

УДК 544.165:615.917

ПРОГНОЗИРОВАНИЕ ВРЕМЕНИ БИОТРАНСФОРМАЦИИ ОРГАНИЧЕСКИХ СОЕДИНЕНИЙ НА ПРИМЕРЕ АВЕРМЕКТИНОВ

О.В. Тиньков^{1,2*}, В.Ю. Григорьев³, Л.Д. Григорьева⁴

(¹ Кафедра фармакологии и фармацевтической химии медицинского факультета Приднестровского государственного университета имени Т.Г. Шевченко, Тирасполь, Приднестровье, Молдова; ² Военный институт имени генерал-лейтенанта А.И. Лебедея Министерства обороны, Тирасполь, Приднестровье, Молдова; ³ Институт физиологически активных веществ Российской академии наук, Черноголовка, Россия; ⁴ Факультет фундаментальной физико-химической инженерии МГУ имени М.В. Ломоносова, Москва, Россия; *e-mail: oleg.tinkov.chem@mail.ru)

В настоящее время распространение коронавируса SARS-CoV-2 представляет собой существенную проблему для всего человечества. Один из перспективных агентов для борьбы с коронавирусом SARS-CoV-2 – ивермектин, представляющий собой комплекс полусинтетических производных природных авермектинов, которые на протяжении многих лет эффективно используются в медицине и сельском хозяйстве как противопаразитарные препараты. Однако экспериментальных данных по экотоксикологической оценке индивидуальных авермектинов до сих пор крайне мало. Цель настоящего исследования состояла в разработке математической модели, которая позволила бы провести достоверный прогноз способности к биотрансформации природных и полусинтетических авермектинов, а также выявить структурные фрагменты молекул, которые больше всего влияют на данный вид биологической активности. Основой для построения модели послужила структурно разнородная выборка, содержащая органические соединения с экспериментальными значениями периода полубиотрансформации (biotransformation half-life, KmHL). С использованием интернет-платформы «ОСНЕМ» (<https://ochem.eu>), в которой реализован расчет PyDescriptor-дескрипторов, а также метод случайного леса и трансформаторной сверточной нейронной сети, была построена удовлетворительная ($R^2_{\text{test}} = 0,81$) модель количественной связи структура – активность (quantitative structure – activity relationship, QSAR). Проведенные расчеты показали, что в организмах рыб быстрее происходит биотрансформация природных авермектинов, чем полусинтетических авермектинов. Определены структурные фрагменты, которые повышают и понижают скорость биотрансформации органических соединений.

Ключевые слова: макролиды, молекулярные дескрипторы, машинное обучение, QSAR.

В настоящее время одним из наиболее перспективных ингибиторов коронавируса SARS-CoV-2 является ивермектин, который в исследованиях *in vitro* позволил снизить долю вирусной РНК на 99,98% в течение 48 ч при концентрации 5 мкМ, считающейся безопасной для клеток [1]. Недавно было показано, что ивермектин успешно снижает вирусную нагрузку на организм больных COVID-19 [2]. Ранее было отмечено снижение репликации вирусов ВИЧ-1 и Денге с помощью ивермектина [3]. Ивермектин, продажи которого на протяжении 20 лет оцениваются на сумму свыше 1 млрд долл. США в год [4], представляет собой комплекс ивермектина B1a (~90%) и ивермектина B1b (~10%) – полу-

синтетических производных природных авермектинов. Природные авермектины на протяжении многих лет крайне эффективно используются против эндо- и эктопаразитов растений, животных и человека [5]. Авермектины также перспективны с точки зрения разработки новых противоопухолевых препаратов [6–8]. Культура *Streptomyces avermitilis* синтезирует комплекс из восьми компонентов, представляющих собой 16-членные макролиды. Химические структуры природных и полусинтетических авермектинов представлены на рис. 1–5.

Антипаразитарные свойства авермектинов обусловлены их воздействием на рецепторы γ -аминомасляной кислоты (ГАМК) и глутаматные

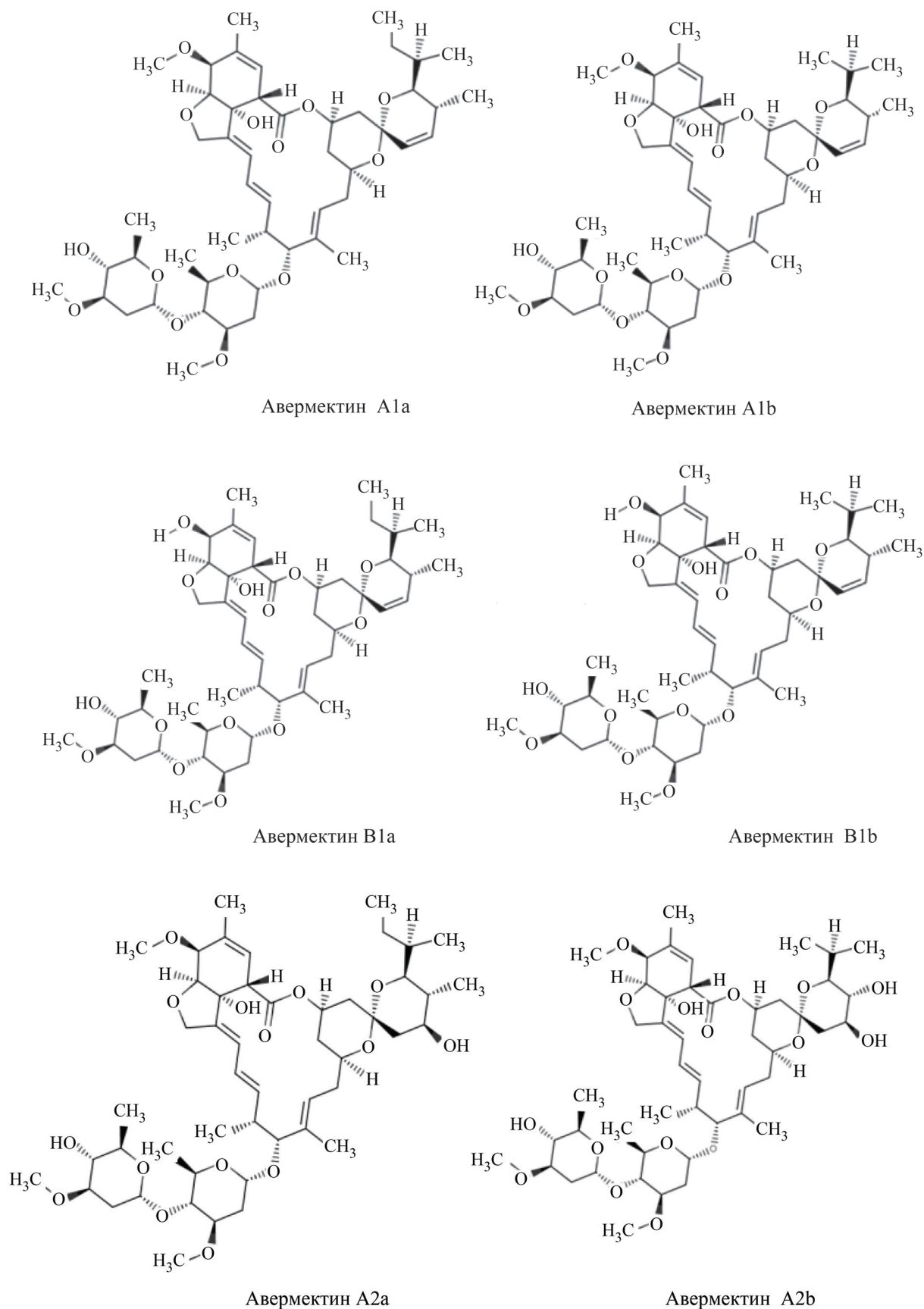


Рис. 1. Химическая структура природных авермектинов: Авермектина А1а, Авермектина А1b, Авермектина В1а, Авермектина В1b, Авермектина А2а, Авермектина А2b

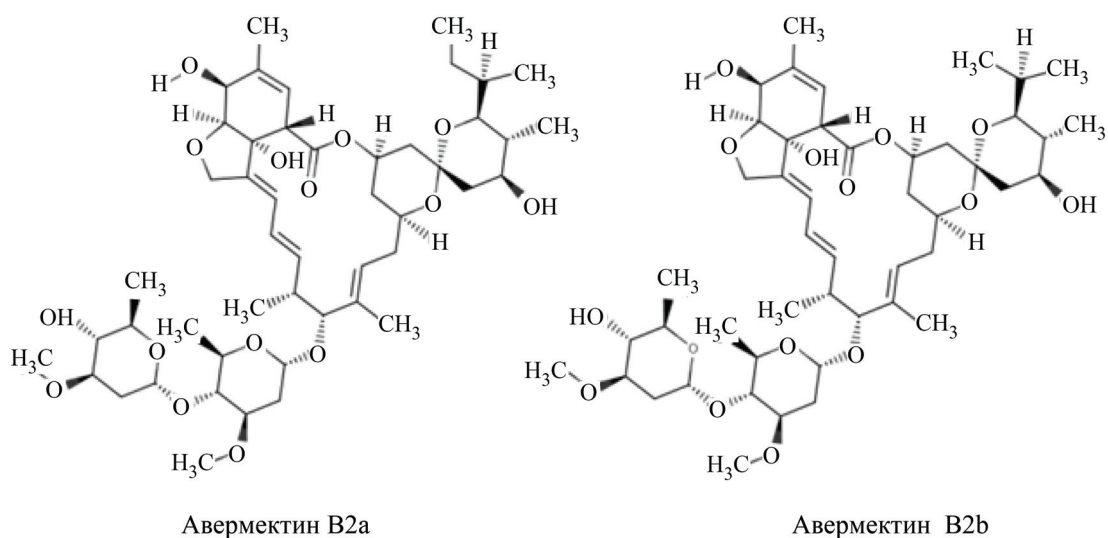
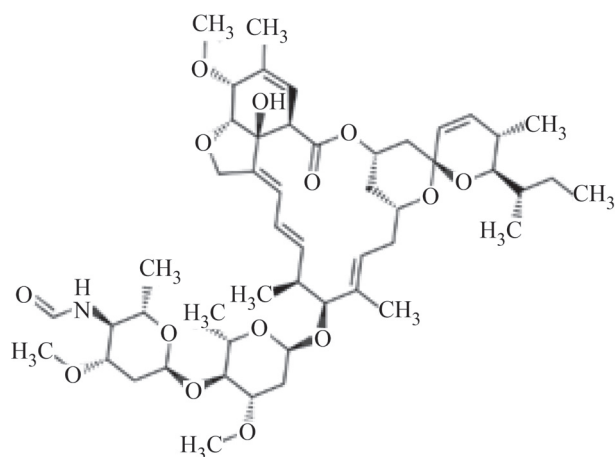
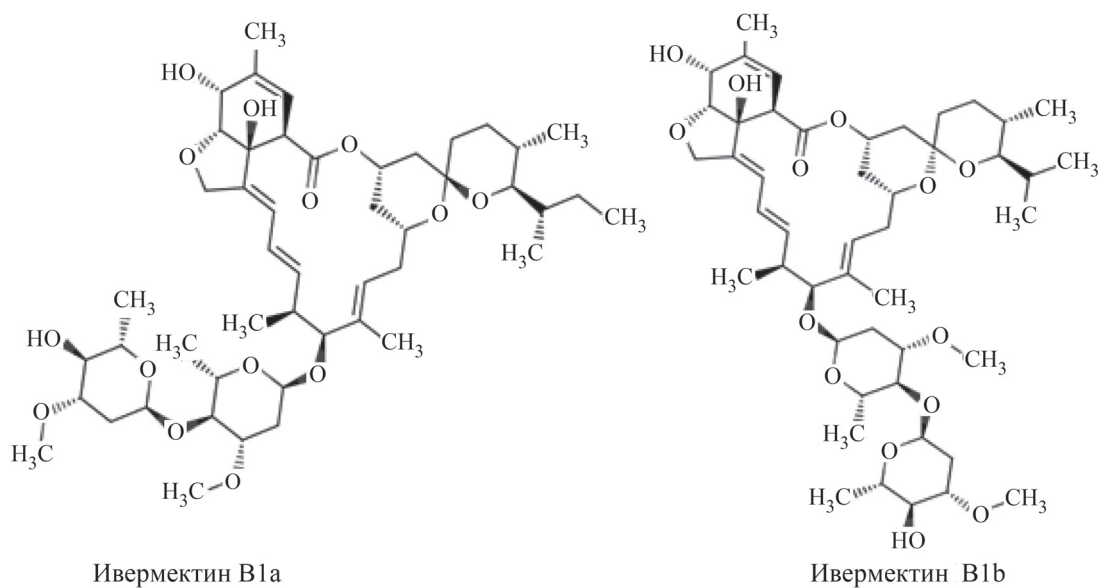


Рис. 2. Химическая структура природных авермектинов: Авермектина В2а и Авермектина В2b



4''-(ацетиламино)-5-О-диметил-4''-диоокси-авермектин А1а

Рис. 3. Химическая структура полусинтетических авермектинов: Ивермектина В1а, Ивермектина В1b, 4''-(ацетиламино)-5-О-диметил-4''-диоокси-авермектина А1а

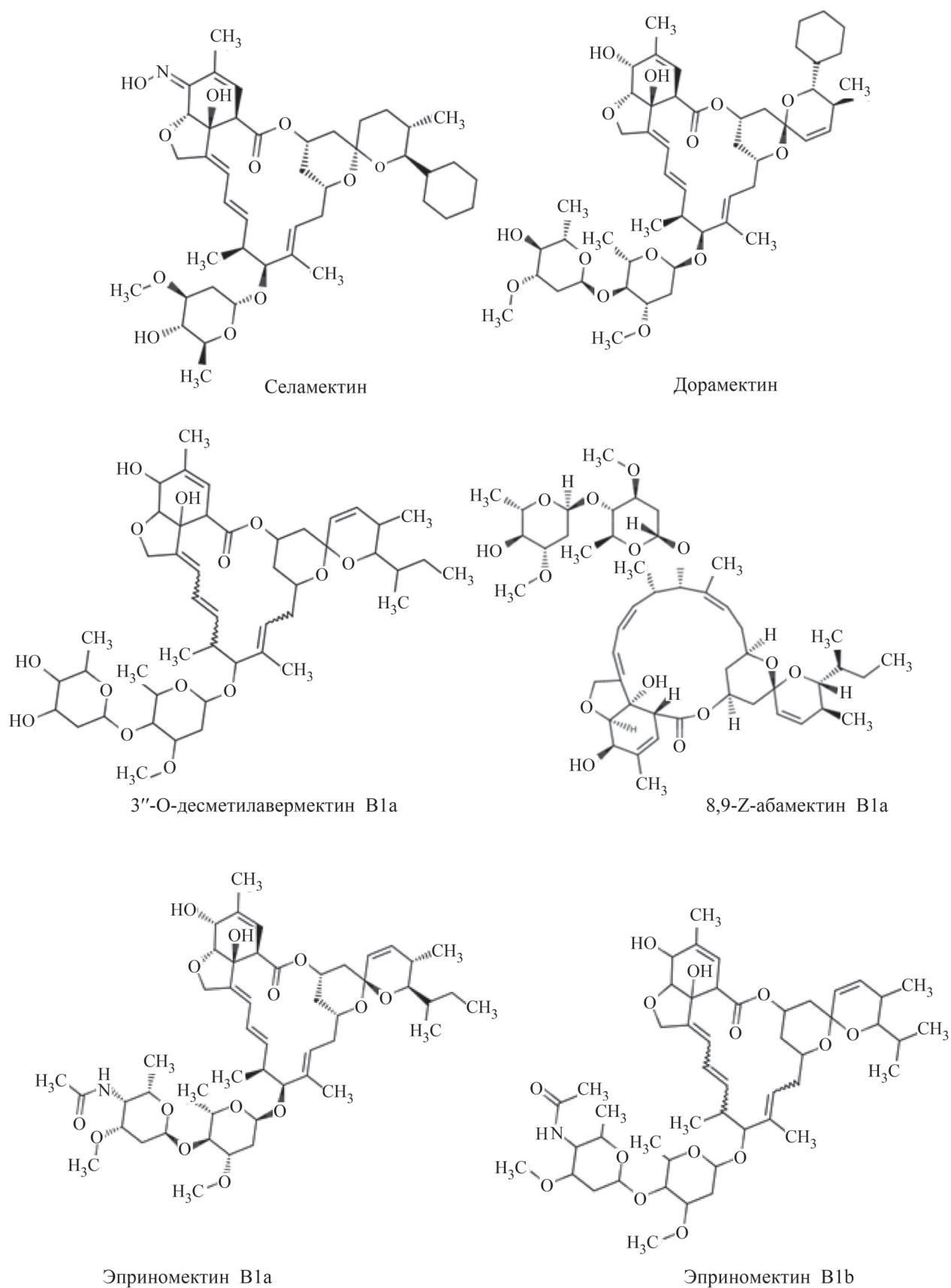


Рис. 4. Химическая структура полусинтетических авермектинов: Селамектина, Дорамектина, 3''-О-десметилавермектина В1а, 8,9-Z-абамектина В1а, Эприномектина В1а, Эприномектина В1b

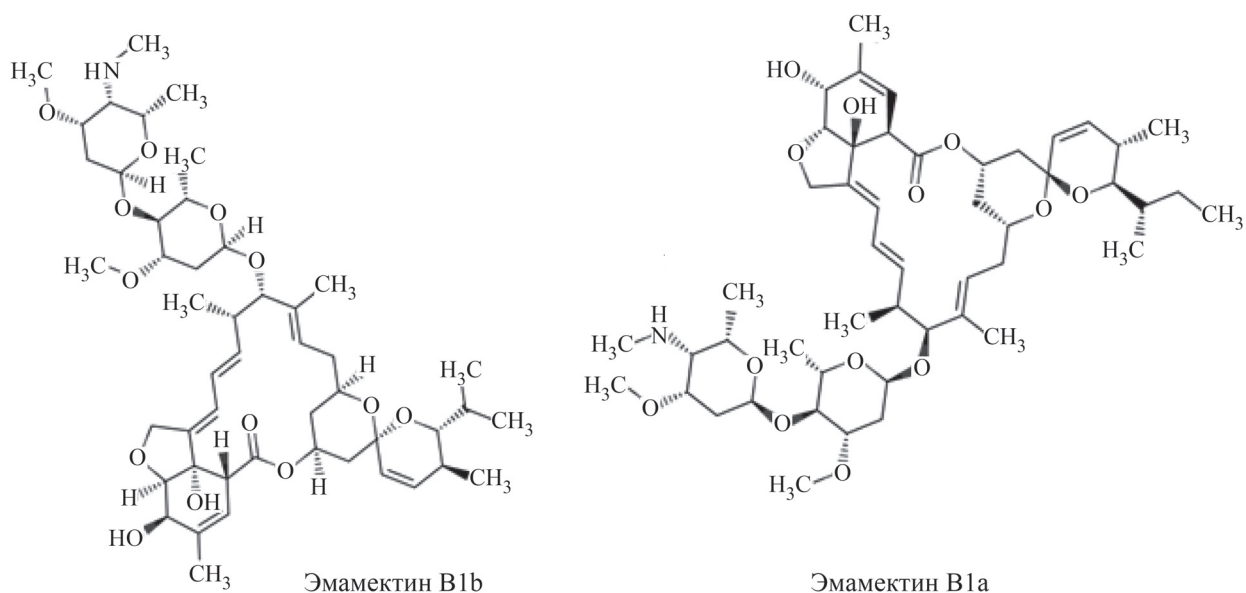


Рис. 5. Химическая структура полусинтетических авермектинов: Эмамектин В1b, Эмамектин В1а

рецепторы. К счастью, клетки нервной системы млекопитающих благодаря гематоэнцефалическому барьеру (ГЭБ) в некоторой степени защищены от воздействия авермектинов [9].

В большинстве европейских стран законодательно определена необходимость предоставления информации об основных экологических параметрах (например, биodeградации) химических веществ, производимых или импортируемых свыше 1 т в течение одного года [10]. Повсеместное использование авермектинов определяет потребность в их детальной экотоксикологической оценке. Основные экологические риски авермектинов связаны с их воздействием на нецелевые организмы. Например, при выделении пролонгированной формы ивермектина с фекалиями телят было отмечено уменьшение популяции насекомых, необходимых для естественной переработки навоза, что послужило причиной отмены в 2004 г. данной формы ивермектина в некоторых странах [9]. Таким образом, крайне важно при разработке новых лекарственных препаратов на основе авермектинов оценивать время их биodeградации, биотрансформации, а также иные экологические показатели потенциальной опасности. С учетом того, что большую часть нашей планеты занимает водная среда, логичным выглядит оценка такого экологического показателя авермектинов как биотрансформация для водных организмов, в частности для рыб, которые составляют существенную часть рациона человека. В обзорной публикации [11] по экологической оценке авер-

мектинов отмечается, что экспериментальных данных по экотоксикологической оценке, к сожалению, крайне мало, несмотря на распространенность указанных веществ по всему миру и их эффективное использование.

В настоящее время ввиду значительных финансовых и временных затрат при оценке экспериментальных показателей химических веществ, а также с учетом требований по этическому отношению к животным значительное внимание исследователей привлекают методы количественной связи структура – активность/свойство (QSAR/QSPR) [12, 13].

С учетом актуальности вопроса исследования биотрансформации органических веществ в ряде публикаций [14–16] предложены удовлетворительные QSAR-модели, количественно описывающие связь между структурой соединений и их способностью к биотрансформации в организмах рыб. Значительная часть этих моделей интегрирована в известные программы, например «EPI Suite V.4.1» [17] и «CompTox Chemistry Dashboard» [18], разработанные специалистами Агентства по охране окружающей среды США (United States Environmental Protection Agency, U.S. EPA). Так, интегрированная в «EPI Suite V.4.1» модель, подробно освещенная в публикации [14], имеет хорошие статистические параметры, но авторами модели, к сожалению, четко не определен алгоритм расчета вхождения соединений в область применимости (applicability domain, AD) модели [19], что является обязательным условием

QSAR-моделирования в соответствии с принципами, предложенными экспертной группой OECD (Organization for Economic Cooperation and Development). К сожалению, природные и полусинтетические авермектины не входят в область применимости другой приемлемой QSAR-модели [16], реализованной в свободном доступе через интернет-приложение «CompTox Chemistry Dashboard» [18].

Кроме того, в соответствии с пятым принципом QSAR-моделирования, разработанным экспертной группой OECD, желательна интерпретация полученных моделей. Для значительной части QSAR-моделей отсутствует структурная интерпретация, что не позволяет проводить рациональный молекулярный дизайн для получения соединений с заданной степенью биотрансформации. В рассмотренной выше публикации [16] проведена краткая структурная интерпретация, заключающаяся в сравнительном анализе использованных 12 дескрипторов и закономерностей, выявленных в рукописи [20], посвященной исследованию влияния структурных факторов на пути биотрансформации амидов. Авторы [10] пришли к выводу, что кольцевые структуры понижают способность соединений к биотрансформации, в то время как неразветвленные алкильные цепи, содержащие более 3 атомов углерода, а также бензильные и толильные группы способствуют окислению вместо реакций N-деалкилирования. Выявленные структурные особенности достаточно ограничены, что связано с незначительным структурным разнообразием в рассмотренном выше исследовании [20] и небольшим числом дескрипторов, использованных в публикации [16].

Таким образом, в задачи данного исследования входило:

1) построение QSAR-модели биотрансформации органических соединений, область применимости которой позволяет надежно прогнозировать данную активность для природных и полусинтетических авермектинов;

2) структурная интерпретация построенной QSAR-модели в целях выявления молекулярных фрагментов, наиболее влияющих на скорость биотрансформации органических соединений в организмах рыб.

Методы исследований

Находящиеся в свободном доступе обучающая (ws) и внешняя контрольная (ts) выборки с экспериментальными значениями биотрансформа-

ции органических соединений в организмах рыб были экспортированы из вышеуказанной публикации [16], где подробно описана процедура проверки исходной информации, размещенной в базе данных «PHYSPROP (The Physical Properties Database)». В качестве способности к биотрансформации органических соединений в организмах рыб авторами был выбран такой количественный показатель, как период полубиотрансформации (biotransformation half-life, KmHL) – число дней, необходимых для биотрансформации половины исследуемого вещества. Этот параметр рассчитывается по константе скорости биотрансформации (biotransformation rate constant, kM). Подробное описание использованных фармакокинетических параметров приведено в рассмотренной выше публикации [15].

Обучающий набор включает 405 соединений, внешняя контрольная выборка – 136 соединений. Диапазон периода полубиотрансформации (KmHL), выраженный через десятичный логарифм (Log KmHL), составляет от –1,6 до 3,0. Необходимость нелинейного преобразования моделируемого параметра (KmHL) обусловлена тем, что экспериментальные данные в исследуемой выборке более равномерно распределены в логарифмическом масштабе (рис. 6, 7), что также находится в соответствии с общепринятой практикой математической предобработки данных при построении моделей «структура – активность/свойство» [21]. Молярная масса исследуемых соединений варьирует в пределах от 68,8 до 959,2 г/моль.

Набор данных структурно неоднороден и включает в себя разнообразные соединения, такие как галогенуглеводороды (полихлорированные дифенилы, диоксины), алифатические и ароматические углеводороды (полициклические ароматические углеводороды), амины, имиды, спирты, фенолы, эфиры, кетоны и сложные эфиры. Для предварительной оценки корректности применения разрабатываемых QSAR-моделей в целях прогнозирования способности авермектинов к биотрансформации было проанализировано химическое пространство обучающей выборки. Одиннадцать соединений обучающей выборки, имеющие максимальное структурное сходство с авермектинами, представлены в табл. 1. В качестве референтного вещества среди авермектинов было выбрано соединение Авермектин А1а. Структурное сходство соединений обучающей выборки определяли путем расчета индексов Танимото и Дайса, используя топологические

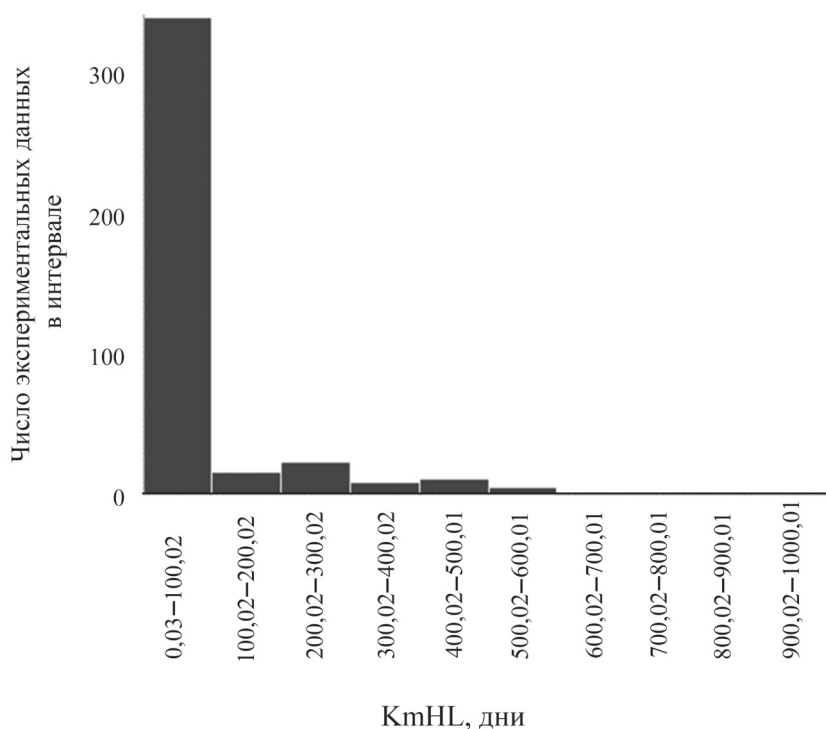


Рис. 6. Гистограмма распределения экспериментальных значений KmHL

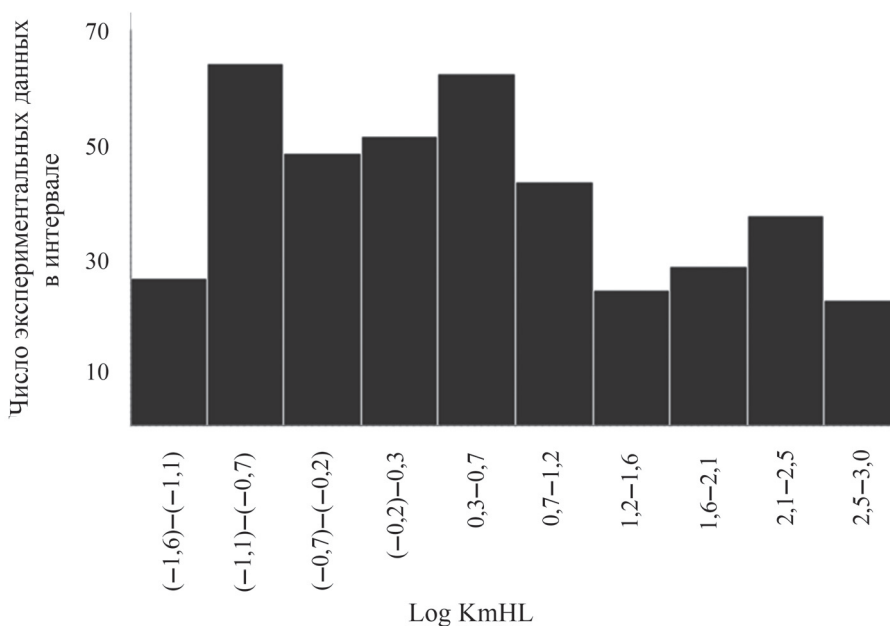


Рис. 7. Гистограмма распределения экспериментальных значений KmHL в логарифмической шкале

«отпечатки пальцев» (fingerprint), представляющие собой разновидность бинарных фрагментных дескрипторов. Расчеты осуществляли с помощью пакета RDKit для языка программирования Python [22]. Ранее было показано, что индексы Танимото и Дайса являются наиболее адекватной мерой определения молекулярного подобия [23].

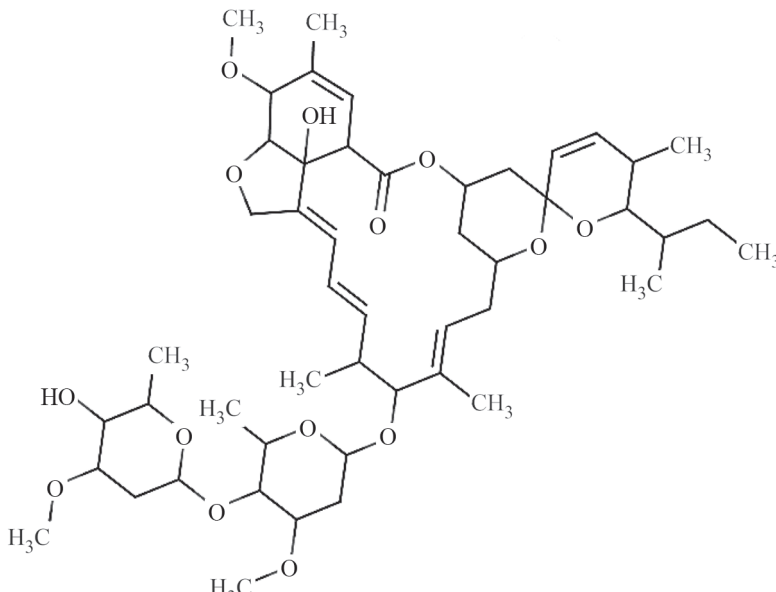
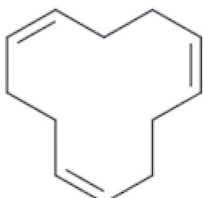
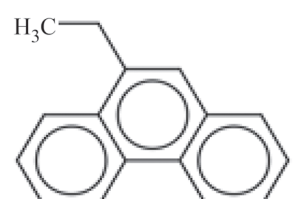

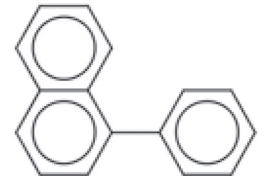
Из табл. 1 видно, что в обучающей выборке присутствуют макрополициклическое соеди-

нение 1,5,9-циклодекатриен, ряд полициклических ароматических углеводородов, а также 2,2',4,4',5,6-гексахлордифениловый эфир, в связи с чем можно предположить, что структурное пространство обучающей выборки позволит адекватно учитывать молекулярные особенности авермектинов при QSAR-моделировании.

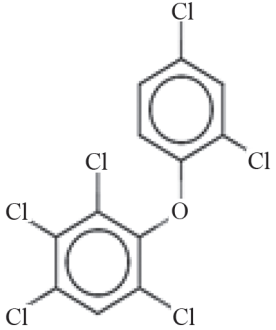
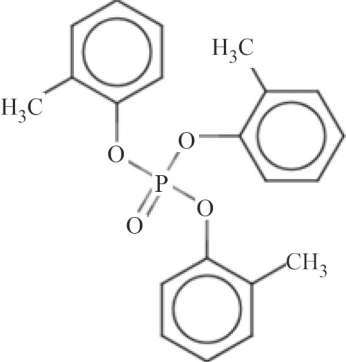
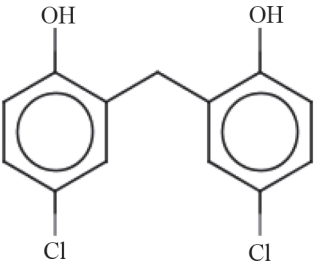
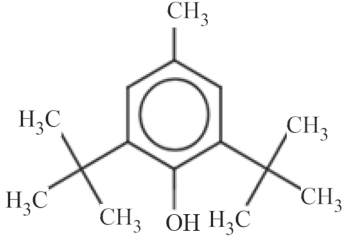
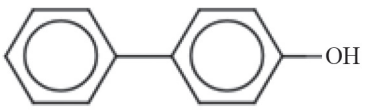
Нами было также проанализировано изменение скорости биотрансформации при

Т а б л и ц а 1

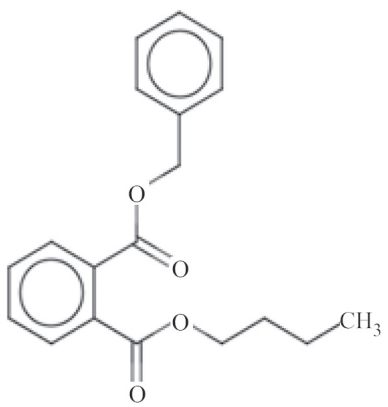
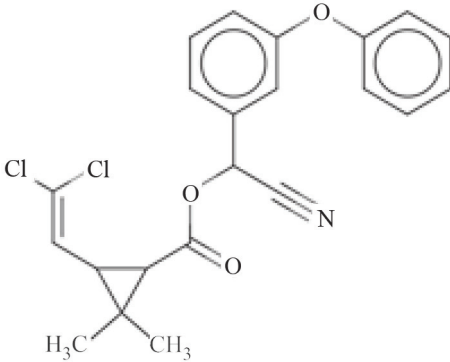
Соединения обучающей выборки, имеющие максимальное структурное сходство с Авермектином А1а

Структура соединения	CAS-номер	Индекс	
		Танимото	Дайса
			
	4904-61-4	0,41	0,58
	3674-75-7	0,53	0,69
	17088-22-1	0,53	0,69
	605-02-7	0,52	0,68

Продолжение табл. 1

Структура соединения	CAS-номер	Индекс	
		Танимото	Дайса
	106220-83-1	0,50	0,66
	78-30-8	0,45	0,62
	97-23-4	0,40	0,57
	128-37-0	0,39	0,57
	92-69-3	0,38	0,55

Окончание табл. 1

Структура соединения	CAS-номер	Индекс	
		Танимото	Дайса
	85-68-7	0,32	0,48
	52315-07-8	0,31	0,48

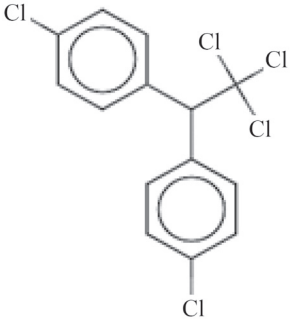
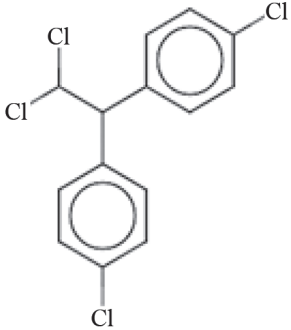
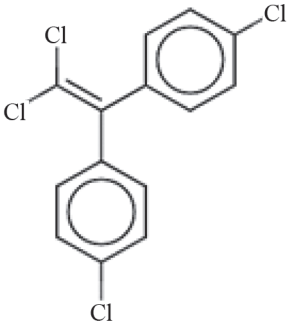
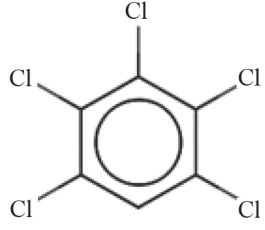
небольших структурных модификациях среди соединений обучающей выборки на примере производных дихлордифенилтрихлорметилметана и 1,2,3,4,5-пентахлорбензола (табл. 2). Так, для первого ряда рассматриваемых соединений (CAS-номера 50-29-3, 72-54-8, 72-55-9), незначительная модификация 4,4'-дихлордифенилтрихлорэтана в 4,4'-дихлордифенилтрихлорэтан приводит к уменьшению скорости биотрансформации более чем в три раза. При этом дегидрирование, в результате которого 4,4'-дихлордифенилтрихлорэтан превращается в 4,4'-дихлордифенилдихлорэтилен, наоборот, увеличивает скорость биотрансформации более чем в пять раз. Для другого ряда изучаемых соединений (CAS-номера 108-67-8, 117-18-0, 527-60-6) введение гидроксильной группы в 1,2,3,4,5-пентахлорбензол или замена хлор-группы в третьем положении данного соединения на нитрогруппу позволяет значительно увеличить скорость биотрансформации. Таким образом, можно надеяться, что эта обучающая

выборка позволит корректно спрогнозировать гипотетическое изменение скорости биотрансформации авермектинов, детерминированное незначительными различиями в структурах изучаемых 16-членных макролидов.

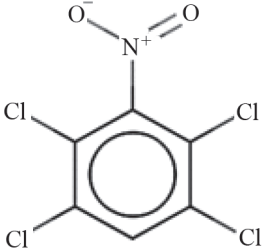
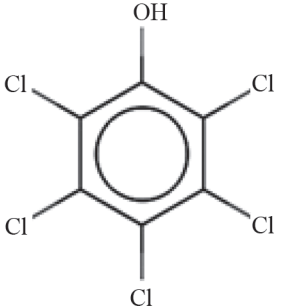
QSAR-моделирование проводили с помощью интернет-ресурса «OCHEM» [24]. Данный сайт содержит значительный объем информации о структурах химических соединений, а также о различных типах биологической активности и физических свойствах, позволяет создавать надежные QSAR/QSPR-модели с использованием различных методов машинного обучения и молекулярных дескрипторов. Для исследуемого набора данных наилучшие результаты были получены на основе дескрипторов PyDescriptor [25], а также алгоритмов случайного леса (Random forest, RF) [26] и трансформаторной сверточной нейронной сети (Transformer convolutional neural network, Trans-CNN) [27]. Предварительно все соединения были стандартизированы с помощью встроенного в «OCHEM» плагина ChemAxon

Т а б л и ц а 2

Изменения скорости биотрансформации в рядах производных дихлордифенилтрихлорметилметана и 1,2,3,4,5-пентахлорбензола

Структура соединения	CAS-номер	Экспериментальное значение KmHL, число дней
 <chem>ClC1=CC=C(C=C1)C(CCl)C(Cl)C2=CC=C(Cl)C=C2</chem>	50-29-3	85,11
 <chem>ClC1=CC=C(C=C1)C(CCl)C(Cl)C2=CC=C(Cl)C=C2</chem>	72-54-8	275,42
 <chem>ClC1=CC=C(C=C1)C=C(Cl)C(Cl)C2=CC=C(Cl)C=C2</chem>	72-55-9	51,29
 <chem>ClC1=C(Cl)C(Cl)=C(Cl)C1Cl</chem>	108-67-8	74,13

Окончание табл. 2

Структура соединения	CAS-номер	Экспериментальное значение KmHL, число дней
	117-18-0	3,09
	527-60-6	0,63

Standardizer [28] в соответствии с общепринятыми протоколами [29, 30]. Программный пакет PyDescriptor позволяет вычислять набор из 11 145 разнообразных молекулярных дескрипторов, при этом на 1D-уровне представления молекулярной структуры вычисляются молекулярный вес, число атомов определенного типа, а также ряд других широко используемых параметров при данном уровне детализации химических соединений; на 2D-уровне представления молекулярной структуры рассчитываются различные физико-химические дескрипторы, например дескрипторы водородной связи, а также разнообразные фрагментарные дескрипторы, в том числе «отпечатки пальцев»; на 3D-уровне описания молекулярной структуры генерируются разнообразные дескрипторы, отражающие особенности пространственной структуры (например, взвешенную величинами атомных зарядов положительно заряженную поверхность). При расчете дескрипторов 3D-уровня детализации структуры соединений были оптимизированы с помощью модуля программы Corina [31], также интегрированного в платформу OCHEM. Все рассчитанные дескрипторы были отфильтрованы с использованием протокола предварительной фильтрации,

предложенного в [32]: 1) удалены дескрипторы, которые являются постоянными среди всех исследуемых веществ; 2) исключены нормализованные дескрипторы, имеющие дисперсию $<0,01$; 3) сгруппированы дескрипторы, которые были взаимно коррелированы с коэффициентом линейной корреляции $R^2 > 0,95$, и только первый дескриптор из группы был выбран для разработки модели. В результате был сформирован ряд, состоящий из 1222 дескрипторов.

Для внутренней валидации обучающей выборки использовали процедуру перекрестного контроля (5-fold cross-validation). С учетом структурного разнообразия обучающей выборки было принято решение первоначально провести расчеты с использованием различных методов, а затем использовать данные модели для консенсусного QSAR-моделирования. Ранее было показано, что данный подход имеет в среднем более высокую надежность, чем любая из отдельных моделей [33, 34]. Разработка консенсусной модели осуществлялась путем усреднения прогнозов отдельных моделей.

Оценку AD проводили с использованием концепции «расстояния до модели», в частности подхода «CONSENSUS-STD» (Стандартное откло-

нение ансамблевых прогнозов) [35]. Предсказание целевой активности для молекулы внешней контрольной выборки в данном подходе рассчитывается как среднее значение, полученное в результате применения ансамбля моделей. При этом расстояние до модели (d_{STD}) определяется следующим образом:

$$d_{STD}(B) = \sqrt{\frac{\sum (y_i(B) - \bar{y})^2}{N - 1}}. \quad (1)$$

где $y(B) = \{y_i(B), i = 1 \dots N\}$ – набор предсказаний для соединения B , заданный набором из N обученных моделей; \bar{y} – среднее значение спрогнозированной активности для соединения B .

В соответствии с алгоритмом «CONSENSUS-STD», надежные прогнозы имеют низкий уровень d_{STD} . Этот подход был признан наиболее надежным для оценки точности

прогнозов в ряде сравнительных исследований [35–37].

Качество прогнозов на основе QSAR-моделей оценивалось коэффициентом детерминации для перекрестного контроля (R^2_{cv}), квадратом коэффициента корреляции для внешней контрольной выборки (R^2_{test}) и среднеквадратичной ошибкой (RMSE), расчет которых реализован в «OCHEM». Подробное описание использованных методов валидации, алгоритма определения области применимости QSAR-модели, а также формулы расчета указанных статистических критериев приведены в руководстве пользователя «OCHEM» [24].

Результаты и обсуждение

Результаты QSAR-моделирования показаны в табл. 3, где также в целях сравнительного анализа приведены статистические показатели QSAR-модели, построенной для исследуемой выборки специалистами U.S. EPA [16]. Полученные нами

Т а б л и ц а 3

Статистические характеристики QSAR-моделей

Модель	R^2_{cv}	RMSE _{cv}	R^2_{test}	RMSE _{test}
RFR_PyDescriptor	0,78	0,56	0,76	0,58
Trans-CNN	0,78	0,55	0,79	0,55
Consensus	0,80	0,53	0,81	0,52
kNN_PaDEL [16]	0,83	0,49	0,73	0,62

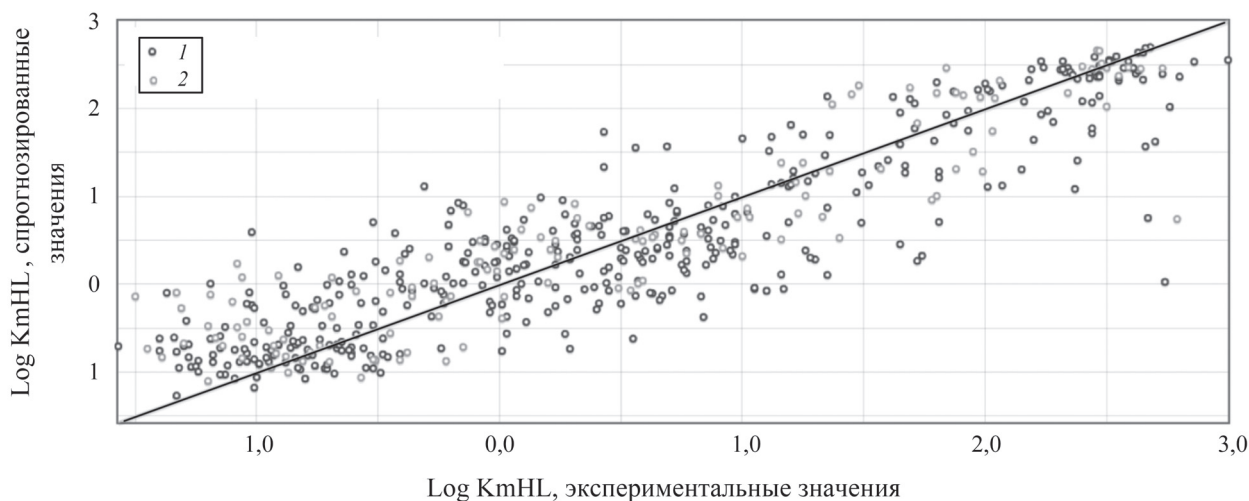

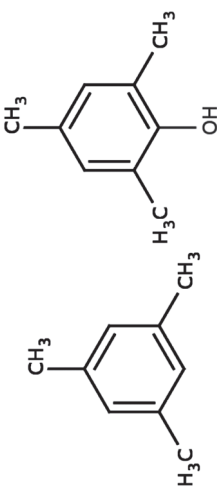

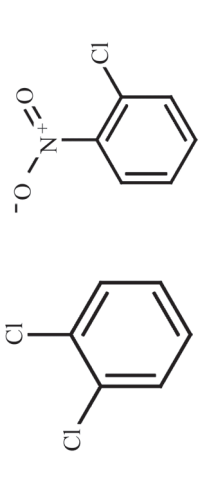
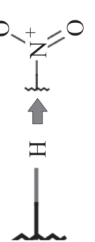
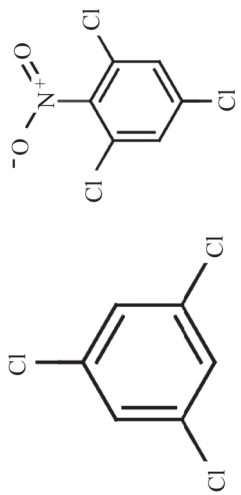



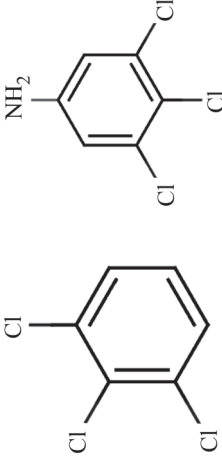
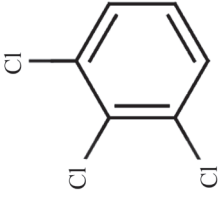


Рис. 8. Сравнение наблюдаемых (экспериментальных) и спрогнозированных консенсусной моделью значений биотрансформации соединений обучающей выборки (1) в рамках перекрестного контроля и соединений тестовой выборки (2)

Таблица 4

Результаты структурной интерпретации

Молекулярная трансформация	Пример пары соединений из обучающей выборки и их экспериментальные значения активности	N	Δmean	SEM	$ t_{\text{calc}} $	SMIRKS
Молекулярные трансформации, повышающие скорость биотрансформации соединений в организмах рыб						
	 <p>Log KmHL = 0,4 Log KmHL = -1,3</p>	10	-1,00	0,26	3,9	*[H] → O*
	 <p>Log KmHL = 0,5 Log KmHL = -0,58</p>	14	-0,90	0,24	3,7	*Cl → [O-][N+](*)=O
	 <p>Log KmHL = 0,92 Log KmHL = -0,19</p>	14	-0,58	0,23	2,5	*[H] → [O-][N+](*)=O

Окончание табл. 4

Молекулярная трансформация	Пример пары соединений из обучающей выборки и их экспериментальные значения активности	N	Δmean	SEM	$ r_{\text{calc}} $	SMIRKS
	  Log KmHL = 0,92 Log KmHL = 0,28	11	-0,52	0,24	2,1	*[H] → N*
Молекулярные трансформации, понижающие скорость биотрансформации соединений в организмах рыб						
	$\text{H}_3\text{C}-(\text{CH}_2)_{10}-\text{CH}_3$ $\text{H}_3\text{C}-(\text{CH}_2)_{10}-(\text{CH}_2)_2-\text{CH}_3$ Log KmHL = 0,26 Log KmHL = 1,49	18	0,40	0,12	3,2	*C* → *CCCC*
	$\text{H}_3\text{C}-(\text{CH}_2)_{10}-\text{CH}_3$ $\text{H}_3\text{C}-(\text{CH}_2)_{10}-(\text{CH}_2)_4-\text{CH}_3$ Log KmHL = 0,26 Log KmHL = 1,47	20	0,74	0,09	8,5	*CC* → *CCCCCC*

Примечание: N – число молекулярных пар, для которых выполняется молекулярная трансформация; Δmean – средняя разница значений Log KmHL при проведении молекулярной трансформации; SEM – стандартная ошибка средней величины Log KmHL; SMIRKS – формат записи молекулярной трансформации (<https://www.daylight.com/dayhtml/doc/theory/theory.smirks.html>).

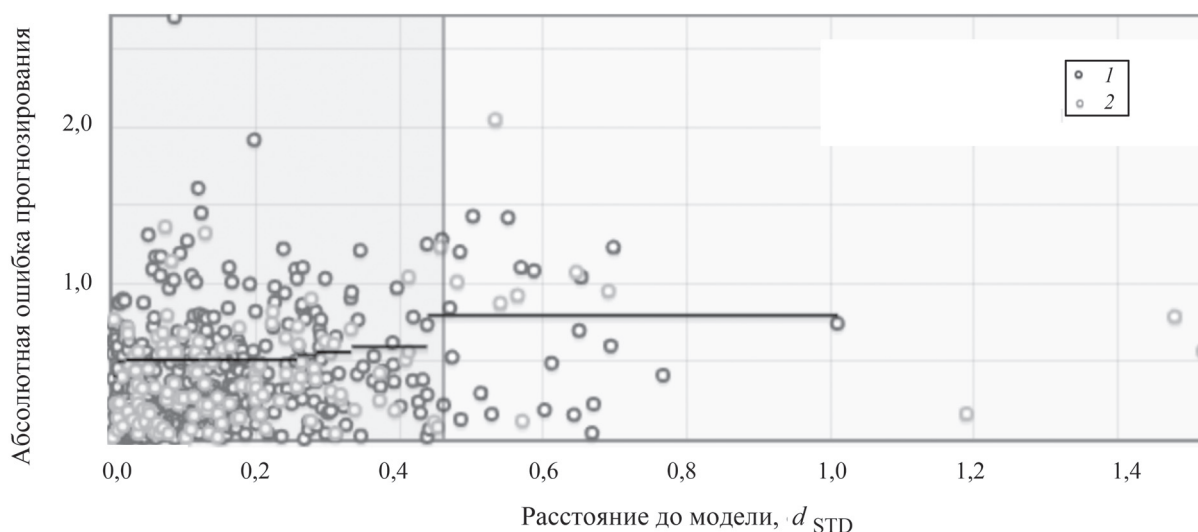


Рис. 9. Диаграмма зависимости абсолютного значения ошибки прогнозирования от «расстояния до модели» для соединений обучающей выборки (1) в рамках перекрестного контроля и соединений тестовой выборки (2)

модели обладают сопоставимой предсказательной способностью с предложенной ранее моделью, которая была разработана с использованием PaDEL-дескрипторов и метода взвешенных k -ближайших соседей. Сравнение экспериментальных и предсказанных значений консенсусной QSAR-моделью представлено на рис. 8.

Все соединения тестовой выборки вошли в область применимости консенсусной QSAR-модели. На рис. 9 представлена диаграмма, подтверждающая использование адекватной меры для определения области применимости модели. Абсцисса на данной диаграмме соответствует расстоянию до модели, ордината – абсолютному значению ошибки прогнозирования для каждого из анализируемых веществ, а каждая точка – одному из химических веществ. Для визуализации зависимости между данными параметрами полный диапазон изменения значений «расстояния до модели» разбивают автоматически на несколько интервалов, содержащих одинаковое число объектов, и для каждого из них проводят горизонтальный отрезок на высоте, равной среднему значению абсолютной ошибки прогнозирования для входящих в него химических объектов. Подтверждением адекватности используемой меры может быть расположение данных отрезков в виде «лестницы», в которой каждый правый отрезок («ступенька») должен находиться выше левого. При использовании неадекватной меры определения области применимости, как правило, на диаграммах данного типа все отрезки («ступеньки») находятся примерно на одном уровне [21].

Для полученной консенсусной модели была проведена структурная интерпретация с помо-

щью метода молекулярных пар, также реализованного в «ОСЧЕМ» [38]. Результаты интерпретации, в частности молекулярные трансформации, которые в наибольшей степени повышают или понижают биотрансформацию органических соединений в организмах рыб, приведены в табл. 4. Степень значимости (существенности) и устойчивости выявленных молекулярных трансформаций определяли с помощью одновыборочного t -критерия Стьюдента в соответствии с предложенной ранее методикой [39, 40] путем сравнения рассчитанного по формуле (2) t -критерия (t_{calc}) с критической величиной (t_{crit}) при уровне значимости $\alpha = 0,05$ и соответствующем числе степеней свободы.

$$t_{calc} = \Delta mean / SEM, (2)$$

где $\Delta mean$ – средняя разница значений Log KmHL при проведении молекулярной трансформации, SEM – стандартная ошибки средней величины Log KmHL.

Согласно [39], если $|t_{calc}| > t_{crit}$, то считается, что выявленная молекулярная трансформация является существенной и статистически значимой. Авторы [39] отмечают, что в большинстве практически важных случаев $t_{crit} \approx 2$ при $\alpha = 0,05$. В этой связи принято считать, что молекулярная трансформация будет статистически значима, если величина $\Delta mean$ по модулю в два раза больше, чем значение SEM, что также указано в публикации [40].

На основании результатов интерпретации можно предположить, что наибольший вклад в повышение биотрансформации происходит при замене атомов водорода и хлора

Т а б л и ц а 5

Прогнозируемые значения способности индивидуальных авермектинов к биотрансформации в организмах рыб

Название	CAS-номер	КмНЛ, число дней	d_{STD}	RMSE
Природные авермектины				
Авермектин А1а	65195-51-9	4,1	0,11	0,51
Авермектин А1б	65195-52-0	4,2	0,13	0,51
Авермектин В1а	65195-55-3	4,4	0,07	0,51
Авермектин В1б	65195-56-4	4,5	0,07	0,51
Авермектин А2а	65195-53-1	3,3	0,11	0,51
Авермектин А2б	65195-54-2	3,3	0,17	0,51
Авермектин В2а	65195-57-5	3,9	0,10	0,51
Авермектин В2б	65195-58-6	3,1	0,00	0,4
Полусинтетические авермектины				
Ивермектин В1а	71827-03-7	7,7	0,30	0,56
Ивермектин В1б	70209-81-3	8,0	0,28	0,54
Дорамектин	117704-25-3	6,9	0,23	0,51
Селамектин	220119-17-5	7,3	0,13	0,51
3''-О-Десметилавермектин В1а	99965-56-7	14,1	0,07	0,51
8,9-Z-абамектин В1а	113665-89-7	7,0	0,22	0,51
4''-(ацетиламино)-5-О-диметил-4''-диокси- Авермектин А1а	172926-97-5	6,8	0,23	0,51
Эмамектин В1б	121424-52-0	6,5	0,26	0,54
Эмамектин В1а	121124-29-6	7,9	0,28	0,54
Эприномектин В1а	1142337-10-7	7,6	0,31	0,56
Эприномектин В1б	133305-89-2	10,0	0,00	0,40

на гидроксильную, нитро- и аминогруппу. Вероятно, это связано с тем, что наиболее частой реакцией в первой фазе метаболизма является окисление, при этом гидроксильное окисление ксенобиотиков с алифатическими радикалами приводит к образованию спиртов, которые также как и амины при дальнейшем окислении трансформируются в альдегиды и кетоны [41]. Таким образом, соединения, содержащие гидроксильную группу и аминогруппу, проходят меньше этапов окисления, что ускоряет их биотрансформацию. В соответствии с результатом структурной интерпретации увеличение длины алифатических неразветвленных радикалов понижает скорость биотрансформации органических соединений в организмах рыб.

Разработанная нами консенсусная модель доступна по ссылке <https://ochem.eu/model/16426620> и может быть использована при прогнозировании биотрансформации новых органических соединений в организмах рыб. С помощью этой модели были рассчитаны приведенные в табл. 5 показатели биотрансформации индивидуальных авермектинов в организмах рыб. Все соединения вошли в область применимости, что может свидетельствовать о достоверности прогнозируемых значений. На основании приведенных в табл. 5 данных можно предположить, что в организмах рыб биотрансформация природных авермектинов происходит быстрее, чем полусинтетических.

Так, при гидрировании природного Авермектина В1b (KmHL составляет примерно четыре с половиной дня) по двойной связи в положение С22, С23 образуется полусинтетическое производное – Ивермектин В1b, для которого значение периода полубиотрансформации увеличивается до 8 дней. Аминопроизводное природного Авермектина В1b, полусинтетический авермектин Эприномектин В1b, также обладает более продолжительным периодом полубиотрансформации (10 дней). На основании вышесказанного можно предположить, что введение ацетамидной группы в природные авермектины позволит увеличить продолжительность биотрансформации полусинтетических производных. К увеличению периода полубиотрансформации приводит также замена в природных авермектинах алифатического радикала в положении С25 на циклогексанный фрагмент (дорамектин, селамектин), что согласуется с закономерностями, изложенными в публикации [20]. И наоборот, замена гидроксильной группы в дорамектине на гидроксилламинную (селамектин) приводит к снижению времени биотрансформации.

Таким образом, использование природных индивидуальных авермектинов в качестве новых лекарственных средств более предпочтительно, так как они обладают более высокой скоростью биотрансформации в организмах рыб, соответственно, уменьшается вероятность их случайного употребления при использовании рыбы в пищу животных и человека.

Выводы

Разработана адекватная консенсусная QSAR-модель биотрансформации органических соединений в организмах рыб, обладающая сопоставимой предсказательной способностью с моделью, предлагаемой специалистами Агентства по охране окружающей среды США. Выбранные нами дескрипторы позволили расширить область применимости модели, что дало возможность провести более достоверное прогнозирование биотрансформации в организмах рыб индивидуальных авермектинов. Согласно полученным расчетным значениям, в организмах рыб биотрансформация природных авермектинов происходит быстрее, чем полусинтетических.

Проведенная структурная интерпретация консенсусной QSAR-модели позволила выявить общие закономерности влияния структуры на скорость биотрансформации, выделив молекулярные трансформации, которые существенно повышают и понижают скорость биотрансформации органических соединений. При анализе индивидуальных авермектинов определены структурные фрагменты, повышающие и понижающие скорость биотрансформации.

Результаты проведенного исследования могут способствовать снижению финансовых, временных и трудовых затрат при определении стратегии разработки новых лекарственных средств на основе индивидуальных авермектинов, в частности при оценке экотоксикологических параметров авермектинов. В дальнейших исследованиях целесообразно провести экспериментальную валидацию полученной модели, а также дать оценку другим экотоксикологическим характеристикам, например острой токсичности по отношению к водным организмам.

Часть работы выполнена в рамках Государственного задания ИФАВ РАН 2020 года (тема № 0090-2019-0004).

Конфликта интересов нет.

Дополнительных материалов нет.

Дополнительной информации нет.

СПИСОК ЛИТЕРАТУРЫ

1. Caly L., Druce J.D., Catton M.G., Jans D.A., Wagstaff K.M. // *Antiviral Res.* 2020. Vol. 178. P.104787.
2. Chaccour C., Casellas A., Blanco-Di Matteo A., Pineda I., Fernandez-Montero A., Ruiz-Castillo P., Richardson M.-A., Rodriguez-Mateos M., Jordan-Iborra C., Brew J., Carmona-Torre F., Giraldez M., Laso E., Gabaldon-Figueira J. C., Dobano C., Moncunill G., Yuste J. R., Del Pozo J. L., Rabinovich R.N., Schoning V., Hammann F., Reina G., Sadaba B., Fernandez-Alonso M. // *EClinicalMedicine.* 2021. Online first, 100720 (<https://doi.org/10.1016/j.eclinm.2020.100720>)
3. Wagstaff K.M., Sivakumaran H., Heaton S.M., Harrich D., Jans D.A. // *Biochem. J.* 2012. Vol. 443. P. 851
4. Crump A., Omura S. // *Proc. Jpn. Acad. Ser. B, Phys. Biol. Sci.* 2011. Vol. 87. N 2. P.13.
5. Zhuo Y., Zhang T., Wang Qi, Cruz-Morales P., Zhang B., Liu M., Barona-Gómez F., Zhang L. // *Biotechnol. J.* 2014. Vol. 9. P. 316.
6. Мосин В.А., Кругряк Е.Б., Стерлина Т.С., Корыстов Ю.Н., Шапошникова В.В., Кублик Л.Н., Левитман М.Х., Викторов А.В., Дриняев В.А. // *Антибиотики и химиотерапия.* 1999. Т. 44. № 6. С. 16.
7. Дриняев В.А., Мосин В.А., Кругряк Е.Б., Стерлина Т.С., Новик Т.С., Ермакова Н.В., Кублик Л.Н., Левитман М.Х., Шапошникова В.В., Корыстов Ю.Н. // *Антибиотики и Химиотерапия.* 2004. Т. 49. № 6. С. 3.
8. Juarez M., Schcolnik-Cabrera A., Dueñas-Gonzalez A. // *Am. J. Cancer Res.* 2018. Vol. 8. P.317.
9. Prichard R., Ménez C., Lespine A. // *Int. J. Parasitol. Drugs Drug Resist.* 2012. Vol. 2. P. 134.
10. Mansouri K., Ringsted T., Ballabio D., Todeschini R., Consonni V. Quantitative structure-activity relationship models for ready biodegradability of chemicals // *J. Chem Inf Model.* 2013. Vol. 53. P. 867.
11. Bai S.H., Ogbourne S. // *Chemosphere.* 2016. Vol. 154. P. 204.
12. Muratov E.N., Bajorath J., Sheridan R.P., Tetko I.V., Filimonov D., Poroikov V., Oprea T.I., Baskin I.I., Varnek A., Roitberg A., Isayev O., Curtalolo S., Fourches D., Cohen Y., Aspuru-Guzik A., Winkler D.A., Agrafiotis D., Cherkasov A., Tropsha A. // *Chem Soc Rev.* 2020. Vol. 49. N 11. P. 3525.
13. Тиньков О.В., Григорьев В.Ю., Полищук П.Г., Ярков А.В., Раевский О.А. // *Биомедицинская химия.* 2019. Т. 65. С. 123.
14. Arnot J.A., Meylan W., Tunkel J., Howard P., Mackay D., Bonnell M., Boethling R.S. // *Environ Toxicol Chem.* 2009. Vol. 28. P.1168.
15. Papa E., van der Wal L., Arnot J.A., Gramatica P. // *Sci Total Environ.* 2014. Vol. 470. P.1040.
16. Mansouri K., Grulke C.M., Judson R.S., Williams A.J. // *J. Cheminform.* 2018. Vol. 10. P.1.
17. URL: <https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface>.
18. URL: <https://comptox.epa.gov/dashboard>.
19. Jaworska J., Nikolova-Jeliazkova N., Aldenberg T. // *Altern Lab Anim.* 2005. Vol. 33. N 5. P.445.
20. Helbling D.E., Hollender J., Kohler H.P.E., Fenner K. // *Environ Sci Technol.* 2010. Vol. 44. N 17. P. 6628
21. Баскин И.И., Маджидов Т.И., Варнек А.А. Введение в хемоинформатику. Ч.3. Моделирование «структура – свойство». Казань, 2015. 304 с.
22. URL: <https://www.rdkit.org/docs/GettingStartedInPython.html>
23. Bajusz D., Rácz A., Héberger K. // *J. Cheminform.* 2015. Vol. 7. Article number: 20.
24. URL: <https://ochem.eu/>
25. Masand V.H., Rastija V. // *Chemom. Intell. Lab. Syst.* 2017. Vol. 169. P. 12.
26. Breiman L. // *Machine Learning.* 2001. Vol. 45. P. 5.
27. Karpov P., Godin G., Tetko I.V. // *J. Cheminform.* 2020. Vol. 12. Article number 17.
28. URL: <https://docs.chemaxon.com/display/docs/Standardizer+Command-line+Application>
29. Fourches D., Muratov E., Tropsha A. // *J. Chem Inf Model.* 2010. Vol. 50. P. 1189.
30. Fourches D., Muratov E., Tropsha A. // *J. Chem Inf Model.* 2016. Vol. 56. P. 1243.
31. Sadowski J., Gasteiger J., Klebe G. // *J. Chem. Inf. Comput. Sci.* 1994. Vol. 34. P.1000.
32. Tetko I.V., Lowe MD., Williams A.J. // *J. Cheminform.* 2016. Vol. 8. Article number: 2.
33. Golbraikh A., Muratov E., Fourches D., Tropsha A. // *J Chem Inf Model.* 2014. Vol. 54.
34. Zakharov A.V., Zhao T., Nguyen D.T., Peryea T., Sheils T., Yasgar A., Huang R., Southall N., Simeonov A. // *J. Chem Inf Model.* 2019. Vol. 59. P. 4613.
35. Tetko I.V., Sushko I., Pandey A.K., Zhu H., Tropsha A., Papa E., Oberg T., Todeschini R., Fourches D., Varnek A. // *J. Chem Inf Model.* 2008. Vol. 48. P. 1733.
36. Sushko I., Novotarskyi S., Körner R., Pandey A.K., Kovalishyn V.V., Prokopenko V.V., Tetko I.V. // *J. Chemom.* 2010. Vol. 24. P. 202.
37. Sushko I., Novotarskyi S., Körner R., Pandey A.K., Cherkasov A., Li J., Gramatica P., Hansen K., Schroeter T., Muller K.R., Xi L., Liu H., Yao X., Oberg T., Hormozdiari F., Dao P., Sahinalp C., Todeschini R., Polishchuk P., Artemenko A., Kuz'min V., Martin T.M., Young D.M., Fourches D., Muratov E., Tropsha A., Baskin I., Horvath D., Marcou G., Muller C., Varnek A., Prokopenko V.V., Tetko I.V. // *J. Chem Inf Model.* 2010. Vol. 50. P. 2094.
38. Sushko Y., Novotarskyi S., Korner R., Vogt J., Abdelaziz, A., Tetko I. // *J. Cheminf.* 2014. Vol. 6. P. 48.
39. Маджидов Т.И., Баскин И.И., Варнек А.А. Введение в хемоинформатику. Ч. 6. Химическое пространство и виртуальный скрининг. Казань, 2019. 238 с.
40. Griffen E., Leach A.G., Robb G.R., Warner D.J. // *J. Med Chem.* 2011. Vol. 54. P. 7739.
41. Вергейчик Т.Х. Токсикологическая химия. М., 2012. 432 с.

Поступила в редакцию 11.03.2021
Получена после доработки 14.03.2021
Принята к публикации 20.03.2021

PREDICTION OF THE BIOTRANSFORMATION TIME OF ORGANIC COMPOUNDS BY THE EXAMPLE OF AVERMECTINS

O.V. Tinkov^{1,2*}, V.Y. Grigorev³, L.D. Grigoreva⁴

*(¹Department of Pharmacology and Pharmaceutical Chemistry, Medical Faculty, Shevchenko Transnistria State University, Tiraspol, Moldova; ²Military Institute of the Ministry of Defense, Tiraspol, Moldova; ³Institute of Physiologically Active Compounds of the Russian Academy of Science, Chernogolovka, Russia; ⁴Department of Fundamental Physical-Chemical Engineering, Lomonosov Moscow State University; *e-mail: oleg.tinkov.chem@mail.ru)*

Currently, the spread of the SARS-CoV-2 coronavirus is a significant problem for all of humanity. One of the promising agents for the fight against the SARS-CoV-2 coronavirus is ivermectin, which is a complex of semi-synthetic derivatives of natural avermectins that have been effectively used in medicine and agriculture for many years as antiparasitic drugs. However, experimental data on the ecotoxicological assessment of individual avermectins are still very small. In this regard, the purpose of this study was to develop a mathematical model that would allow for a reliable prediction of the biotransformation ability of natural and semi-synthetic avermectins, as well as to identify the structural fragments of molecules that most affect the manifestation of this property. The basis for the construction of the model was a structurally heterogeneous set represented by an organic compound with experimental values of the biotransformation half-life (KmHL). Using the web platform OCHEM (<https://ochem.eu>), in which the calculation of PyDescriptor descriptors and the Random Forest method, Transformer-CNN are implemented, a satisfactory ($R^2_{\text{test}} = 0.81$) Quantitative Relationship Structure – Activity (QSAR) model was developed. The calculations showed that the biotransformation of natural avermectins in fish occurs on average faster than semi-synthetic ones. In addition, structural fragments that increase and decrease the rate of biotransformation have been identified.

Key words: macrolides, molecular descriptors, machine learning, QSAR.

Сведения об авторах: Тиньков Олег Викторович – доцент кафедры фармакологии и фармацевтической химии медицинского факультета Приднестровского государственного университета; заместитель начальника учебной части Военного института им. генерал-лейтенанта А.И. Лебедея Министерства обороны ПМР, канд. хим. наук (oleg.tinkov.chem@mail.ru); Григорьев Вениамин Юрьевич – зав. отделом компьютерного молекулярного дизайна ИФАВ РАН, докт. хим. наук (beng@ipac.ac.ru); Григорьева Людмила Дмитриевна – заместитель декана по учебной работе, доцент факультета фундаментальной физико-химической инженерии МГУ имени М.В. Ломоносова, канд. физ.-матем. наук (ldg@physchem.msu.ru).