Available online at www.sciencedirect.com
SCIENCE DIRECT
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 16 (2006) 2306–2311

QSAR study of mosquito repellents using Codessa Pro

Alan R. Katritzky,[a,*] Dimitar A. Dobchev,[a,b] Indrek Tulp,[a,b] Mati Karelson[c] and David A. Carlson[d]

[a]*Center for Heterocyclic Compounds, Department of Chemistry, University of Florida, Gainesville, FL 32611, USA*
[b]*Department of Chemistry, University of Tartu, 2 Jakobi Street, Tartu 51014, Estonia*
[c]*Department of Chemistry, Tallinn University of Technology, Ehitajate tee 5, Tallinn 19086, Estonia*
[d]*Department of Agriculture, Agricultural Research Service, Center for Medical Agricultural and Veterinary Entomology, Gainesville, FL, USA*

Received 14 October 2005; revised 23 November 2005; accepted 28 November 2005
Available online 20 February 2006

Abstract—Protection times provided by 31 synthetic repellents against Aedes aegypti mosquitoes were correlated with the chemical structures of these repellents using Codessa Pro software. Two statistically significant quantitative models with R^2 values of ca. 0.80 are presented and discussed.

Repellents are materials that disrupt the natural behavior of blood-seeking insects and other organisms; repellents provide personal protection and represent the first line of defense for humans and animals against biting. A well-known standard repellent is *N*,*N*-diethyl-3-methylbenzamide or *N*,*N*-diethyl-*m*-toluamide (DEET, compound **7**, Table 1).[1] However, it has become urgent to locate repellents which are more effective than DEET.

Few attempts have previously been made to apply QSAR modeling to repellent activities. One reason for this is that most of the extensive testing that has been carried out[1] has yielded only semi-quantitative data. An exception is the work of Suryanarayana et al.[2] who measured a set of 31 repellents and proposed the correlation Eq. 1, where $\log P$, $\log V_p$, and ML are lipophilicity, vapor pressure, and molecular length, respectively, and *a*–*d* are constants.

$$PT = a\log P + b\log V\mathrm{p} + c\log \mathrm{ML} + d. \qquad (1)$$

However, Eq. 1 has a low correlation coefficient R at 0.551 (corresponding to a R^2 of 0.304) and in addition, one of the descriptors is the measured vapor pressure which has to be obtained before Eq. 1 can be used to predict the activity of unknown compounds.

Other authors[3] have also suggested that vapor pressure and boiling point were related to repellent activity; repellency is lost at vapor concentrations below a certain minimum.[4,5] Factors including evaporation from human skin, skin absorption, and penetration clearly influence repellent bioassays.[6] Test-related factors (such as the mosquito species utilized the cage size and the mosquito density) also affect repellent bioassays.[7]

Ma et al.[4] discussed the Suryanarayana's data set and postulated that amide group made an important contribution for potent repellent activity, but reported no numerical correlation. The same group explored molecular similarity between insect juvenile hormone and DEET analogues but they did not explore any quantitative correlation with structure.[8]

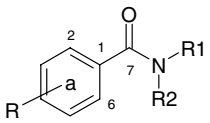
The present QSAR study correlates mosquito repellent activity (protection time, PT) as reported by Suryanarayana et al.[2] with theoretical molecular descriptors; we have also examined repellency using vapor pressure as an external descriptor in view of the importance attributed to it by earlier workers.

Methodology for a general QSAR approach has previously been incorporated in the Codessa Pro[9] software package which enables the calculation of numerous

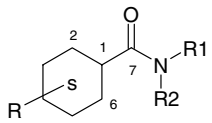
Keywords: QSAR; Mosquito repellents; DEET; Vapor pressure; Multilinear regression.
* Corresponding author. Tel.: +1 352 392 0554; fax: +1 352 392 9199; e-mail: katritzky@chem.ufl.edu

0960-894X/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2005.11.113

Table 1. DEET (compound **7**) and DEPA analogues



a = aromatic ring



s = saturated ring

ID	Compound	PT (h)	Ring	R	R1 = R2	CAS nr.	Exp V_p (Torr)	Pred. log V_p
1	*o*-Chlorobenzamide	5	a	2-Cl	CH_3	6526-67-6	9.63E-04	−2.94
2	Cyclohexamide	3	s	H	CH_3	17566-51-7	0.0225	−1.77
3	*m*-Toluamide	3	a	3-CH_3	CH_3	6935-65-5	2.18E-03	−2.42
4	*o*-Ethoxylbenzamide	2.83	a	2-OC_2H_5	CH_3	90526-02-6	1.65E-04	−3.75
5	Benzamide	1.67	a	H	CH3	611-74-5	0.0157	−2.02
6	*p*-Anisamide	1	a	4-OCH_3	CH_3	7291-00-1	4.38E-04	−3.35
7	*m*-Toluamide	5	a	3-CH_3	C_2H_5	134-62-3	1.35E-03	−3.25
8	Benzamide	4	a	H	C_2H_5	1696-17-9	2.30E-03	−2.79
9	Cyclohexamide	4	s	H	C_2H_5	5461-52-9	2.92E-03	−2.34
10	*o*-Ethoxylbenzamide	3.5	a	2-OC_2H_5	C_2H_5	3688-82-2	2.58E-05	−4.61
11	*p*-Toluamide	2.83	a	4-CH_3	C_2H_5	2728-05-4	3.65E-04	−3.25
12	*p*-Anisamide	1	a	4-OCH_3	C_2H_5	7465-86-3	7.22E-05	−4.17
13	Benzamide	3	a	H	*i*-C_3H_7	14657-86-4	3.14E-04	−3.62
14	*m*-Toluamide	2.67	a	3-CH_3	*i*-C_3H_7	5448-37-3	5.10E-05	−4.11
15	Cyclohexamide	2	s	H	*i*-C_3H_7	67013-94-9	3.70E-04	−3.60
16	*p*-Anisamide	1.17	a	4-OCH_3	*i*-C_3H_7	349397-58-6	1.07E-05	−5.02
17	*o*-Ethoxybenzamide	1.08	a	2-OC_2H_5	*i*-C_3H_7	5442-04-6	3.70E-06	−5.49
18	*o*-Chlorobenzamide	1	a	2-Cl	*i*-C_3H_7	349397-59-7	2.39E-05	−4.53
19	*p*-Toluamide	0.5	a	4-CH_3	*i*-C_3H_7	5448-37-3	5.10E-05	−4.12
					R1 R2			−3.16
20	*m*-Toluamide	0.67	a	3-CH_3	HC_2H_5	26819-07-8	1.85E-03	−2.86
21	Benzamide	0.58	a	H	HC_2H_5	614-17-5	3.95E-04	−3.07
22	Cyclohexamide	0.5	s	H	HC_2H_5	138324-59-1	8.66E-04	−3.08
23	*p*-Toluamide	0.08	a	4-CH_3	HC_2H_5	26819-08-9	7.67E-04	−3.98
24	*p*-Anisamide	0.08	a	4-OCH_3	HC_2H_5	7403-41-0	1.29E-04	−4.33
25	*o*-Ethoxybenzamide	0.08	a	2-OC_2H_5	HC_2H_5	99985-68-9	5.68E-05	−3.76
					N, R1, R2			−3.66
26	Benzamide	3	a	H	Piperidine	776-75-0	3.16E-04	−4.19
27	Cyclohexamide	2	s	H	Piperidine	7103-46-0	1.56E-04	−4.66
28	*m*-Toluamide	1.42	a	3-CH_3	Piperidine	13290-48-7	4.41E-05	−4.21
29	*o*-Chlorobenzamide	1	a	2-Cl	Piperidine	22342-21-8	1.94E-05	−5.12
30	*p*-Toluamide	1	a	4-CH_3	Piperidine	13707-23-8	4.90E-05	−2.94
31	*p*-Anisamide	0.75	a	4-OCH_3	Piperidine	57700-94-4	8.81E-06	−1.77

quantitative descriptors from the molecular structural formula.[10,11] Codessa Pro has previously correlated successfully numerous physical properties[12] including chromatographic retention times and response features, melting and boiling points, solvent scales, and refractive indexes.[13] Recent examples include correlations for: (i) binding energies for 1:1 complexation systems of organic guests and β-cyclodextrin,[14] (ii) the in vitro minimum inhibitory concentration (MIC) of 3-aryloxazolidin-2-one antibacterials[15], and (iii) partition coefficients of medicinal drugs between human breast milk and plasma.[16]

We correlated the 31 protection times (PT) determined by Suryanarayana et al.[2] by testing the compounds at a dose of 1 mg/cm^2 onto the external surface of a human hand followed by exposure to 200 female (5–7 days old) *Aedes aegypti* mosquitoes. The PT is defined as the period of protection in minutes until two consecutive bites are made within a 30 min interval. The reported protection times represent averages of multiple determinations. The compound dataset represents 31 amide analogues of *N*,*N*-diethyl-*m*-toluamide (DEET) and *N*,*N*-diethylphenylacetamide (DEPA) (see Table 1).

Conformational searches were carried out over all 31 structures using the AMBER2 force field method in molecular mechanics (MM) optimization encoded in HyperChem software[17] in our attempts to obtain the lowest energy conformer within a reasonable computational time. Depending on the number of free torsion angles in each molecule numerous conformers (between 100 and 200) were found by MM optimizations. These optimizations were concluded when a gradient of 0.01 kcal/(Å mol) was reached for a certain conformer. The lowest energy conformer for a given molecule was then subjected to the quantum-mechanical semi-empirical AM1 calculations[18] in order to calculate the molecular characteristics. These optimized structures were loaded in Codessa Pro and more than 740 theoretical descriptors were calculated. These descriptors can be classified into several groups: (i) constitutional, (ii)

topological, (iii) geometrical, (iv) thermodynamic, (v) quantum chemical, and (vi) charge-related. The stepwise regression algorithm[19] encoded in Codessa Pro software was used to select significant descriptors for building multilinear QSAR models. The treatment started with the reduction in the number of molecular descriptors. If two descriptors intercorrelated highly with each other, then only one of them was selected; descriptors with insignificant variance for the data set treated were also rejected. This helps to speed up the descriptor selection and reduce the probability of including unrelated descriptors by chance. The 'best multilinear regression' (BMLR) approach encoded in Codessa Pro provides a QSAR equation that best fits the experimental data in terms of the Fisher criterion and the cross-validation coefficient R^2_{cv}.

A major decision in developing successive QSAR is when to stop adding descriptors to the model during the stepwise regression procedure. A simple technique to control the model expansion is the so-called 'breaking point' in the improvement of the statistical quality of the model, by analyzing the plot of the number of descriptors involved in the obtained models versus squared correlation coefficient values corresponding to those models. Frequently, the statistical improvement of the regression model is less significant ($\Delta R^2 < 0.02$–0.04) after a certain number of independent variables in the model ('breaking point'). Consequently, the model corresponding to the breaking point is considered the best/optimum model.

Another important step in the QSAR modeling is to validate the obtained model. Internal validation was carried out for the best model obtained by Codessa Pro as follows: (i) the parent data points (31) were divided into three subsets (A–C): the first, fourth, seventh, etc., data points go into the first subset (A), the second, fifth, eight, etc., into the second subset (B), and the third, sixth, ninth, etc., into the third subset (C), (ii) the three sets A–C were prepared as the combinations of two training subsets (A and B), (A and C), and (B and C), respectively. The remaining subsets (A, B, and C, respectively) become the corresponding test sets then, and (iii) a correlation equation was derived for each of the training sets with the same descriptors (but different regression coefficients). Next, the equation obtained was used to predict the protection time values for the compounds from the corresponding test set.

Another validation that was used in this study is leave-one-out approach.[20] This validation was performed for the main models. Thus, the efficiency of the QSAR equations to predict protection time was estimated based on the comparison of criterion such as the cross-validation coefficient R^2_{cv}. We also divided the parent data set to provide an external test set consisting of every fifth compound; we used the remaining 26 compounds as a training set to obtain a 4-descriptor model ($R^2 = 0.83$), where the external set was tested. It gave a satisfactory $R^2_{pred} = 0.76$.

The best statistical model obtained by using Codessa Pro descriptors for the PT data is shown in Table 2. This model includes 4-descriptors that are ordered by descending order according to their statistical significance (t test). In Table 2, X and ΔX are the regression coefficients and their standard errors. The co-linearity of any pair of the descriptors is less than $r^2_{col} = 0.42$.

Therefore, the model descriptors can be considered sufficiently orthogonal. The number of parameters was selected according to the breaking point rule for the improvement of R^2 as demonstrated in Figure 1.

There are several treatments of V_p in the literature suggesting that the vapor pressure can be correlated well with the protection time.[3,6] Because of the importance of the vapor pressure indicated by the previous workers, we tested whether the use of V_p as a descriptor would improve the correlation. It was clearly not appropriate to use measured V_p since an equation including such a descriptor could not be used conveniently for predictive

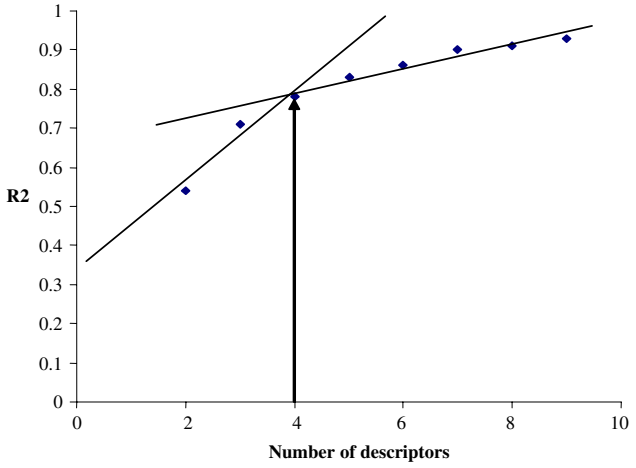


Figure 1. Breaking point rule for determination of the number of the descriptors.

Table 2. The best 4-descriptor QSAR model with $R^2 = 0.78$, $N = 31$, $F = 23.9$ and $s^2 = 0.51$

Descriptor no.	X	$\pm\Delta X$	t test	R^2	R^2_{cv}	s^2	Descriptor[a]
0	21.1	2.09	10.1				Intercept
1	−86.2	10.5	−8.19	0.16	0.08	1.78	Principal moment of inertia A, D_1
2	−0.93	0.13	−6.89	0.54	0.46	1.00	Structure information content (0), D_2
3	−0.99	0.23	−4.28	0.71	0.63	0.65	Kier and Hall index (order 2), D_3
4	−2.66	0.91	−2.91	0.79	0.70	0.51	Tot hybrid. comp. of molec. dipole, D_4

[a] Descriptor definitions are given in Supplementary material.

purposes. The dependence of V_p on structure depends somewhat on the type of compound considered, therefore a new QSPR model for the vapor pressure was derived specifically for the data set of 31 DEET analogues.

Experimental vapor pressures (V_p) of these compounds were taken from the SciFinder catalog.[21] Using Codessa Pro in the normal manner gave the 3-descriptor model for the vapor pressure shown in Table 3 with the following statistical characteristics: $R^2 = 0.956$, $R^2_{cv} = 0.940$, $F = 198.82$, and $s^2 = 0.041$.

The most significant descriptor according to the t test in Table 3 is D_5, the gravitation index calculated over all bonds of the molecule. As can be noted, with descriptor D_5 alone the model R^2 is already 0.80. In addition, the combination of the descriptors D_5 and D_6 shows that the equation is similar to the more general model developed in work (22) which is based on 411 compounds. All these suggests that the QSPR equation in Table 3 is reliable and can be used for adequate prediction of the vapor pressure.

Next, values of the vapor pressure predicted by Table 3 relationship (see Table 1) were used as an external descriptor in the common descriptor pool to try to improve the model of Table 2 for the protection times. Two functional relations were constructed from the predicted V_p, that is, (i) $\log V_p$ and (ii) $(\log V_p)^2$. After loading these descriptors in the whole Codessa Pro storage, the BMLR algorithm was run again in order to build the models. The best 4-descriptor model found among 744 descriptors is shown in Table 4. This equation included the $(\log V_p)^2$ as an independent variable. Moreover, a model with two descriptors (including $(\log V_p)^2$) already gave a significantly high correlation $R^2 = 0.70$ as can be seen from Table 4.

Tables 2 and 4 show that these two models are close from a statistical point of view. However, the equation in Table 4 is better than that of Table 2 in terms of R^2, s^2, and F. Also, the values of the statistical parameters show that the models are robust and describe well the experimental data. The Table 5 values of the protection time collects predicted from each of the models

Table 3. Three-parameter model for the vapor pressure (V_p) based on 31 compounds

Descriptor no.	X	$\pm\Delta X$	t test	R^2	R^2_{cv}	s^2	Descriptor[b]
0	10.49	1.01	10.41				Intercept
1	0.01	3e-4	−23.19	0.80	0.77	0.18	Gravitation index (all bonds), D_5
2	−26.51	2.88	−9.21	0.88	0.85	0.11	H-donors FPSA (version 2), D_6
3	0.43	0.06	7.01	0.96	0.94	0.04	Tot molecular 2-center resonance energy/no. of atoms, D_7

[b] Descriptor definitions are given in Supplementary material.

Table 4. The best 4-parameter model with calculated descriptor $(\log V_p)^2$: $R^2 = 0.80$, $F = 26$, and $s^2 = 0.47$

Descriptor no.	X	$\pm\Delta X$	t test	R^2	R^2_{cv}	s^2	Descriptor[c]
0	41.10	11.74	3.49				Intercept
1	−77.09	8.61	−8.95	0.165	0.08	1.77	Principal moment of inertia A, D_1
2	−0.25	0.03	−8.86	0.70	0.64	0.65	$(\log V_p)^2$, D_8
3	0.41	0.10	3.95	0.77	0.67	0.51	HA-dependent HDSA-2(Zefirov), D_9
4	−44.62	15.08	−2.95	0.80	0.72	0.47	Minimum atomic orbital electronic population, D_{10}

[c] Descriptor definitions are given in Supplementary material.

Table 5. Predicted protection times (PT) in hours

ID	Exp. PT	Pred. PT-2	Pred. PT-4	ID	Exp. PT	Pred. PT-2	Pred. PT-4
1	5	4.49	4.24	17	1.08	0.71	1.00
2	3	2.13	2.98	18	1	1.88	1.21
3	3	3.37	2.80	19	0.5	1.73	1.47
4	2.83	3.61	3.50	20	0.67	0.95	1.68
5	1.67	2.69	2.16	21	0.58	0.74	0.71
6	1	0.92	1.32	22	0.5	0.84	0.97
7	5	3.66	3.71	23	0.08	0.12	0.31
8	4	3.84	3.32	24	0.08	0	−0.30
9	4	3.45	3.81	25	0.08	−0.59	−1.28
10	3.5	2.84	3.24	26	3	2.96	2.19
11	2.83	2.93	3.16	27	2	1.2	1.28
12	1	1.65	2.24	28	1.42	2.03	1.39
13	3	2.65	2.93	29	1	2.15	1.63
14	2.67	1.71	2.15	30	1	1.61	1.57
15	2	2.35	2.52	31	0.75	0.57	0.92
16	1.17	0.59	0.57				

Predicted PT-2 using the model in Table 2.
Predicted PT-4 using the model in Table 4.

given in Tables 2 and 4. Graphical presentations of these predictions are provided in Figures 2 and 3.

It can be noted from both figures and Table 5 above that the PT of the compounds with ID 24 and 25 was predicted as a negative value. However, these compounds are not outliers according to the model errors (standard deviation). Since the BMLR method is not a constrained algorithm by the experimental values, this is possible.

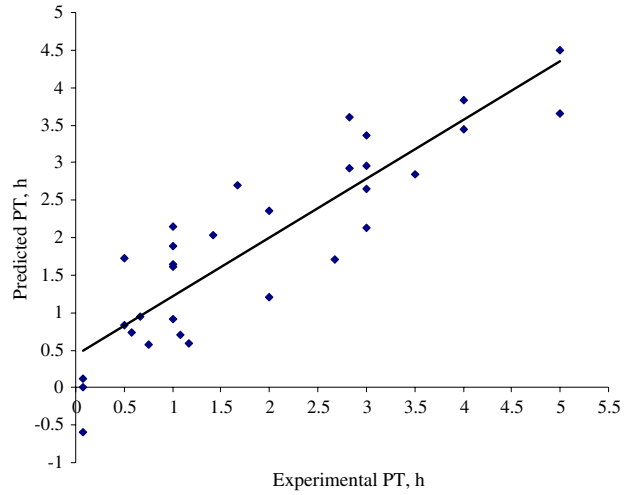


Figure 2. Experimental versus predicted PT according to the model in Table 2.

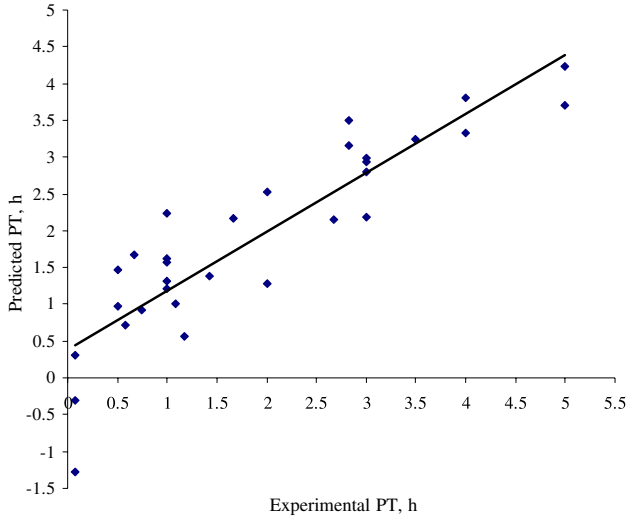


Figure 3. Experimental versus predicted PT according to the model in Table 4.

In order to test the predictive power of the models an internal threefold cross-validation was performed for the current data set. The results of this testing are shown in Table 6. The data sets A–C are divided as was explained previously. The superior robustness of the Table 4 model is also evident from Table 6.

The descriptors involved in the models could be possibly explained as follows: (i) D_1 and D_2, that are molecular shape related descriptors, represent the repellent fit into a receptor active center, (ii) D_5 describes the repellent chemical reaction with a receptor active center. Basically, the repellent activity quantified by the protection time can be assigned to the influence of three main molecular interactions. First, vaporization is connected with the duration of time when a mosquito can have contact with the repellent. As shown above and in previous models of vapor pressure,[22] the molecular size and shape descriptors such as D_1 and D_2 play a determining role for vapor pressure of compounds. The second important characteristic is the structural fit on an unknown active receptor center. The third kind of interaction should relate to the chemical reaction with a receptor, resulting in the act of repelling. Again, this should be directly related to the shape and size descriptors of this QSPR model (D_1 and D_2). The interaction between the active compound and its biological counterpart can be also reflected by D_3, that is connected to the shape and branching of the compound, and to descriptor D_4 that characterized the charge distribution in the compound. The dipole moment indicates the intrinsic polarity of the molecule. Its magnitude is also a good indicator of lipophilicity and hydrophobicity; the larger its magnitude, the higher is its hydrophilicity.[4]

Regarding functional groups and structural correlations, all compounds in the data set include an O atom. The descriptors D_9 and D_{10} are connected to the hydrogen donor capabilities of the molecule and the orbital electronic population. In turn, it could possibly influence the protection time of the repellent.[3] The most active compounds seem to be compounds with the aromatic ring bearing one substituent (CH_3 or Cl). The examina-

Table 6. Internal validation of the QSAR models

Training set	N	R^2 (fit)	R^2_{cv} (fit)	s^2 (fit)	Test set	N	R^2 (pred)	s^2 (pred)
				Model in Table 2				
A + B	20	0.84	0.72	0.43	C	11	0.71	0.80
A + C	21	0.83	0.73	0.51	B	10	0.60	0.97
B + C	21	0.72	0.67	0.59	A	10	0.82	0.62
Average		0.80	0.71	0.51			0.71	0.79
				Model in Table 4				
A + B	20	0.78	0.74	0.58	C	11	0.91	0.51
A + C	21	0.82	0.72	0.54	B	10	0.87	0.72
B + C	21	0.82	0.70	0.39	A	10	0.83	1.02
Average		0.81	0.72	0.51			0.87	0.75

tion of the respective descriptor values for such compounds showed a tendency for D_8 and D_1 values to be low. Change of the aromatic via alicyclic was usually slightly deleterious to PT. These structural criteria could be used for guidance for the synthesis of active repellents.

The equations possess one common descriptor: Principal moment of inertia A. It is likely that this descriptor related linearly to the experimental data and is important for the mass distribution of the molecule (see Supplementary material for the descriptor definition). Finally, the Principal moment of inertia A possesses the largest t values in both equations (t values define the statistical significance of a descriptor).

The model in Table 4 (in contrast to the model in Table 2) R^2 of 0.70 with just two descriptors. The addition of the external descriptor $(\log V_p)^2$ drastically improves the quality of the fit. We also tested as an additional external descriptor the lipophilicity $\log P$ (octanol–water partition coefficient), however, its inclusion did not lead to a better QSAR model.

Two QSAR models were developed for the description of mosquito repellent protection times PT with satisfactory statistical characteristics. The models include 4-descriptors revealing the linear relationship with the protection times of 31 repellents. External descriptors such as $\log V_p$ and its square function were added to the descriptor pool since the vapor pressure is important factor for the PT.

An additional QSPR model was developed for $\log V_p$ and thus no experimental data are needed for predictions of the PT from this data set. The examination of this descriptor space revealed that the descriptor $(\log V_p)^2$ is statistically significant and improves the model quality significantly. In addition, the descriptors that appeared in the models and feature the shape and volume as well as the charge distribution of compounds are likely important for determining the activity of the repellents. The PT predicted by both models (see Table 5) are slightly higher for compounds **8** and **9** that possess experimental PT values lower than DEET (**7**). However, the main prediction trend of these equations follows the experimental data within the error limits.

The success of the present work suggests that a general QSPR treatment of repellents could be of great benefit in synthetic efforts in the effort to discover better compounds for practical use.

Acknowledgment

The Estonian Science Foundation grant No. 4548 is acknowledged for the partial support of this work, as is the United States Department of Defense under SCA No. 58-6615-4-238 to the University of Florida.

References and notes

1. Plimmer, J. *Encyclopedia of Agrochemicals*; John Wiley and Sons, 2003.
2. Suryanarayana, M.; Pandey, K.; Prakash, S.; Raghuveeran, C.; Dangi, R.; Swamy, R.; Rao, K. *J. Pharm. Sci.* **1991**, *80*, 1055.
3. Davis, E. *J. Med. Entomol.* **1985**, *22*, 327.
4. Ma, D.; Bhatacharjee, A.; Gupta, R.; Karle, J. *Am. J. Trop. Med. Hyg.* **1999**, *60*, 1.
5. Peterson, C.; Coast, J. *Pesticide Outlook* **2001**, 154.
6. Gabel, M. L.; Spencer, T. S.; Akers, W. A. *Mosquito News* **1976**, *36*, 141.
7. Barnard, D.; Posey, K.; Smith, D.; Schreck, C. *Med. Vet. Entimol.* **1998**, *12*, 39.
8. Bhatarcharjee, A. K.; Gupta, R. K.; Ma, D.; Karle, J. M. *J. Mol. Recognit.* **2000**, *13*, 213.
9. Codessa Pro software, University of Florida, 2002, www.codessa-pro.com.
10. Katritzky, A. R.; Karelson, M.; Lobanov, V. S. *Pure Appl. Chem.* **1997**, 245.
11. Katritzky, A. R.; Fara, D.; Petrukhin, R.; Tatham, D.; Lomaka, A.; Karelson, M. *Curr. Top. Med. Chem.* **2002**, *2*, 1333.
12. Karelson, M.; Maran, U.; Wang, Y.; Katritzky, A. R. *Collect. Czech. Chem. Commun.* **1999**, *64*, 1551.
13. http://ark.chem.ufl.edu/pages/Research/qspr_2000/qspr_files/frame.htm.;
14. Katritzky, A. R.; Fara, D. C.; Yang, H.; Karelson, M.; Suzuki, T.; Solov'ev, V. P.; Varnek, A. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 529.
15. Katritzky, A. R.; Fara, D. C.; Karelson, M. *Bioorg. Med. Chem.* **2004**, *12*, 3027.
16. Katritzky, A. R.; Dobchev, D. A.; Fara, D. C.; Karelson, M. *Bioorg. Med. Chem.* **2005**, *13*, 1623.
17. HyperChem 7.05, Gainesville, Florida 2004, www.hyper.com.
18. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
19. Karelson, M. *Molecular Descriptors in QSAR/QSPR*; Wiley-Interscience: New York, 2000.
20. Stone, M. *J. R. Stat. Soc. B.* **1977**, *38*, 44.
21. SciFinder Scholar, American Chemical Society, 2003.
22. Katritzky, A. R.; Wang, Y.; Sild, S.; Tamm, T.; Karelson, M. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 720.