JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

QSAR Evaluation of Monoterpenoids' Insecticidal Activity

JUSTIN A. GRODNITZKY AND JOEL R. COATS*

Pesticide Toxicology Laboratory, Department of Entomology, Iowa State University, Ames, Iowa 50011

Monoterpenoids are naturally occurring compounds that are found in higher-order plants. These compounds are secondary metabolites that seem to play no major role in the metabolic functioning of the plants. One role of monoterpenoids in the plants is to defend against plant-directed pathogens, herbivores, or competing plant species. These compounds are good leads for synthesis or isolation of more effective insecticides. To accomplish these goals, we developed quantitative structure−activity relationships (QSARs) in order to predict insect toxicity of monoterpenoids and derivatives that have not yet been synthesized or experimentally tested. Correlations were found between toxicity and certain quantum and traditional chemical parameters. We found a linear relationship between LD_{50} values for house fly toxicity and Mulliken populations in aromatic monoterpenoids. Multiple linear regression of an E-State descriptor and a GETAWAY (GEometry, Topology and Atomic Weights AssemblY) descriptor also showed a relationship with house fly toxicity for a wide range of monoterpenoids.

KEYWORDS: Quantitative structure−activity relationship (QSAR); monoterpenoid; insecticide; Mulliken population; quantum descriptor; and electrotopological state

INTRODUCTION

Monoterpenoids are components of essential oils found in many higher-order plants. These compounds give plants their unique odoriferous properties. For example, limonene is primarily responsible for the scent in oranges, and thymol gives thyme its unique odor. These compounds are often found in perfumes and other cosmetics and are commonly used as food additives and therapeutic drugs (*1*).

Monoterpenoids are secondary plant metabolites that consist of two isoprene units. These compounds contain 10 carbons and seem to play no major role in the basal metabolic functioning of the plant. However, monoterpenoids are important to plants because they can attract beneficial insects to the plants, which can aid in pollination, and they can help plants defend against pathogens and herbivores such as insects (*2*). The natural insecticidal properties of some monoterpenoids make them excellent lead compounds for the development of safe, effective, and fully biodegradable insecticides.

Monoterpenoids have been shown to possess insecticidal activity, and a few of these compounds are currently being used commercially as pesticides or repellents (*d*-limonene, menthol, citronellal, and linalool) (*1*). Although these monoterpenoids are being used commercially, the mode of action of monoterpenoids is still not well understood. In addition, quantitative structure−activity relationships (QSAR) have not been determined, so the chemical basis for their insecticidal properties is not yet known.

* To whom correspondence should be addressed. Phone: 515-294-4776. Fax: 515-294-8027. E-mail: jcoats@iastate.edu.

In this paper, the order of toxicity of 30 natural monoterpenoids and their derivatives to *Musca domestica* (house fly) is examined, using some data reported previously from this laboratory (*3*). The toxicities of these compounds were then used to develop QSARs. Developing these relationships may facilitate the design of more effective insecticidal monoterpenoids and provide insight into the structural properties that are responsible for their toxicity. The QSAR development investigated a variety of parameters that help explain receptor−ligand interactions. Steric and electronic descriptors were used to help encode information about the important characteristics of monoterpenoids that are responsible for their toxic effects. Geometry, topology and atomic weights assembly descriptors (GETAWAY) were used to capture information about the three-dimensional structures of the molecules and encode information about the structural and steric requirements of these compounds (*4*). This descriptor assigns higher values for atoms that are distal from the molecule's geometric center. Higher values can be interpreted as the atoms that are more accessible to external interactions, and encode information about the molecule's size and shape. Two descriptors were used to explain electronic features important for eliciting a toxic response. Mulliken populations and an electrotopological state descriptor were used to represent electron density around certain atoms in the molecule. Electrotopological state descriptors have been shown to be highly correlated with Mulliken population and encode information about an atom's electron accessibility for external interactions (*5*). These descriptors have been useful in other QSARs to help explain the importance of electronic properties of molecules for various biological and physicochemical effects

10.1021/jf0201475 CCC: $22.00
Published on Web 07/02/2002

(*6*−*9*). In this study, these descriptors were used to develop effective models to explain and predict insect toxicity of monoterpenoids and their derivatives.

MATERIALS AND METHODS

Synthesis of Monoterpenoid Esters. A total of 30 monoterpenoids and their derivatives were examined in this study. Parent monoterpenoids (alcohols or phenols) were used to synthesize their ester and ether derivatives (*10*). (*l*)-Carveol (*l* at the 4-position) and its derivatives were constructed from a mixture of isomers (racemic at the 6-position). (*S*)-(−)Perillyl alcohol, (*R*)-(−)carvone, (*S*)-(−)limonene, (1*R*)-(−)-myrtenol, and mixtures of carvomenthen-4-ol, and α-terpineol isomers were used in this study.

Esters. Parent alcohols or phenols (1 mol) were added to the corresponding anhydride or acid chloride (2 mol) to form ester derivatives in the presence of a catalytic amount of pyridine (2−5 drops). Methylene chloride was used as the solvent, and the reaction was allowed to stir for 24−48 h at room temperature. Reactions were monitored by thin-layer chromatography (TLC) using a 9:1 hexane/acetone mobile phase solution, and visualization of the TLC plate was accomplished with vanillin spray (8 g of vanillin, 1.25 mL of sulfuric acid, brought up to 250 mL with methanol) and heat. The reaction was worked up with four washes of 100 mL of 10% $NaHCO_3$ solution. Methylene chloride was removed using a rotary evaporator. Compounds were purified using silica gel column clean up, using a 19:1 hexane/acetone solvent system. Identities of the esters were determined using TLC, comparing R_f values of the parent alcohols or phenols against reaction products, and were confirmed using ^{1}H NMR on a Varian VXR 300 MHz spectrometer. All ^{1}H NMR spectra performed used $DCCl_3$ as the solvent.

Ethers. Ether reactions were carried out using thymol and the corresponding alkyl halide in the presence of a phase-transfer catalyst, benzyltributylammonium bromide (BTAB). Thymol (5 g) was dissolved in 50 mL of CH_2Cl_2 together with the alkyl halide (10 mL) and BTAB (0.535 g). Then NaOH (2.0 g) dissolved in H_2O (175 mL) was added onto the organic layer. The reaction was allowed to stir for two weeks. The reactions were worked up and structure identities were determined as described for esters.

Bioassay. LD_{50} values were obtained for 30 monoterpenoids. A topical application was used to apply 1 μL of various concentrations of monoterpenoids to the pronotum of house flies. Ten treated house flies where placed in a jar. House flies used in this study were a mix of males and females. For each concentration, three replications of 10 monoterpenoid-treated flies were used. Six treatment concentrations were used for each chemical tested. Acetone was used as the solvent for each treatment. Controls were run for each of the treatments. A 1-μL aliquot of solvent was applied to the pronotum of the flies. At the end of the 24-hr exposure, mortalities of the house flies were recorded. LD_{50}s of all the monoterpenoids were calculated using the Spearman−Karber method (*11*). Some compounds' LD_{50} values were previously reported from our lab (*3*, *10*). LD_{50} values were converted to be expressed as μmol/fly.

Descriptors. Mulliken population was calculated in GAMESS molecular modeling software (*12*). Geometry and energy of all the molecules were optimized using a split valence basis set and a polarization function (6−31*d) calculation using GAMESS. Hessian runs were performed using 6−31*d calculations using GAMESS to show that all the molecules tested were at an energy-minimum conformation. Electrotopological state descriptors were calculated using E-calc. GETAWAY descriptor HATS2p was calculated by optimizing each structure using AM1 calculations on the CAChe work system (Fujitsu). Atomic coordinates were then input into the DRAGON program to obtain the GETAWAY descriptors.

Regression Analysis. Linear and multiple linear regressions were performed using SAS. The toxicity data was expressed as log(1/*C*), where *C* is the concentration in μmol that produces 50% mortality in house flies. The quality of each of the regression models was evaluated using the square of the correlation coefficient (r^2), and cross-validation (q^2). Q^2 values indicated the predictive power of the multiple regression equation. We examined only regressions with $r^2 > 0.80$. To evaluate the validation of the models, the leave-one-out method was used, which is calculated using the following equations:

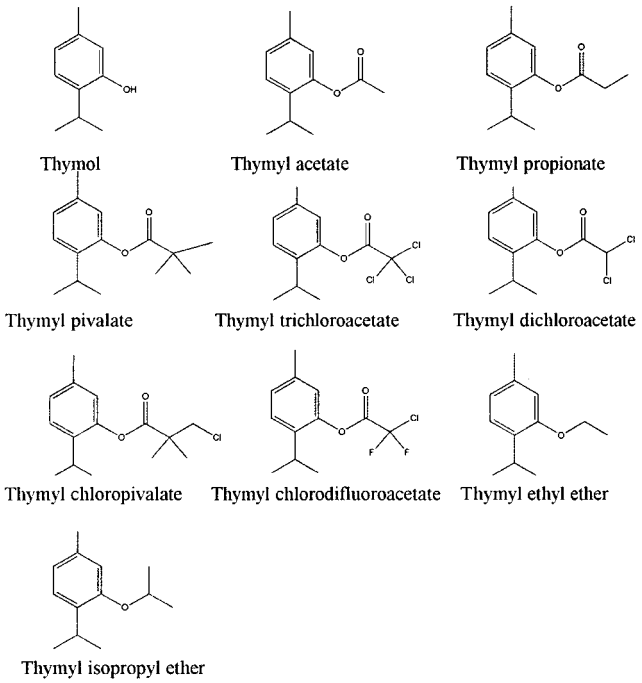

Figure 1. Structures of thymol and its derivatives.

$$\text{Cross-validation } q^2 = 1 - (\text{PRESS/SSTO}) \quad (1)$$

where

$$\text{PRESS} = \sum_y (Y_{\text{predicted}} - Y_{\text{actual}})^2 \quad (2)$$

and SSTO is the sum of squares total.

Cross-validation values greater than 0.40 have been used to imply a nonrandom relationship (*13*).

RESULTS

Because of the diverse structures of the monoterpenoids in this study, they were studied as two groups: the aromatic monoterpenoid, thymol, and nine derivatives (**Figure 1**), and 20 monocyclic and bicyclic (alicyclic) monoterpenoids (**Figure 2**). The set of alicyclic monoterpenoids all contained a methylcyclohexene backbone with the double bond present between carbon atoms 1 and 2 (**Figure 3**).

Twenty alicyclic monoterpenoids were used to construct the first QSAR model. An electronic descriptor and a steric descriptor were used to predict the toxicity of monoterpenoids and their derivatives to house flies. The GETAWAY descriptor was used to account for size and shape requirements that are essential for monoterpenoid toxicity. The electrotopological state (E-state) descriptor on atom 2 (**Figure 3**) was used to represent the important electronic properties necessary for monoterpenoids to exert their toxic effects. An excellent multiple linear correlation was obtained between GETAWAY and electrotopological state descriptors with house fly toxicity ($n = 20$, $s = 0.11$, $F = 32.59$, $r^2 = 0.86$, and $q^2 = 0.72$). The model obtained is as follows:

$$\log(1/LD_{50}) = -30.7(\pm 4.9) + 15.1(\pm 2.4)[\text{E-state}] + 213.8(\pm 36)[\text{GETAWAY}] - 105.8(\pm 17.6)[\text{interaction}] \quad (3)$$

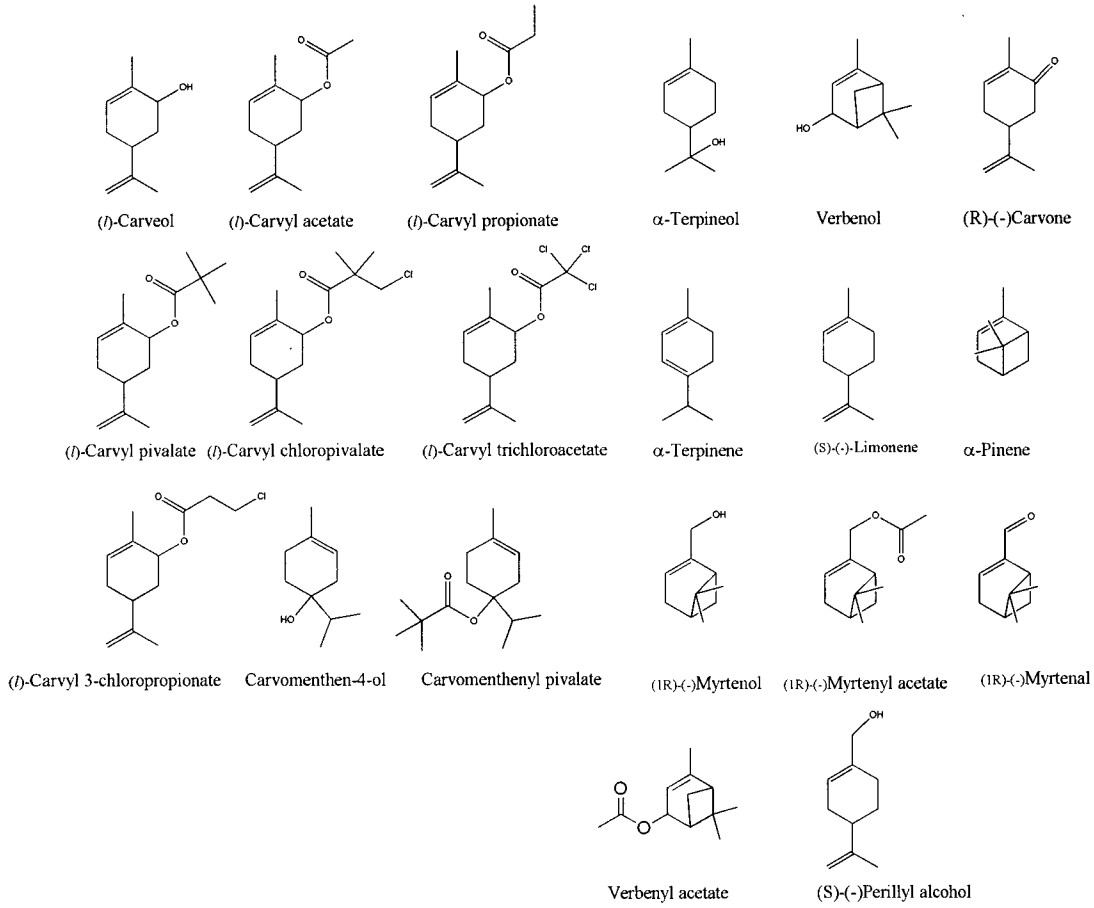

Figure 2. Structures of alicyclic compounds studied.

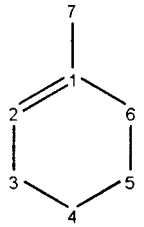

Figure 3. Methyl-cyclohexene carbon skeleton with a double bond present between carbon atoms 1 and 2 is the minimum structural requirement for the molecules in the alicyclic QSAR model.

Experimental and calculated log(1/LD_{50}) values are shown in **Table 1**. The model showed a good correlation between descriptors and their insecticidal activity (**Figure 4**). Equation 3 demonstrates that as the electron population or electronic accessibility increases, the monoterpenoid's toxicity increases, and as GETAWAY values increase, toxicity also increases. In this model, there is an interaction effect between the two descriptors. There is a balance between GETAWAY and electrotopological state descriptors due to the negative interaction term. This indicates that if both of these descriptors' values get too large, toxicity will decrease. If one of the descriptors yields a smaller value, it allows the other descriptor to have a larger value and the compound will still exert a toxic effect.

The other QSAR was developed for thymol and nine of its derivatives. Two of the derivatives were ethers and the other seven were esters. Thymol and the two ether derivatives had the greatest toxicity to the house fly. A linear relationship was obtained between toxicity of the thymol group of compounds and the Mulliken population (electron density) around atom 1, the phenolic carbon ($n = 10$, $s = 0.08$, $F = 68.52$, $r^2 = 0.90$, and $q^2 = 0.84$). The numbers on the atoms of the thymol compounds correspond to the IUPAC nomenclature (**Figure 5**). Experimental and calculated log(1/LD_{50}) values are shown in **Table 2**. The model showed a good correlation between Mulliken population and toxicity for the thymol compounds and

Table 1. House Fly LD_{50} (with 95% Confidence Intervals), Predicted and Residual Values for 20 Alicyclic Monoterpenoids and Their Derivatives

chemical sample	experimental LD_{50} (μmol/fly)	experimental log(1/LD_{50})	predicted log(1/LD_{50})	residual
(*l*)-carveol	1.03 (0.78–1.46)	–0.01	–0.11	0.10
(*l*)-carvyl acetate	0.57 (0.54–0.61)	0.24	0.16	0.08
(*l*)-carvyl propionate	0.99 (0.92–1.06)	0.00	0.23	–0.23
(*l*)-carvyl 3-chloropropionate	1.43 (1.33–1.54)	–0.16	–0.09	–0.06
(*l*)-carvyl trichloroacetate	2.70 (2.38–3.02)	–0.43	–0.42	–0.01
(*l*)-carvyl pivalate	0.37 (0.35–0.40)	0.43	0.39	0.04
(*l*)-carvyl chloropivalate	0.96 (0.85–1.09)	0.02	0.18	–0.16
(*l*)-carvomenthen-4-ol	0.71 (0.67–0.75)	0.15	0.12	0.02
carvomenthen-4-yl pivalate	0.16 (0.13–0.20)	0.79	0.74	0.05
(*R*)-carvone	1.12 (0.68–1.42)	–0.05	–0.20	0.15
α-terpineol	1.29 (0.86–1.45)	–0.11	0.05	–0.16
(*S*)-perillyl alcohol	0.38 (0.32–0.45)	0.42	0.31	0.11
α-terpinene	0.86 (0.62–1.20)	0.07	0.24	–0.17
limonene	0.37 (0.34–0.58)	0.43	0.31	0.13
(*R*)-myrtenol	0.64 (0.57–0.73)	0.19	0.16	0.03
(*R*)-myrtenal	1.54 (1.47–1.62)	–0.19	–0.16	–0.03
(*R*)-myrtenyl acetate	0.36 (0.31–0.41)	0.45	0.45	0.00
verbenyl acetate	0.59 (0.54–0.67)	0.23	0.13	0.10
verbenol	1.51 (1.45–1.56)	–0.18	–0.18	0.00
α-pinene	0.82 (0.62–1.20)	0.09	0.09	0.00

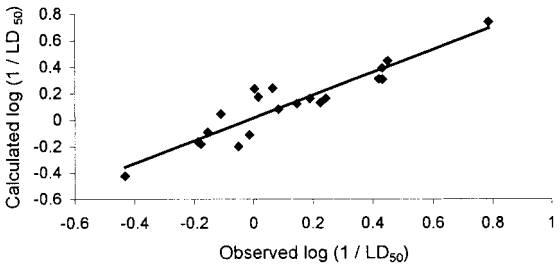

Figure 4. Plot of calculated versus observed toxicity values for 20 alicyclic monoterpenoids and their derivatives.

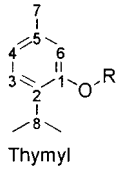

Figure 5. Numbering of the carbon atoms for thymol compounds.

Table 2. House Fly LD_{50} (with 95% Confidence Intervals), Predicted and Residual Values for Thymol and Nine Derivatives

chemical sample	experimental LD_{50} (μmol/fly)	experimental log(1/LD_{50})	predicted log(1/LD_{50})	residual
thymol	0.22 (0.20–0.24)	0.66	0.62	0.04
thymyl acetate	0.49 (0.44–0.54)	0.31	0.38	–0.07
thymyl propionate	0.49 (0.40–0.62)	0.31	0.37	–0.06
thymyl pivalate	0.34 (0.22–0.42)	0.47	0.47	0.00
thymyl chloropivalate	1.12 (0.98–1.27)	–0.05	–0.05	0.00
thymyl dichloroacetate	0.47 (0.31–0.68)	0.33	0.21	0.12
thymyl chlorodifluoroacetate	0.90 (0.70–1.60)	0.05	0.11	–0.06
thymyl trichloroacetate	0.62 (0.56–0.69)	0.21	0.21	0.01
thymyl ethyl ether	0.27 (0.21–0.37)	0.57	0.66	–0.09
thymyl isopropyl ether	0.23 (0.18–0.32)	0.64	0.51	0.13

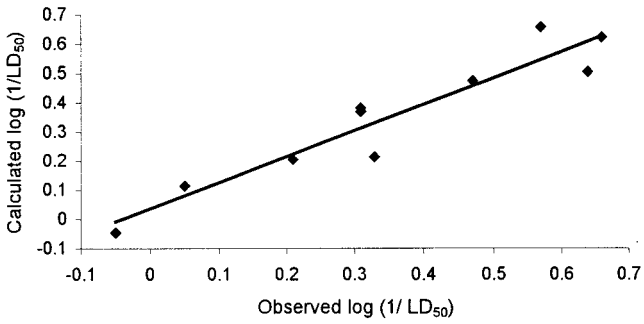

Figure 6. Plot of calculated versus observed toxicity values for thymol and nine derivatives.

is presented in eq 4 (**Figure 6**):

$$\log(1/LD_{50}) = 65.3(\pm 7.9) - 11.6(\pm 1.4)[\text{Mulliken population}] \quad (4)$$

This relationship shows that toxicity decreases as the Mulliken population around the phenolic carbon 1 (atom number in **Figure 5**) increases.

DISCUSSION

Two new QSAR models were developed for insecticidal monoterpenoids. In both models, the electronic characteristics of the molecules play an important role in toxicity. The alicyclic model used an electrotopological state descriptor to represent the electronic properties of the molecules. As the electron population or electron accessibility (rich) increased, toxicity also increased. This relationship might be due to the electrostatic interaction of these compounds to a receptor, and as electron accessibility for the monoterpenoids molecules increases, binding affinity also increases. Electronic properties were also the essential component in the thymol QSAR. However, its effect was the inverse of the alicyclic QSAR. In both models, the electronic properties are very important and indicate the toxicophore regions in the molecules that are important for their toxicity.

Size and shape of the molecule were implicitly expressed in the GETAWAY descriptor. The alicyclic monoterpenoid QSAR used a dynamic range of compounds, including monocyclic and bicyclic monoterpenoids. Because of this dynamic range, our model used the GETAWAY descriptor to account for the minimum required shape and size of these compounds. The use of this descriptor infers that there is an optimum shape and size requirement that monoterpenoids must possess to fit into a site of toxic action.

In the thymol QSAR, the size and shape of the structure had very little effect on toxicity. This is probably due to the fact that these compounds are very closely related in structure, and that the slight changes in steric dimensions have little effect on their activity. This information potentially can be used to aid in the development of a more effective thymol derivative. Also, because changing the size and shape of the structural moiety on the oxygen atom has little effect on toxicity, selected groups can be added as substituents on that atom to influence the electronic properties of the molecules in order to increase their toxicity.

As previously noted, the modes of action of monoterpenoids are not well understood. Researchers have shown that the monoterpenoids linalool, bornyl acetate, cineole, citral, and pulegone have affinity for acetylcholinesterase (*14*); however, high concentrations of monoterpenoids were used in that study. Researchers were unable to correlate the inhibition of enzyme activity with the monoterpenoids' toxicity. Other researchers demonstrated that thujone and some of its metabolites showed good binding affinity for GABA-gated chloride ion channel in mouse brain (*15*). It also showed *Drosophila* of the dieldrin-resistant (Rdl) strain were less susceptible to thujone toxicity. Another researcher has demonstrated that eugenol, α-terpineol, and some other monoterpenoids bind to octopamine receptors (*16*).

Monoterpenoids are a diverse set of compounds whose mode of action may not be represented by one exclusive mechanism. The two individually developed QSAR models (aromatic, and alicyclic) suggest that the monoterpenoids included within each model exert the same mode of action. However, the QSARs derived for each of the two models are somewhat different, suggesting that the compounds in the alicyclic model may have a different mode of action than the compounds in the aromatic model, because of the different structural requirements needed to develop the models. These models provide insight into the important regions of the molecules responsible for their insecticidal properties. We hope that these models will be used in the future to develop new effective alternative insecticides, as well as contributing to a better understanding of their mechanism(s) of action.

ACKNOWLEDGMENT

We are grateful for the technical assistance of Kim Hoover and Erica Simbro for their laboratory assistance.

LITERATURE CITED

(1) Tsao, R.; Coats, J. R. Starting from nature to make better insecticides. *Chemtech* **1995**, *25*, 23−28.

(2) Hallahan, D. L. Monoterpenoid Biosynthesis in Glandular Trichomes of Labiate Plants. In *Plant Trichomes*; Academic Press: San Diego, CA, 2000; pp 77−120.

(3) Rice, P. J.; Coats, J. R. Insecticidal properties of monoterpenoid derivatives to the house fly (Diptera: Muscidae) and red flour beetle (Coleoptera: Tenebrionidae). *Pestic. Sci.* **1994**, *41*, 195−201.

(4) Consonni, V.; Todeschini, T. Pavan, R. Structure/Response Correlations and Similarity/Diversity Analysis by GETAWAY Descriptors: 1. Theory of the Novel 3D Molecular Descriptors. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 682−692.

(5) Kier, L. B.; Hall, L. H. Molecular Structure Description: The Electrotopological State; Academic Press: New York, 1999.

(6) Abou-Shaaban, R. R.; al-Khamees, H. A.; Abou-Auda, H. S.; Simonelli, A. P. Atom level electrotopological state indices in QSAR: Design and testing antithyroid agents. *Pharmacol. Res*. **1996**, *13*, 129−134.

(7) Bualamwini, J. K.; Raghavan, K. M.; Pommies, Y.; Kohn, K. W.; Weinstein, J. N. Application of the electrotopological state index in QSAR. Analysis of flavone derivatives as HIV-1 integrase inhibitors. *Pharmacol. Res*. **1996**, *13*, 1892−1897.

(8) De Gregorio, C.; Kier, L. B.; Hall, L. H. QSAR modeling with the electrotopological state indices: Corticosteroids. *J. Comput.-Aided Mol. Des*. **1998**, *12*, 557−561.

(9) Gough, J. D.; Hall, L. H. Modeling the toxicity of amide herbicides using the electrotopological state. *Environ. Toxicol. Chem*. **1999**, *18*, 1069−1075.

(10) Tsao, R.; Lee, S.; Rice, P. J.; Jensen, C.; Coats, J. R. Monoterpenoids and their synthetic derivatives as leads for new insect-control agents. In *Synthesis and Chemistry of Agrochemicals IV*; American Chemical Society: Washington, DC, 1995; Chapter 28.

(11) Hamilton, M. A.; Russo, R. C.; Thurston, R. V. Trimmed Spearman−Karber method for estimating median lethal concentrations in toxicity bioassays. *Environ. Sci. Technol*. **1977**, *11*, 714−719; Correction. **1978**, *12*, 417.

(12) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, K.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General Atomic and Molecular Electronic Structure System, *J. Comput. Chem*. **1993**, *14*, 1347−1363.

(13) Wold, S. Validation of QSAR's. *Quant. Struct.-Act. Relat*. **1991**, *10*, 191−193.

(14) Ryan, M. F.; Byrne, O. Plant-Insect coevolution and inhibition of acetylcholinesterase. *J. Chem. Ecol.* **1998**, *14*, 1965−1975.

(15) Hold, K. M.; Sirisoma, N. S.; Ikeda, T.; Narahashi, T.; Casida, J. E. α-Thujone (the active component of absinthe): γ-Aminobutyric acid type A receptor modulation and metabolic detoxification. *PNAS* **2000**, *97*, 3826−3831.

(16) Enan, E. Insecticidal activity of essential oils: Octopaminergic sites of action. *Comp. Biochem. Physiol., C Toxicol. Pharmacol*. **2001**, *130*, 325−337.

Received for review February 4, 2002. Revised manuscript received May 27, 2002. Accepted May 29, 2002. This chapter is journal paper J-19708 of the Iowa Agriculture and Home Economics Experiment Station, Ames, IA 50011.

JF0201475