Research Article

Mariia Nesterkina*, Viacheslav Muratov, Luidmyla Ognichenko, Iryna Kravchenko, Victor Kuz'min

# Quantitative structure–activity relationship study on prolonged anticonvulsant activity of terpene derivatives in pentylenetetrazole test

https://doi.org/10.1515/chem-2021-0108
received August 13, 2021; accepted November 22, 2021

**Abstract:** Quantitative structure–activity relationship (QSAR) study has been conducted on 36 terpene derivatives with anticonvulsant activity in timed pentylenetetrazole (PTZ) infusion test. QSAR models for anticonvulsant activity prediction of hydrazones and esters of some monocyclic/bicyclic terpenoids were developed using simplex representation of molecular structure (SiRMS; informational field [IF]) approach based on the SiRMS and the IF of molecule. Four 2D partial least squares QSAR consensus models were developed with the coefficient of determination for test sets $R^2_{test} > 0.62$. Based on the established QSAR models, we found that carvone and verbenone cores possess the most significant contribution to antiseizure action examined on the model of PTZ-induced convulsions at 3 and 24 h after oral administration of terpene derivatives. Moreover, carbonyl and hydroxy group substitution in terpenoid molecules followed by hydrazones and esters formation leads to enhancement and prolongation of antiseizure action due to the contribution of additional molecular fragments. The presented QSAR models might be utilized to predict anticonvulsant effect among terpene derivatives for their oral administration against onset seizures.

**Keywords:** QSAR models, terpenes, anticonvulsant activity, SiRMS approach

## 1 Introduction

The field of drug investigation aimed at seizure prevention during epilepsy has become a cornerstone in medicinal chemistry due to the limitations of existing antiepileptic medications such as ineffectiveness or various side effects: diplopia, dizziness, ataxia, sedation, and so on [1]. In this context, particular attention is paid on natural biologically active substances including glycosides, coumarins, flavonoids, alkaloids, and terpenoids that have limited or no toxic side effects [2]. The latter compounds were found to positively modulate human recombinant γ-aminobutyric acid type A receptors ($GABA_A$) receptors or activate glutamic acid decarboxylase enzyme (a key enzyme in the biosynthesis of GABA) confirming terpenoids' potential for the treatment of epilepsy and seizures [3]. To enhance anticonvulsant properties, terpenes/terpenoids were synthetically converted to miscellaneous derivatives comprising hydrazones [4–6], epoxides [7], and esters [8,9]. The main concept of modification consists in pharmacophore fragments introduction into the molecule leading to increased bonding to receptors. In contrast, the labile bonds of the aforementioned compounds undergo hydrolytic cleavage both *in vitro* and *in vivo* followed by the release of active ingredients [10].

To date, the creation of a new medicine costs more than 1 billion dollars and the costs of this process are growing up steadily [11]. Therefore, different theoretical approaches are used for a facilitation and acceleration of a new drugs development process that is very expensive, multistep, and prolonged [12]. To this end, quantitative structure–activity relationship (QSAR) analysis is widely applied for virtual screening of potential bioactive compounds followed by their structure optimization. QSAR

* **Corresponding author: Mariia Nesterkina,** Department of Organic and Pharmaceutical Technology, Odessa National Polytechnic University, Odessa 65044, Ukraine, e-mail: mashaneutron@gmail.com, tel: +380-93-713-3853
**Viacheslav Muratov, Luidmyla Ognichenko, Victor Kuz'min:** Department of Molecular Structures and Chemoinformatics, A.V. Bogatsky Physical-Chemical Institute NAS of Ukraine, Odessa 65080, Ukraine
**Iryna Kravchenko:** Department of Organic and Pharmaceutical Technology, Odessa National Polytechnic University, Odessa 65044, Ukraine

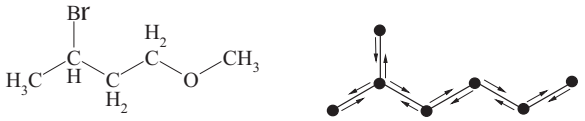
methods allow discovering correlation between numerical physicochemical property or molecular feature of compounds and their biological or pharmacological activity [13]. The present study is a logical continuation of our current research devoted to the QSAR analysis for a number of terpene derivatives [14]. The goal of our study is to develop predictive QSAR models for anticonvulsant activity assessment of terpenoids and their derivatives.

## 2 Materials and methods

The data set consisting of 36 terpene derivatives was used in the current study: hydrazones of menthone, carvone, and verbenone and esters of menthol, thymol, carvacrol, borneol, eugenol, and guaiacol. Minimum effective doses (MEDs) of convulsant pentylenetetrazole (PTZ) that induce clonic–tonic convulsions (DCTC) and tonic extension (dose inducing tonic extension [DTE]) in experimental animals have been registered and used as properties. Anticonvulsant activity of terpene derivatives was estimated at different time points – 3 and 24 h after oral administration according to the described procedure [15]. Thus, four data sets were formed: two of which contain information on PTZ MED causing clonic–tonic seizures determined after 3 and 24 h, DCTC (3 h) and DCTC (24 h), and the other two with data on PTZ MED causing tonic extension, DTE (3 h) and DTE (24 h), respectively.

For all the investigated compounds, the structural descriptors were calculated using the simplex representation of molecular structure (SiRMS; informational field [IF]) approach [16]. QSAR methods on the basis of SiRMS [17,18] along with the IF of molecule [19] already demonstrated their efficiency when dealing with variety of QSAR tasks. Two different methods are united in the system of structural parameters resulting in system (SiRMS (IF)) that exhibits versatility of simplex approach along with sensitivity to the structural features within the scope of IF. Through the use of SiRMS (IF) approach, simplex vertexes might be better differentiated since informational potential (IP) vary according to the impact of near and far surroundings. In SiRMS (IF), the IF characteristics are applied for atom's differentiation in simplexes. Notably, the potentials of the aforementioned IFs weighed by different atomic properties (charge, lipophilicity, electronic polarizability, etc.) might be exploited.

In this study, SiRMS (IF) approach has been applied for the representation of molecular structure at 2D level. According to this model, each molecule can be represented in the form of molecular graph. The vertices of such a graph provide a whole range of information which spreads only on the graph edges. Noteworthy, all potential bindings between every atom pair are recognized:




The topological IP of *i*th atom can be described as [16]

$$\mathrm{IP}_i = w_i \cdot \sum_{j=1}^{n} \left( \frac{\sum_m \mathrm{lb}\left(\frac{r}{2R_{ij}+1}\right)}{m} \right),$$

where $m$ is the number of all possible paths between every atom pair, $n$ is the number of atoms in the given molecule, $R_{ij}$ is the number of bonds between the $i$th and $j$th atoms (path length), $w_i$ is the weighed parameters describing any property ($p$) of the atoms ($w_i = 1$ in the case of unweighted IP), and $r$ is the tuning parameter of the model, $w_i = p_i / \sum_{i=1}^{n} p_i$.

The example of simplex differentiation on the basis of atomic topologic IP for alanine at 2D level is shown in Figure 1.

For the weighed molecular topological IPs in the preliminary stage, the range of all possible values is divided into a certain number of discrete groups. The number of groups is a tuning parameter. The label of an atom is determined from the point of view of its belonging to a particular group. The following schemes were used in the calculation of simplex-informational descriptors:

Unweighted topological IP: $A < 3.45 \le B < 3.84 \le C < 4.22 \le D < 4.60 \le E < 4.99 \le F < 5.37 \le G$.

Topological IP weighed by electronegativity: $A < 3.51 \le B < 3.89 \le C < 4.27 \le D < 4.65 \le E < 5.03 \le F < 5.41 \le G$.

Topological IP weighed by refraction: $A < 3.65 \le B < 4.02 \le C < 4.38 \le D < 4.75 \le E < 5.12 \le F < 5.49 \le G$.

Topological IP weighed by atomic charge: $A < -0.51 \le B < -0.38 \le C < -0.26 \le D < -0.13 \le E < 0 \le F < 0.12 \le G$.

Topological IP weighed by van der Waals radius: $A < 3.75 \le B < 4.11 \le C < 4.47 \le D < 4.83 \le E < 5.2 \le F < 5.56 \le G$.

Topological IP weighed by lipophilicity: $A < 0.58 \le B < 1.32 \le C < 2.06 \le D < 2.81 \le E < 3.55 \le F < 4.29 \le G$.

A total of 2,374 simplex-informational descriptors were calculated for the investigated molecules. $\log P$ (I.log$P$), molecular refraction (I.RF), equalized electronegativity (I.EN), and molecular mass (I.AW) were used as trivial

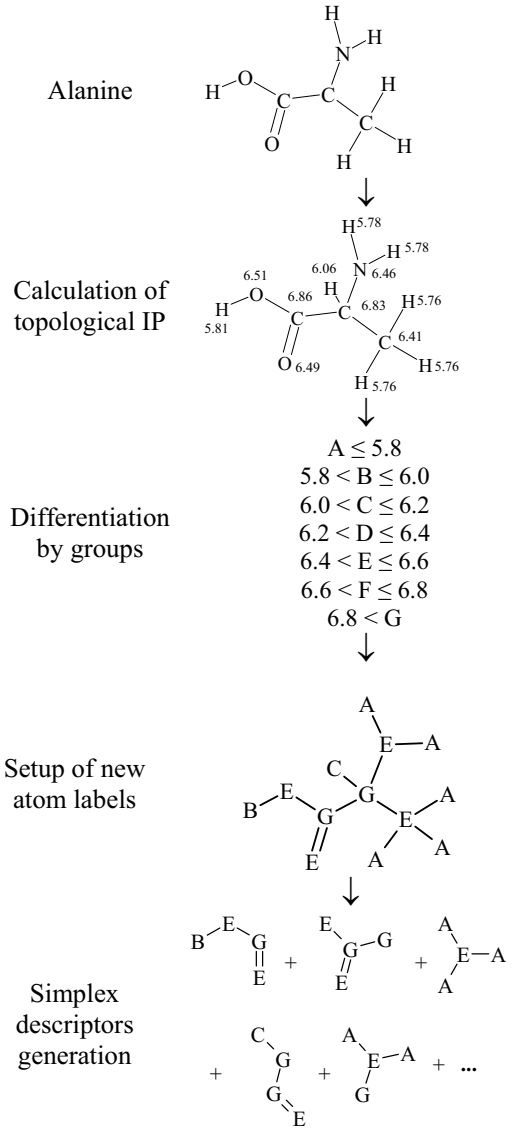



**Figure 1:** Example of simplex descriptor generation using atomic topological IPs for alanine at 2D level.

integral parameters of the molecules [17,18]. The relationships between calculated descriptors and investigated properties were established with common method of partial least squares (PLS) [20]. The removal of highly correlated and constant descriptors, the trend vector method [21], genetic algorithm [22], and the automatic variable selection strategy [18] have all been used to select the descriptors in PLS. Structures of monocyclic/bicyclic terpenoids and their derivatives used in this study and corresponding values of their anticonvulsant activity are collected in Table 1.

QSAR investigations must be used to make predictions for compounds with unknown activity. Five-fold cross-validation was performed for each property to assess the prediction ability. The molecules within each data set are ordered according to the values of their anticonvulsant activity and then every fifth molecule in this row is selected into the test set. Significantly, among the five data sets obtained, each molecule enters in the test set only once. The test set (20% of the total number of molecules) is excluded from the model building process. Based on the training set data, a model used to predict the properties of test set compounds is built. Thus, for each of the four data sets (DCTC [3 h], DCTC [24 h], DTE [3 h], DTE [24 h]), five data sets are formed, each containing training and test sets.

According to the OECD principles of QSAR model validation, in this task, the following domain applicability (DA) procedures were used:

(a) This method is based on the estimation of leverage value $h_i$. It has been visualized as the Williams plot [23]. The new molecule is situated outside of the DA if $h_i > h_{cr} = 3(A + 1)/N$, where $N$ is the number of molecules in a work set.

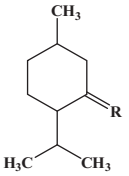
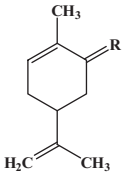
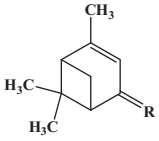
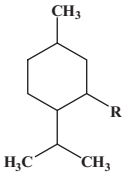
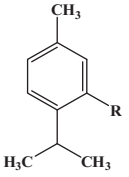
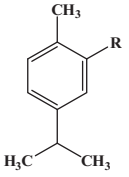
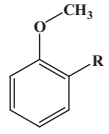
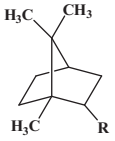
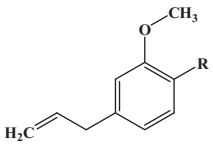
(b) DA ellipsoid that developed [17], where the distribution of a training set of molecules in a space of latent variables (A) $T_1$–$T_A$ (axes of coordinates), can be obtained from the PLS. DA represents an ellipsoid that is built from the molecules at the center of the training set ($T_1 = 0$; $T_2 = 0$) with the semiaxes length $3ST_1$ and $3ST_2$, respectively (where $ST_1$ and $ST_2$ are the root mean square deviations). The prognoses for molecules approximated to the center of the DA are the most reliable.

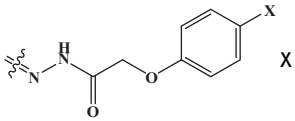
The "randomization" procedure (Y-scrambling) [18] was used to confirm the "nonrandomness" of the developed QSAR models. Y-scrambling procedure is the creation of the models using the method of random selection of coefficients. The essence of this method is that after developing the model on a work set, random mixing of the values of the property is carried out. If the model is retrained, then a fairly good model will be obtained on the Y-randomized set. If the model parameters do not lead to retraining, then after Y-randomization, there will be a significant deterioration in the descriptive ability of the model.

The following formulas were used for calculating of statistical characteristics:

$$\text{Coefficient of determination } R^2 = \frac{\sum_{i=1}^{N}(y_i^{obs} - \bar{y})^2 - \sum_{i=1}^{N}(y_i^{calc} - y_i^{obs})^2}{\sum_{i=1}^{N}(y_i^{obs} - \bar{y})^2},$$

**Table 1:** Structures and anticonvulsant activities of terpenes and their derivatives

| Structure | Core | $R_i$ | 3 h | | 24 h | |
|---|---|---|---|---|---|---|
| | | | DCTC (%) | DTE (%) | DCTC (%) | DTE (%) |
|  | A | $R_1$ | 157 | 197 | 166 | 176 |
| | | $R_2$ | 171 | 182 | 191 | 205 |
| | | $R_3$ | 223 | 261 | 200 | 224 |
| | | $R_4$ | 177 | 203 | 177 | 189 |
| | | $R_5$ | 189 | 232 | 166 | 182 |
| | | $R_6$ | 174 | 179 | 169 | 182 |
|  | B | $R_1$ | 196 | 200 | 130 | 138 |
| | | $R_2$ | 193 | 204 | 240 | 220 |
| | | $R_3$ | 267 | 269 | 204 | 204 |
| | | $R_4$ | 227 | 246 | 204 | 208 |
| | | $R_5$ | 221 | 233 | 209 | 227 |
| | | $R_6$ | 196 | 204 | 204 | 212 |
|  | C | $R_1$ | 179 | 180 | 130 | 138 |
| | | $R_2$ | 218 | 212 | 240 | 220 |
| | | $R_3$ | 278 | 303 | 204 | 204 |
| | | $R_4$ | 207 | 235 | 204 | 208 |
| | | $R_5$ | 216 | 230 | 209 | 227 |
| | | $R_6$ | 160 | 162 | 204 | 212 |
|  | D | $R_7$ | 167 | 186 | 167 | 189 |
| | | $R_8$ | 162 | 192 | 162 | 194 |
| | | $R_9$ | 152 | 186 | 167 | 189 |
|  | E | $R_7$ | 213 | 181 | 214 | 183 |
| | | $R_8$ | 207 | 190 | 208 | 189 |
| | | $R_9$ | 198 | 144 | 198 | 132 |
|  | F | $R_7$ | 197 | 142 | 197 | 127 |
| | | $R_8$ | 235 | 155 | 226 | 152 |
| | | $R_9$ | 154 | 136 | 144 | 133 |
|  | G | $R_7$ | 186 | 157 | 179 | 147 |
| | | $R_8$ | 214 | 207 | 203 | 203 |
| | | $R_9$ | 183 | 114 | 186 | 108 |
|  | H | $R_7$ | 190 | 207 | 183 | 208 |
| | | $R_8$ | 242 | 161 | 267 | 171 |
| | | $R_9$ | 217 | 165 | 217 | 158 |
|  | I | $R_7$ | 240 | 156 | 241 | 155 |
| | | $R_8$ | 210 | 156 | 212 | 159 |
| | | $R_9$ | 168 | 168 | 159 | 176 |

Note: $R_1$: O; $R_2$–$R_6$:  X = H ($R_2$); Br ($R_3$); Cl ($R_4$); $C(CH_3)_3$ ($R_5$); O–Ph ($R_6$); $R_7$: –OOC–$(CH_2)_3$–$NH_2$; $R_8$: –OOC–$CH_2$–$NH_2$; $R_9$: –OH.

DCTC, dose inducing clonic–tonic convulsions; DTE, dose inducing tonic extension.

where $y_i^{obs}$ is the observed value of property for $i$ molecule, $y_i^{calc}$ is the calculated (predicted) value of property for $i$ molecule, and $N$ is the number of molecules, $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i^{obs}$;

$$\text{Root mean square error (RMSE)} = \sqrt{\frac{\sum_{i=1}^{N}(y_i^{calc} - y_i^{obs})^2}{N}},$$

$$\text{Mean absolute error (MAE)} = \frac{1}{N}\sum_{i=1}^{N}|(y_i^{calc} - y_i^{obs})|.$$

To calculate simplex-informational descriptors and integral parameters of the molecules and also to develop QSAR models, software HiTQSAR developed in the Department of Molecular Structures and Chemoinformatics (A.V. Bogatsky Physical-Chemical Institute NAS of Ukraine) was used.

## 3 Results and discussion

At the first stage, in the construction of PLS ratios, all molecules were included in the study (a training set of 36 terpene derivatives). For DCTC (3 h), DCTC (24 h), DTE (3 h), and DTE (24 h) data sets 2D PLS QSAR models M1–M4 were developed. Models M1–M4 had the number of latent variables $A = 2$ and adequate statistical characteristics (Table 2).

The models obtained using the Y-scrambling procedure had low values of $R^2_{(Y\text{-scr})} < 0.29$ and $Q^2_{(Y\text{-scr})} < 0.14$, indicating that the nonrandomness of the established relationship between the structure of compounds and their anticonvulsant activities obtained from the previous models M1–M4.

Five-fold cross-validation was performed for each property to assess the prediction ability. Adequate 2D QSAR models were obtained and used further as basis for the consensus models. For consensus models, the coefficient of determination for test sets was found to be ($R^2_{test}$) > 0.62 (Table 3).

**Table 3:** The statistical characteristics of consensus models (test set)

| Anticonvulsant activity | $R^2$ | RMSE | MAE |
|---|---|---|---|
| DCTC (3 h) | 0.62 | 18.8 | 15.3 |
| DCTC (24 h) | 0.79 | 13.9 | 11.6 |
| DTE (3 h) | 0.75 | 17.6 | 14.3 |
| DTE (24 h) | 0.81 | 14.7 | 11.2 |

Noteworthy, the quality of the models for 24 h is better than those for 3 h that, obviously, may be due to the fact that some compounds fully manifest their activity after time. Relationship between observed and predicted values of anticonvulsant activities for the test set within the consensus models for four sets is presented in Figure 2a–d. As described, significant correlation between observed and predicted values of anticonvulsant activities in DCTC (3 h), DCTC (24 h), DTE (3 h), and DTE (24 h) sets was revealed for the majority of compounds.

Physicochemical and structural interpretation of the models was carried out based on consensus 2D-QSAR models developed for the DCTC (3 h), DCTC (24 h), DTE (3 h), and DTE (24 h) samples. To define the relative influence of the different physical and chemical factors on the character of the molecules interaction with the biological target, it is necessary to sum and compare absolute values of PLS regression coefficients of simplexes for all used groups for atom differentiation. Figure 3 shows the relative influence of some physicochemical factors on the anticonvulsant activity of terpene derivatives at 3 and 24 h after oral administration. As seen, the analysis of the obtained 2D QSAR model using SiRMS (IF) approach confirms that in all cases the contribution of electrostatic factor is predominant (48–63%). Equally important factors are the steric factors that are determined, on the one hand, by values of the van der Waals radii (20–35%), and on the other hand, by topology which is associated with the molecular form (13–22%). Lipophilicity effect in all

**Table 2:** The statistical characteristics of 2D QSAR models

| Model | Anticonvulsant activity | Work set | | | Leave-one-out | | | Y-Scrambling | |
|---|---|---|---|---|---|---|---|---|---|
| | | $R^2$ | RMSE | MAE | $Q^2$ | RMSE | MAE | $R^2$ | $Q^2$ |
| M1 | DCTC (3 h) | 0.88 | 10.6 | 7.9 | 0.81 | 13.5 | 10.1 | 0.26 ± 0.03 | 0.11 ± 0.03 |
| M2 | DCTC (24 h) | 0.91 | 9.2 | 7.4 | 0.81 | 13.5 | 10.6 | 0.17 ± 0.02 | 0.08 ± 0.02 |
| M3 | DTE (3 h) | 0.80 | 15.6 | 12.1 | 0.75 | 18.0 | 14.1 | 0.14 ± 0.02 | 0.04 ± 0.03 |
| M4 | DTE (24 h) | 0.90 | 9.9 | 8.0 | 0.86 | 12.4 | 9.8 | 0.19 ± 0.03 | 0.11 ± 0.03 |

Note: $R^2$ – coefficient of determination; $Q^2$ – coefficient of determination for cross validation (leave-one-out); RMSE – root mean square error; MAE – mean absolute error.

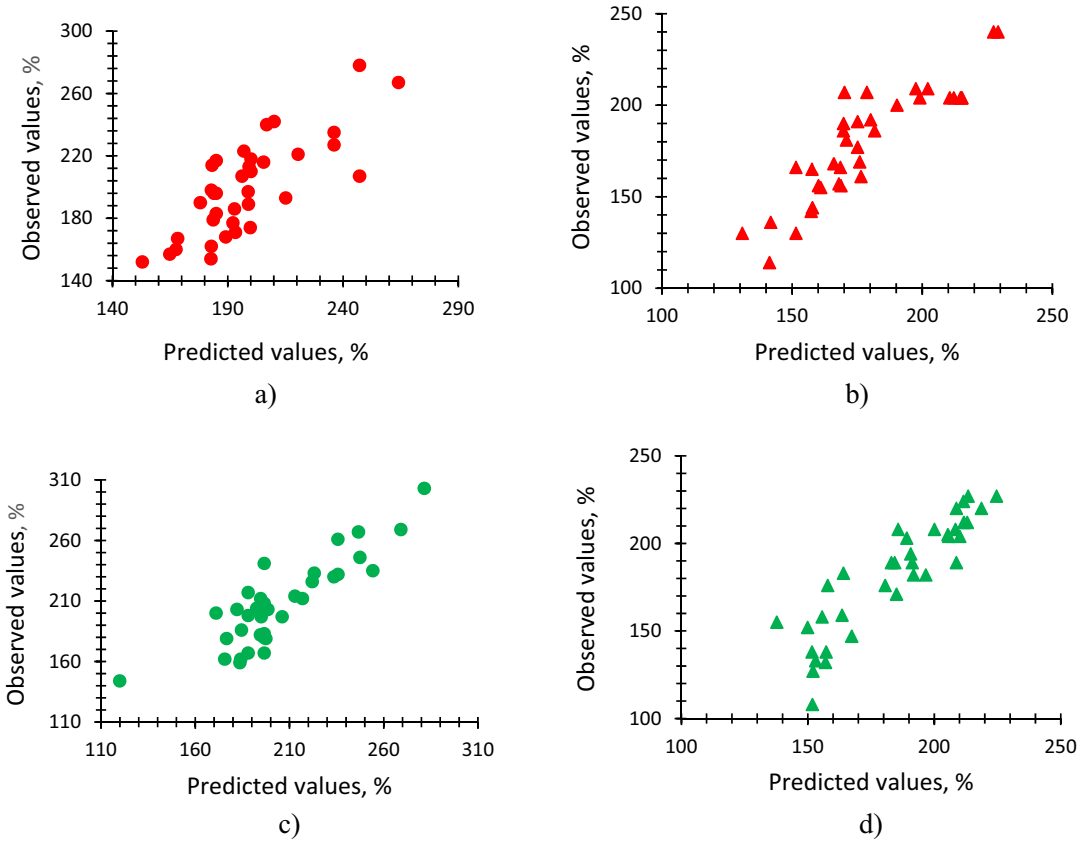



**Figure 2:** Observed versus predicted diagram for anticonvulsant activities of terpene derivatives: (a) DCTC (3 h); (b) DCTC (24 h); (c) DTE (3 h); and (d) DTE (24 h).

instances was minimal, less than 9% from the sum of all considered factors.

One of SiRMS approach advantages, also proper to the SiRMS (IF), is the possibility of appropriate structural interpretation. The influence of each atom (as a part of simplex) into investigated property can be calculated on the basis of the obtained QSAR models [17,18]. The contribution ($C$) of each $j$-atom in the molecule can be defined as the ratio of the sum of PLS regression coefficient ($b_I$) of all simplexes this atom contains ($M$) to the number of atoms ($n$) in the simplex (or fragment): $C_j = \frac{1}{n}\sum_{I=1}^{M} b_I$ (for simplex $n = 4$). According to this formula, the atomic contribution depends on the number of simplexes that include this atom. This value (count of simplexes) is not constant, and it varies in different molecules and depends on other constituents (surroundings); hence, this contribution is nonadditive [16]. The analysis of such information allowed selecting molecular fragments which have positive and negative influence on considered activity.

The entire set consisting of 36 terpene derivatives was divided into nine subgroups in which the effect of R substituents was analyzed. The contribution of the structural cores A–I was also evaluated (Table 4).

The higher the value of the MEDs of seizure agent that induce DCTC and DTE, the greater the anticonvulsant effect of the compound. As demonstrated in Table 3, carvone (B) and verbenone (C) nuclei were discovered to facilitate the manifestation of anticonvulsant activity in the case of DCTC (3 h) and DCTC (24 h). The influence of eugenol nucleus (I) in its derivatives prevents the manifestation of anticonvulsant activity with increase in time after oral administration of compounds. The effects of menthol (D) and menthone (A) nuclei become positive during the transition from 3 to 24 h after the administration of their derivatives; consequently, the presence of the aforementioned nuclei in the molecule structure will contribute to the expression of antiseizure action. This phenomenon might be explained by higher lipophilicity of menthone (log $P$ 2.63) core compared with verbenone (log $P$ 1.97) and carvone (log $P$ 2.26) leading to prolonged action of menthone derivatives and their more pronounced activity over a long time period (24 h).

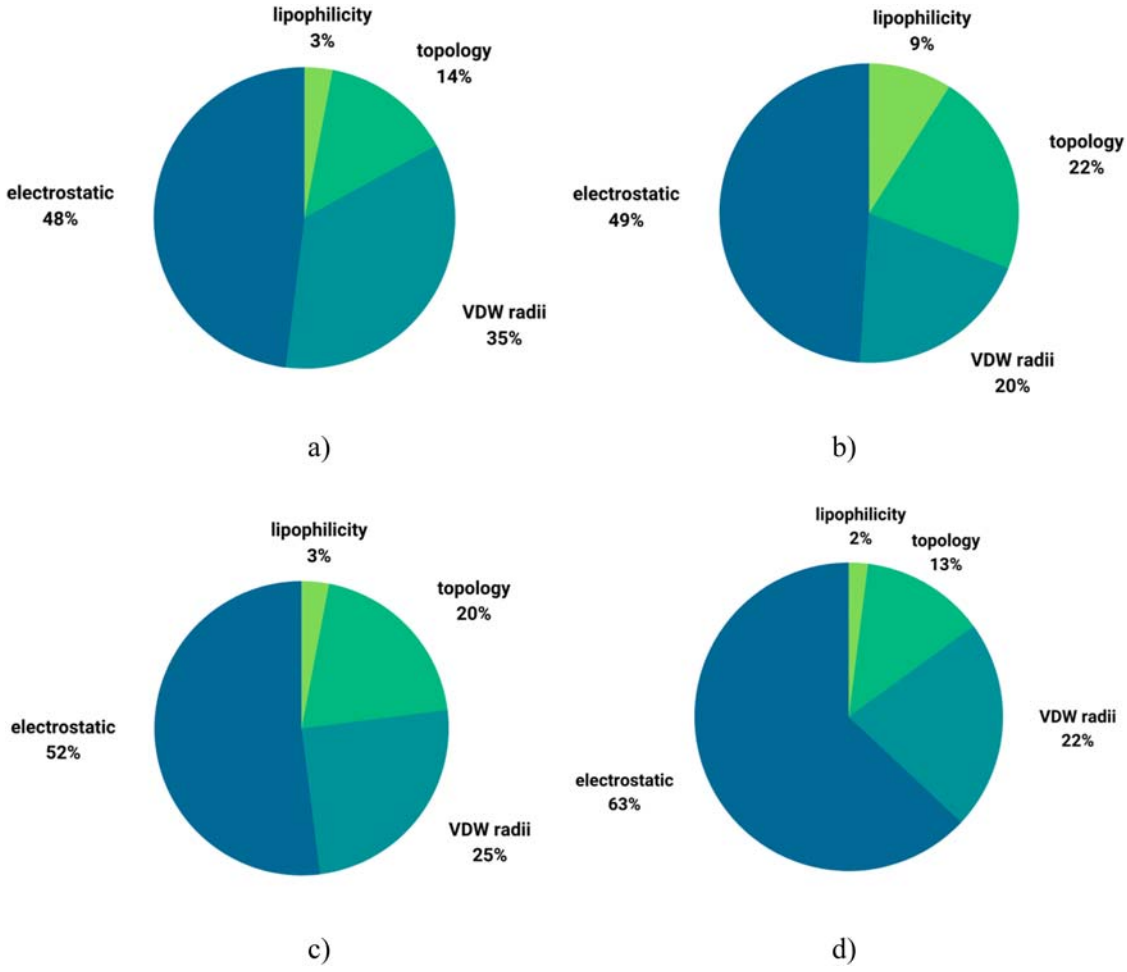



**Figure 3:** Relative influence of some structural factors on anticonvulsant activity of terpene derivatives estimated on the basis of consensus 2D-QSAR models: (a) DCTC (3 h); (b) DCTC (24 h); (c) DTE (3 h); and (d) DTE (24 h).

In the case of DTE (3 h), the nuclei of verbenone (C), carvone (B), and borneol (H) derivatives were the most conducive to the manifestation of anticonvulsant activity. Similarly to DCTC, the presence of menthol nucleus (D) initially prevents the expression of effect; however, over time its effect becomes reversed leading to enhanced seizure protection due to gradual enzymatic hydrolysis of ester bond in menthol derivatives. Generally, the presence of carvone (B), verbenone (C), and borneol (H) nuclei in the structure of terpenoid derivatives at all time periods (3 and 24 h) promote the demonstration of anticonvulsant activity on the model of PTZ-induced seizures. Comparative analysis of substituent (R) contribution to anticonvulsant effect was also conducted and its results are summarized in Table 5.

When analyzing the influence of substituents ($R_1$–$R_9$) on the manifestation of anticonvulsant activity, it was found that the replacement of the carbonyl group linked to the cyclic terpene fragment with hydrazone substituents as well as replacement of the hydroxyl group with GABA or glycine residues resulted in strengthening of antiseizure action.

**Table 4:** Relative contribution of structural cores on anticonvulsant activity of terpenoids and their derivatives

| Cores | DCTC (3 h) | DCTC (24 h) | DTE (3 h) | DTE (24 h) |
|---|---|---|---|---|
| A | −1.3 | 4.8 | 8.5 | 7.5 |
| B | 17.7 | 18.7 | 18.8 | 9.6 |
| C | 14.8 | 13.1 | 19.8 | 5.9 |
| D | −18.4 | 1.2 | −12.8 | 23.8 |
| E | 6.6 | −2.9 | 4.8 | 1.2 |
| F | 8.6 | −15.4 | −0.1 | −16.8 |
| G | 1.2 | −12.9 | −0.6 | −9.5 |
| H | 5.5 | 0.5 | 17.9 | 8.7 |
| I | 15.3 | −4.9 | 11.1 | −5.0 |

**Table 5:** Relative contribution of molecular fragments on anticonvulsant activity of terpenoids and their derivatives

| Structural cores | Fragments (R) | DCTC (3 h) | DCTC (24 h) | DTE (3 h) | DTE (24 h) |
|---|---|---|---|---|---|
| A | $R_1$ | 0.1 | 0 | 0 | 0.4 |
| | $R_2$ | 3.1 | 14.1 | −2.9 | 33.8 |
| | $R_3$ | 20.3 | 10.2 | 32.7 | 34.6 |
| | $R_4$ | 7.2 | 8.8 | 4.9 | 33.7 |
| | $R_5$ | 8.4 | 1.9 | 30.3 | 20.0 |
| | $R_6$ | −6.3 | 4.3 | −10.2 | 10.2 |
| B | $R_1$ | 0.1 | 0 | 0 | 0.4 |
| | $R_2$ | 9.1 | 21.6 | 0.1 | 40.3 |
| | $R_3$ | 24.6 | 14.9 | 30.0 | 35.7 |
| | $R_4$ | 15.8 | 17.5 | 11.8 | 37.3 |
| | $R_5$ | 25.1 | 15.5 | 38.9 | 48.0 |
| | $R_6$ | 4.2 | 16.1 | −3.2 | 40.3 |
| C | $R_1$ | 0.1 | 0 | 0 | 0.4 |
| | $R_2$ | 12.9 | 24.9 | 3.4 | 39.8 |
| | $R_3$ | 25.2 | 18.1 | 37.2 | 35.9 |
| | $R_4$ | 15.8 | 19.6 | 27.8 | 36.0 |
| | $R_5$ | 18.8 | 17.5 | 33.1 | 47.4 |
| | $R_6$ | −8.3 | 14.5 | −7.2 | 38.8 |
| D | $R_7$ | −3.9 | 4.9 | −2.3 | 5.7 |
| | $R_8$ | −7.6 | 6.5 | −4.2 | 7.6 |
| | $R_9$ | −5.1 | 1.3 | −0.5 | 1.3 |
| E | $R_7$ | 5.9 | 3.6 | 4.4 | 3.2 |
| | $R_8$ | 6.6 | 4.1 | 13.6 | 7.0 |
| | $R_9$ | −0.3 | 0.2 | 0 | 0.5 |
| F | $R_7$ | 25.7 | −2.3 | 15.2 | 4.4 |
| | $R_8$ | 4.6 | −2.0 | 6.6 | 1.2 |
| | $R_9$ | −4.6 | −0.02 | −0.1 | 0.5 |
| G | $R_7$ | 3.9 | 18.6 | 2.3 | 26.9 |
| | $R_8$ | 0.3 | 2.2 | 2.4 | 8.4 |
| | $R_9$ | −0.6 | −0.2 | −0.8 | 0.5 |
| H | $R_7$ | 24.9 | −1.7 | 31.2 | 4.4 |
| | $R_8$ | 3.9 | 8.1 | −1.4 | 12.7 |
| | $R_9$ | 2.0 | −0.3 | 3.7 | 0.5 |
| I | $R_7$ | 14.5 | 1.9 | 5.6 | 6.3 |
| | $R_8$ | 8.1 | 2.3 | 12 | 6.0 |
| | $R_9$ | 0.1 | 0.9 | −1.5 | 2.2 |

DCTC, dose inducing clonic–tonic convulsions; DTE, dose inducing tonic extension.

# 4 Conclusion

In the present study, QSAR models for anticonvulsant activity prediction of hydrazones and esters based on monocyclic/bicyclic terpenoids were developed using SiRMS (IF) approach, which captured the essence of simplex representation for molecular structure and its IF. Physicochemical and structural interpretation of the models was carried out. The contribution of electrostatic factor on the anticonvulsant activity of terpene derivatives at 3 and 24 h after oral administration is predominant (48–63%). Equally important factors are the steric factors: values of the van der Waals radii (20–35%) and topology (13–22%). Molecular fragments responsible for the manifestation of anticonvulsant action defined on the basis of the developed models. In particular, we found that carvone, verbenone cores possess the most significant contribution to antiseizure action examined on the model of PTZ-induced convulsions at 3 and 24 h after oral administration of terpene derivatives. Moreover, carbonyl and hydroxy group substitution in terpenoid molecules followed by hydrazones and esters formation leads to enhancement and prolongation of antiseizure action due to the contribution of additional molecular fragments. The presented QSAR models contribute significantly to design of novel terpene derivatives for oral administration against onset seizures.

# Abbreviations

| | |
|---|---|
| DA | domain applicability |
| DCTC | dose inducing clonic–tonic convulsions |
| DTE | dose inducing tonic extension |
| MAE | mean absolute error |
| MED | minimum effective dose |
| PTZ | pentylenetetrazole |
| QSAR | quantitative structure–activity relationship |
| RMSE | root mean square error |
| SiRMS | simplex representation of molecular structure |
| IF | informational field |
| IP | informational potential |
| PLS | partial least squares |

**Funding information:** No funding to declare.

**Author contributions:** All authors contributed to the study conception and design. M.N., L.O., and V.M. performed material preparation, data collection, and writing of original draft. I.K. and V.K.: supervision, editing of original draft. All authors read and approved the final manuscript.

**Conflict of interest:** The authors declare no potential conflict of interest.

**Ethical approval:** The conducted research is not related to either human or animal use.

**Data availability statement:** The data sets generated during the present study are available from the corresponding author on reasonable request.

# References


[1] St Louis EK. Minimizing AED adverse effects: improving quality of life in the interictal state in epilepsy care. Curr Neuropharmacol. 2009;7(2):106–14. doi: 10.2174/157015909788848857.

[2] He LY, Hu MB, Li RL, Zhao R, Fan LH, He L, et al. Natural medicines for the treatment of epilepsy: bioactive components, pharmacology and mechanism. Front Pharmacol. 2021;12:604040. doi: 10.3389/fphar.2021.604040.

[3] Manayi A, Nabavi SM, Daglia M, Jafari S. Natural terpenoids as a promising source for modulation of GABAergic system and treatment of neurological diseases. Pharmacol Rep. 2016;68(4):671–9. doi: 10.1016/j.pharep.2016.03.014.

[4] Nesterkina M, Barbalat D, Konovalova I, Shishkina S, Atakay M, Salih B, et al. Novel (–)-carvone derivatives as potential anticonvulsant and analgesic agents. Nat Prod Res. 2021;35(23):4978–87. doi: 10.1080/14786419.2020.1756804.

[5] Nesterkina M, Barbalat D, Kravchenko I. Design, synthesis and pharmacological profile of (–)-verbenone hydrazones. Open Chem. 2020;18(1):943–50. doi: 10.1515/chem-2020-0103.

[6] Nesterkina M, Barbalat D, Zheltvay I, Rakipov I, Atakay M, Salih B, et al. 2*S*,5*R*)-2-Isopropyl-5-methylcyclohexanone hydrazones. Molbank. 2019;2019(2):M1062. doi: 10.3390/M1062.

[7] Salgado PR, da Fonsêca DV, Braga RM, de Melo CG, Andrade LN, de Almeida RN, et al. Comparative anticonvulsant study of epoxycarvone stereoisomers. Molecules. 2015;20(11):19660–73. doi: 10.3390/molecules201119649.

[8] Nesterkina M, Kravchenko I. Synthesis and pharmacological properties of novel esters based on monocyclic terpenes and GABA. Pharmaceuticals. 2016;9(2):32. doi: 10.3390/ph9020032.

[9] Nesterkina M, Kravchenko I. Synthesis and anticonvulsant activity of menthyl γ-aminobutyrate. Chem Nat Compd. 2016;52:237–9. doi: 10.1007/s10600-016-1604-9.

[10] Angelova V, Karabeliov V, Andreeva-Gateva PA, Tchekalarova J. Recent developments of hydrazide/hydrazone derivatives and their analogs as anticonvulsant agents in animal models. Drug Dev Res. 2016;77(7):379–92. doi: 10.1002/ddr.21329.

[11] Ooms F. Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry. Curr Med Chem. 2000;7(2):141–58. doi: 10.2174/0929867003375317.

[12] Xu J, Hagler A. Chemoinformatics and drug discovery. Molecules. 2002;7(8):566–600. doi: 10.3390/70800566.

[13] Bellera CL, Talevi A. Quantitative structure–activity relationship models for compounds with anticonvulsant activity. Expert Opin Drug Discov. 2019;14(7):653–65. doi: 10.1080/17460441.2019.1613368.

[14] Nesterkina M, Ognichenko L, Shyrykalova A, Kravchenko I, Kuz'min V. QSAR models for analgesic activity prediction of terpenes and their derivatives. Struct Chem. 2020;31:947–54. doi: 10.1007/s11224-019-01479-7.

[15] Nesterkina MV, Alekseeva EA, Kravchenko IA. Synthesis, physicochemical properties, and anticonvulsant activity of the GABA complex with a calix[4]arene derivative. Pharm Chem J. 2014;48:82–4. doi: 10.1007/s11094-014-1052-4.

[16] Ognichenko L, Kuz'min V, Artemenko A. New structural descriptors of molecules on the basis of symbiosis of the informational field model and simplex representation of molecular structure. QSAR Comb Sci. 2009;28(9):939–45. doi: 10.1002/qsar.200860073.

[17] Kuz'min VE, Artemenko AG, Muratov EN. Hierarchical QSAR technology based on the Simplex representation of molecular structure. J Comput Aided Mol Des. 2008;22(6–7):403–21. doi: 10.1007/s10822-008-9179-6.

[18] Kuz'min VE, Artemenko AG, Ognichenko LN, Hromov AI, Kosinskaya AP, Stelmakh SI, et al. Simplex representation of molecular structure as universal QSAR/QSPR tool. J Struct Chem. 2021;32:1365–92. doi: 10.1007/s11224-021-01793-z.

[19] Kuz'min V, Ognichenko L, Artemenko A. Modeling of the informational field of molecules. J Mol Model. 2001;7:278–85. doi: 10.1007/s008940100036.

[20] Rannar S, Lindgren F, Geladi P, Wold SA. PLS kernel algorithm for data sets with many variables and fewer objects. Part 1: theory and algorithm. J Chemometrics. 1994;8(2):111–25. doi: 10.1002/cem.1180080204.

[21] Carhart R, Smith DH, Venkataraghavan R. Atom pairs as molecular features in structure–activity studies: definition and applications. J Chem Inf Comput Sci. 1985;25(2):64–73. doi: 10.1021/ci00046a002.

[22] Hasegawa K, Miyashita Y, Funatsu K. GA strategy for variable selection in QSAR studies: GA-based PLS analysis of calcium channel antagonists. J Chem Inf Comput Sci. 1997;37(2):306–10. doi: 10.1021/ci960047x.

[23] Meloun M, Militku J, Hill M. Crucial problems in regression modelling and their solutions. Analyst. 2002;127:433–50. doi: 10.1039/b110779h.