

Prediction of an Organic Compound's Biotransformation Time: A Study Using Avermectins

O. V. Tinkov^{a, b, *}, V. Y. Grigorev^c, and L. D. Grigoreva^d

^a Department of Medicine, Shevchenko Transnistria State University, Tiraspol, Moldova

^b Lebed Military Institute, Ministry of Defense, Tiraspol, Moldova

^c Institute of Physiologically Active Compounds, Russian Academy of Science, Chernogolovka, Russia

^d Department of Fundamental Physical-Chemical Engineering, Moscow State University, Moscow, Russia

*e-mail: oleg.tinkov.chem@mail.ru

Received March 11, 2021; revised March 14, 2021; accepted March 20, 2021

Abstract—The current spread of the SARS-CoV-2 coronavirus is a challenge for the entire world. Ivermectin is a promising agent, which could be used to combat the SARS-CoV-2 coronavirus. It represents a complex of semisynthetic derivatives of natural avermectins that have been taken advantage of for a long time in medicine and agriculture as antiparasitic drugs. However, the experimental ecotoxicology assessment data for individual avermectins are still scarce. In relation to this, the aim of this study is to develop a mathematical model that would allow reliably predicting the biotransformation ability of natural and semisynthetic avermectins and identifying the structural fragments of avermectin molecules that have the largest impact on this biological activity. The base for the model construction was a structurally heterogeneous set including organic compounds with experimentally determined biotransformation half-life periods (KmHL). Using the OCHEM web platform (<https://ochem.eu>) with the implemented PyDescriptor plugin for the descriptor calculation and Random Forest and Transformer-CNN algorithms, a satisfactory ($R_{\text{test}}^2 = 0.81$) Quantitative Relationship Structure—Activity (QSAR) model was developed. The subsequent calculations have shown that natural avermectins undergo on average faster biotransformation in fish than the semisynthetic ones. In addition, structural fragments that increase and decrease the biotransformation rate are identified.

Keywords: macrolides, molecular descriptors, machine learning, QSAR

DOI: 10.3103/S0027131421040088

Currently, ivermectin is considered one of the most promising inhibitors of the SARS-CoV-2 coronavirus, making it possible to reduce the viral RNA fraction by 99.98% within 48 h when used at a concentration of 5 μM considered safe for cells in the *in vitro* studies [1]. It has been recently demonstrated that ivermectin successfully decreased the virus load in COVID-19 patients [2]. Low levels of HIV-1 and Dengue virus replication in the presence of ivermectin has been previously reported [3]. Ivermectin, whose sales are estimated to be in excess of \$1 billion per year during the past 20 years, represents a combination of ivermectin B1a (~90%) and ivermectin B1b (~10%), which are the semisynthetic derivatives of natural avermectins. Natural avermectins have been long used to efficiently combat endo- and ectoparasites of plants, animals, and human [5]. Avermectins are also promising in terms of development of new antitumor drugs [6–8]. *Streptomyces avermitilis* synthesizes the eight-component complex of 16-membered macrolides. The chemical structures of the natural and semisynthetic avermectins are provided in Figs. 1–5.

The antiparasitic properties of avermectins are due to their effects on γ -aminobutyric acid (GABA) and glutamate receptors. Fortunately, mammalian nervous system cells are somewhat protected from avermectin exposure because of the blood-brain barrier (BBB) [9].

In most European countries, there is a legal requirement to provide information on the main environmental parameters (e.g., biodegradation) of the chemicals produced or imported in excess of one ton a year [10]. The ubiquitous use of avermectins entails the need for their detailed ecotoxicological assessment. The key ecological risks associated with avermectins are related to their possible effects on the non-target organisms. For example, when a prolonged ivermectin form was excreted with calf feces, a decrease in the insect population required for processing natural manure was reported, which caused this form of ivermectin to be discontinued in some countries in 2004 [9]. Therefore, when new medical preparations are developed based on avermectins, it is important to assess their biodegradation and biotransformation periods along with other parameters of their potential negative ecological impact. Given that most of the

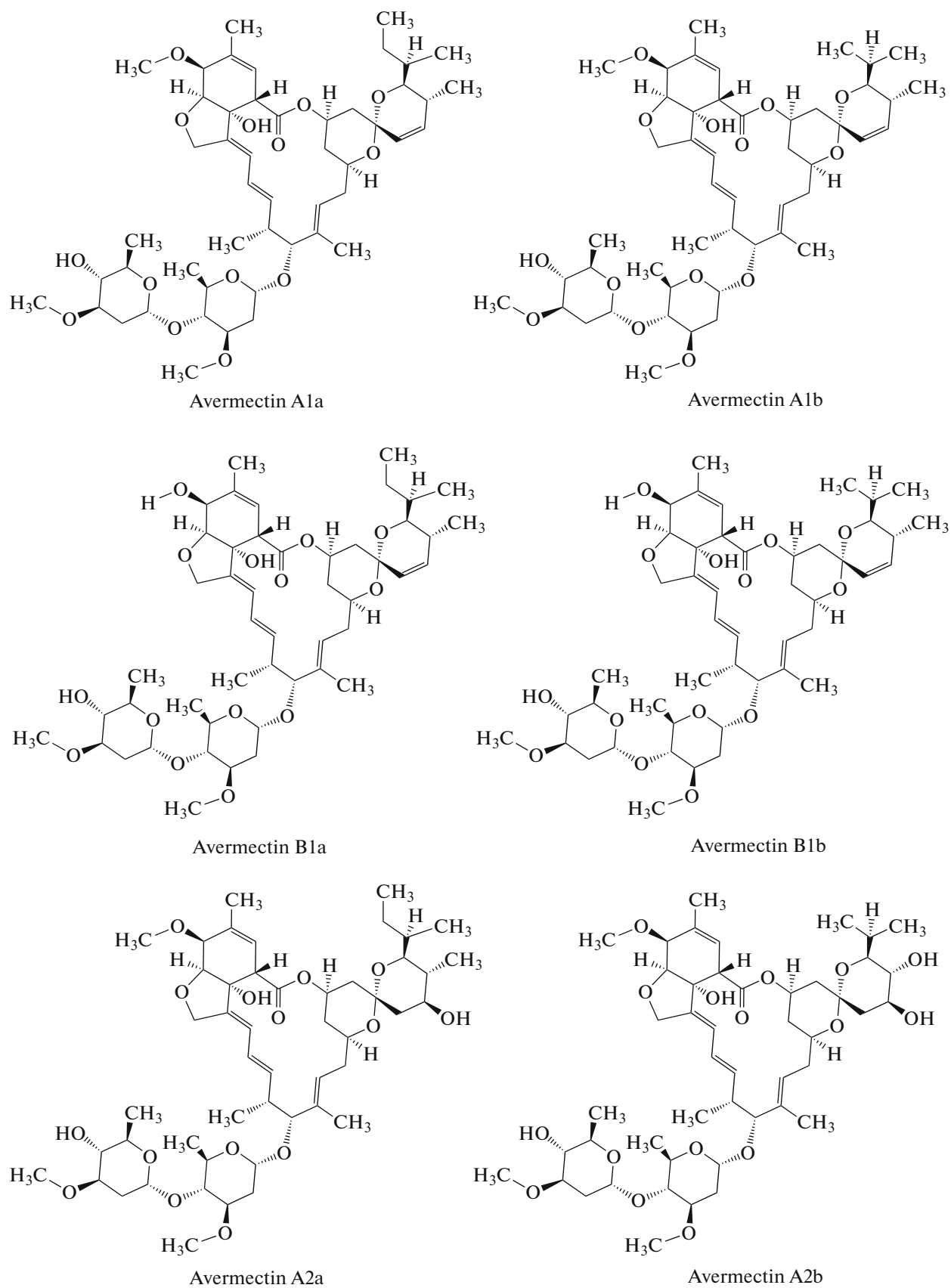


Fig. 1. Chemical structure of natural avermectins: Avermectin A1a, Avermectin A1b, Avermectin B1a, Avermectin B1b, Avermectin A2a, and Avermectin A2b.

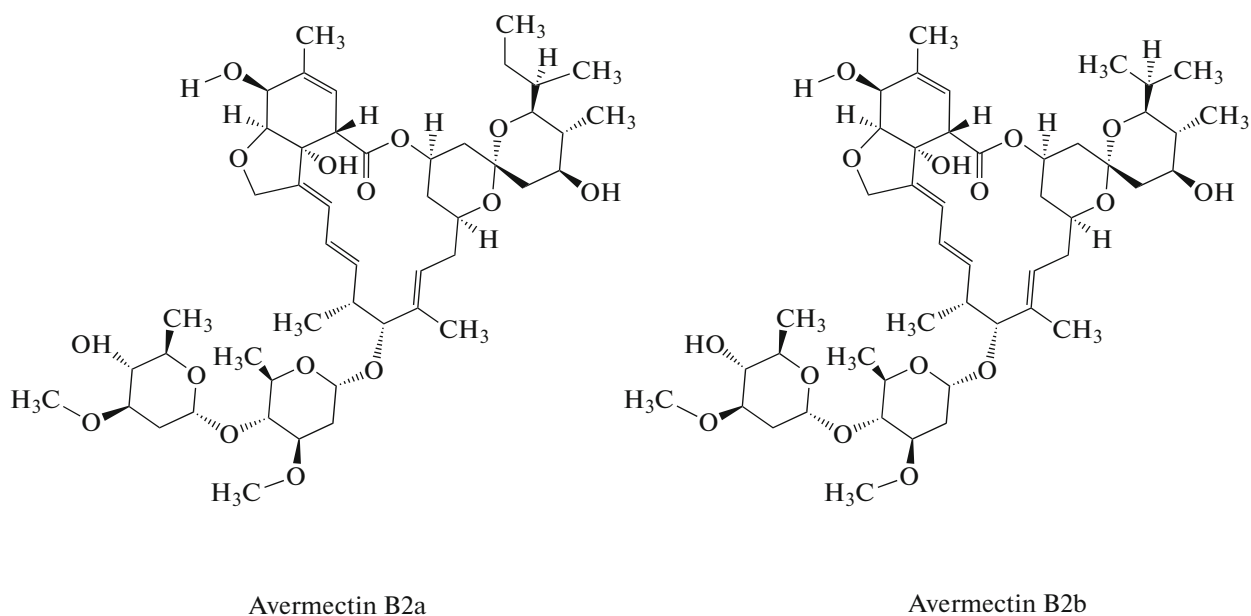


Fig. 2. Chemical structure of natural avermectins: Avermectin B2a and Avermectin B2b.

Earth is occupied by water, it seems logical to assess the environmental parameters of avermectins such as biotransformation in aquatic organisms, in particular, fish, which constitute a significant part of the human diet. The review [11], which discusses the ecological evaluation of avermectins, notes that, unfortunately, notwithstanding the intensive use of these compounds worldwide, the experimental data on their ecotoxicological assessment is scarce.

Since the experimental assessment of the chemical compounds' properties is costly and time-consuming and taking into account the requirements for the ethical treatment of animals, the quantitative structure-activity/property relationship (QSAR/QSPR) methods are attracting the attention of researchers [12, 13].

Considering the current importance of studying the biotransformation of organic compounds, a number of works [14–16] have suggested satisfactory QSAR models which can quantitatively describe the relationship between the structure of a compound and its capacity for biotransformation in fish. A vast majority of these models is integrated into the well-known software such as the EPI Suite V.4.1 [17] and CompTox Chemistry Dashboard [18] developed by IT specialists from the United States Environmental Protection Agency (EPA). In particular, the model integrated into EPI Suite V.4.1, which is described in detail in [14], has good statistic parameters. Unfortunately, the authors of this model have not clearly defined the algorithm for calculating whether a compound is included or not in the applicability domain (AD) of the model [19], which is prerequisite for QSAR modeling according to the principles suggested by the expert group of the Organization for Economic Coop-

eration and Development (OECD). Unfortunately, natural and semisynthetic avermectins are not included in the applicability domain of another useful QSAR model [16] publicly available as a part of the CompTox Chemistry Dashboard web-application [18].

Additionally, the interpretation of the developed models would be recommended according to the fifth principle of QSAR modeling worked out by the OECD expert group. Most QSAR models lack structural interpretation, which makes it impossible to make a rational molecular design in order to obtain compounds with the desired biotransformation rates. The above-mentioned work [16] provides a short structural interpretation consisting in the comparative analysis of the twelve descriptors used and the patterns described in the work [20], in which the impact of the structural factors on amide biotransformation pathways were studied. The authors of the work [16] came to the conclusion that the presence of ring structures reduces the capacity of the corresponding compounds for biotransformation, while unbranched alkyl chains containing more than 3 carbon atoms, as well as the presence of benzyl and tolyl groups, favor oxidation rather than N-dealkylation. The described structural particularities are quite limited, which is mostly due to the low structural diversity used in the above-mentioned work [20] and the small number of descriptors used in the study [16].

Thus, the aims of this study are as follows:

(1) to build a QSAR model of an organic compound biotransformation whose applicability domain would allow to reliably predict this activity in natural and semisynthetic avermectins;

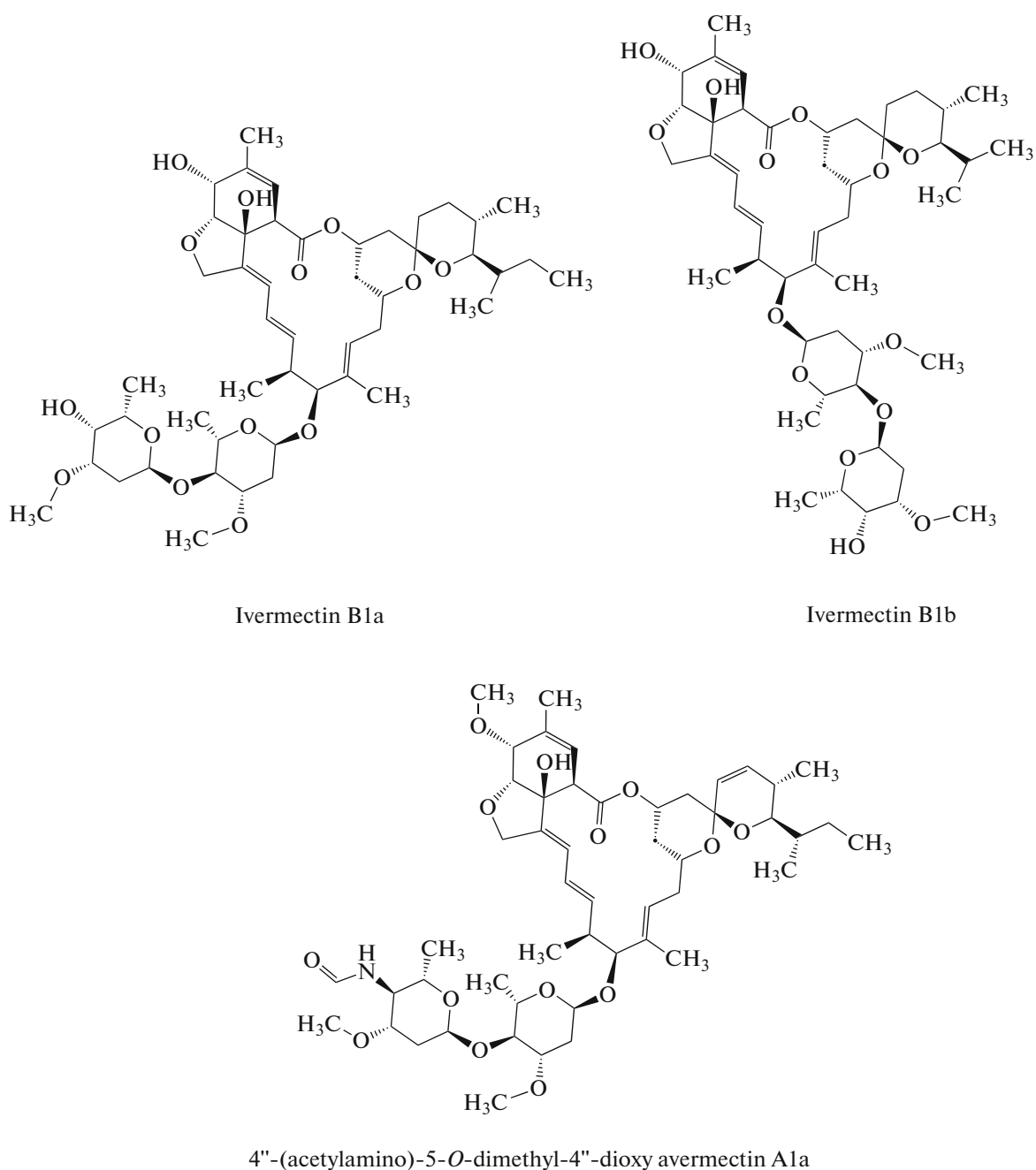


Fig. 3. Chemical structure of semisynthetic avermectins: Ivermectin B1a, Ivermectin B1b, and 4''-(acetylamino)-5-O-dimethyl-4''-dioxy avermectin A1a.

(2) to perform the structural interpretation of the obtained QSAR model in order to define those molecular fragments which have the strongest impact on the organic compound biotransformation rate in fish.

EXPERIMENTAL

The publicly available training (ws) and external test (ts) sets with the experimentally calculated biotransformation parameters in fish were exported

from the above-mentioned work [16], which provides a detailed description of the verification procedure of the initial data available in the PHYSPROP (The Physical Properties) database. The quantitative parameter, the biotransformation half-life period (KmHL), which is the number of days during which half of the studied compound undergoes biotransformation, was chosen as the indicator of the organic compound's biotransformation ability in fish. This parameter is calculated based on the biotransforma-

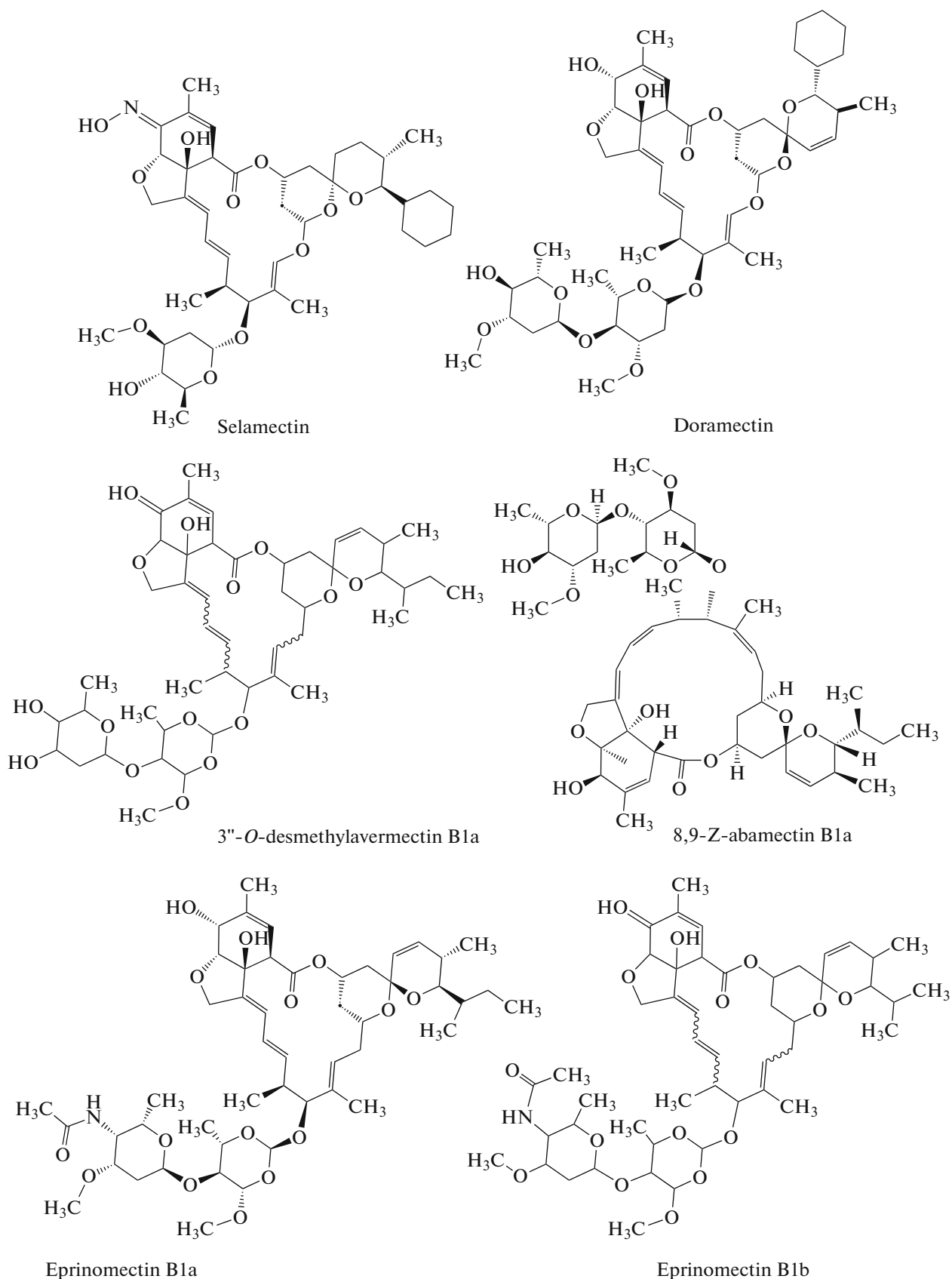


Fig. 4. Chemical structure of semisynthetic avermectins: Selamectin, Doramectin, 3''-O-desmethylavermectin B1a, 8,9-Z-abamectin B1a, Eprinomectin B1a, and Eprinomectin B1b.

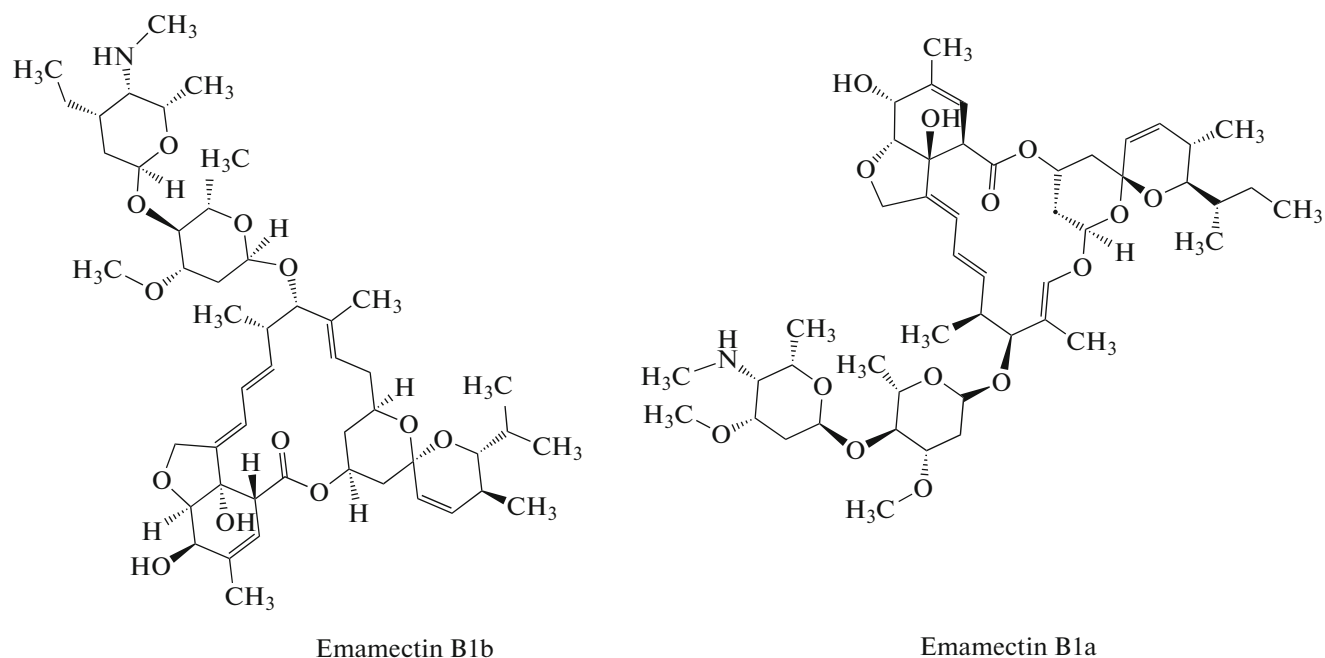


Fig. 5. Chemical structure of semisynthetic avermectins: Emamectin B1b and Emamectin B1a.

tion rate constant (k_M). The detailed characterization of the used pharmacokinetic parameters can be found in the above-mentioned work [15].

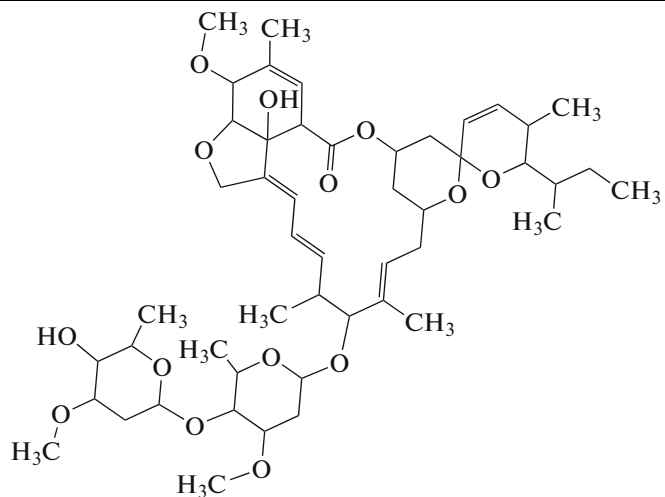
The training set included 405 compounds and the external test set of 136 compounds. The biotransformation half-life period (K_{mHL}) expressed as a decimal logarithm varied is in the range from -1.6 to 3.0 . A nonlinear transformation of the modeled parameter (K_{mHL}) was necessary since the experimental data in the analyzed sample was more evenly distributed on a logarithmic scale (Figs. 6, 7), which also corresponds well with the common practice of mathematical data processing prior to the construction of structure–activity/property models [21]. The molar mass of the analyzed compounds varied between 68.8 g/mol and 959.2 g/mol.

The dataset used was structurally heterogeneous and comprised various compounds such as halohydrocarbons (polychlorinated biphenyls and dioxins), aliphatic and aromatic hydrocarbons (polycyclic aromatic hydrocarbons), amines, imides, alcohols, phenols, ethers, ketones, and esters. The chemical space of the training set was analyzed to preliminarily evaluate the applicability of the developed QSAR models for predicting the biotransformation ability of avermectins. Eleven compounds within the training set with the chemical structure most similar to that of avermectins are presented in Table 1. Avermectin A1a was selected as the reference avermectin. The structural similarity of the compounds within the training set was estimated by calculating the Tanimoto and Dice indices using the topological fingerprints, which are a type of a binary fragment descriptor. The calculations were made

using the RDKit package for Python [22]. It has been previously demonstrated that the Tanimoto and Dice indices are the most adequate metrics for molecular similarity calculations [23].

It can be seen from Table 1 that the training set contains a macropolycyclic compound 1,5,9-cyclododecatriene, a number of polycyclic aromatic hydrocarbons, and 2,2',4,4',5,6-hexachlorodiphenyl ether, which suggests that the structural space of the training set allows us to properly consider the molecular properties of avermectins in QSAR modeling.

The change in the biotransformation rate following slight structural modifications introduced to the compounds within the training set was also analyzed using dichlorodiphenyltrichloromethylmethane and 1,2,3,4,5-pentachlorobenzene derivatives (Table 2). In particular, in the analyzed compounds from the first series (CAS numbers 50-29-3, 72-54-8, and 72-55-9), the slight transformation of 4,4'-dichlorodiphenyltrichloroethane into 4,4'-dichlorodiphenyldichloroethane results in a decrease by a factor of more than three in the biotransformation rate. At the same time, conversely, dehydrogenation resulting in 4,4'-dichlorodiphenyldichloroethane conversion into 4,4'-dichlorodiphenyldichloroethylene leads to a more than five fold increase in the biotransformation rate. In the case of the compounds from the other series (CAS numbers 108-67-8, 117-18-0, and 527-60-6), the introduction of a hydroxyl group into 1,2,3,4,5-pentachlorobenzene, or the substitution of the chloro group at position 3 by a nitro group, makes it possible to significantly increase its biodegradation rate. Hence, it may be hoped that the chosen training set allows us to make

Table 1. Compounds within the training set most structurally similar to Avermectin A1a

Compound structure	CAS number	index	
		Tanimoto	Dice
	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
	106220-83-1	0.50	0.66
	78-30-8	0.45	0.62

Table 1. (Contd.)

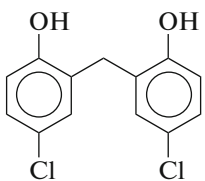
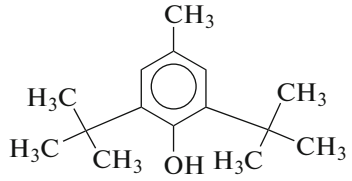
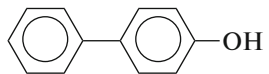
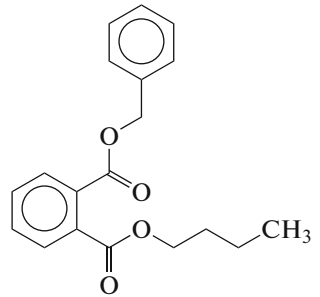
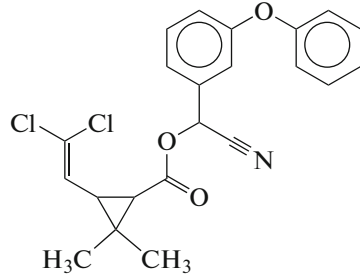
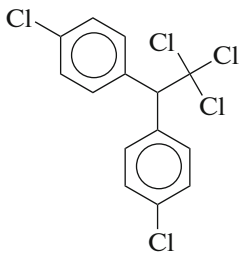
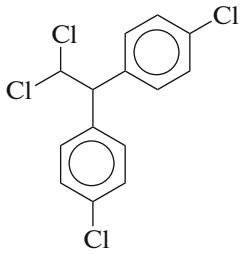
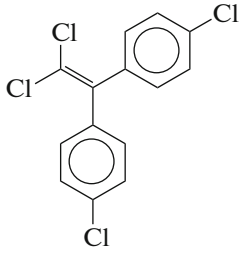
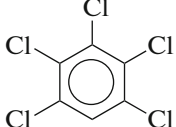
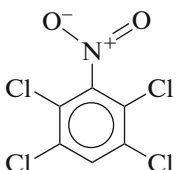
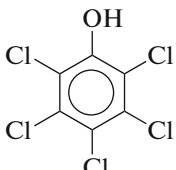
Compound structure	CAS number	index	
		Tanimoto	Dice
	97-23-4	0.40	0.57
	128-37-0	0.39	0.57
	92-69-3	0.38	0.55
	85-68-7	0.32	0.48
	52315-07-8	0.31	0.48

Table 2. Changes in biotransformation rates observed within the dichlorodiphenyltrichloromethylmethane and 1,2,3,4,5-penta-chlorobenzene derivative series

Compound structure	CAS number	Experimental KmHL, number of days
	50-29-3	85.11
	72-54-8	275.42
	72-55-9	51.29
	108-67-8	74.13
	117-18-0	3.09
	527-60-6	0.63

an adequate prediction of the hypothetical changes in the avermectin biodegradation rates determined by the slight differences observed between the structures of the analyzed 16-member macrolides.

QSAR modeling was performed at the OCHEM website [24]. This site contains a significant amount of data on the structures of chemical compounds, as well as on various types of biological activity and physical

properties and allows us to create reliable QSAR/QSPR models using various machine learning methods and molecular descriptors. The best results were obtained for the analyzed sample when PyDescriptor descriptors [25] and Random forest (RF) [26] and Transformed convolutional neural network (Trans-CNN) [27] algorithms were used. All the analyzed compounds were standardized using the

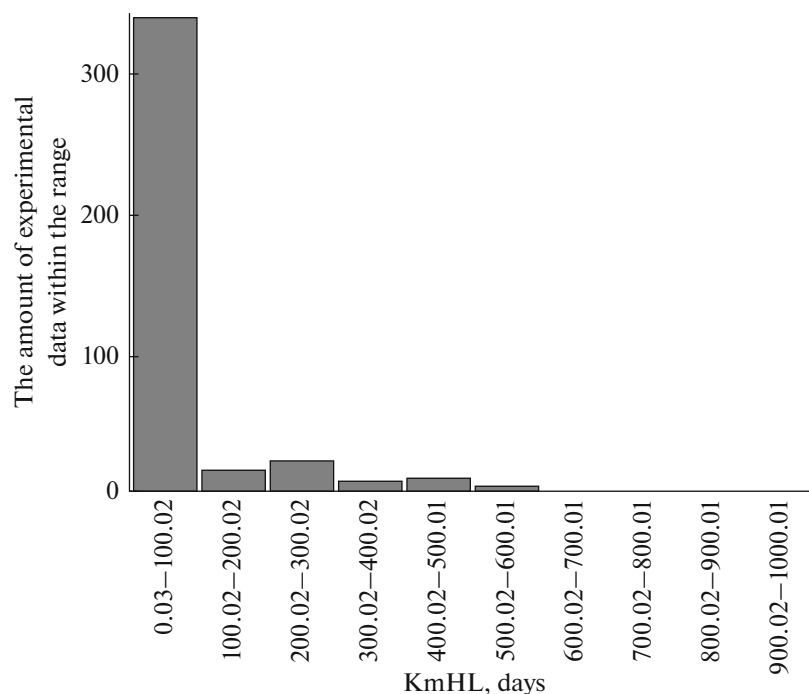


Fig. 6. Experimental KmHL value distribution histogram.

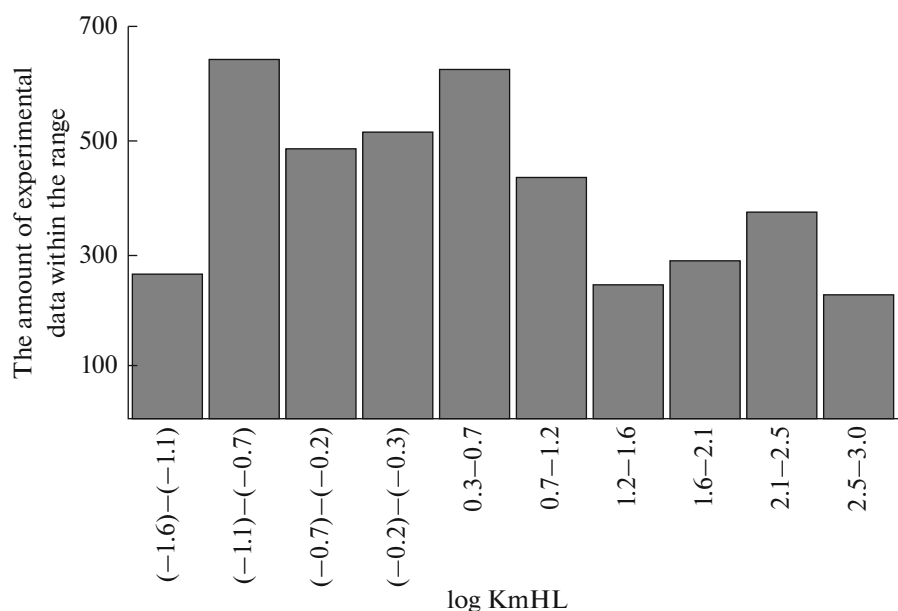


Fig. 7. Histogram of experimental KmHL value distribution on the logarithmic scale.

built-in ChemAxon Standardizer plugin [28] according to the commonly accepted protocols [29, 30].

The PyDescriptor software package allows us to calculate the set of 11 145 different molecular descriptors. At the 1D molecular structure level, molecular weight, the number atoms of a specific type, and some other parameters commonly used at this level of detail for chemical compounds are calculated. At the 2D

level, various physicochemical descriptors are calculated, such as, for example, hydrogen bond descriptors, as well as various fragment descriptors including fingerprints. At the 3D level, descriptors characterizing the spatial structure of the molecule are generated, for example, an atomic charge weighted positively charged surface. To calculate the 3D level descriptors, the structures of the analyzed compounds were opti-

Table 3. Statistical parameters of QSAR models

Model	R_{cv}^2	RMSE _{cv}	R_{test}^2	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

mized using the Corina software module [31] also integrated into the OCHEM platform. All the calculated descriptors were filtered using the prefiltration protocol suggested in [32], namely, (1) descriptors which were the same for the analyzed compounds were removed; (2) normalized descriptors with the variance < 0.01 were excluded; and (3) descriptors with the pairwise linear correlation coefficient $R^2 > 0.95$ were grouped together, and only the first descriptor in the group was taken to develop the model. As a result, a series of 1222 descriptors were obtained.

Internal validation of the training set was carried out using the 5-fold cross-validation procedure. Given the structural diversity of the training set, it was decided to perform the initial calculations using different methods, and then to use the obtained models for consensus QSAR modeling. It has been previously demonstrated that this approach is more reliable on average than each of the models alone [33, 34]. The consensus model was developed by averaging the predictions obtained using individual models.

AD was estimated using the distance-to-model concept, in particular, the CONSENSUS-STD approach (standard deviation of predictions of the ensemble of models) [35]. The predicted target activity for the external test set molecules in this case is calculated as the average for the ensemble of models. The

distance to model (d_{STD}) is calculated using the following formula:

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

where $y(B) = \{y_i(B), i = 1 \dots N\}$ is the set of predictions for the compound B obtained using the set of N trained models and \bar{y} is the average predicted activity for the compound B .

The CONSENSUS-STD algorithm assumes that reliable predictions have low d_{STD} . This approach was shown to be the most reliable in terms of the prediction accuracy assessment in a number of comparative studies [35–37].

The accuracy of predictions obtained with the QSAR models was assessed using the coefficient of determination of the cross validation (R_{cv}^2), squared coefficient of correlation for the external test set (R_{test}^2), and root mean squared error (RMSE), the calculation of which is implemented as a part of OCHEM. The validation methods used, the algorithm for calculating the applicability domains of QSAR models, and the formulas used to calculate the above-mentioned statistical parameters are described in detail in the OCHEM manual [24].

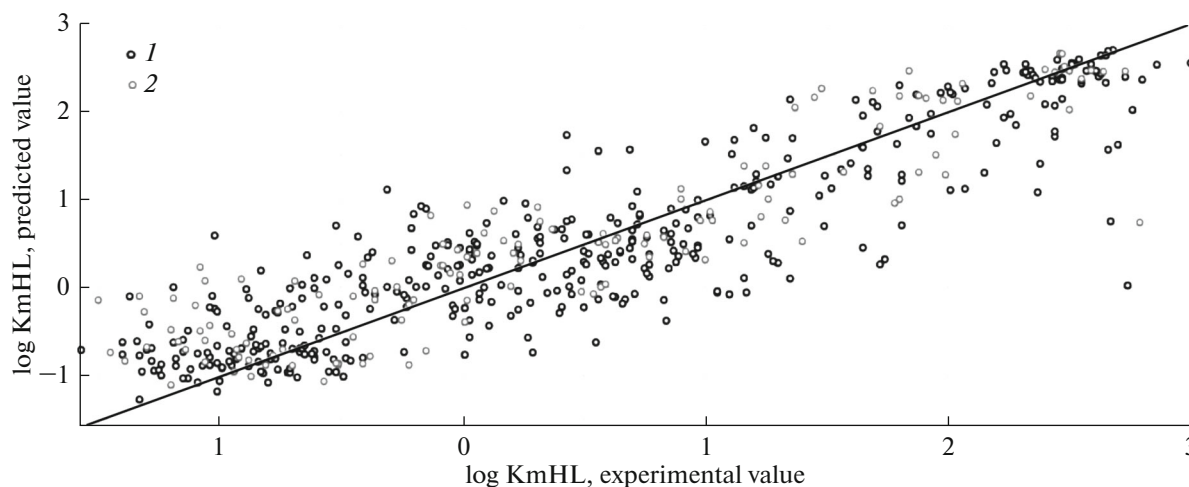


Fig. 8. The comparison of the observed (experimental) and predicted by the consensus model biotransformation parameters for the compounds in the training set (1) (cross control) and in the test set (2).

Table 4. Results of structural interpretation

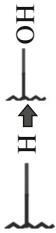
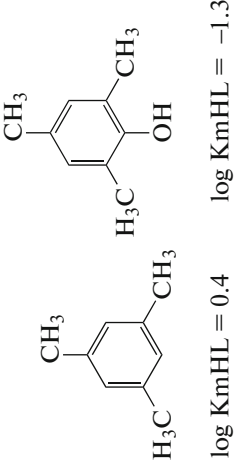

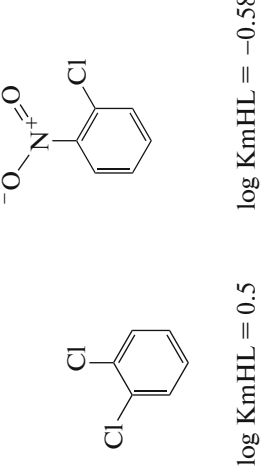

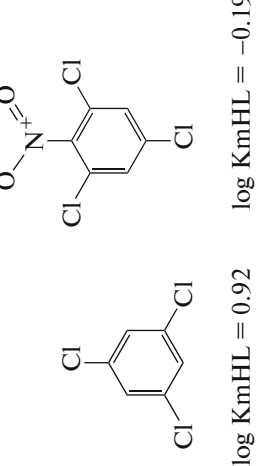
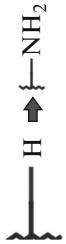
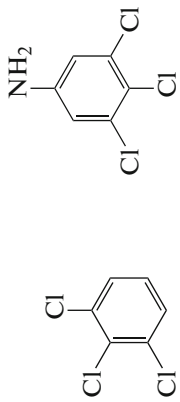


Molecular transformation	The example of a compounds pair from the training set and their experimental activity values	<i>N</i>	Δ_{mean}	SEM	$ r_{\text{calc}} $	SMIRKS
		10	-1.00	0.26	3.9	*[H] → O*
		14	-0.90	0.24	3.7	*Cl → [O-][N+](*)=O
		14	-0.58	0.23	2.5	*[H] → [O-][N+](*)=O

Table 4. (Contd.)

Molecular transformation	The example of a compounds pair from the training set and their experimental activity values	N	Δmean	SEM	t _{calc}	SMIRKS
	 <p>log KmHL = 0.92 log KmHL = 0.28</p>	11	-0.52	0.24	2.1	*[H] → N*
Molecular transformations decreasing the rate of chemical compound biotransformation in fish						
	<p>H₃C-(CH₂)₁₀-CH₃ H₃C-(CH₂)₁₀-(CH₂)₂-CH₃ log KmHL = 0.26 log KmHL = 1.49</p>	18	0.40	0.12	3.2	*C → *CCC*
	<p>H₃C-(CH₂)₁₀-CH₃ H₃C-(CH₂)₁₀-(CH₂)₄-CH₃ log KmHL = 0.26 log KmHL = 1.47</p>	20	0.74	0.09	8.5	*CC* → *CCCCC*

N is the number of molecular pairs in which molecular transformation takes place; Δmean is the average difference between the log KmHL values resulting from a molecular transformation; SEM is the standard error of mean log KmHL; and SMIRKS is the format of molecular transformation representation (<https://www.daylight.com/dayhtml/doc/theory/theory.smirks.html>).

Table 5. Predicted biotransformation abilities of individual avermectins in fish

Name	CAS number	KmHL, number of days	d_{STD}	RMSE
Natural avermectins				
Avermectin A1a	65195-51-9	4.1	0.11	0.51
Avermectin A1b	65195-52-0	4.2	0.13	0.51
Avermectin B1a	65195-55-3	4.4	0.07	0.51
Avermectin B1b	65195-56-4	4.5	0.07	0.51
Avermectin A2a	65195-53-1	3.3	0.11	0.51
Avermectin A2b	65195-54-2	3.3	0.17	0.51
Avermectin B2a	65195-57-5	3.9	0.10	0.51
Avermectin B2b	65195-58-6	3.1	0.00	0.4
Semisynthetic avermectins				
Ivermectin B1a	71827-03-7	7.7	0.30	0.56
Ivermectin B1b	70209-81-3	8.0	0.28	0.54
Doramectin	117704-25-3	6.9	0.23	0.51
Selamectin	220119-17-5	7.3	0.13	0.51
3''- <i>O</i> -Desmethylavermectin B1a	99965-56-7	14.1	0.07	0.51
8,9- <i>Z</i> -abamectin B1a	113665-89-7	7.0	0.22	0.51
4''-(Acetylamino)-5- <i>O</i> -dimethyl-4''-dioxavermectin A1a	172926-97-5	6.8	0.23	0.51
Emamectin B1b	121424-52-0	6.5	0.26	0.54
Emamectin B1a	121124-29-6	7.9	0.28	0.54
Eprinomectin B1a	1142337-10-7	7.6	0.31	0.56
Eprinomectin B1b	133305-89-2	10.0	0.00	0.40

RESULTS AND DISCUSSION

The results of QSAR modeling are summarized in Table 3 together with the statistical parameter for the QSAR model developed by the US EPA specialists for the analyzed sample [16] provided with the purpose of a comparative analysis. The models obtained in the present study have the prediction ability comparable with that of the earlier model developed using PaDEL descriptors and the weighted k -nearest neighbors algorithm. The comparison of the experimentally obtained values and those predicted by the consensus QSAR model is presented in Fig. 8.

All compounds in the training set were included in the applicability domain of the consensus QSAR model. Figure 9 presents a plot confirming that an adequate measure was used to determine the applicability domain of the model. The x -axis shows the distance to model, and the y -axis shows the absolute prediction error for the analyzed compounds, while the single dots correspond to individual chemical compounds. To visualize the dependence between these two parameters, the full range of the distance-to-model variation is automatically divided into several intervals containing the same number of objects, and for each of them a horizontal segment is drawn at the height equal to the average value of the absolute prediction error for the chemical compounds included in

the interval. The obtained segments forming a ladder, where each segment located to the right (a step) is higher than the segment located to the left, may be the confirmation of the adequacy of the measure used. When an inadequate measure is used, all segments (steps) in this type of plot are located almost at the same level [21].

For the obtained consensus model, structural interpretation using the matched molecular pairs analysis approach also implemented in OCHEM [38] was performed. The results, in particular, those molecular transformations which increase or decrease the biotransformation rates of the organic compounds in fish to the largest extent are indicated in Table 4. The level of significance and stability of the identified molecular transformations was determined using the one sample Student's t test as previously suggested [39, 40] by comparing the t value calculated according to formula (2) (t_{calc}) with the critical value (t_{crit}) at the significance level $\alpha = 0.05$ with the corresponding degree of freedom.

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

where Δmean is the average difference in the Log KmHL values for a molecular transformation and SEM is the standard error of the average value of Log KmHL.

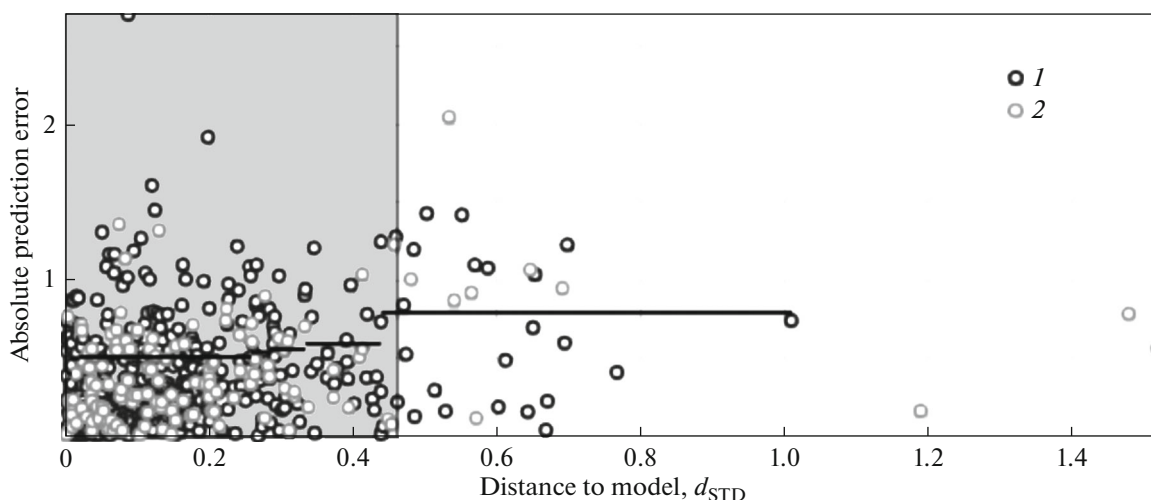


Fig. 9. Absolute prediction error–distance to model dependence plot for the compounds in the training set (1) (cross control) and the compounds in the test set (2).

According to [39], if $|t_{\text{calc}}| > t_{\text{crit}}$, then the identified molecular transformation is considered to be statistically significant. The authors of the work [39] noted that in most practically relevant cases, $t_{\text{crit}} \approx 2$ at $\alpha = 0.05$. In relation to this, it is accepted that a molecular transformation is statistically significant when the absolute Δmean value is double the SEM value, which is also mentioned in the work [40].

The structural interpretation results allow us to suggest that the substitution of hydrogen and chlorine atoms by hydroxy, nitro, and amino groups makes the largest contribution to the increase in the biotransformation rate. This is probably due to the fact that oxidation is the most frequent reaction in the first phase of metabolism, while hydroxylation of xenobiotics containing aliphatic radicals leads to the formation of alcohols, which like amines are transformed into aldehydes and ketones as a result of further oxidation [41]. Hence, the compounds containing hydroxy and amino groups undergo a fewer number of oxidation cycles, which results in higher biotransformation rates for them. According to the results of structural interpretation, the increase in the length of aliphatic unbranched radicals slows the rate of organic compound biotransformation in fish.

The consensus model developed in this study is available at <https://ochem.eu/model/16426620> and can be used to predict the biotransformation rates of new organic compounds in fish. We used this model to calculate the biotransformation parameters for individual avermectins in fish, which are presented in Table 5. All compounds were included in the applicability domain, which may be evidence of the reliability of the predicted values. Based on the data presented in Table 5, it may be assumed that the biotransformation of natural avermectins in fish is faster than the biotransformation of semisynthetic avermectins.

For example, hydrogenation of the natural Avermectin B1b (KmHL is approximately four-and-a-half days) at the double bond between C22 and C23 produces a semisynthetic derivative, Ivermectin B1b, the biotransformation half-life period for which is increased to eight days. The amino derivative of the natural Avermectin B1b, the semisynthetic avermectin Eprinomectin B1b, also has a longer biotransformation half-life period (10 days). Based on these observations, it may be suggested that the introduction of the acetamide group into natural avermectins will increase the biotransformation half-life period of the semisynthetic derivatives. The replacement of the aliphatic radical at C25 with a cyclohexane fragment (doramectin and selamectin) in natural avermectins also leads to an increase in the biotransformation half-life period, which agrees with the trends outlined in the work [20]. Conversely, the replacement of the hydroxy group in doramectin with a hydroxylamine group (selamectin) leads to a decrease in the biotransformation time.

Hence, the use of individual natural avermectins as new drugs is preferable due to their higher biotransformation rate in fish, and consequently, lower probability of their possible consumption with fish by animals or humans.

CONCLUSIONS

The adequate consensus QSAR model of an organic compound biotransformation in fish was developed, which showed the predictive ability comparable to the model proposed by the US Environmental Protection Agency. The descriptors selected allowed us to expand the applicability domain of the model, which made it possible to more reliably predict the biotransformation of individual avermectins in fish. According to the calculated values, the biotrans-

formation of natural avermectins in fish occurs faster than the biotransformation of semisynthetic avermectins.

The structural interpretation of the consensus QSAR model allowed us to identify the general effects of the structure on the biotransformation rate, distinguishing those molecular transformations which significantly increase and decrease the biotransformation rate of organic compounds. The analysis of individual avermectins identified those structural fragments which increase and decrease the biotransformation rate.

The results of the present study can help us reduce the financial, time, and labor costs associated with identifying strategies for the development of new drugs based on individual avermectins, in particular when evaluating the ecotoxicological parameters of avermectins. In future studies, experimental validation of the model obtained and evaluation of other ecotoxicological characteristics, such as acute toxicity to aquatic organisms, may prove to be useful.

FUNDING

This study was partially supported by a state task for the Institute of Physiologically Active Compounds of the Russian Academy of Science for 2020 (project no. 0090-2019-0004).

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1. Caly, L., Druce, J.D., Catton, M.G., Jans, D.A., and Wagstaff, K.M., *Antiviral Res.*, 2020, vol. 178, 104787.
2. Chaccour, C., Casellas, A., Blanco-Di, MatteoA., Pineda, I., Fernandez-Montero, A., Ruiz-Castillo, P., Richardson, M.-A., Rodríguez-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., and Fernandez-Alonso, M., *Clin. Med.*, 2021, 100720. <https://doi.org/10.1016/j.eclinm.2020.100720>
3. Wagstaff, K.M., Sivakumaran, H., Heaton, S.M., Harrich, D., and Jans, D.A., *Biochem. J.*, 2012, vol. 443, p. 851.
4. Crump, A. and Omura, S., *Proc. Jpn. Acad., Ser. B*, 2011, vol. 87, no. 2, p. 13.
5. Zhuo, Y., Zhang, T., Wang Qi, Cruz-Morales, P., Zhang, B., Liu, M., Barona-Gomez, F., and Zhang, L., *Biotechnol. J.*, 2014, vol. 9, p. 316.
6. Mosin, V.A., Kruglyak, E.B., Sterlina, T.S., Korystov, Yu.N., Shaposhnikova, V.V., Kublik, L.N., Levitman, M.Kh., Viktorov, A.V., and Drinyaev, V.A., *Antibiot. Khimioter.*, 1999, vol. 44, no. 6, p. 16.
7. Drinyaev, V.A., Mosin, V.A., Kruglyak, E.B., Sterlina, T.S., Novik, T.S., Ermakova, N.V., Kublik, L.N., Levitman, M.Kh., Shaposhnikova, V.V., and Korystov, Yu.N., *Antibiot. Khimioter.*, 2004, vol. 49, no. 6, p. 3.
8. Juarez, M., Schcolnik-Cabrera, A., and Dueñas-Gonzalez, A., *Am. J. Cancer Res.*, 2018, vol. 8, p. 317.
9. Prichard, R., Menez, C., and Lespine, A., *Int. J. Parasitol.: Drugs Drug Resist.*, 2012, vol. 2, p. 134.
10. Mansouri, K., Ringsted, T., Ballabio, D., Todeschini, R., and Consonni, V., *J. Chem. Inf. Model.*, 2013, vol. 53, p. 867.
11. Bai, S.H. and 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., and Tropsha, A., *Chem. Soc. Rev.*, 2020, vol. 49, no. 11, p. 3525.
13. Tin'kov, O.V., Grigor'ev, V.Yu., Polishchuk, P.G., Yarkov, A.V., and Raevskii, O.A., *Biomed. Khim.*, 2019, vol. 65, p. 123.
14. Arnot, J.A., Meylan, W., Tunkel, J., Howard, P., Mackay, D., Bonnell, M., and Boethling, R.S., *Environ. Toxicol. Chem.*, 2009, vol. 28, p. 1168.
15. Papa, E., van der Wal, L., Arnot, J.A., and Gramatica, P., *Sci. Total. Environ.*, 2014, vol. 470, p. 1040.
16. Mansouri, K., Grulke, C.M., Judson, R.S., and Williams, A.J., *J. Cheminf.*, 2018, vol. 10, p. 1.
17. United States Environmental Protection Agency. www.epa.gov/tsca-screening-tools/episuitetm-estimation-program-interface.
18. CompTox Chemicals Dashboard. <https://comptox.epa.gov/dashboard>.
19. Jaworska, J., Nikolova-Jeliazkova, N., and Aldenberg, T., *Altern. Lab. Anim.*, 2005, vol. 33, no. 5, p. 445.
20. Helbling, D.E., Hollender, J., Kohler, H.P.E., and Fenner, K., *Environ. Sci. Technol.*, 2010, vol. 44, no. 17, p. 6628.
21. Madzhidov, T.I., Baskin, I.I., Antipin, I.S., and Varnek, A.A., *Vvedenie v khemoinformatiku* (Introduction to Chemoinformatics), vol. 3: *Modelirovanie "struktura-svoystvo"* (Modeling "Structure-Property"), Kazan: Kazan. Gos. Univ., 2015.
22. RDKit: Open-Source Cheminformatics Software. www.rdkit.org/docs/GettingStartedIn-Python.html.
23. Bajusz, D., Racz, A., and Heberger, K., *J. Cheminf.*, 2015, vol. 7, 20.
24. OCHEM. <https://ochem.eu/>.
25. Masand, V.H. and Rastija, V., *Chemom. Intell. Lab. Syst.*, 2017, vol. 169, p. 12.
26. Breiman, L., *Mach. Learn.*, 2001, vol. 45, p. 5.
27. Karpov, P., Godin, G., and Tetko, I.V., *J. Cheminf.*, 2020, vol. 12, 17.
28. ChemAxon. <https://docs.chemaxon.com/display/docs/Standardizer+Command-line+Application>.
29. Fourches, D., Muratov, E., and Tropsha, A., *J. Chem. Inf. Model.*, 2010, vol. 50, p. 1189.
30. Fourches, D., Muratov, E., and Tropsha, A., *J. Chem. Inf. Model.*, 2016, vol. 56, p. 1243.
31. Sadowski, J., Gasteiger, J., and Klebe, G., *J. Chem. Inf. Comput. Sci.*, 1994, vol. 34, p. 1000.
32. Tetko, I.V., Lowe, M.D., and Williams, A.J., *J. Cheminf.*, 2016, vol. 8, 2.

33. Golbraikh, A., Muratov, E., Fourches, D., and Tropsha, A., *J. Chem. Inf. Model.*, 2014, vol. 54, p. 1.
34. Zakharov, A.V., Zhao, T., Nguyen, D.T., Peryea, T., Sheils, T., Yasgar, A., Huang, R., Southall, N., and 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., and 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., and 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., Müller, C., Varnek, A., Prokopenko, V.V., and Tetko, I.V., *J. Chem. Inf. Model.*, 2010, vol. 50, p. 2094.
38. Sushko, Y., Novotarskyi, S., Korner, R., Vogt, J., Abdelaziz, A., and Tetko, I., *J. Cheminf.*, 2014, vol. 6, p. 48.
39. Madzhidov, T.I., Baskin, I.I., Antipin, I.S., and Varnek, A.A., *Vvedenie v khemoinformatiku* (Introduction to Chemoinformatics), vol. 6: *Khimicheskoe prostranstvo i virtual'nyi skring* (Chemical Space and Virtual Screening), Kazan: Kazan. Gos. Univ., 2019.
40. Griffen, E., Leach, A.G., Robb, G.R., and Warner, D.J., *J. Med. Chem.*, 2011, vol. 54, p. 7739.
41. Vergeichik, T.Kh., *Toksikologicheskaya khimiya* (Toxicological Chemistry), Moscow, 2012.

Translated by E. Martynova