

В. В. Поройков, Д. А. Филимонов, Ю. В. Бородина,
Т. А. Глориозова, В. Б. Ситников, С. В. Садовников, А. В. Соснов

АНАЛИЗ КОЛИЧЕСТВЕННЫХ ЗАВИСИМОСТЕЙ “СТРУКТУРА – ОТСТАВЛЕННАЯ НЕЙРОТОКСИЧНОСТЬ” МЕТОДОМ САМОСОГЛАСОВАННОЙ РЕГРЕССИИ НА ОСНОВЕ PASS

ГУ НИИ биомедицинской химии им. В. Н. Ореховича РАМН, Москва;
ГНЦ “ГосНИИОХТ”, Москва

Широкое применение различных химических соединений в промышленности, сельском хозяйстве и быту с особой остротой ставит проблему их безопасности. При санитарном нормировании используемых в народном хозяйстве России веществ отдаленная нейротоксичность, как правило, не оценивается. В то же время выявлен круг веществ, способных вызывать развитие двигательных параличей верхних и нижних конечностей в отставленные после воздействия сроки (2 – 3 недели и более). Такого рода эффекты проявляют, в частности, фосфорорганические пестициды [1, 2] и пиретроиды [3]. Основной мишенью нейротоксического действия фосфорорганических соединений является белковый комплекс мембраны, степень влияния на который оценивают по ингибированию нейротоксической эстеразы (НТЭ) [4, 5].

Механизмы отставленной нейротоксичности (ОНТ) до конца не выяснены, диагностика данной патологии

затруднена, классификация в зависимости от структуры вещества отсутствует. Поэтому актуальной задачей является разработка специализированной базы данных по веществам, проявляющим ОНТ, и создание компьютерной системы анализа взаимосвязей “структура – отставленная нейротоксичность”, позволяющей осуществлять прогнозирование отставленной нейротоксичности для новых веществ.

Разработка базы данных DEL-NEUROTOX

С целью оценки современного состояния исследований отставленной нейротоксичности нами проведен информационный поиск в базах данных и полнотекстовых электронных журналах, доступных локально или через Интернет (PharmaProjects, MDDR, Medline, ScienceDirect, E-Library и др.). Оказалось, что, несмотря на значительное внимание, уделяемое данной проблематике в официальных документах [6, 7], специализированной базы данных по отставленной нейроток-

The screenshot displays the DEL-NEUROTOX database interface. At the top, the title bar reads "ISIS/Base - [DELAYNEIROTOXITY.DB/Form1]". The menu bar includes "File", "Edit", "Options", "Object", "Database", "Search", "List", "Window", and "Help". The toolbar contains "Forms", "Query", "Browse", and "Update". The main window shows a search result for "100125" with the title "Ингибирование других ферментов." (Inhibition of other enzymes). A chemical structure of a phosphorus compound is shown on the left. Below the structure are several data tables:

K	Pi50	Ki	Enzym	Source	Treatment	NTR
	-4.91		NTE	MFNC	in vitro	150
	-4.30		NTE	ANC	in vitro	42
	-2.74		AChE	MFNC	in vitro	
	-2.54		AChE	ANC	in vitro	

Below this are tables for "Процент угнетения НТЭ." (Percentage of NTE inhibition), "T1/2 - полупериод старения" (T1/2 - aging half-life), "LD50" (Minimal dose causing attack), and "pI₅₀ НТЭ." (pI₅₀ NTE). At the bottom, there are tables for "Многократное введение" (Multiple administration) and "Литература." (Literature), along with a box containing "Dichlorvos, DDVP" and a box with the number "3".

Рис. 1. Пример экрана базы данных DEL-NEUROTOX.

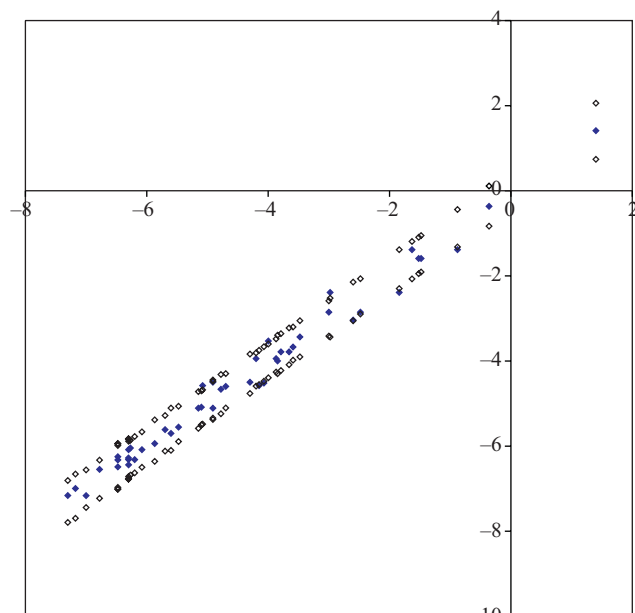


Рис. 2. Визуализация зависимости “структура – PI_{50} ” для ингибиторов НТЭ, полученной на основе самосогласованной регрессии. По оси абсцисс отложены экспериментальные, а по оси ординат — расчетные значения. Светлые точки соответствуют экспериментальным значениям \pm оценки стандартного отклонения.

№	Название вещества
100-003	O,O-Дипентил-O-(2-хлорвинил)фосфат
100-004	O-(2,2-дихлорвинил)-O,O-дипропилфосфат
100-005	O-(2,2-дихлорвинил)-O,O-дифенилфосфат
100-016	O-(4-нитрофенил)-O,O-дибутилфосфат
100-017	O-(4-нитрофенил)-O,O-дипентилфосфат
100-018	O-(4-нитрофенил)-O,O-бис(2-хлорэтил)фосфат
100-021	O-(2-хлорвинил)-O,O-диэтилфосфат
100-024	O-(2-хлорэтил)-O,O-дибутилфосфат
100-028	O,O-диметил-O-(4-иод-2,5-дихлорфенил)фосфат
100-038	O-(3-хлор-4-метилгумарин-7-ил)-O,O-бис(2-хлорэтил)фосфат
100-043	O-(4-нитрофенил)-O-пентилпентилфосфонат
100-046	O-(4-нитрофенил)-O-этилбензилфосфонат
100-047	O-(4-нитрофенил)-O-этил-3-фенилпропилфосфонат
100-048	O-(4-нитрофенил)-O-этилфенилфосфонат
100-050	O-(2,4-дихлорфенил)-O-метилметилфосфонат
100-053	O-(2,4,5-трихлорфенил)-O-этилэтилфосфонат
100-056	O-(4-бром-2,5-дихлорфенил)-O-метилфенилфосфонат
100-105	O-этилоктилфторфосфонат
100-108	2-[Метилтио(метокси)фосфорилтио]диэтилсукцинат
100-110	2-[Метилтио(метокси)фосфорилтио]диэтилсукцинат P S, 2 R
100-111	2-оксо-2-октил-4Н-1,3,2-бензодиазепин-5-ил-фосфорин (R)
100-114	L-Ala-CS*
100-115	L-Val-CS
100-117	D-Norval-CS
100-118	L-Phe-CS
100-119	L-Norleu-CS
100-120	L-Met-CS
100-122	O-(4-нитрофенил)-O,O-диэтилфосфат
100-123	2-(Диметоксифосфорилтио)диэтилсукцинат
100-124	O-(3,5,6-хлорпиридил)-O,O-диэтилфосфат
100-125	O,O-диметил-O-(2,2-дихлорвинил)фосфат
100-126	O,O-диметил-1-гидрокси-2,2,2-трихлорэтилфосфонат
100-127	O,O-диизопропилфторфосфат
100-128	2-оксо-2-(3-метилфенилокси)-4Н-1,3,2-бензодиазепин-5-ил-фосфорин
100-129	2-оксо-2-фенилокси-4Н-1,3,2-бензодиазепин-5-ил-фосфорин
100-130	N,N'-диизопропиламидофосфорфторид
100-132	O,O-диоктил-O-(2,2-дихлорвинил)фосфат

Примечание. CS — карбонилсульфонат.

сичности, которая могла бы быть использована в качестве обучающей выборки для системы прогноза, в настоящее время не существует.

Выбор системы управления базами данных (СУБД) для создания базы данных (БД) по веществам, проявляющим отставленную нейротоксичность, был продиктован необходимостью ввода в нее информации по структуре, свойствам и биологической активности. СУБД ISIS/Base, разработанная фирмой MDL Information Systems, Inc. [8], в полной мере удовлетворяла поставленным требованиям, обеспечивая ввод, хранение, поиск и обработку данных по структуре и различным свойствам химических соединений, включая возможности генерации разнообразных отчетных форм. ISIS/Base может быть использована на отдельных (stand-alone) персональных компьютерах, либо как программа-клиент для более мощной СУБД ISIS/Host, которая инсталлирована на сервере с сетевым доступом. Преимуществами использования ISIS/Base являются: (1) представление структурных данных в формате MOL- или SD-файлов является де-факто стандартом на представление структурной информации в компьютере; (2) возможность ввода структурной информации, содержащей неклассические особенности.

Была разработана логическая структура базы данных по веществам, проявляющим отставленную нейротоксичность. В число параметров, характеризующих отставленную нейротоксичность веществ, включены следующие экспериментальные характеристики: способность веществ ингибировать выделенную нейротоксическую эстеразу (НТЭ) *in vitro*; I_{50} мкМ и PI_{50} ; К ингибирования $M^{-1} \text{ мин}^{-1}$; время старения фермента; $I_{50} \text{ АХЭ} / I_{50} \text{ НТЭ}$. Кроме того, необходимо учитывать определенные параметры, относящиеся к опытам *in vivo*: это доза, которая вызывает или не вызывает атаксию, и остаточная активность НТЭ; минимальная доза, вызывающая атаксию; наличие атаксии при многократном введении.

Нами были созданы соответствующие формы для заполнения БД, разработана технология ввода информации, осуществлены сбор из доступных источников, анализ и заполнение базы данных. На момент подготовки публикации в БД содержалась информация по 132 веществам, пример записи для вещества с регистрационным номером 131 приведен на рис. 1.

Анализ количественных характеристик, содержащихся в БД DEL-NEUROTOX, показал, что наиболее адекватным параметром для построения количественных зависимостей являются величины логарифма константы ингибирования НТЭ (PI_{50}), надежные значения которых имелись для 37 веществ. Названия веществ представлены в таблице.

Данные для этих веществ были экспортированы в SD-файл, содержащий ID, структурную формулу и значение PI_{50} для каждого вещества.

Анализ зависимостей “структура – константа ингибирования НТЭ”.

Для построения количественных моделей зависимости константы ингибирования НТЭ от структуры вещества был использован метод самосогласованной регрессии (ССР). ССР основана на принципе статистической регуляризации, имеет строгое математическое обоснование и позволяет строить устойчивые регрессионные модели при произвольном соотношении объёма обучающей выборки и числа независимых переменных. В качестве независимых переменных использовались результаты прогноза спектра биологической активности с помощью программы PASS. Подробное описание данного алгоритма приведено в нашей публикации [9].

Из данных, содержащихся в SD-файле (экспорт из базы данных), программа PASS [10] формирует двухмерный массив, содержащий значения величины, для которой предполагается построить модель зависимости, и результаты прогноза каждого вида биологической активности в виде разностей оценок вероятностей наличия и отсутствия данного вида активности Pa-Pi. Для 37 веществ обучающей выборки выполнялся прогноз 783 видов биологической активности программой PASS, Pa-Pi для которых использовались в качестве независимых переменных. Этот массив использовался как входная информация программным модулем, выполняющим расчёт самосогласованной регрессии.

В модуле самосогласованной регрессии можно выбрать либо режим автоматической оценки точности данных по результатам расчета, либо режим с заданием этой точности пользователем. Во втором случае, если реальная точность данных выше, в результате расчета получается более грубая, но и более устойчивая к случайным возмущениям модель зависимости. С учетом небольшого объема выборки мы выбрали второй вариант со значением величины ошибки для логарифма константы ингибирования НТЭ 0,5, что соответствует возможным отклонениям значения самой константы около 3 раз.

Алгоритм самосогласованной регрессии после 127 итераций закончил свою работу с регрессионной моделью, содержащей 23 независимых переменных из 783 для выборки из 63 химических соединений со значениями логарифма константы ингибирования НТЭ.

Эффективная размерность модели равна 21,9. Качество модели характеризуется следующими величинами:

Коэффициент корреляции — 0,983, оценка стандартного отклонения — 0,478, что хорошо согласуется с заданным значением 0,5 и указывает на некоторую загруженность модели, значение критерия Фишера — 76,17. Такие статистические характеристики позволяют считать построенную модель вполне удовлетворительной.

На рис. 2 приведена диаграмма зависимости расчетного значения, полученного на основе построенной модели, от исходного значения логарифма константы ингибирования НТЭ.

Учитывая сложность экспериментального определения этой характеристики и разнородность веществ обучающей выборки, можно считать, что полученные результаты указывают на возможность построения количественных моделей связи “структура – отставленная нейротоксичность”. Использование внеэкспериментального прогнозирования для новых химических соединений позволит установить приоритетность экспериментального тестирования веществ, для которых значительна расчетная вероятность отставленной нейротоксичности.

Выражаем искреннюю признательность МНТЦ за финансовую поддержку данной работы (грант № 574).

ЛИТЕРАТУРА

1. S. C. Bondy, *Applied Toxicology*, **6**(2), 208 – 217 (1986).
2. J. G. Hollingshaus and T. R. Fukuto, in: *Effect of chronic exposures to pesticides on animal systems*, Raven Press, New York (1982), pp. 85 – 120.
3. W. N. Aldridge, *Crit. Rev. Toxicol.*, **21**(2), 89 – 104 (1990).
4. M. K. Johnson, *Biochemical J.*, **114**, 711 – 717 (1969).
5. M. K. Johnson, *Crit. Rev. Toxicol.*, **3**, 289 – 295 (1975).
6. U. S. EPA 712-C-96-237 June 1996 “Prevention, Pesticides and Toxic Substances”.
7. U. S. PDA Redbook 2000 “Toxicological Principles for the Safety of Food Ingredients”.
8. <http://www.mdli.com>
9. Д. А. Филимонов, Д. В. Акимов, А. А. Лагунин, В. В. Пороиков, *Хим.-фарм. журн.*, **38**(1), 21 – 24 (2004).
10. V. V. Poroikov, D. A. Filimonov, Yu. V. Borodina, et al., *J. Chem. Inf. Comput. Sci.*, **40**(6), 1349 – 1355 (2000).

Поступила 30.06.03