки квадрат коэффициента линейной корреляции и стандартное отклонение (в единицах  $\log(1/LD_{50})$ ) составили 0.5 и 0.45 соответственно.

**Ключевые слова:** острая оральная токсичность; QSAR; HYBOT**DOI:** 10.18097/BMCRM00066**ВВЕДЕНИЕ**

Острая оральная токсичность у млекопитающих является сложным и недостаточно хорошо понимаемым явлением, состоящим из множества неспецифических взаимодействий химических соединений с биологическими мембранами, а также специфического связывания с большим числом биологических мишеней внутри организма. В связи с этим создание адекватных моделей количественных взаимосвязей “структура - токсичность” затруднено. Так, например, детальный анализ QSAR моделей острой токсичности, проведенный в [1] показал, что многие модели острой токсичности у мышей имеют ограниченное применение вследствие их скромных статистических параметров или ввиду того, что модели были построены на небольшом количестве данных. Более того, в большинстве публикаций острая токсичность коррелирует с гидрофобностью, хотя в настоящее время установлено, что таких простых корреляций для описания этого сложного явления недостаточно [2, 3].

В публикации [4] нами была предложена оригинальная локальная QSAR модель Арифметическое Среднее Свойство (АСС), основанная на структурном и физико-химическом сходстве. Эта модель была успешно использована для оценки острой внутривенной токсичности на мышах [4]. В работе рассмотрены результаты детального исследования острой оральной токсичности химическими соединениями у мышей с применением АСС модели.

**МАТЕРИАЛЫ И МЕТОДЫ**

В настоящей работе использованы экспериментальные данные по острой оральной токсичности химическими соединениями у мышей из SYMYX Toxicity Database [5]. Отобранные экспериментальные значения  $LD_{50}$  в мг/кг живого веса были пересчитаны в  $\log(1/LD_{50})$  (ммоль/кг) и использованы как количественная мера острой токсичности. Общее число отобранных нейтральных

химических соединений составило 7490. Все эти соединения были ранжированы по величине  $LD_{50}$ . Каждое пятое соединение было включено в тестовую выборку. Остальные соединения были использованы в качестве обучающей выборки. Для каждого соединения было рассчитано по 214 физико-химических дескриптора. Расчёты выполнены программами VolSurf [6], HYBOT [7] и DRAGON [8].

*Алгоритм создания АСС модели*

**Шаг 1.** Для каждой анализируемой молекулы из обучающего и тестового ряда определяли величину структурного сходства (в виде коэффициента Танимото,  $T_c$ ) с остальными молекулами. В качестве соседей при этом выступали только молекулы обучающей выборки. Величины  $T_c$  были упорядочены по убыванию. Первичные кластеры ближайших структурных соседей, сформированы путём задания минимального ( $T_{c,min}$ ) и максимального ( $T_{c,max}$ ) граничного значения  $T_c$ . В настоящей работе использовали  $T_{c,min}=0.00$  и  $T_{c,max}=1.00$ .

**Шаг 2.** Для каждой молекулы рассчитывали набор из 214 дескрипторов. В результате были сформированы дескрипторные обучающая и тестовая матрицы.

**Шаг 3.** С использованием каждого дескриптора в отдельности и первичного структурного кластера для каждой молекулы был рассчитан вторичный кластер. Из молекул первичного кластера выбирали такие пары ближайших соседей, которые имели значение дескриптора больше и меньше, чем у анализируемой молекулы. При этом использовали только одиночные дескрипторы, то есть общее число комбинаций было равно 214. Размер вторичного кластера регулировали путём задания числа молекулярных пар. В данной работе использовали 3 молекулярные пары.

**Шаг 4.** Токсичность анализируемой молекулы рассчитывали как арифметическое среднее токсичностей молекулярных пар.



Схематическое объяснение модели АСС представлено на рисунке 1. В этом примере молекулярная поляризуемость ( $\alpha$ ) используется как дескриптор для формирования вторичного кластера соседей на основе двух молекулярных пар. Анализируемое соединение ( $\log(1/LD_{50})=1.96$ ) имеет 10 ближайших структурных соседей (первичный кластер). По сравнению с анализируемым соединением, 8 соседей имеют меньшую величину  $\alpha$  (1 группа), а 2 соседа – большую величину  $\alpha$  (2 группа). Для образования первой пары берётся сосед с максимальным значением  $T_c$  ( $\log(1/LD_{50})=2.46$ , 1 группа). Ближайшим соседом из 2 группы оказывается соединение с  $\log(1/LD_{50})=3.76$ . Для формирования второй пары берётся ещё неиспользованное соединение с максимальным значением  $T_c$  ( $\log(1/LD_{50})=0.71$ , 2 группа). Ближайшим соседом из 1 группы оказывается соединение с  $\log(1/LD_{50})=0.30$ . На основе двух выбранных молекулярных пар можно оценить величину токсичности для анализируемого соединения:  $\log(1/LD_{50})=(2.46+3.76+0.71+0.30)/4=1.81$ .

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В таблице 1 представлены отдельные результаты расчётов, выполненные в рамках модели АСС. Они представляют собой выборку лучших (в статистическом плане) данных, полученных при анализе 214 моделей, использующих по одному единственному дескриптору.

Попытки улучшить параметры моделей путём увеличения числа используемых дескрипторов (и, соответственно, ростом числа их комбинаций) не увенчались успехом. Тем не менее, из таблицы 1 следует, что лучшие модели, использующие одиночные дескрипторы, имеют вполне удовлетворительные статистические параметры: коэффициенты линейной корреляции больше 0.5 и стандартное отклонение 0.45 (в единицах  $\log(1/LD_{50})$ ). Следует отметить, что статистические характеристики полученных АСС моделей вполне соответствуют опубликованным в литературе данным с близким числом анализируемых соединений. Так, в работе [9], посвященной исследованию 8972 соединений с острой оральной токсичностью у крыс с использованием модели PASS, величины  $Q^2(R^2)$  находились в интервале  $0.50\div 0.59$ , а стандартное отклонение составляло  $0.56\div 0.62$ .

Таким образом, в настоящей работе установлено, что подход АСС является вполне адекватным при моделировании острой оральной токсичности по отношению к мышам в большом ряду органических соединений. Полученные регрессионные модели обладают удовлетворительными статистическими характеристиками и могут быть использованы для оценки острой токсичности новых соединений.

## БЛАГОДАРНОСТИ

Работа выполнена в рамках государственного задания на 2018 год (тема № 0090-2017-0020).

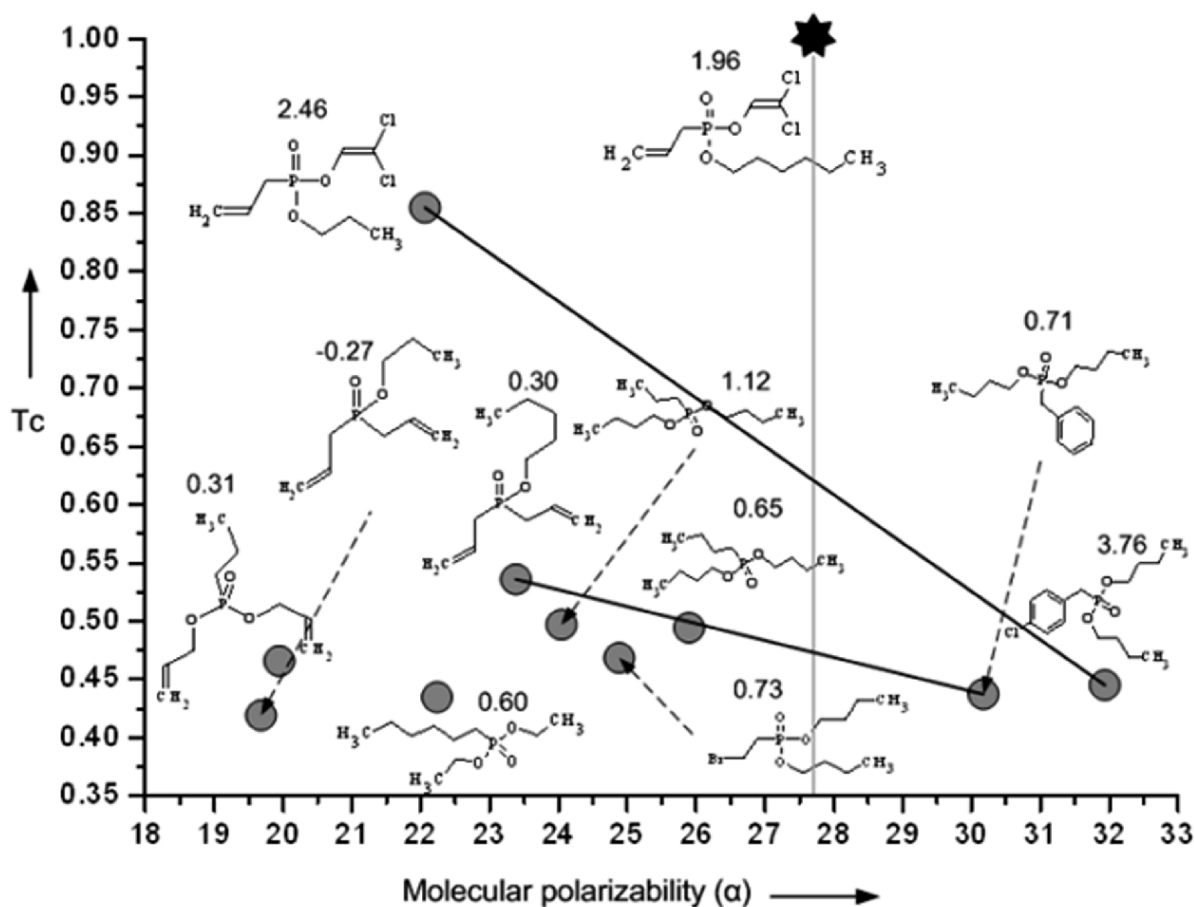


Рисунок 1. Схема выбора соседей в методе АСС. \* - анализируемое соединение, • - соседи.

**Таблица 1.** Статистические характеристики моделей острой токсичности

№	Descriptor	Learning set			Test set		
		n	Q <sup>2</sup>	s <sub>sv</sub>	n	R <sup>2</sup>	s
1	DPSA2	5957	0.500	0.48	1504	0.513	0.45
2	PPSA2	5984	0.499	0.49	1506	0.504	0.47
3	FPSA2	5919	0.500	0.49	1504	0.509	0.47
4	V	5941	0.484	0.49	1504	0.496	0.47
5	PPSA1	5939	0.496	0.49	1502	0.498	0.48
6	WPSA3	5899	0.484	0.49	1502	0.506	0.47
7	Volume	5961	0.477	0.49	1518	0.497	0.46
8	MolVolume	5954	0.476	0.49	1517	0.496	0.46

Примечание. n – число соединений; R<sup>2</sup> – коэффициент линейной корреляции; s – стандартное отклонение; Q<sup>2</sup> – коэффициент линейной корреляции в условиях перекрестного контроля с выбором по одному; s<sub>sv</sub> – стандартное отклонение в условиях перекрестного контроля с выбором по одному; DPSA2 = PPSA2 - PNSA2, PPSA2 – парциальная положительная площадь поверхности, взвешенная на основе общего заряда; PNSA2 – парциальная отрицательная площадь поверхности, взвешенная на основе общего заряда; FPSA2 = PPSA2/A<sub>total</sub>, A<sub>total</sub> – общая молекулярная площадь поверхности; V – молекулярный объём за исключением воды; PPSA1 – парциальная положительная площадь поверхности; WPSA3 = PPSA3\*A<sub>total</sub>/1000, PPSA3 – парциальная положительная площадь поверхности, взвешенная на основе атомных зарядов; Volume, MolVolume – молекулярные объёмы.

## ЛИТЕРАТУРА

1. Devillers, J., & Devillers, H. (2009) Prediction of acute mammalian toxicity from QSARs and interspecies correlations. SAR QSAR Environ. Res., 20(5-6), 467-500. DOI: 10.1080/10629360903278651
2. Lipnick, R.L. (1991) Outliers: Their origin and use in the classification of molecular mechanisms of toxicity. Sci. Total Environ., 109/110, 131-153. DOI: 10.1016/0048-9697(91)90175-E
3. Verhaar, H.J.M., VanLeeuwen, C.J., & Hermens, J.L.M. (1992) Classifying environmental pollutants. Chemosphere, 25(4), 471-491. DOI: 10.1016/0045-6535(92)90280-5
4. Raevsky, O.A., Grigor'ev, V.Ju., Modina, E.A., & Worth, A.P. (2010) Prediction of acute toxicity to mice by the Arithmetic Mean Toxicity (AMT) modelling approach. SAR QSAR Environ. Res., 21(3-4), 265-275. DOI: 10.1080/10629361003771025
5. SYMYX Toxicity Database. Retrieved August 15, 2018, from <http://accelrys.com/products/pdf/dg-pharmacology-package-ds.pdf>
6. Volsurf. Retrieved August 15, 2018, from <http://www.moldiscovery.com/software/vsplus/>
7. Raevsky, O.A., Grigor'ev, V.Ju., Trepalin, S.V. (1999) HYBOT program package, Registration by Russian State Patent Agency No. 990090 of 26.02.99.
8. DRAGON. Retrieved August 15, 2018, from [http://www.taletе.mi.it/products/dragon\\_projects.htm](http://www.taletе.mi.it/products/dragon_projects.htm)
9. Lagunin, A., Zakharov, A., Filimonov, D., & Poroikov V. (2011) QSAR modelling of rat acute toxicity on the basis of PASS prediction. Mol. Inf., 30, 241-250. DOI: 10.1002/minf.201000151

Поступила: 28. 08. 2018.  
Принята к публикации: 07. 09. 2018.

## QSAR MODELING OF MAMMAL ACUTE TOXICITY BY ORAL EXPOSURE

*O.A. Raevsky\*, V.Yu. Grigorev, A.V. Yarkov, O.V. Tinkov*

Institute of Physiologically Active Compounds of the Russian Academy of Sciences,  
1 Severny proezd, Moscow region, Chernogolovka, 142432 Russia; \*e-mail: raevsky@ipac.ac.ru

7490 organic compounds exhibiting acute oral toxicity in mice were studied. Regression models with satisfactory statistical characteristics have been created using the original AMP (arithmetic mean property) approach. The best models using the training and test sets were characterized by the squared linear correlation coefficient and the standard deviation of 0.5 and 0.45 (in log(1/LD<sub>50</sub>) units).

**Key words:** acute oral toxicity; QSAR; HYBOT

## ACKNOWLEDGMENTS

The work was performed within the framework of the state task for 2018 (topic number 0090-2017-0020).