COMPUTER-BASED PREDICTION OF TOXICITY USING THE ELECTRON-CONFORMATIONAL METHOD. APPLICATION TO FRAGRANCE ALLERGENS AND OTHER ENVIRONMENTAL POLLUTANTS

Dedicated to the 60th birthday of Academician Gheorghe Duca

Natalia N. Gorinchoy[a]* and Isaac B. Bersuker[a,b]

[a]*Institute of Chemistry, Academy of Sciences of Moldova, Academiei str. 3, MD 2028 Kishinev, Republic of Moldova*
[b]*Department of Chemistry & Biochemistry, The University of Texas at Austin, Austin, Texas 78712, USA*
*E-mail: ngorinchoy@yahoo.com; Phone 373 22 739675

Abstract: The electron-conformational (EC) method is employed for the toxicophore (Tph) identification and quantitative prediction of toxicity using the training set of 24 compounds that are considered as fragrance allergens. The values of $a=LD_{50}$ in oral exposure of rats were chosen as a measure of toxicity. EC parameters are evaluated on the base of conformational analysis and *ab initio* electronic structure calculations (including solvent influence). The Tph consists of four sites which in this series of compounds are represented by three carbon and one oxygen atoms, but may be any other atoms that in the compound under consideration exhibit the same electronic and geometric features within the tolerance limits. The regression model taking into consideration the Tph flexibility, anti-Tph shielding, and influence of out-of-Tph functional groups predicts well the experimental values of toxicity ($R^2 = 0.93$) with a reasonable leave-one-out cross-validation. The methodology worked out in this paper can be applied to other environmental pollutants.

Keywords: Structure-activity relationships, Electron-conformational method, Toxicity screening and prediction, fragrance allergens.

Abbreviations: Tph –toxicophore
EC – electron-conformational
ECMC – electron-conformational matrix of congruity
ECSA – electron-conformational sub-matrix of activity
QSAR – quantitative structure-activity relationships

1. Introduction

Fragrance chemicals are of increased interest to both the fragrance industry and environmental protection agencies, as well as to dermatologists, due to a number of allergies caused by these chemicals in cosmetics and toiletries. For this reason, use of fragrance chemicals is subjected to strict restrictions. Considerable attention is paid to safety, stability and innocuity of the substances (natural or artificial) used as fragrance ingredients. Checking the latter on toxicity using animal testing is extremely expensive and of long duration. Therefore alternative, computer-based methods of screening and prediction of toxicity for large number of the fragrance chemicals based on limited numbers of experimentally studied ones, becomes very attractive.

Computer-based toxicology is presently a well recognized and rapidly developing trend in environmental chemistry. The QSAR (Quantitative Structure-Activity Relationship) methodology is usually used to reveal relationships between the chemical structure of the compound and its specific toxicity in order to predict the latter in new chemicals. There are a number of QSAR systems and commercial programs, which allow prediction of a wide range of biological endpoints and toxicity (see, e.g., [1-5] and references therein). The main problem in this approach is to choose the molecular features (descriptors) that properly represent the possible interaction of the toxicant with the bioreceptor to produce the toxicity, and to correlate the descriptors with the toxicity by means of some regression relationships. There is a common shortcoming in all the approaches to QSAR problems: the choice of molecular descriptors deemed to be responsible for the toxicity is not directly based on first principles, meaning it is arbitrary, some (or all) descriptors being thus artifacts with no physical meaning implied in their initial choice (see also [2]).

Distinguished from the traditional QSAR approaches, the electron-conformational (EC) method [1, 2] does not employ arbitrary descriptors and statistics in evaluation of their weight. Instead, the electronic structure and topology of the molecule evaluated by quantum-chemical calculations is used as one (a unique) descriptor, presented in a computer

friendly digital-matrix form. The comparison of these matrices with a certain type of toxicities allows one to reveal a group of matrix elements that are common to the active compounds under consideration, and represent the numerical picture of toxicophore (Tph). This approach has been applied successfully to study several types of biological activities (see, e.g., [6-10]) and to predict aquatic toxicity to fish [11].

In this paper the predictive power of the EC method and its efficiency is demonstrated by revealing the origin of toxicity in a series of 24 compounds which are identified as fragrance allergens. Using the EC method we performed conformational analysis (optimization of geometries of the low-energy conformers) and electronic structure calculations (by *ab initio* method including solvent influence) for these allergens and constructed the Electron-Conformational Matrix of Congruity (ECMC) for each of their (room-temperature) conformations. Then the toxicophore (Tph), the group of atoms (atomic characteristics) responsible for the toxicity, is revealed by evaluation of the EC sub-matrix of activity (ECSA), a sub-matrix with matrix elements common to all the active compounds under consideration within minimal tolerances.. Starting with only five most toxic compounds, their ECSA (toxicophore) was found to consist of a 4x4 matrix (four sites with certain electronic and topologic characteristics) which was shown to be present in all the most active compounds. In addition to the Tph identification, a regression structure-toxicity model was employed to take into account the influence of Tph flexibility, anti-Tph shielding (ATS), and other out-of-Tph functional groups (auxiliary groups, AG) on the quantitative value of toxicity. The results predict well ($R^2 = 0.93$) the experimental values of toxicity with a reasonable leave-one-out cross-validation. The methodology worked out in this paper can be applied to screening and prediction of toxicity of other environmental pollutants.

2. Data set and method

The chosen data set contains 24 compounds that are included in the list of fragrance contact allergens by the European Union scientific committee SCCNFP's [12]. The values of $a=LD_{50}$ in oral exposure of rats were chosen as a measure of toxicity. These values were obtained from the site ChemIDplus of the United States National Library of Medicine [13]. Structural formulas for chemicals studied, along with the experimental LD_{50} values, are given in the Table 1.

Molecular structures and experimental toxicities (LD_{50}) of 24 fragrance contact allergens

Structure	LD_{50} (mg·kg^{-1})	Structure	LD_{50} (mg·kg^{-1})	Structure	LD_{50} (mg·kg^{-1})
1 Coumarin	196	**9** Cinnamyl alcohol	2000	**17** Amylcinnamyl alcohol	4000
2 Anisyl alcohol	1200	**10** Cinnamal	2200	**18** Linalool	4180
3 Benzyl alcohol	1230	**11** Benzyl salicylate	2227	**19** d-Limonene	4400
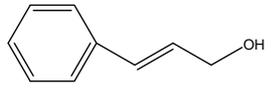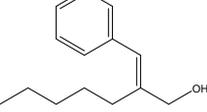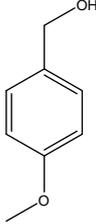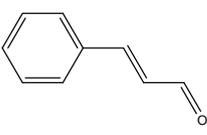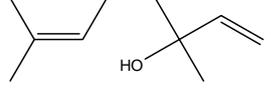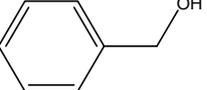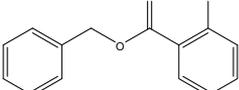
4 Lilial	1390	**12** Hexylcinnamaldehyde	3100	**20** Citral	4960
5 Methyl heptine carbonate	1530	**13** Lyral®	3250	**21** benzyl cinnamate	>5000
6 Isoeugenol	1560	**14** Citronellol	3450	**22** Isomethyl α-ionone	>5000
7 benzyl benzoate	1700	**15** Geraniol	3600	**23** Hydroxycitronellal	>5000
8 Eugenol	1930	**16** Amyl cinnamal	3730	**24** Farnesol	>5000
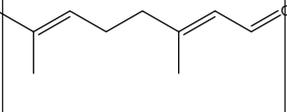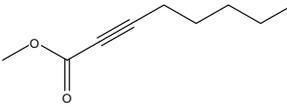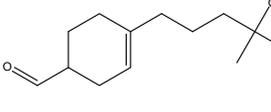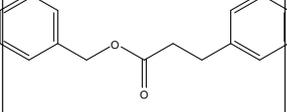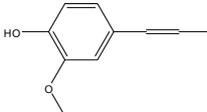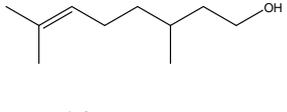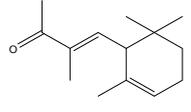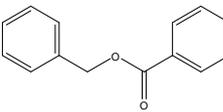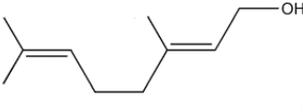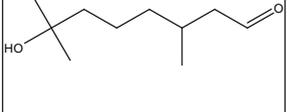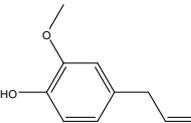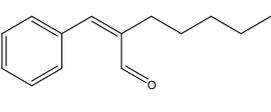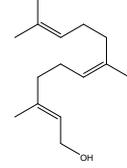
The electron-conformational method used in this study of structure-toxicity relations is outlined in detail elsewhere [1, 2]. It includes as a first part the conformational analysis and quantum chemical calculations of the equilibrium geometries and electronic structure of compounds under consideration. Next, using results of calculations (molecular orbital population analysis, Mulliken atomic charges, and bond orders), the so-called Electron-Conformational Matrices of Congruity (ECMCs) are constructed for the low-lying conformations of molecules of the training set. These ECMCs contain a rather full description of both the geometry and electronic properties of molecules presented by the corresponding matrix elements.

Figure 1 illustrates an example of the ECMC calculated for the lowest conformation of compound **3**. The hydrogen atoms are excluded from consideration here for simplicity. Diagonal matrix elements reproduce atomic interaction indices (*II*) [1], which are measures of electron-donor properties of the corresponding atoms in the molecule:

$$II^A = g^A \exp\left(-R_0 \sqrt{2\,\text{VOIP}^A}\right) \tag{1}$$

Here g^A is the Mulliken electron population of the outermost orbital of the atom A (g^A for np-elements is equal to one-third of the total occupancy of valence p-orbitals, p_x, p_y, and p_z, of the atom), and VOIP^A in atomic units refers to the Valence Orbital Ionization Potential of this atom-in-molecule orbital calculated as a function of the Mulliken charge and the electronic configuration of the atom using the reference data [14]. A value of $R_0 = 1.51$ Bohr radii (0.8 Å) is conventional [9]. Off-diagonal matrix elements represent Mulliken bond orders for chemically bonded atoms and interatomic distances for non-bonded pairs (see Figure 1).

	C1	C4	C2	C6	C5	C3	C7	O1
C1	0.13							
C4	2.78	0.16						
C2	1.39	2.41	0.16					
C6	1.39	1.40	2.41	0.17				
C5	2.41	1.45	2.78	1.39	0.16			
C3	2.41	1.39	1.44	2.78	1.43	0.16		
C7	1.50	4.29	0.93	2.51	3.78	3.79	0.11	
O1	2.36	4.92	2.94	3.44	4.60	4.24	1.41	0.38
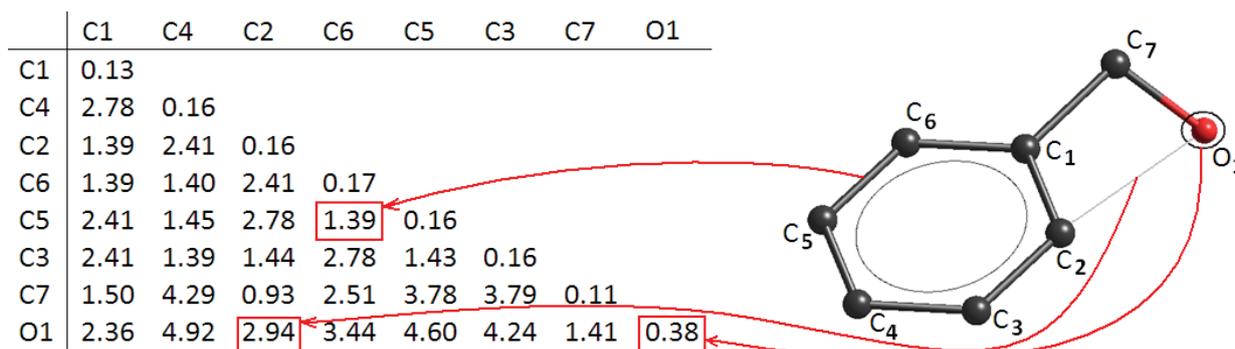
Figure 1. The electron-conformational matrix of congruity for molecule 3. Hydrogen atoms are omitted for simplicity. The diagonal elements refer to the atomic (e.g., O_1 in the picture) interaction indices calculated by Eq. 1, while the off-diagonal elements reproduce Mulliken's bond orders for chemically bonded pairs of atoms (e.g., C_5-C_6) and interatomic distances for non-bonded pairs (e.g., C_2-O_1).

The next step in the EC method is the one-by-one comparison of the matrix elements of different ECMCs in order to reveal the so-called electron-conformational submatrix of activity (ECSA), namely the set of matrix elements (geometry and electronic parameters) that, within some tolerances, are common for all the toxic compounds. The ECSA describes a group of atoms in the molecule, which are characterized by means of interatomic distances, bond orders, and interaction indices, the toxicophore (Tph).

Numerical values of toxicities can be evaluated using a general formula for biological activity [1]:

$$a_i = a_{ref} \exp\left(-\frac{E_i^{Tph} - E_{ref}^{Tph}}{k_B T}\right)\exp\left(-S_i[R]\right), \tag{2}$$

$$S_i[R] = \sum_{j=1}^{P} k_j \left(R_j^{(i)} - R_j^{ref}\right), \tag{3}$$

where a_i and a_{ref} stand for numerical values of activity of the *i*-th compound and the reference compound, respectively, E^{Tph} is the relative energy of the lowest energy conformer that contains the Tph, and $S_i(R)$ is a function of the electronic and geometric parameters of the substrate molecule, parameters R stand for the toxicophore flexibility and ATS/AG influence in the substrate molecule (see below)..

Regression coefficients k in Eq. 3 are calculated by means of a least-squares fit of the values of toxicity a_i (calculated by Eq. 2 for all the compounds in the training set, which contain the Tph) to corresponding experimental values a_i^{exp}. The quality of the linear regression can be evaluated by the analysis of variance [15]:

$$SE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}}, \qquad \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{4}$$

$$F\left(\frac{\nu_1}{\nu_2}, \alpha\right) = \frac{R^2}{1-R^2}\frac{\nu_2}{\nu_1}, \tag{5}$$

where SE is a standard error, $x_i = a_i^{\exp} - a_i^{\text{theory}}$, and R is the correlation coefficient. The F-statistics value can be compared to tabulated values [15] for F with the confidence α and the degrees of freedom $v = P$, $v = N - P - 1$ (P is the number of descriptors R in Eq. 3, N is the number of molecules that have the toxicophore).[1]The E-statistics values are calculated as a ratio:

$$E = \frac{PRESS_P}{PRESS_{P-1}}, \qquad PRESS = \sum_{i=1}^{N} \left(a_i^{\exp} - a_i^{theory} \right), \tag{6}$$

The E-values characterize the weight of each eliminated parameter R in the numerical value of toxicity. If $E > 0.4$ for any (eliminated) parameter R in Eq. 6, the latter is considered to be statistically irrelevant [16], meaning this parameter does not influence significantly the activity.

3. Results and discussion

Construction of electron-conformational matrices of congruity. Conformational analysis for all compounds from Table 1 was performed using the methods of molecular mechanics with Merck force field [17] and the Monte-Carlo randomized search method with the help of the SPARTAN package [18, 19]. The total energies of conformers in their equilibrium geometry were calculated by *ab initio* RHF method in the 6-31G* basis sets and corrected for the aqueous solvation effect by the polarizable continuum model. This part of calculations was carried out using the PC GAMESS version [20] of the GAMESS (US) QC package [21]. The results of single-point *ab initio* calculations of the electronic structure (molecular orbital population analysis, Mulliken atomic charges, and bond orders) were used then for construction of the ECMCs for all conformers that are significantly populated at room temperature.

Pharmacophore Identification. The electron-conformational submatrix of toxicity was revealed first using 5 compounds **2-6** with statistically significant values of toxicity (see Table 1). The Tph consists of four molecular sites with specific interaction indices, which in our set are occupied by one oxygen and three carbon atoms, but they can equally well be occupied by any other atoms, that have the same electronic structure parameters and geometric positions in the compound within the limits of tolerances (Figure 2).

	A_1	A_2	A_3	A_4
A_1	0.151 ± 0.034			
A_2	2.791 ± 0.754	0.166 ± 0.030		
A_3	2.209 ± 0.689	1.423 ± 0.102	0.154 ± 0.033	
A_4	2.620 ± 0.833	4.903 ± 1.195	3.840 ± 1.130	0.332 ± 0.091

Figure 2. The electron-conformational submatrix of toxicity obtained via identification of the toxicophore in the series of compounds under consideration. The tolerances reflect the toxicophore flexibility.

Next, we examined the rest of the molecules from the series to reveal the ECSA for their low-lying conformers with energies below 1 kcal/mol (read off the lowest energy one). We found that in addition to 5 compounds with the Tph, all other compounds except **19** also contain the Tph. Examples of molecular structures of some of the toxic compounds and their ECSAs are shown in Figure 3.

	C_1	C_2	C_3	O_1
C_1	0.17			
C_2	2.79	0.16		
C_3	2.42	1.39	0.14	
O_1	2.33	3.71	4.16	0.38

	C_1	C_2	C_3	O_1
C_1	0.18			
C_2	3.55	0.17		
C_3	2.48	1.39	0.13	
O_1	2.42	5.00	4.31	0.30

	C_1	C_2	C_3	O_1
C_1	0.14			
C_2	2.79	0.17		
C_3	2.41	1.39	0.17	
O_1	2.34	4.89	4.40	0.36
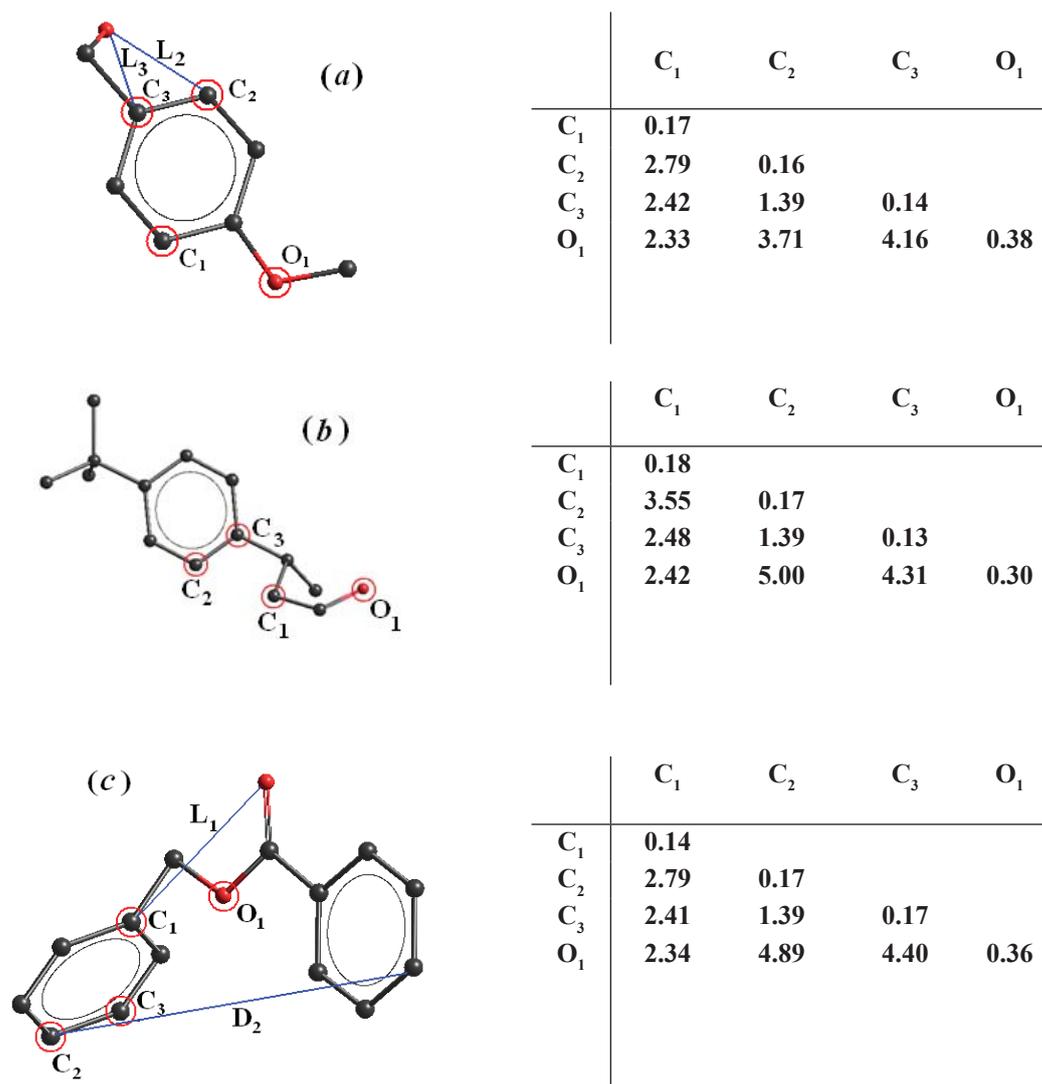
Figure 3. Electron-conformational submatrices of toxicity and molecular structures of compounds 2 (a), 4 (b), and 7 (c). Hydrogen atoms are omitted for simplicity. The four toxicophore atoms are marked by circles. The interatomic distances D_2, L_2 and L_2 are given as an example of descriptors used below for quantitative toxicity prediction.

Quantitative prediction of toxicity. For quantitative prediction of the values of toxicity of compounds containing the Tph the general formula for biological activity (2) and (3) can be rewritten as follows

$$\log(LD_{50})_i = \log(LD_{50})_{ref} - 2.30259\left\{\frac{E_i - E_{ref}}{kT} + S_i[R]\right\} \qquad (7)$$

We performed parameterization of the Tph flexibility and ATS/AG influence with a training set of 19 compounds containing the Tph (see Table 1 for **1–18** and **20**). Compounds **21-24** were considered inactive due to their statistically low experimental toxicities (LD$_{50}$>5000 mg/kg) and were not included in this training set. Compound **1** with the most significant parameter of toxicity was used as the reference one.

A regression model for quantitative prediction of toxicity was constructed. It is composed of as few as six descriptors, selected to parameterize the toxicophore flexibility and ATS/AG influence using Eq. (3) with the coefficients determined from the best possible correlation of the theoretical values of LD$_{50}$ with the experimental data. Three parameters describing the toxicophore flexibility, $R(A_1–A_2)$, $R(A_1–A_3)$ and $R(A_2–A_3)$, are just the distance between A_i

and A_j sites of the toxicophore (Figure 2). For instance, the sites A_i (i=1–4) appear in the case of **2, 4** and **7** as atoms C(1), C(2), C(3), and O(1), respectively (Figure 3). The interaction index $II(A_1)$ is a function of the Mulliken charge and the electronic configuration of the atom C(1) in the molecule and stand for the donor-acceptor activity of this atom..To describe the anti-toxicophore shielding, ATS groups we introduce the parameter D_2 which is the distance (in Å) from the toxicophore site A_2 to some non-hydrogen entity defined as the out-of-Tph atom *most distant* from the geometric center of the Tph (Figure 3, c) [22]. Large values of D_2 correspond to relatively large (mainly hydrophobic) groups (for instance, phenylalkyl groups in **7**, **9, 11**, **21** or long alkyl chains in **5, 13, 14, 15** and **17**) that obstruct the toxicity.

To describe the auxiliary out-of-toxicophore AG groups we introduce parameters l_i (in Å^{-1}):

$$l_i = \frac{II^B}{L_i}, \qquad (8)$$

where L_i is the distance from the Tph site A_i to the *nearest* non-hydrogen atom B with the interaction index II^B belonging to the auxiliary (mostly electron donor) group (Fig. 3, a, c). The influence of such groups increases with the growth of the electronic density near the atom B (larger value of II^B) and with the decrease of L_i. The function l_i reflects the electrostatic potential induced by a point charge B at the distance L_i from the charge.

Table 2 shows the optimal values of the linear regression coefficients k_i (Eq. 3) corresponding to the six descriptors. Cross-validation and E-statistics (Eq. 6) of regression parameters indicate that the most statistically significant descriptors are two values of $R(A_i–A_j)$, followed by less important $II(A_1)$ and D_2, and statistically irrelevant l_2; removal of the latter results in only a slight loss of accuracy with the value of R^2 dropping to 0.85 (Table 2).

Table 2

Regression Coefficients k_i and Cross-Validation Statistics [a]

Descriptor	k_i	R^2	E-statistics
$II(A_1)$	-2,71	78,37	0,31
$R(A_1\text{-}A_2)$	0,17	83,39	0,41
$R(A_1\text{-}A_3)$	-0,19	76,63	0,29
$R(A_2\text{-}A_3)$	-1,26	40,72	0,11
l_2	0,68	85,20	0,46
D_2	-0,01	80,05	0,34

[a] Dimensions of the regression coefficients k_i ($i = 1 – 6$) are reciprocal to the dimensions of corresponding descriptors. Values of the correlation coefficient R and the E-statistics are calculated by means of consequent removal of one of the six descriptors out of the regression analysis.

Theoretical values of toxicity (LD$_{50}$) calculated by the Eq. (7) with the regression coefficients from Table 2 are presented in Table 3. The plot of predicted versus experimental data is shown in Figure 4. The quality of the model is characterized by following values of statistical parameters: square correlation coefficient $R^2 = 0.93$, SE = 0.029 (standard error, Eq. 4), and a value of $F\left(\dfrac{6}{12}, \alpha\right) = 26.6$ for F-statistics (Eq. 5) with 95% of confidence α. Cross-validated "leave-one-out" values for toxicities (LOO in Table 3) are calculated by means of consequent removal of one of the compounds from the regression analysis and then calculating the predicted activity for the removed compound. These LOO values still predict the experimental data rather well (Table 3), the value of $R^2 = 0.93$ being lowered to $R^2 = 0.79$.
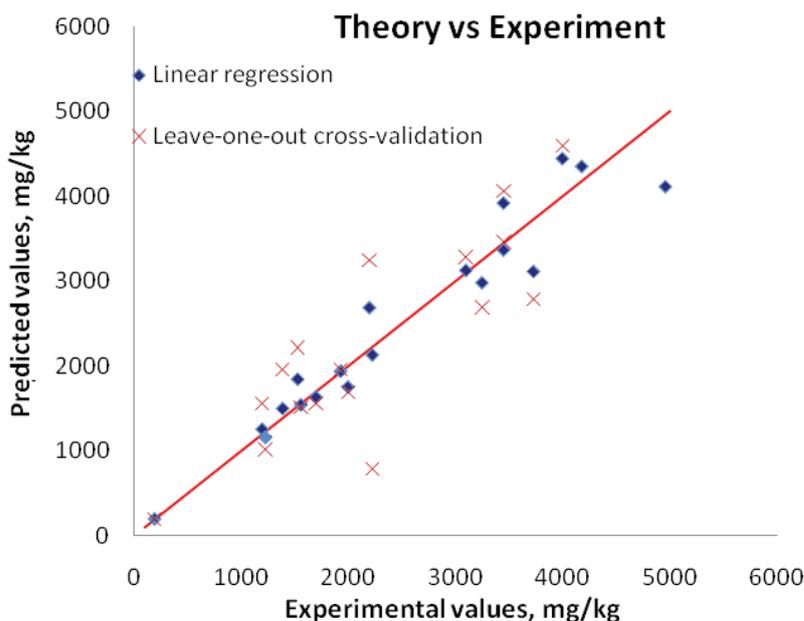
Figure 4. Theoretical versus experimental values of LD$_{50}$ (mg/kg). Linear regression (points marked by squares) results in the value of $R^2 = 0.93$ for the correlation coefficient R. Leave-one-out cross-validation (points marked by crosses) predicts $R^2 = 0.79$.

Table 3

Correlation Between Experimental (Exp) and Calculated (Theory) Values of LD$_{50}$ (mg/kg) and Prediction of Toxicity in the "Leave-One-Out" Cross-Validation Scheme for the Training Set of 19 Compounds [a]

Cpd	Exp	Theory	LOO	Cpd	Exp	Theory	LOO
1 (Ref)	196	196	196	**11**	2227	1893	1663
2	1200	1242	1305	**12**	3100	3122	3167
3	1230	1126	960	**13**	3250	3189	3164
4	1390	1489	1713	**14**	3450	3840	4035
5	1530	1850	2246	**15**	3450	3298	3338
6	1560	1644	1609	**16**	3730	3109	2786
7	1700	1718	1767	**17**	4000	4404	4527
8	1930	1996	2035	**18**	4180	4331	4521
9	2000	1734	1674	**20**	4960	4058	3796
10	2200	2697	3142				

[a] Experimental data (exp) on acute toxicity to rats are quoted from ref [13]. Only 19 compounds with the Tph (Tables 1) are included in the linear regression. Theoretical values of LD$_{50}$ (Theory) calculated with use of regression coefficients k and corresponding descriptors R, Π, D, and l (Table 2) are compared with experimental data.

Screening of new compounds. Next, we applied the obtained above results of the toxicophore identification and toxicity predictions to a new set of compounds that are included in the third European Union scientific committee SCCNFP's list of 78 fragrance chemicals that may be used subject to certain restrictions. Only for 20 compounds from them listed in Table 4 the chemical structures and the toxicity data were found in the literature, but the reliability of these data is uncertain. The electron-conformational matrices for their low-lying conformers (below 1 kcal/mol) were calculated and checked whether or not they contain the electron-conformational submatrix of activity, the Tph. Then for the compounds that have the Tph theoretical values of toxicity (LD$_{50}$) were calculated by Eq. (7) with the regression

coefficients from Table 2. The results are shown in Table 4. From the 20 compounds 17 have the Tph and three (**6**, **15** and **18** inTable 3) have not.

Table 4

Molecular structures and toxicities (LD$_{50}$) of the set of new 20 compounds

Structure	LD$_{50}$ [a] (mg·kg^{-1})	Structure	LD$_{50}$ [a] (mg·kg^{-1})
1 Allyl hexanoate	218 2500	**11** Allyl 3,5,5-trimethylhexanoate	1400 424
2 Allyl isovalerate	230 322	**12** allyl cinnamate	1520 1240
3 Allyl butyrate	250 2490	**13** 3-Propylidenephthalide	1650 1760
4 Allyl phenoxyacetate	475 409	**14** 4-(prop-1-en-2-yl) cyclohex-1-enecarbaldehyde	1720 3200
5 Allyl phenoxyacetate	500 2630	**15** 3-Methyl-2(3)-nonenenitrile	1720 no Tph
6 Allyl octanoate	570 no Tph	**16** Amylcyclopentenone	2200 2160
7 Allyl cyclohexanepropionate	585 407	**17** Methyl octine carbonate	2220 1780
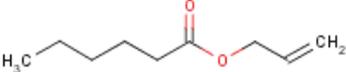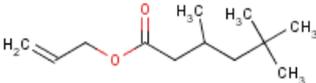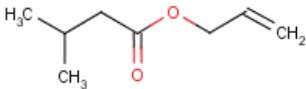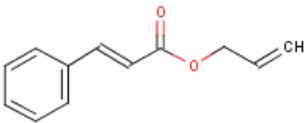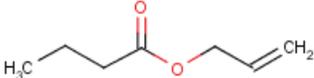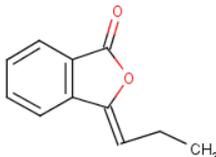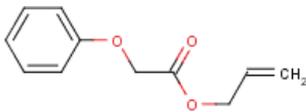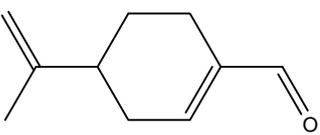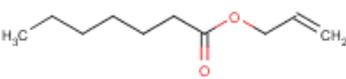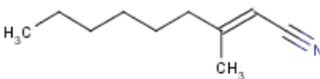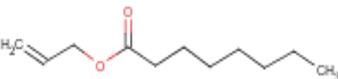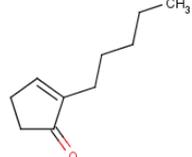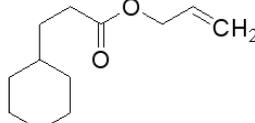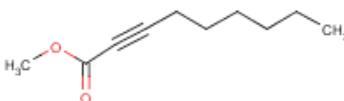
153
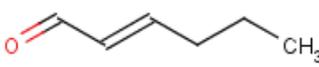
8 trans-Hexen-2-al	780 2530	**18** p-tert-butyldihydrocinnamaldehyde	2700 no Tph
9 1-Octen-3-yl acetate	850 700	**19** Methyl N-methylanthranilate	3380 2960
10 Allyl cyclohexaneacetate	900 3270	**20** Menthadiene-7-methyl formate	3568 646
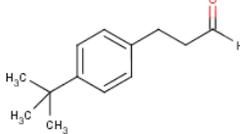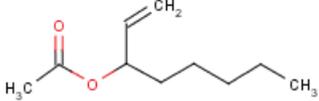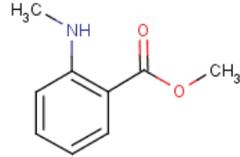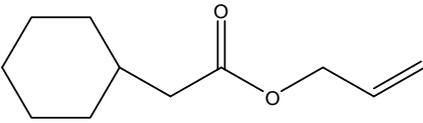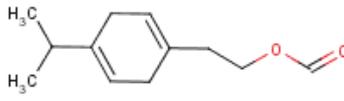
[a)]The first row in the second and fourth columns presents the experimental values of LD_{50} (mg·kg^{-1}), while the calculated values are in the second row.

It is seen from Table 4 that for 9 compounds of this series (**2, 4, 7, 9. 12, 13, 16, 17, 19**) calculated values of toxicities agree rather well with the experimental data (triangles in Fig.5); note that they are all of higher toxicity where the experimental measurements are more reliable. For these compounds, taking into account the uncertainties in the toxicity measurements, the theoretical predictions are excellent. For the rest 8 compounds (**1**, **3**, **5**, **8**, **10**, **11, 14, 20**) the theory predicts that they are toxic, but there is a significant discrepancy between the quantitative values of the toxicity predicted by the theory and the experimental data. Three compounds (**6, 15, 18**) have no Tph. The remaining discrepancies between the EC theory and experiment when tried on a different series of compounds may have reasonable explanation in the inaccuracies of both experimental measurements of toxicities and chemical structures; it may stimulate additional investigation of the experimentally measured toxicity of these compounds, as well as their chemical composition.
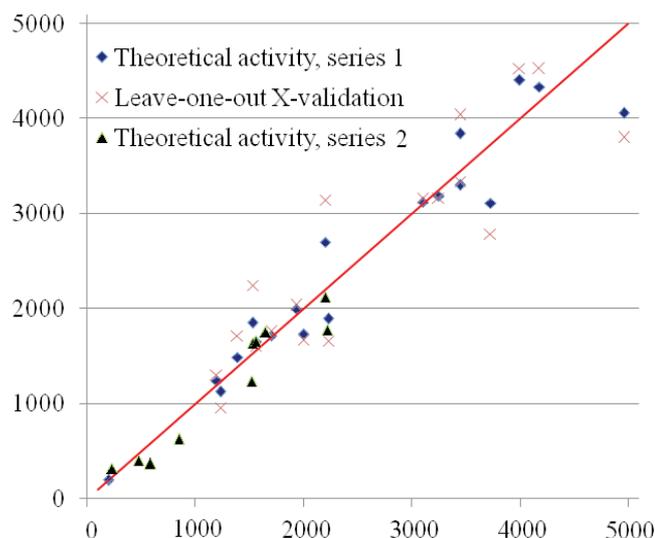
Figure 5. Calculated versus experimental toxicities in two series of chemical allergens. Green triangles mark the calculated values of LD_{50} (mg/kg) for the testing set.

4. Conclusions

The electron-conformational method applied to the acute toxicity of 24 fragrance allergens reveals that this toxicity is controlled by a toxicophore that consists of four sites with certain electronic and topologic characteristics. In the series of compounds under consideration they are occupied by three carbon and one oxygen atoms, but may be substituted with any other atoms that have the same electronic structure parameters and geometric positions within the derived tolerances. A model composed of six parameters describing the toxicophore flexibility and the influence of anti-toxicophore shielding and other out-of-Tph groups reproduces rather well the experimental data [13].

The method worked out in this paper can be used in a similar way for screening and prediction of toxicity of other environmental pollutants.

5. References

[1]. Bersuker, I. B. Current Pharmaceutical Design., 2003, 9, 1575-1606.

[2]. Bersuker, I. B. J. Comp.-Aid. Mol. Des., 2008, 22, 423.

[3]. Guner O.F., Editor, Pharmacophore Perception, Development, and Use in Drug Design, International University Line: La Jolla, 2000.

[4]. Kubinyi, H.; Folkers, G.; Martin, Y. C. 3D QSAR in Drug Design. Volume 3, Recent Advances; Kluwer Academic Publishers: New York, 2002.

[5]. Marshall G.R. In: Wolff M.E., Ed. Burger's medicinal chemistry and drug discovery: principles and practice. 5th ed. Vol. 1. New York: John Wiley, 1995, 573-659.

[6]. Bersuker, I. B.; Dimoglo,A. S.; Gorbachov, M. Yu.; Vlad, P. F. and Pesaro, M., New J. Chem., 199,15, 307-320.

[7]. Bersuker, I. B.; Bahceci, S.; Boggs, J. E.; Pearlman, R. S. J. Computer-aided Mol. Design 1999, 13, 419–434; Bersuker, I. B.; Bahceci, S.; Boggs, J. E. J. Chem. Inf. Comput. Sci., 2000, 40, 1363–1376.

[8]. Makkouk, A. H.; Bersuker, I. B.; and Boggs, J. E. Int. J. Pharm. Med., 2004, 18, 81-89; Rosines, E.; Bersuker, I. B.; Boggs, J. E.Quantitative Struct.-Activity relationships, 2001, 20, 327-334.

[9]. Marenich, A.V.; Pei-Han Yong; Bersuker, I. B.; Boggs, J. E. J. Chem. Inf. Model., 2008, 48, 556-568.

[10]. Sahin, K.; Sarıpınar, E.; Yanmaz, E.; Geçen, N. SAR QSAR Environ Res.,2011, 22, 217-238; Saripinar, E.; Geçen, N.; Sahin, K.; Yanmaz, E. Eur. J. Med. Chem., 2010, 45, 4157-4168; Yanmaz, E.; Sarıpınar, E.; Şahin, K.; Geçen, N.; Çopur, F. Bioorg. Med. Chem., 2011, 19, 2199-2210.

[11]. Gorinchoy, N. N.; Ogurtsov, I. Ya.; Tihonovschi, A.; Balan, I.; Bersuker, I. B.; Marenich, A.; Boggs, J. E. Chem. J. Mold., 2008, 3, 94-104.

[12]. The Scientific Committee on Cosmetic Products and Non-Food Products Intended for Consumers, Fragrance Allergy in Consumers, SC-CNFP/0017/98 Final, 8 December 1999.

[13]. http://chem.sis.nlm.nih.gov/chemidplus/

[14]. Basch, H.; Viste, A.; Gray, H. B. Valence Orbital Ionization Potentials from Atomic Spectral Data. Theoret. chim. Acta (Berl.) 1965, 3, 458–464.

[15]. Livingstone, D. Data Analysis for Chemists: Applications to QSAR and Chemical Product Design; Oxford University Press: Oxford, NY, 1995.

[16]. Wold, S. Cross-Validatory Estimation of Number of Components in Factor and Principal Components Models. Technometrics, 1978, 20, 397–405.

[17]. Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. J. Comp. Chem., 1996, 17, 490–519.

[18]. Spartan'02; Wavefunction, Inc.: Irvin, CA. Except for molecular mechanics and semi-empirical models, the calculation methods used in Spartan'02 have been documented in: Kong, J.; White, C. A.; Krylov, A. I.; Sherrill, D.; Adamson, R. D.; Furlani, T. R.; Lee, M. S.; Lee, A. M.; Gwaltney, S. R.; Adams, T. R.; Ochsenfeld, C.; Gilbert, A. T. B.; Kedziora, G. S.; Rassolov, V. A.;Maurice,D. R.; Nair, N.; Shao, Y.; Besley, N. A.; Maslen, P. E.; Dombroski, J. P.; Daschel, H.; Zhang, W.; Korambath, P. P.; Baker, H.; Byrd, E. F. C.; Voorhis, T. V.; Oumi, M.; Hirata, S.; Hsu, C.-P.; Ishikawa, N.; Florian, J.; Warshel, A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. Q-Chem 2.0: A High-Performance Ab Initio Electronic Structure Program Package. J. Comput. Chem. 2000, 21, 1532–1548.

[19]. Hehre, W. J. Guide to Molecular Mechanics and Quantum Chemical Calculations; Wavefunction, Inc.: Irvin, CA, 2003.

[20].   Alex A. Granovsky, www http://classic.chem.msu.su/gran/gamess/index.html

[21].   Schmidt, M.W.; Baldridge, K.K.; Boatz, J.A.; Elbert, S.T.; Gordon, M.S.; Jensen, J.H.; Koseki, S.; Matsunaga, N.; Nguyen, K.A.; Su, S.; Windus, T.L.; Dupuis, M.; Montgomery, J.A. J.Comput.Chem. 1993, 14, 1347-1363.

[22].   The geometric center of the Tph (C) was defined as averaged values X, Y, Z of the Cartesian coordinates of all the toxicophore sites in the molecule. To apply the parameter D_2 only for description of large anti-toxicophore shielding groups with respect to the toxicophore site A_2, we excluded all "small" groups with reducing D_2 to zero when the distance from any of the A_i sites to the outmost non-toxicophore atom in the molecule was shorter than the corresponding C–A_i distance (the index i runs over all the toxicophore sites, i=1÷4).