Journal of Siberian Federal University. Chemistry 1 (2012 5) 95-104
~ ~ ~

УДК 544.165: 615.917

Исследование острой токсичности и физико-химических свойств органических соединений

О.В. Тиньков[а]*, П.Г. Полищук[б] ,
А.Г.Артеменко[б], В.Е. Кузьмин[а,б]
[а] Приднестровский государственный университет им. Т.Г. Шевченко,
ул. 25 Октября, 128, Тирасполь MD-3300,
Республика Молдова, Приднестровье
[б] Физико-химический институт им. А.В. Богатского НАН Украины,
Украина 65080, Одесса, Люсторфская дорога, 86 [1]

Received 2.03.2012, received in revised form 9.03.2012, accepted 16.03.2012

Загрязнение окружающей среды является глобальной проблемой. Основным фактором загрязнения нашей планеты служат химические соединения.
Цель исследования состоит в определении зависимости между структурой и токсичностью органических соединений. Разработана база данных «Toxic v.1.1.5» по острой токсичности и физико-химическим свойствам органических соединений. Исследовано влияние молекулярной структуры различных органических соединений на острую токсичность и физико-химические свойства при помощи 2D симплексного представления молекулярной структуры и метода Random Forest. Были получены адекватные QSAR (количественное соотношение структура–свойство) модели. Кроме того, относительное влияние некоторых физико-химических факторов на изменение острой токсичности было оценено на основе QSAR моделей.

Ключевые слова: база данных, токсичность, моделирование.

Введение

Ускоренное развитие химической промышленности в XX в. объясняется стремлением компенсировать нехватку традиционно используемых природных материалов и продуктов, а также создать новые синтетические вещества, превосходящие природные соединения по потребительским свойствам либо отличающиеся более широким спектром областей применения. Эти вполне естественные в условиях непрерывного роста населения планеты устремления сво-

* Corresponding author E-mail address: tinkov84@mail.ru
1 © Siberian Federal University. All rights reserved

ей обратной стороной имеют увеличение степени отрицательного воздействия на окружающую среду [1].

В настоящее время значительная часть человечества в той или иной мере подвержена действию различных химических соединений. Они проникают в организм с вдыхаемым воздухом, водой и продуктами питания. Очевидным последствием такого воздействия стало увеличение уровня заболеваемости, особенно связанной с нарушением иммунного статуса. Поэтому важнейшим требованием к вновь синтезируемым соединениям является их незначительная, приемлемая степень токсичности.

Традиционно для выявления степени токсичности служили эксперименты на лабораторных животных. Данные токсикологические исследования – дорогостоящий, длительный процесс, альтернативой которому может выступить внеэкспериментальный скрининг с использованием экспертных систем. Данные системы базируются на выявленных ранее количественных закономерностях структура–токсичность для уже известных соединений и позволяют прогнозировать токсические свойства для новых, еще не синтезированных химических соединений.

Различают среднесмертельные дозы (LD_{50}, LC_{50}), абсолютно смертельные ($LD_{90\text{-}100}$), минимально смертельные ($LD_{0\text{-}10}$), среднеэффективные (ED_{50}) – вызывающие определенные токсические эффекты, пороговые (PD_{50}) и др. Цифры в индексе – вероятность в % появления определенного токсического эффекта – смерти, порогового действия и др. Наиболее часто используют величины LD_{50}, LC_{50}, которые статистически более достоверны по сравнению с другими [2,3]. Основное отличие LC_{50} от LD_{50} заключается в том, LC_{50} – величина среднесмертельной дозы при ингаляционном способе введения токсиканта, в то время как LD_{50} измеряется при других способах введения (пероральном, внутривенном и др.).

Сегодня в мировой литературе опубликовано много материалов по вопросу действия химических веществ на живые организмы [2-5]. Выпущены фундаментальные издания, посвященные как общим, так и частным вопросам токсикологии [6-7]. Существуют сайты в Интернете, публикующие данные по токсичности химических соединений для животных и человека (табл. 1).

Следует отметить, что много данных такого рода содержится в отечественных журналах, которые недоступны в электронном виде через всемирную сеть Интернет. Создание электронной базы позволит интегрировать зарубежные и отечественные данные, увеличив общее количество информации по токсичности химических соединений. Важно, чтобы в базе данных была реализована возможность формирования выборок химических соединений, которые будут использоваться для построения моделей, описывающих количественную связь структура–активность (QSAR – Quantitative Structure Activity Relationship) и количественную связь структура–свойство (QSPR – Quantitative Structure Property Relationship).

Поэтому задачами данной работы являются:

1) анализ и селекция информации по токсичности органических соединений различных классов;

2) представление полученных данных в виде наглядной электронной базы данных, позволяющей формировать выборки химических соединений для их последующего анализа и выявления количественной зависимости структура-свойство;

3) оценка острой токсичности, а также таких физических свойств, как липофильность и растворимость в воде, от которых в значительной мере зависит экологическая опасность соединений.

Таблица 1. Основные электронные ресурсы по токсичности, доступные в сети Internet

Название ресурса и его краткое описание	Электронный адрес в сети Internet
Национальная медицинская библиотека США Поисковая система для установления наличия искомого вещества в одной из многочисленных токсикологических баз данных системы TOXNET. Division of Specialized Information Services, NLM. ChemIDplus Chemical Search Input Page	http://chem.sis.nlm.nih.gov/chemidplus/
Cheminfo. База токсикологических данных Канадского центра по профессиональной безопасности и здоровью (CCOHS) Содержит информацию по токсичности, экологической опасности химических соединений	http://www.ccohs.ca/products/databases/cheminfo.html
Envirofacts Master Chemical Integrator (EMCI) Электронный ресурс Агентства США по охране окружающей среды, предназначенный для выявления ссылок на интересующее вещество в различных базах данных и перечнях веществ, регулируемых Агентством США по охране окружающей среды	http://www.epa.gov/enviro/html/emci/chemref/60297.html
Вермонтский университет. Vermont SIRI MSDS Collection Электронная коллекция карт безопасности для около 180 тысяч химических веществ	http://hazard.com/msds/index.html
ChemExper Бельгийская поисковая система карт химической безопасности. Содержит информацию о более 70 тысячах химических веществ	http://www.chemexper.com/

Материалы и методы

Информация для заполнения базы данных была получена из советских и российских журналов различных годов выпуска (1920-2005 гг.): «Гигиена труда» (СССР), «Гигиена и санитария» (СССР), «Токсикологический вестник» (Россия); справочных пособий: «Вредные вещества в промышленности» (СССР) [7] и т.д. Помимо этого, часть информации в базу данных была экспортирована с сайтов National Library of Medicine of National Institutes of Health (http://chem.sis.nlm.nih.gov) и U.S. Environmental Protection Agency (http://www.epa.gov/ncct/toxcast/data_sets.html), данные на которых представлены в свободном доступе.

Для хранения и управления полученной информацией нами была разработана электронная система управления базами данных «Toxic v.1.1.5», работающая под управлением операционных систем Windows XP, Vista, Windows 7.

Для описания молекулярной структуры были использованы 2D фрагментарные симплексные дескрипторы, рассчитанные при помощи программы HiT QSAR [8] и 2D интегральные Dragon дескрипторы (http://www.talete.mi.it).

В качестве статистического метода получения QSAR моделей был выбран метод случайного леса (RF – Random Forest), который хорошо зарекомендовал себя при решении QSAR задач [9].

Каждая из исследуемых выборок была разбита на обучающую и тестовую. Как в обучающую, так и тестовую выборки включались органические соединения различных классов. В каждом случае в тестовую выборку было отобрано 20 % из всех соединений с различной токсичностью, все оставшиеся соединения выступали в качестве обучающей выборки.

Качество прогнозов на основе QSAR/QSPR моделей оценивалось коэффициентом детерминации для внешней тестовой выборки

$$R^2 = 1 - \frac{\sum_i (Y_{i,obs} - Y_{i,pred})^2}{\sum_i (Y_{i,obs} - Y_{mean})^2},$$

где R^2 – коэффициент детерминации для внешней тестовой выборки, $Y_{i,obs}$ – наблюдаемое значение активности для i-го соединения тестовой выборки, $Y_{i,pred}$ – предсказанное значение активности для i-го соединения тестовой выборки, Y_{mean} – среднее значение активности для всех соединений обучающей выборки.

Подробное и наглядное описание QSAR методов, использованных в данной работе, приведено в работах [8, 9].

Результаты и обсуждение

Хранение и управление информацией в базе данных «Toxic»

На рис. 1. представлен вид оболочки базы данных по токсичности. База данных содержит информацию по острой и пренатальной токсичности, физическим свойствам для 2067 различных органических соединений.

Одной из особенностей базы данных является возможность формирования выборок соединений для заданного вида токсичности, вида организма и пути введения. В ходе экспорта выборки происходит пересчет величины токсичности из одной шкалы в другую, более подходящую для построения на ее основе QSAR-моделей. Например, для токсичности величины, выраженные в мг/кг, приводятся к шкале моль/кг.

QSAR-анализ токсичности органических соединений

В ходе исследования был получен ряд адекватных QSAR-моделей (табл. 2).

На основе QSAR-моделей, построенных с использованием симплексных дескрипторов и обладающих наилучшей прогнозирующей способностью (выделены жирным шрифтом в табл. 2), проведен анализ относительных влияний физико-химических характеристик соединений на их острую токсичность, результаты которого представлены на рис. 2.

Анализ относительного влияния некоторых физико-химических характеристик на изменение токсичности, оцененного из полученных QSAR моделей, показал следующее:

- при одинаковом способе введения (внутривенно) для схожих по метаболизму видов организмов (мыши и крысы) относительные влияния физико-химических характеристик атомов практически совпадают (рис. 2 в, с);

- при различных способах введения происходят изменения влияний физико-химических характеристик атомов на токсичность соединений. При всех способах введения доминирую-

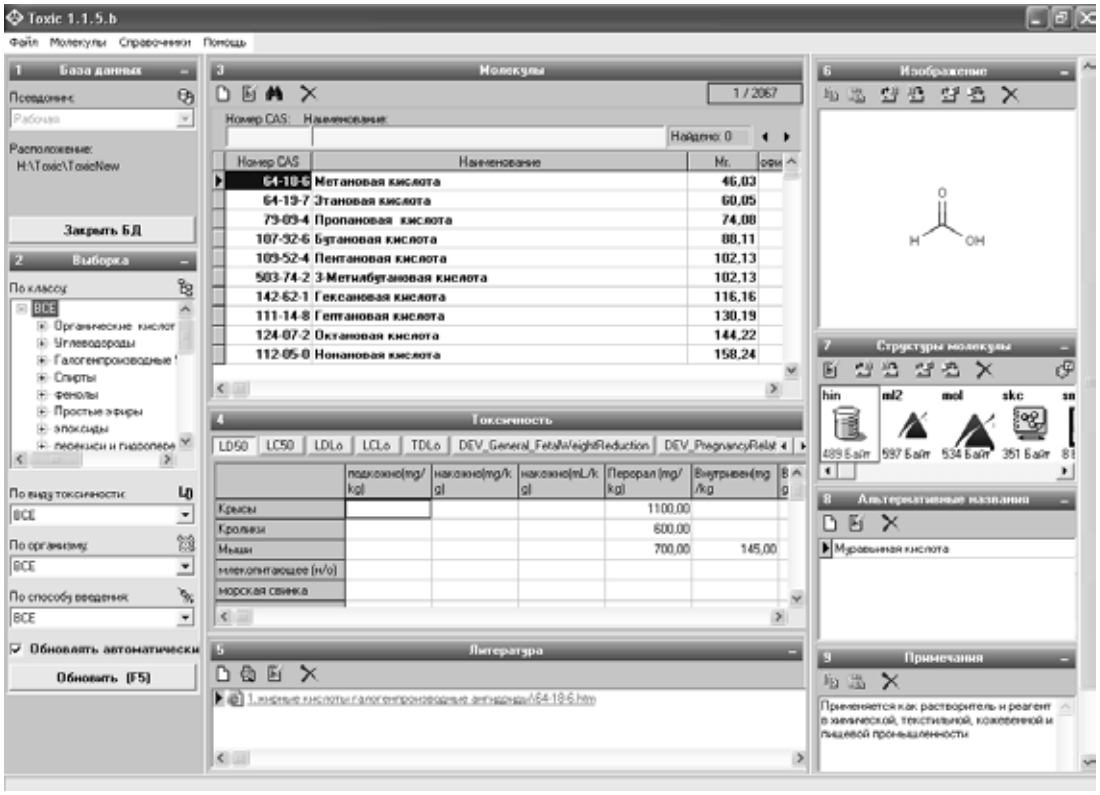

Рис. 1. Общий вид базы данных по токсичности «Toxic v.1.1.5»

Таблица 2. Статистические характеристики QSAR-моделей по острой токсичности

№	Путь введения	Организм	Кол-во соединений в выборке (обучающая / тестовая)	R^2 test	Программный комплекс, генерирующий дескрипторы
1	**перорально**	**крысы**	**1072(858/214)**	**0,55**	**HiT QSAR**
2	перорально	крысы	1072(858/214)	0,53	Dragon
3	**внутривенно**	**крысы**	**132(105/27)**	**0,67**	**HiT QSAR**
4	внутривенно	крысы	132(105/27)	0,74	Dragon
5	**внутривенно**	**мыши**	**307(246/61)**	**0,61**	**HiT QSAR**
6	внутривенно	мыши	307(246/61)	0,59	Dragon
7	**ингаляционно**	**мыши**	**216(172/44)**	**0,59**	**HiT QSAR**
8	ингаляционно	мыши	216(172/44)	0,74	Dragon

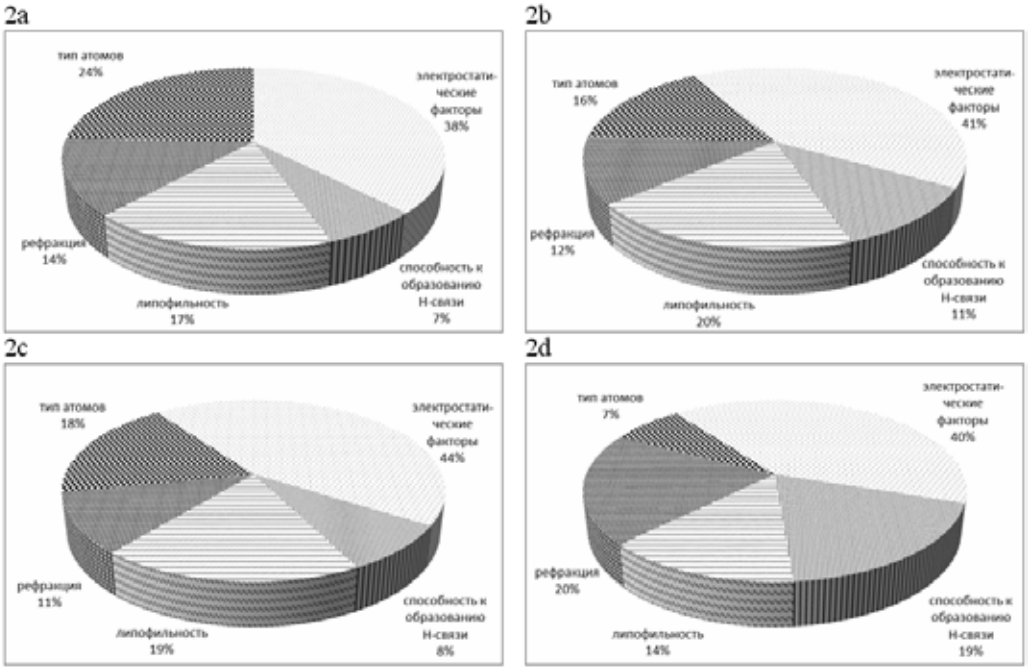

Рис. 2. Относительное влияние физико-химических характеристик исследуемых соединений на значения их среднесмертельных доз при пероральном введении для крыс (2a), при внутривенном введении для крыс (2b), при внутривенном для мышей (2c), при ингаляционном для крыс (2d)

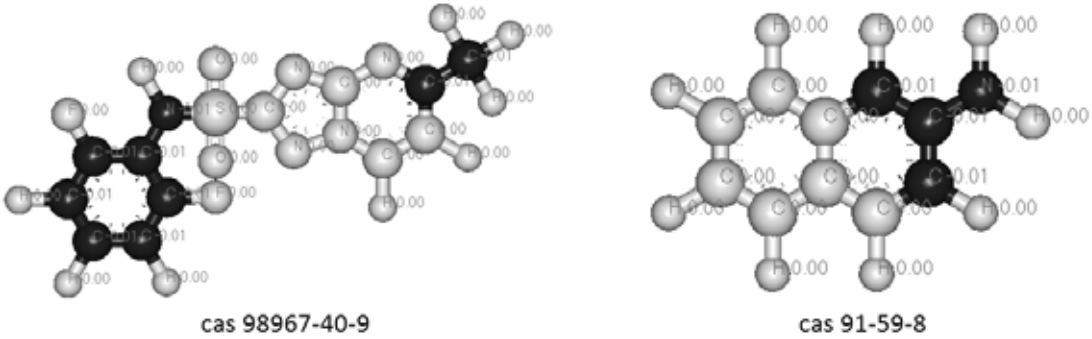

Рис. 3. Пример влияния фрагментов, кодированных цветом

щую роль играют электростатические факторы взаимодействия токсикантов с мишенью. Помимо этого, высока роль липофильности, что, вероятно, связано с процессом пассивной диффузии соединений через мембрану клеток при всасывании в органах и тканях. Кроме того, чем выше растворимость вещества в липидах, тем хуже оно выводится из организма [13]. При переходе от перорального и внутривенного к ингаляционному способу введения увеличивается значимость водородных связей. Это может быть связано с тем, что при ингаляционном способе введения увеличивается образование ассоциатов токсикантов с фармакодинамически активными структурами организма (ферментами, рецепторами и т.д) за счет водородных связей.

Рис. 4. Потенциальные токсикофоры

Следует отметить, что в программном комплексе «HiT QSAR» существует возможность цветового кодирования атомов, положительно или отрицательно влияющих на изучаемое свойство, что позволяет достаточно наглядно представить полученные результаты и визуально определить группы атомов, ответственные за проявление активности. В качестве примера на рис. 3 представлены соединения, для которых при пероральном способе введения для крыс наиболее темные фрагменты способствуют повышению токсичности.

Обобщая указанную информацию по всем выборкам можно выделить фрагменты, определяющие проявление токсичности (токсикофоры). На рис. 4 приведены некоторые токсикофоры для всех рассмотренных видов токсичности, сгруппированные по структурным типам. Относительные вклады указанных фрагментов в исследуемую токсичность не отображены, так как величины вкладов существенно зависят от положения в молекуле и структурного окружения соответствующих токсикофоров. Эта информация позволит проводить предварительную селекцию потенциально опасных соединений.

QSPR-анализ физико-химических свойств органических соединений

Помимо токсичности важнейшими факторами, влияющими на экологическую опасность химических соединений, являются такие физические свойства, как липофильность и растворимость в воде.

Растворимость токсиканта в воде – необходимое условие его резорбции во внутренние среды организма: corpora non agunt nisi soluta (что не растворяется, то не действует). Для того чтобы достичь структуры-мишени, токсикант должен попасть в водную фазу, так как вода – основа межклеточной жидкости организма.

Растворимость в липидах (липофильность) имеет основное значение для процессов проникновения и распределения больших молекул токсикантов в организме. Высока роль липофильности в накоплении (депонировании) токсикантов в живых организмах. Нерастворимые в жирах молекулы могут попасть в организм из окружающей среды лишь в том случае, если они проходят через поры биологических мембран либо переносятся через барьеры с помощью специальных механизмов, например пиноцитоза и т.д.

Таблица 3. Статистические характеристики QSAR-моделей по физическим свойствам органических соединений

№	Физическое свойство	Кол-во соединений в выборке (обучающая / тестовая)	R2 test
1	log P (октанол-вода)	925(740/185)	0,874
2	Растворимость в воде	979(784/195)	0,840

Большие, нерастворимые в липидах вещества, как правило, относятся к числу малотоксичных.

Жирорастворимость и водорастворимость – связанные между собой свойства. Чем полярнее молекула вещества, тем лучше она растворяется в воде и хуже – в липидах. Наибольшей биологической активностью, как правило, обладают токсиканты с промежуточной растворимостью. Важной характеристикой для экологической опасности является коэффициент распределения вещества в среде: октанол-вода (log P).

В связи с этим в рамках создания вышеуказанной экспертной системы разрабатываются QSAR модели, которые позволят предсказывать липофильность и растворимость в воде для новых, еще не синтезированных структур. Построенные нами QSAR-модели на основе 2D симплексного представления молекулярной структуры имеют вполне приемлемые статистические характеристики (табл. 3) и пригодны для внеэкспериментального скрининга.

Заключение

Проведенные исследования позволили:

– разработать базу данных «Toxic v.1.1.5», которая в некоторой степени интегрировала имеющиеся данные зарубежных и отечественных исследователей по токсичности, некоторым физическим свойствам химических соединений. Внедренный в базу данных алгоритм позволяет хранить информацию по токсичности в наглядной форме, дающей возможность оперативно управлять информацией. Одним из видов управления информацией является возможность формирования выборки соединений для дальнейшего анализа количественной зависимости структура-токсичность;
– получить ряд адекватных QSAR-моделей на основе дескрипторов, сгенерированных программными комплексами «HiT QSAR» и «Dragon», с привлечением статистического метода обработки информации Random Forest. Данные модели в целом обладают приемлемыми статистическими показателями и предсказывающей способностью, что особенно важно с учетом возрастающей роли экспертных систем для внеэкспериментального скрининга экологической опасности химических соединений;
– на основе некоторых полученных QSAR-моделей провести анализ относительных влияний физико-химических характеристик атомов на токсичность органических соединений. Это способствовало созданию ряда предположений о фармакокинетических механизмах токсичности при разных путях введения;
– определить структурные фрагменты, устойчиво повышающие токсичность, что может быть полезным при проведении молекулярного дизайна новых соединений с невысокой степенью токсичности.

Данная работа не претендует на исчерпывающее раскрытие вопроса, было бы целесообразно ее продолжение в направлении построения QSAR-моделей для наиболее представительных классов органических соединений с целью выявления важнейших токсикофоров в гомологичных рядах токсикантов.

Авторы выражают особую благодарность А.Е. Шустикову за оказанную помощь в разработке базы данных «Toxic v.1.1.5».

Список литературы

1. Исидоров В.А. Введение в химическую экотоксикологию. СПб.: Химиздат, 1999. 4 с.

2. Курляндский Б.А., Филов В.А. Общая токсикология. М.: Медицина, 2002.

3. Крамаренко В.Ф. Токсикологическая химия. Киев: Выща школа, 1989.

4. Елизарова О.Н., Жидкова Л.В., Кочеткова Т.А. Пособие по токсикологии для лаборантов. М.: Медицина, 1974.

5. Куценко С.А. Основы токсикологии. СПб., 2002.

6. Линг Л.Дж., Кларк Р.Ф., Эриксон Т.Б., Трестрейл Д.Х. III. Секреты токсикологии. СПб., 2006.

7. Лазарева Н.В., Левина Э.Н. Вредные вещества в промышленности: В 3 т. Л.: Химия, 1976.

8. Kuz'min V.E., Artemenko A.G., Muratov E.N., Polischuk P.G., Ognichenko L.N., Liahovsky A.V., Hromov A.I. and Varlamova E.V. Virtual screening and molecular design based on hierarchical QSAR technology «Recent Advances in QSAR Studies», Springer, London, 2010, pp. 127-176.

9. Polishchuk P.G., Muratov E.N., Artemenko A.G., Kolumbin O.G., Muratov N.N., Kuz'min V.E. Application of Random Forest Approach to QSAR Prediction of Aquatic Toxicity// J. Chem. Inf. Model. 2009. 49.

The Investigation of Acute Toxicity and Physical-Chemical Properties of Organic Compounds

Oleg V. Tinkov[a], Pavel G. Polischuk[b], Anatoly G. Artemenko[b] and Victor E. Kuz'min[a,b]

[a] T.G. Shevchenko Transdniestria State University, 107, 25 October str., Tiraspol, MD-3300, Moldova, Transdniestria

[b] A.V. Bogatsky Physical-Chemical Institute NAS of Ukraine 86 Lustdorfskaya Doroga, Odessa, 65080 Ukraine

Nowadays the pollution of the environment is a global problem. The main factor of pollution is different chemical compounds.

The aim of the study is to determine dependence between the chemical structure and the toxicity of organic substances. The data base «Toxic v.1.1.5.» for acute toxicity and physical-chemical properties of organic substances has been prepared. The investigation of influence of the molecular structure of different organic compounds on acute toxicity has been carried out by 2D simplex representation of molecular structure with help of approaches Random Forest. The quite satisfactory QSAR (Quantitative Structure Activity Relationship) models has been obtained. In addition relative influence of some physical-chemical factors on variation of acute toxicity estimated on the base of QSAR models.

Keywords: database, toxicity, modeling.