

УДК 544.165: 615.917

Изучение и прогнозирование острой токсичности углеводов и их галогенпроизводных с помощью метода QSAR

О. В. ТИньков¹, П. Г. ПОЛИЩУК², В. Е. КУЗЬМИН^{1,2}¹Приднестровский государственный университет им. Т. Г. Шевченко, ул. 25 Октября, 128, Тирасполь MD-3300, Приднестровье (Молдова)

E-mail: tinkov84@mail.ru

²Физико-химический институт им. А. В. Богатского НАН Украины, Люсторфская дорога, 86, Одесса 65080 (Украина)

(Поступила 27.03.12; после доработки 27.04.12)

Аннотация

Исследовано влияние молекулярной структуры углеводов и их галогенпроизводных на острую токсичность с использованием 2D-симплексного представления молекулярной структуры и метода частичных наименьших квадратов (Partial Least Squares). Получены адекватные модели QSAR, на основании которых выявлены структурные фрагменты, повышающие токсичность. С помощью моделей QSAR оценено относительное влияние некоторых физико-химических факторов на изменение острой токсичности.

Ключевые слова: острая токсичность, углеводороды и их галогенпроизводные, моделирование

ВВЕДЕНИЕ

Углеводороды и их галогенпроизводные широко применяются в промышленности в качестве растворителей, хладагентов, промежуточных продуктов в химических реакциях, фумигантов и пестицидов. При определении стратегии синтеза и биоиспытаний наряду с целевым эффектом важно оценивать токсическое действие, в частности острую токсичность (LD_{50}). В мире известно более 2 млн химических соединений, и для значительной их части нет сведений о важнейших токсикологических характеристиках, необходимых для определения их опасности. Затраты на экспериментальное определение не всегда целесообразны и доступны. Вместе с тем с помощью математического моделирования и вычислительных методов можно выявить наиболее эффективные и безопасные препараты на ранних стадиях исследования. Одновременно снижается риск позднего выяв-

ления негативного воздействия этих препаратов на человека и окружающую среду [1].

В связи с изложенным актуальное значение имеют исследования на основе методов QSAR, в ходе которых моделируется зависимость токсичности органических соединений от их структуры. Далее полученные QSAR-модели используются для прогноза токсических свойств новых, еще не синтезированных химических соединений.

В данной работе выполнено построение математических моделей прогноза острой токсичности углеводов и их галогенпроизводных и исследовано влияние структурных фрагментов на острую токсичность углеводов и их галогенпроизводных.

МАТЕРИАЛЫ И МЕТОДЫ

Информация по острой токсичности и химической структуре углеводов и их га-

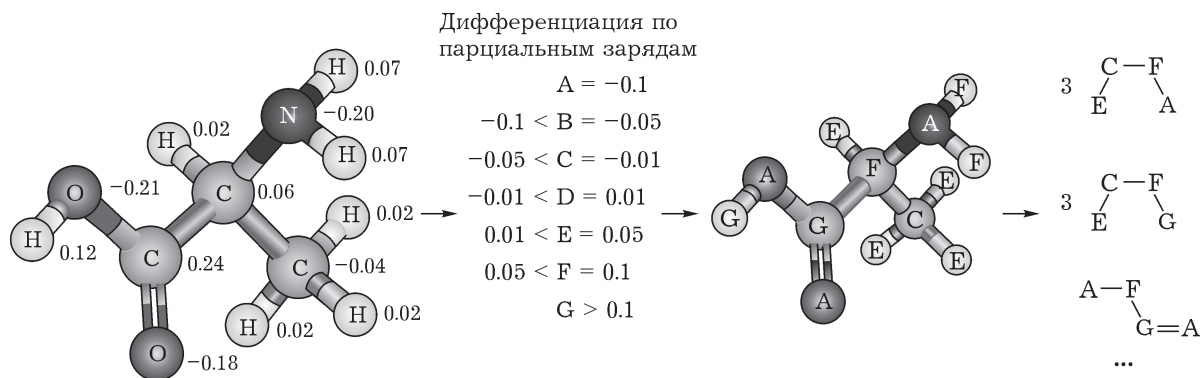


Рис. 1. Пример 2D-генерации симплексных дескрипторов для аланина на 2D-уровне при использовании дифференциации атомов по их парциальным зарядам.

логенпроизводных для построения математических моделей получена при помощи обработки электронной системы управления базами данных Toxic v.1.1.5 [2]. На основании имеющихся данных сформирована выборка из 115 соединений с известными данными по острой токсичности (LD_{50}) при пероральном введении крысам.

Для описания молекулярной структуры использованы 2D-симплексные дескрипторы [3], расчет которых реализован в программном комплексе TheorChem. В рамках симплексного представления молекулярной структуры молекула рассматривается как система различных симплексов – четырехатомных молекулярных фрагментов фиксированной структуры (рис. 1). Дескриптором в этом случае служит количество симплексов одного вида. На 2D-уровне атомы (вершины симплекса) дифференцируют не только по химическим элементам атомов, но и с учетом различных физико-химических свойств – парциального заряда на атоме, его липофильности, рефракции, способности выступать донором или акцептором водорода при образовании водородной связи.

Адекватные модели QSAR, используемые для предсказания активности еще не исследованных молекул, должны иметь высокие статистические характеристики. В то же время хороший набор статистических характеристик еще не означает, что модель будет обладать высокой предсказывающей способностью. Для оценки такой способности используется метод формирования внешней тестовой выборки (test set, или ts). В ходе этой процедуры часть молекул (до 50 %) исключается

из процесса построения модели, а из оставшихся молекул формируется обучающая выборка (work set, или ws). После построения модель используется для прогноза активности соединений тестовой выборки. Ввиду небольшого количества изучаемых сложных эфиров и их структурного разнообразия проведена пятикратная внешняя кросс-валидация. С этой целью все соединения обучающей выборки разделяли на пять групп таким образом, чтобы соединения внутри каждой группы распределялись схожим образом по величине изучаемой токсичности. Для этого все соединения предварительно отсортировали по величине их токсичности и выделили каждое пятое соединение в отдельную группу. Затем четыре полученных группы из пяти объединяли и для полученной выборки строили QSAR-модель; оставшуюся пятую группу соединений использовали в качестве тестовой выборки. Эту процедуру повторяли пятикратно таким образом, чтобы все соединения один раз попали в тестовую выборку.

Для каждого из пяти сформированных наборов обучающей и тестовой выборок были построены математические модели, объединенные далее в консенсусную модель. В этой модели вычисленные значения активностей соответствуют средним значениям активностей, рассчитанным с помощью описанных выше моделей.

Для оценки устойчивости получаемых моделей использовали процедуру скользящего контроля, в ходе которой каждое соединение исключается из выборки и для него предсказывается значение активности. Коэффициент детерминации Q^2 , вычисленный в рам-

ках процедуры скользящего контроля, рассчитывался по формуле

$$Q^2 = \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2 - \sum_{i=1}^m (y_i - \hat{y}_i^{CV})^2}{\sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (1)$$

где m – количество молекул в выборке; y_i – заданное значение активности; \hat{y}_i – вычисленное значение активности; \hat{y}_i^{CV} – значение активности, вычисленное в условиях скользящего контроля для i -й молекулы.

Качество прогнозов на основе QSAR-моделей оценивалось посредством коэффициента детерминации для внешней тестовой выборки $R_{\text{тест}}^2$:

$$R_{\text{тест}}^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \hat{y}_{i\text{cp}})^2} \quad (2)$$

где y_i – наблюдаемое значение активности для i -го соединения тестовой выборки; \hat{y}_i – предсказанное значение активности для i -го соединения тестовой выборки; $\hat{y}_{i\text{cp}}$ – среднее значение активности для всех соединений обучающей выборки.

Стандартная ошибка предсказания (S) определяется как

$$S = \sqrt{\frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{m - 1}} \quad (3)$$

Для более строгой оценки статистической достоверности моделей определены коэффициенты детерминации с применением процедуры рандомизации. Последняя предполагает создание выборки, для которой случайным образом перемешиваются значения токсичности, а значения дескрипторов остаются прежними. После этого выполняется построение 100 моделей. Значительно более низкие средние статистические показатели этих моделей по сравнению с оригинальной свидетельствуют о статистической достоверности полученных оригинальных моделей.

В качестве статистического метода построения моделей QSAR выбран метод частичных наименьших квадратов (Partial Least Squares, PLS) [4].

Расчет вкладов структурных фрагментов на величину острой токсичности LD_{50} выполнен с помощью программы Lattice & Simplex Modelling (LSM), входящей в программный комплекс TheorChem.

Подробное описание используемых QSAR-методов приведено в работах [5, 6].

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В ходе исследования получен ряд адекватных QSAR-моделей (табл. 1), описывающих математическую зависимость между структурой и острой токсичностью углеводородов и

ТАБЛИЦА 1

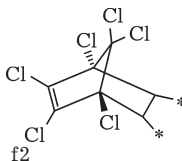
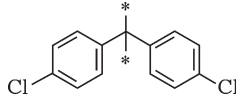

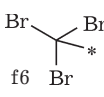
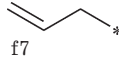
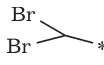
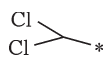
Статистические характеристики моделей QSAR по острой токсичности

Модели	R^2	Q^2	$R_{\text{тест}}^2$	S_{WS}	S_{CV}	S_{TS}	A	R^{2*}
1	0.837	0.722	0.678	0.335	0.441	0.525	3	0.164
2	0.902	0.836	0.673	0.257	0.335	0.482	3	0.210
3	0.907	0.816	0.791	0.253	0.361	0.367	3	0.206
4	0.826	0.775	0.618	0.347	0.396	0.497	3	0.091
5	0.864	0.804	0.733	0.306	0.370	0.421	3	0.129
6 (консенсусная)	0.885	–	0.700	0.228	–	0.451	–	–

Примечание. R_2 – коэффициент детерминации; Q^2 – коэффициент детерминации, вычисленный в рамках процедуры скользящего контроля; $R_{\text{тест}}^2$ – коэффициент детерминации для внешней тестовой выборки; S_{WS} – стандартная ошибка предсказания для рабочей выборки; S_{CV} – стандартная ошибка, вычисленная в рамках процедуры скользящего контроля; S_{TS} – стандартная ошибка предсказания для внешней тестовой выборки; A – количество латентных переменных; R^{2*} – средний коэффициент детерминации, вычисленный в рамках процедуры рандомизации.

ТАБЛИЦА 2

Вклады потенциальных токсокофоров в острую токсичность углеводов и их галогенпроизводных

Фрагменты	Число соединений с этим фрагментом	Вклад		
		Минимальный	Максимальный	Средний
I —* f1	1	-1.342	-1.342	-1.342
 f2	8	-1.673	-0.816	-1.311
 f3	4	-1.371	-0.915	-1.178
 f4	1	-0.990	-0.990	-0.990
Br —* f5	10	-0.879	-0.435	-0.624
 f6	1	-0.609	-0.609	-0.609
 f7	2	-0.756	-0.323	-0.539
 f8	2	-0.502	-0.502	-0.502
 f9	9	-0.518	-0.233	-0.321
Cl —* f10	17	-0.332	-0.027	-0.170

Примечания. 1. Наименьшее значение вклада соответствует более токсичным фрагментам.
2. "Звездочкой" обозначено место связи данного фрагмента с любой частью молекулы.

их галогенпроизводных при пероральном способе введения крысам.

С целью выявления молекулярных фрагментов, способствующих повышению токсичности химических соединений, на основании консенсусной модели определены вклады отдельных фрагментов в величину LD₅₀. Далее из них выбраны фрагменты, однозначно способствующие повышению токсичности, т. е. выступающие в роли потенциальных токсокофоров (табл. 2). Приведенная информация позволит проводить предварительную селекцию потенциально опасных соединений.

Согласно полученным результатам, значительный вклад в проявление токсичности углеводов вносят фрагменты, содержащие галогены. При этом среди структурно схожих галогенсодержащих фрагментов наибольший вклад оказывают фрагменты, содержащие йод. Величина вклада в токсичность углеводородных фрагментов, содержащих фтор, хлор, бром и йод, возрастает в ряду: F < Cl < Br < I. Это подтверждается данными [7], согласно которым механизм токсического действия этих соединений заключается в их биологическом дегалогенировании в орга-



Рис. 2. Относительное влияние физико-химических характеристик исследуемых соединений на их среднесмертельные дозы при пероральном введении для крыс.

низме. Дегалогенирование определяется прочностью связи галогена с углеродом, которая уменьшается при переходе от фтора к йоду. Соединения, содержащие галогенуглеродные фрагменты, образуют свободные радикалы (ферментным или неферментным путем), которые алкилируют активные группы белков и ферментов, приводя к возникновению соответствующей патологии.

Из приведенных результатов анализа видно, что фрагменты, содержащие полихлорированные полициклические диеновые радикалы (фрагмент f2) и дихлордифенилметановые радикалы (фрагмент f3), характеризуются высоким вкладом в токсическое действие химических соединений. Соединения, содержащие фрагменты f2 и f3, принято относить к хлорорганическим экотоксикантам, в частности к полихлорированным пестицидам, которые оказывают устойчивое отрицательное воздействие на живые организмы [8]. Данные вещества проявляют острое политоксическое действие на организм, нарушая работу различных систем и органов животных, человека [9]. Согласно данным [10], соединения, содержащие фрагменты f2 и f3, проявляют иммунотоксичность. Вероятный механизм действия полихлорированных пестицидов связан с их алкилирующим действием на основные биомишени – аминокислоты, пептиды, белки, нуклеиновые кислоты, липиды.

На основании полученной консенсусной модели проведен анализ относительных влияний физико-химических характеристик уг-

леводородов и их галогенпроизводных на острую токсичность (рис. 2).

В ходе проведенного анализа влияния физико-химических характеристик атомов исследуемых углеводородов и их галогенпроизводных на значения острой токсичности выявлена доминирующая роль электростатических факторов взаимодействия токсикантов с мишенью. Кроме того, высока роль липофильности, что, вероятно, связано с процессом пассивной диффузии токсикантов через мембрану клеток при всасывании в органах и тканях.

ЗАКЛЮЧЕНИЕ

В результате проведенных исследований получен ряд адекватных QSAR-моделей на основе 2D-симплексных дескрипторов с привлечением статистического метода частичных наименьших квадратов (Partial Least Squares). Установлено, что данные модели в целом обладают приемлемыми статистическими показателями и предсказывающей способностью. На основе QSAR-моделей определены фрагменты, повышающие токсичность углеводородов и их галогенпроизводных. Полученная информация может быть полезной при проведении молекулярного дизайна новых соединений с невысокой степенью токсичности. Для некоторых молекулярных фрагментов, повышающих токсичность, рассмотрены возможные механизмы токсического патогенеза.

Данная работа не претендует на исчерпывающее раскрытие вопроса. Целесообразно продолжить ее в направлении расширения выборки изучаемых соединений для построения QSAR-моделей с целью выявления важнейших токсикофоров.

СПИСОК ЛИТЕРАТУРЫ

- 1 Кирлан А.В. Планирование направления синтеза биологически активных окси- и амидосодержащих гетероциклических соединений с учетом токсичности: Автореф. дис. ... канд. хим. наук. Уфа, 2003г. 24 с.
- 2 Тиньков О.В., Полищук П.Г., Артеменко А.Г., Огниченко Л.Н., Кузьмин В.Е. // Вестн. Приднестр. гос. ун-та. 2011. № 2. С. 112–118.
- 3 Кузьмин В.Е., Артеменко А.Г., Челомбитко В.А., Муратов Е.Н., Желтвай А.И., Мещеряков А.К., Ляховский А.В. // Тр. науч. семинара “Связь “структу-

- ра – активность” биологически активных веществ”. Гурзуф, 2002. С. 22–26.
- 4 Rannar S., Lindgren F., Geladi P., Wold S. // *J. Chemometrics*. 1994. Vol. 8. P. 111–125.
- 5 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. // *Recent Advances in QSAR Studies* / T. Puzyn, J. Leszczynski, M. T. D. Cronin (Eds.). London: Springer, 2010. P. 127–176.
- 6 Polishchuk P. G., Muratov E. N., Artemenko A. G., Kolumbin O. G., Muratov N. N., and Kuz'min V. E. // *J. Chem. Inf. Model*. 2009. Vol. 49 P. 2481–2488.
- 7 Лазарева Н. В., Левина Э. Н. Вредные вещества в промышленности. Т. I. Л.: Химия, 1976. С. 190.
- 8 Исидоров В. А. Введение в химическую экотоксикологию. СПб.: Химиздат, 1999. С. 90–97.
- 9 Линг Л. Дж., Кларк Р. Ф., Эрикссон Т. Б., Трестрейл П. Д. Х. Секреты токсикологии: Пер. с англ. М.–СПб.: БИНОМ–Диалект, 2006. С. 244–245.
- 10 Курляндский Б. А., Филов В. А. Общая токсикология. М.: Медицина, 2002. С. 368–369.