 Bull. Chem. Soc. Jpn., 74, 2451–2461 (2001) 

# Artificial Neural Network for Predicting the Toxicity of Organic Molecules

**Gary Admans, Yoshimasa Takahashi,* Satoshi Ban, Hiroaki Kato, Hidetsugu Abe, and Sosuke Hanai†**

Department of Knowledge-based Information Engineering, Toyohashi University of Technology, 1-1 Hibarigaoka, Tempaku-cho, Toyohashi 441-8580

†Chemical Safety Management Center, Japan Chemical Industry Association, 3-2-4 Kasumigaseki, Chiyoda, Tokyo 100-0013

(Received May 28, 2001)

Structure–activity relationships for aquatic toxicity were studied using neural networks and linear regression analysis. The structural features contributing to toxicity were identified in molecules exhibiting a level of toxicity greater than that of non-reactive organic molecules. A neural network was trained for the toxicity of non-polar narcotics, polar narcotics, or reactive toxicants. Quantitative structure–activity relationships (QSARs) were developed, relating a molecular aquatic toxicity to its log *P* and to a set of 16 structural descriptors based upon the presence of selected structural features. The inclusion of these structural descriptors into a QSAR was found to enhance the correlation of the equation, and thus to improve its ability for predicting aquatic toxicity.

Chemists, toxicologists, and regulatory authorities need to be able to predict the biological effects of a new chemical to estimate the danger it poses to the human population and to the environment. The first step in this process involves identifying any potential hazards presented by the new chemical. Once the hazards have been identified, then a quantitative estimate of the danger posed by the molecule can be performed by comparing the properties of the new molecule with those of similar molecules that were tested previously. This approach is increasingly important as a key element for priority setting in chemical risk assessment.

Many studies have been performed to rationalize the structure-activity relationships for aquatic toxicants using QSAR, Quantitative Structure–Activity Relationships.[1] This method has become accepted as the most promising of the various methods used for predicting biological activity. In a QSAR study, one attempts to find a relationship between the measured biological activity and the molecular descriptors related to a molecular structure, electronic properties, and other measured or calculated physical properties. The limitations of QSAR are that relationships, and hence their predictions, can only be made for similar molecules, usually with a common mode of toxic action, and that the prediction is only reliable when the data from a new molecule is contained within the spanned substituent space of the data used to create the QSAR relationship. The utility of QSAR in aquatic toxicity was demonstrated by Könemann,[2] who conducted a study on 72 non-reactive industrial pollutants and found that a QSAR based solely upon log *P*, a molecular octanol–water partition coefficient, gave a good estimation of the toxicity for most of the tested molecules within the group. The relationship was linear with an upper limit of about 6, at which point the low water solubility of hydrophobic molecules would reduce their availability within the medium. Könemann concluded that the lethal effect was probably caused by membrane perturbation, and that it seemed to be a minimum effect. This proved to be a very effective method for predicting the toxicity of new molecules because there are various computational methods available for estimating log *P*, thus avoiding the need for laboratory tests.[3] Later studies of non-reactive chemicals produced similar equations to that of Könemann, and they have come to be known as baseline toxicity QSAR's.[4,5] Because the baseline QSAR is thought to predict the minimum toxicity for an unreactive chemical, more reactive chemicals usually have an activity higher than what would be predicted by the baseline QSAR. Various baseline QSARs were found to be highly similar, even when comparing those from 19 different aquatic species.[6] Veith and co-workers[4] observed that a bilinear equation gave a better fit to the toxicity data for non-reactive molecules, particularly at higher log *P*.

The aquatic toxicity produced by unreactive molecules is known as narcosis, and it's biological effects can be likened to the effects of anaesthetics on mammals. Narcosis acts by a non-specific reversible disturbance of membrane function, caused by the accumulation of the toxicant in the hydrophobic phases within an organism.[7] A distinction was made between the types of molecules exhibiting narcotic toxicity; consequently, the majority of narcotics were termed non-polar narcotics, and a subgroup of more polar molecules exhibiting slightly higher toxicity was termed polar-narcotics. Most unreactive organic chemicals fit the non-polar narcosis model, for example alcohols, ketones, ethers, alkyl halides, and some benzene derivatives.

Veith and Broderius[8] found that a baseline toxicity QSAR did not accurately predict the toxicity of some polar compounds despite the appearance of the symptoms of toxicity be-

ing narcotic. Non-polar and polar narcotics have been distinguished using additivity experiments with octanol and phenol, which are themselves recognised as being non-polar and polar narcotics, respectively. The types of molecules that were identified as polar narcotics are aromatic amides, amines, phenols, and some nitrogen-containing heterocycles. Non-polar narcosis may be the result of hydrophobic bonding of the molecule to an enzyme or membrane, whereas polar narcosis may result from the presence of a strong hydrogen bonding group in the molecule. Veith and Broderius concluded that molecules with log *P* values less than 2.7 were polar narcotics, whereas those with higher log *P* values were often jointly additive with both octanol and phenol, suggesting that their activity could be modeled using the baseline toxicity QSAR. Other molecules have toxicities that are many times higher than that which would be predicted by a baseline toxicity QSAR. These molecules are said to have a "reactive toxicity", and the physiological responses of some types have allowed them to be distinguished from chemicals that are toxic by narcosis.[9] Their structures are diverse, and many, such as electrophiles,[10] are chemically reactive, but have a non-specific biological reactivity, whereas others possess a more specific biological activity, such as organophosphorus pesticides, which are known to be acetyl choline esterase inhibitors.[11] The extent to which a molecules is more toxic than what would be predicted by the baseline QSAR is known as the excess toxicity (Te); it can be calculated using the following equation:[5]

$$\text{Te} = \text{predicted toxicity/observed toxicity.} \quad (1)$$

Those molecules possessing an excess toxicity greater than 2 have been termed "excess toxicants". Many were identified by Lipnick[12] who classified them according to their chemical structures and possible mechanisms of biological activity. Similarities can be observed in the magnitude of excess toxicity for molecules possessing similar chemical structures. Baseline toxicity QSARs have been used to predict the minimum toxicity for molecules possessing chemically diverse structures; however, for reactive chemicals, sometimes other means are necessary to obtain a more accurate prediction of their elevated toxicities. In the case of aldehydes, the QSAR developed was very similar to a baseline QSAR in that only log *P* of a molecule was needed to estimate it's toxicity.[13] In other cases, however, additional factors, such as molecular descriptors related to structural or electronic properties were needed to create a more accurate QSAR. For example, the reaction rate constants from a chemical reaction with a standard nucleophile greatly improved the QSAR's for reactive halides[14] and epoxides.[15] Indeed, a QSAR has even been created with good predictive ability for many different types of electrophiles by including a parameter for the average nucleophile superdelocalizability.[16]

Most of the methods mentioned above give a prediction of toxicity that is limited to a small group of molecules. In order to apply these QSAR equations to the prediction of toxicity of new molecules, methods are needed to classify them into known groups. Verhaar and co-workers[17] proposed a scheme that divides molecules into four classes based upon both the structural features identified within the molecule and the molecular log *P*. The four classes identified in this scheme are inert chemicals (non-polar narcotics), less inert chemicals (polar narcotics), reactive chemicals, and specifically active chemicals. The scheme used structural features to determine the classification of a new molecule. Once a molecular class had been identified, it's toxicity could be calculated from a baseline toxicity QSAR for class 1 and by applying range factors to the baseline toxicity prediction for classes 2, 3, and 4. The range factors, the equivalent of excess toxicity scores, were obtained from a study of the excess toxicity exhibited by molecules belonging to each class. (The range factors used were a minimum of 5 to a maximum of 10 for class 2, and a minimum of 10 to a maximum of 10000 for classes 3 and 4.) This approach was successful in many respects, allowing the classification of 48% of a database of 2000 chemicals, affording a good estimate of the toxicity of 230 class-1 molecules, and worst-case estimates for the other classes. Those chemicals that remained unclassified were typically inorganics, mixtures, or polymers.[18] Other researchers have aimed to find a method for discriminating between baseline toxicants and excess toxicants using only structural features. Hermens[10] reviewed chemically reactive substructures for electrophiles and indicated instances where QSAR relationships have been developed for electrophiles exhibiting aquatic toxicity. Lipnick[12] highlighted the similarity between the structures of "outliers" or excess toxicants that were responsible for aquatic toxicity $LC_{50}$ and rat oral $LD_{50}$; he classified the molecules according to proposed mechanisms for their activity. Jäckel and Nendza[19] analyzed the ERL-D database, and identified 122 (24.4%) of the molecules as excess toxicants; the remaining molecules fitted a baseline toxicity model. They found that 14 substructural indicators accounted for the excess toxicity of 90 (73.4%) of the excess toxicants. Examples of these substructural indicators were, for example, allylic/propargylic primary/secondary alcohol, catechol, and imidazole. The application of these substructures, in combination with a baseline QSAR, allowed correct predictions of baseline and excess toxicities for 86.4% of the chemicals analyzed. On the other hand, Eldred et al.[20] investigated the ability of a neural network for modeling the QSARs of aquatic toxicity of diverse organic molecules. However, the descriptors used in their work were often difficult, or impossible, to interpret the actual relationship discovered by the network.

In our study, we chose to use neural networks for classifying highly reactive aquatic toxicants, and also for creating some of the QSAR's used to predict the toxicity of new molecules using only the calculated structural features Their advantages have been extolled as the ability to create non-linear relationships and the ability to identify any patterns in a data set without the need to make assumptions about the underlying reasons for those patterns. They can allow the discovery of relationships starting from a varied data set containing both relevant and irrelevant data, even when the data are noisy or partially erroneous.

## Methodology

The data used for our studies were taken from the U.S.-EPA Environment Research Laboratory-Duluth database (ERL-D), which in the format published by Nendza and Russom con-

tained the results from 618 tests on 485 molecules in the 96-hr $LC_{50}$ of the fathead minnow (*Pimephales promelas*).[21] In instances where there were duplicated tests results for one molecule, an average was taken, and used for our study. Structures of the molecules were found using standard chemistry reference books, drawn in 2D using ISIS draw, and later exported as .mol files. The .mol files were then converted into a format suitable for input to our own software. Log *P*, MW (molecular weight), and MR (molar refractivity) were calculated using Tutors/Procalc.[22] The program SSCheck[23] was used to analyze the substructures contained within the molecular structures, producing a list of types and number of occurrences per molecule. Of the 485 molecules, seven were excluded from our study because of problems encountered in the estimating their log *P*, resulting in a database of 478 molecules. These excluded molecules were iodoform, tetrabutyltin, tetraethyltin, tricaine (MS-222), amphetamine sulfate, nicotine sulfate and maitus yellow dihydrate. For salts with a single counter ion, such as a sodium cation or a chloride anion, both of which were assumed to be non-toxic in dilute solution, the structure of the protonated or deprotonated organic compound was used without a counter ion in place of the original structure. Thus, organic acid salts were processed by using the structure of the protonated free acid, and quaternary ammonium salts were processed using the structure of the free tertiary amine. The baseline toxicity and excess toxicity for the molecules were calculated using the program ExTox, written in C, on a PC equipped with Microsoft Windows 95 and Visual C++. The QSAR model by Veith and Broderius[8] was used to calculate the baseline toxicity, and the excess toxicity was calculated using Eq. 1. For consistency and to afford a comparison with other data sets, the $LC_{50}$ values which we used were in units of mol/L. A computational neural network was utilized for model building. The computational program used for training the artificial neural network was developed by the authors following a method presented by Aoyama and co-workers.[24]

## Results and Discussion

In an analysis of the chemical structures, 2D structural descriptors were found to be more effective than 3D descriptors for the purpose of distinguishing between biologically active and inactive compounds.[25] To identify 2D structural descriptors relevant to aquatic toxicity, both the ERL-D database and a second database, containing a collection of aquatic toxicity tests performed on the guppy (*Poecilia reticulata*),[17] were used. The excess toxicities were calculated using ExTox and the excess toxicants were identified as those molecules possessing an excess toxicity greater than 2 units. Their molecular structures were examined, and they were classified into groups according to their chemical composition, as shown in Table 1.

Structural features that are potentially responsible for excess toxicity were identified from the molecules in these groups of excess toxicants. There were two types of structural features: those that individually appear to be the causative factors for excess toxicity, and those which cause excess toxicity only when other groups are present within a molecule. Examples of the former are chemically reactive substructures, such as aldehyde, epoxide, and $\alpha,\beta$-unsaturated carbonyl groups. Examples of the later group that are active by a combination of functional groups are aromatic halogens, aniline, phenol, and nitrobenzenes. To increase the predictive ability of our system, additional structural features were added. These were obtained either by extrapolation based upon knowledge of the chemical reactivity or taken from the literature.[10,12,17,19] Finally, a collection of over 100 structural features was produced (see Fig. 1).

The program SSCheck already contained 97 structural descriptors which described both the ring structures and functional groups that had been compiled for a general analysis of organic compounds. From the substructures identified in this study, 59 more were added to compliment those already in SSCheck. Using all 156 substructures, their frequencies within the molecules in the ERL-D database were determined and the results of this analysis were used to eliminate those substructures found only a few times in the whole data set. This was done because single data points would add unnecessary noise to the system and detract from the training of the neural networks, which require the repetition of patterns. Therefore, the 30 most frequently found substructures were selected to provide input to our neural network systems.

The molecules within the ERL-D database were divided into 3 types of aquatic toxicants using the following procedure. Molecules previously used in published studies[2,4,17,26] of nonpolar narcotics were identified and classed as group 1. Molecules identified as polar narcotics in studies[8,9] of binary mixtures were classed as group 2. Those molecules possessing an excess toxicity greater than 10 were categorized as group 3. The remaining molecules were classified by examining both their functional groups and their values of excess toxicity. Those molecules that were structural analogues of already classified molecules and had a similar excess toxicity value were placed in the same class. Molecules possessing no reactive functional groups or an excess toxicity below 2 were placed in group 1; the remainder were placed in group 3. Thus, three groups of 252, 61, and 165 molecules were produced for classes 1 to 3, respectively. We intended to create relationships for both the whole data set and for the individual classes in the hope that the prediction ability of our system could be improved for the classified molecules. The data set was further divided into a training set of 384 molecules (80%) and a testing set of 94 molecules (20%). The molecules in the testing set were specifically selected to represent typical examples of all the frequently found molecular types. Table 2 gives the results of this manual classification with the whole data set.

We initially used a linear regression analysis to create a QSAR relating the aquatic toxicity to log *P*. This analysis facilitated a comparison of our results with other studies, and also provided a benchmark for our work. Using all 478 molecules in the data set, the following equation was produced:

$$\log 1/LC_{50} = 0.825 \log P + 1.194, \quad (2)$$
$$n = 478,\ r = 0.685,\ s = 0.825.$$

This equation evidently has only a fair correlation for the whole data set, and compares unfavorably with the baseline QSAR, in following equation 3 created by Nendza and Russom, who modeled 147 non polar narcotic molecules taken from the same data set:[21]

Table 1. Classification of Excess Toxicants

| Chemical Class | Sub-division |
|---|---|
| 1. Aldehydes | |
| 2. Strained 3-membered rings containing a heteroatom, and their precursors. | a)Epoxides |
| | b) Imines |
| | c) Precursors: 1,2 halohydrins, ? -halo ethers, nitrogen, and sulfur mustards |
| 3. Halides | a) Allylic |
| | b) Propargylic halides and nitrilic halides |
| | c) Benzylic halides |
| | d) $\alpha$-halo esters, amides, ketones, and aldehydes |
| | e) others: e.g. alkyl dihalides, multi halo aromatic rings |
| 4. Michael acceptors: | a) $\alpha,\beta$-unsaturated aldehydes, ketones, esters, amides |
| | b) Nitrostyrenes |
| | c) $\alpha,\beta$-unsaturated nitriles |
| | d) Vinyl sulfones |
| | e) Quinones |
| 5. Allylic and propargylic alcohols | |
| 6. Allyl ethers | |
| 7. Nitriles | a) cyanohydrins |
| | b) $\alpha$-amino nitriles |
| | c) ester precursor to a cyanohydrin |
| | d) $\gamma$-allylic and gamma propargylic nitriles |
| | e) thiocyanates |
| 8. Nitroaromatics (particularly multi nitrobenzenes) | |
| 9. Aromatic alcohols | a)phenols |
| | b) catechols |
| | c) precursors to catechols (methyl ethers, esters) |
| 10. Amines | a) primary aromatic amines (anilines) |
| | b) aliphatic amines (particularly cyclic) |
| 11. Esters (particularly diesters) | |
| 12. Carbamates | |
| 13. Thiophosphoric acid esters | |
| 14. Quaternary ammonium salts | |
| 15. Heterocyclic rings | a) 5-membered: e. g. furans, imidazoles, thiazoles |
| | b) Pyridines (particularly halogen substituted) |
| 16. Sulfur-containing molecules | a) disulfides |
| | b) sulfides |
| | c) thiols |
| | d) disulfurans |
| 17. Anhydrides | |
| 18. Nitrosoaromatics | a) with *ortho*-or para-amino or hydroxy groups |
| | b) precursors to *ortho*- or *para*-nitrosoanilines |
| 19. Diazo-compounds | |
| 20. Haloacetamides | |
| 21. Organometallics | |
| 22. Benzene derivative with salicylate type 1,2 functionality | |
| 23. Conjugated double bonds | |

$$\log 1/LC_{50} = 0.79 \log P - 1.35, \quad (3)$$
$$n = 147,\ r = 0.92\ s = 0.40.$$

Although Eq. 2 gives a poor correlation between the toxicity and log $P$, this was the expected result because of the heterogenous nature of the structures in the data set. In comparison, the

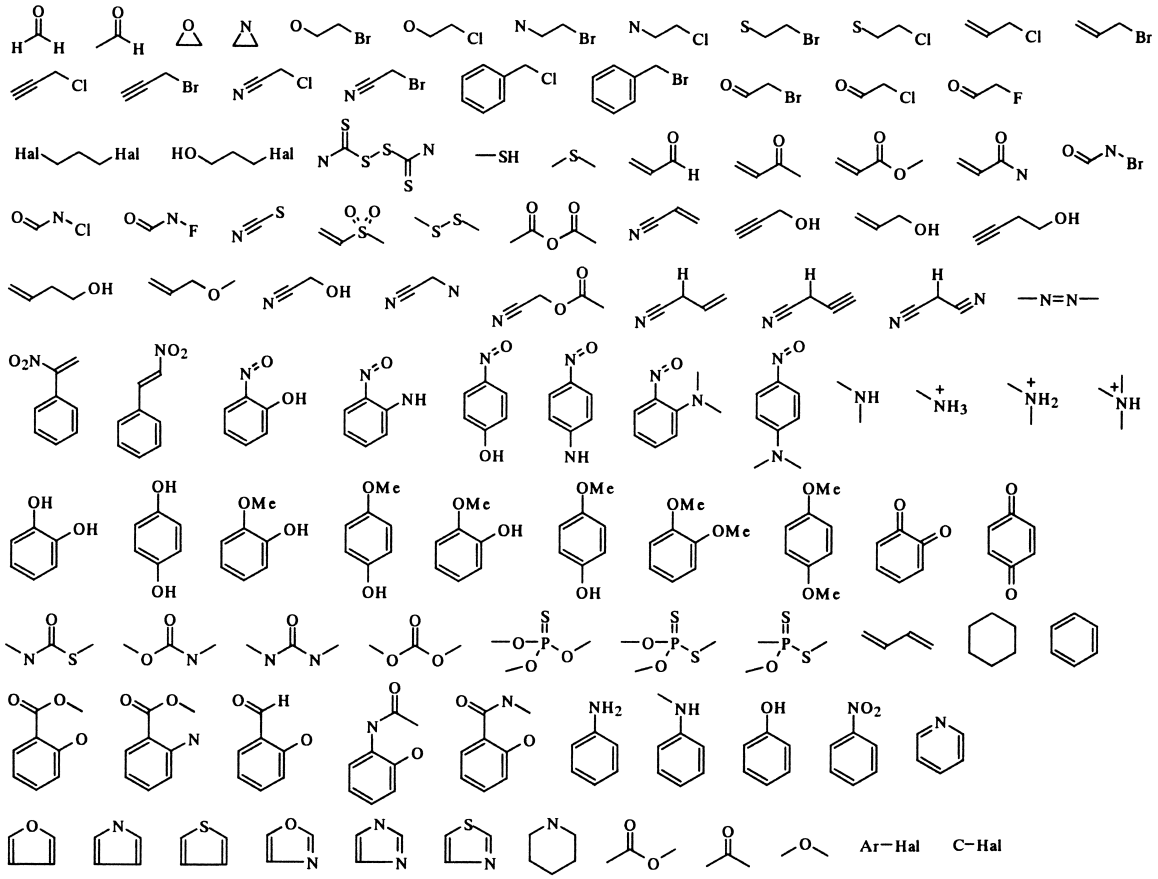


Fig. 1. Substructural fragments found in aquatic toxicants.

baseline QSAR (3) was created using only selected molecules from class 1, affording a much higher correlation. To afford a later comparison between the methods of regression analysis and neural network modeling, relationships were created using all of the training sets and tested with their associated training sets. The results for the modeling of the combined classes and for classes 1 to 3 individually are given in Table 3.

From these results, it is evident that a much better correlation was obtained when the classes of toxicant were modeled individually. For classes 1 to 3, both the correlation increased and the error decreased compared with the model for all the data. The correlation for the testing set is higher for classes 1 and 2, which suggests that while choosing the testing set we had identified molecules that fit the models better than those in the training set.

An artificial neural network was created for a non-linear multiple regression analysis with the training set of 384 molecules. It consisted of an input layer of 31 neurons, a single hidden layer of 15 neurons, and an output layer of a single neuron. Starting from the QSAR model relating only log $P$ to toxicity, 30 structural descriptors were added sequentially, and the correlation of the new relationship was observed to rise gradually from around 0.68 to 0.78. To reduce the possibility of a chance correlation, only the 16 structural descriptors with highest correlation were selected for further modeling. These descriptors were chosen to keep the highest correlation between the observed toxicity and the calculated toxicity for the testing set. The results of the trainings and the predictions are summarized in Table 4.

Using this approach, some very favorable results were produced, and for each of the training sets the correlation has improved over those produced using the previous model. However, for the testing set a lower correlation was obtained for classes 2 and 3, suggesting that over-training had occurred during creation of the neural net. Correlation plots for the training set and the testing set for the combined classes are shown in Fig. 2.

For a comparison we repeated this study using a multiple linear regression analysis; the following equation was produced for all molecules in the training set:

$$\begin{aligned}\log 1/LC_{50} = {} & 0.776 \log P + 0.130\ SS7 - 0.199\ SS13 \\ & + 0.224\ SS21 + 0.242\ SS\ 25 - 1.357\ SS\ 35 \\ & + 0.405\ SS\ 48 + 0.110\ SS\ 49 + 0.275\ SS\ 54 \\ & - 0.163\ SS\ 65 - 0.116\ SS\ 68 + 0.539\ SS\ 72 \\ & + 0.683\ SS\ 89 + 0.434\ SS\ 90 + 0.617\ SS\ 91 \\ & - 0.514\ SS\ 94 + 0.222\ SS\ 96 + 0.992, \qquad (4) \\ & n = 384,\ r = 0.788,\ s = 0.718.\end{aligned}$$

Applying this equation to predict for the testing set, a correlation of 0.791 and a standard error of 0.667 were achieved. This was, in fact, the best result obtained for modeling of the combined classes. For a comparison of this model with that obtained just using log $P$, the two correlation plots are shown in Fig. 3.

An examination of the substructures found to have the high-

Table 2. Compound Name, Chemical Class, Experimental Toxicity and Calculated Toxicity by Artificial Neural Network Model and Multiple Linear Regression Analysis

| No. | Compound Name | Class[a] | Toxicity[b] | NN[c] | MLR[d] | No. | Compound Name | Class[a] | Toxicity[b] | NN[c] | MLR[d] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | <Training Set> | | | | | 98 | 3,4-Dimethyl-1-pentyn-3-ol | 1 | 2.734 | 2.468 | 2.845 |
| 1 | Adamantane | 3 | 5.687 | 3.623 | 3.915 | 99 | 3.6-Dimethyl-1-heptyn-3-ol | 1 | 3.457 | 3.443 | 3.368 |
| 2 | Cyclohexane | 3 | 4.269 | 3.231 | 3.706 | 101 | 2,3-Dimethylvaleraldehyde | 3 | 3.854 | 3.872 | 3.910 |
| 3 | Hexane | 3 | 4.537 | 3.889 | 4.049 | 102 | 2,4,5-Trimethoxybenzaldehyde | 3 | 3.598 | 4.037 | 3.811 |
| 4 | 1,9-Decadiene | 1 | 5.678 | 4.603 | 4.419 | 104 | 2,4-Dihydroxybenzaldehyde | 3 | 4.023 | 4.409 | 3.384 |
| 5 | 2,3-Dimethyl-1,3-butadiene | 1 | 4.075 | 2.781 | 3.395 | 105 | 2,4-Dimethoxybenzaldehyde | 3 | 3.917 | 4.119 | 3.885 |
| 6 | 2,4-Hexadiene | 1 | 3.614 | 3.186 | 3.680 | 107 | 2-Chloro-6-fluorobenzaldehyde | 3 | 4.227 | 4.958 | 4.749 |
| 7 | 2,5-Dimethyl-2,4-hexadiene | 1 | 4.465 | 3.590 | 3.898 | 108 | 2-Methylbutyraldehyde | 3 | 3.936 | 3.389 | 3.380 |
| 8 | Norbornylene | 1 | 3.974 | 2.715 | 3.330 | 109 | 2-Methylvaleraldehyde | 3 | 3.727 | 3.679 | 3.668 |
| 9 | 1,1,1-Trichloroethane | 1 | 3.448 | 3.087 | 3.690 | 110 | 3,5-Dibromosalicylaldehyde | 3 | 5.518 | 5.400 | 5.290 |
| 10 | 1,1,2,2-Tetrachloroethane | 1 | 3.917 | 3.349 | 3.946 | 112 | 4,6-Dimethoxy-2-hydroxybenzaldehyde | 3 | 4.832 | 4.178 | 3.560 |
| 11 | 1,1,2-Trichloroethane | 1 | 3.213 | 3.244 | 3.774 | 113 | 4-(Diethylamino)benzaldehyde | 1 | 3.870 | 4.292 | 4.715 |
| 12 | 1,2-Dichloroethane | 1 | 2.862 | 3.199 | 3.601 | 114 | 4-(Diethylamino)salicylaldehyde | 3 | 4.557 | 4.482 | 4.395 |
| 13 | 1,2-Dichloropropane | 1 | 2.949 | 3.315 | 3.900 | 115 | 4-Chlorobenzaldehyde | 3 | 4.805 | 4.810 | 4.648 |
| 14 | 1,3-Dibromopropane | 3 | 4.980 | 4.372 | 3.696 | 116 | 4-(Hexyloxy)-M-anisaldehyde | 3 | 4.947 | 4.829 | 5.200 |
| 15 | 1,3-Dichloropropane | 1 | 3.002 | 3.204 | 3.639 | 117 | 5-Bromo-2-nitrovanillin | 3 | 3.576 | 3.758 | 4.806 |
| 16 | 1,4-Dichlorobutane | 1 | 3.391 | 3.365 | 3.966 | 119 | 5-Bromovanillin | 3 | 3.588 | 3.477 | 4.434 |
| 17 | 1,5-Dichloropentane | 1 | 3.746 | 3.695 | 4.251 | 120 | 5-Chlorosalicylaldehyde | 3 | 5.308 | 5.207 | 4.326 |
| 20 | 1-Bromohexane | 1 | 4.680 | 4.057 | 4.317 | 121 | 5-Hydroxy-2-nitrobenzaldehyde | 3 | 3.601 | 3.729 | 4.079 |
| 21 | 1-Bromooctane | 1 | 5.363 | 4.369 | 4.812 | 122 | $\alpha,\alpha,\alpha$-Trifluoro-m-tolualdehyde | 3 | 5.269 | 4.250 | 4.668 |
| 22 | 1-Bromopropane | 1 | 3.262 | 3.521 | 3.466 | 123 | Benzaldehyde | 3 | 4.017 | 3.667 | 4.031 |
| 24 | Hexachloroethane | 1 | 5.221 | 4.733 | 4.777 | 124 | Butanal | 3 | 3.650 | 2.824 | 2.970 |
| 26 | Trans-1,2-Dichlorocyclohexane | 1 | 3.920 | 4.138 | 4.505 | 125 | Hexanal | 3 | 3.745 | 3.567 | 3.547 |
| 27 | Chloroform | 1 | 3.228 | 2.918 | 3.379 | 127 | *o*-Fluorobenzaldehyde | 3 | 4.963 | 3.773 | 4.132 |
| 29 | 3-Chloro-2-chloromethyl-1-propene | 3 | 5.818 | 3.279 | 3.844 | 128 | *o*-Nitrobenzaldehyde | 3 | 4.016 | 3.055 | 4.403 |
| 30 | Hexachloro-1,3-butadiene | 3 | 6.462 | 4.865 | 4.823 | 129 | *o*-Tolualdehyde1 | 1 | 3.356 | 3.994 | 4.369 |
| 32 | Trichloroethylene | 1 | 3.474 | 3.421 | 4.028 | 130 | *o*-Vanillin | 3 | 4.802 | 3.659 | 3.635 |
| 34 | Diethyl ether | 1 | 1.462 | 2.529 | 2.471 | 131 | *p*-Dimethylaminobenzaldehyde | 1 | 3.514 | 3.861 | 4.223 |
| 35 | Isopropyl ether | 1 | 2.367 | 2.541 | 3.074 | 132 | *p*-Ethoxybenzaldehyde | 1 | 3.728 | 3.735 | 4.096 |
| 36 | Pentyl ether | 1 | 4.703 | 4.339 | 4.277 | 133 | *p*-Isopropyl benzaldehyde | 1 | 4.350 | 4.446 | 4.886 |
| 38 | 4-Nitrophenyl phenyl ether | 1 | 4.910 | 4.888 | 5.077 | 134 | *p*-Phenoxybenzaldehyde | 1 | 4.634 | 4.071 | 5.175 |
| 40 | 4-4′-Dihydroxydiphenyl ether | 1 | 4.542 | 4.739 | 4.094 | 135 | Pentafluorobenzaldehyde | 3 | 5.251 | 4.138 | 4.536 |
| 41 | *p*-Chlorophenyl-o-nitrophenyl ether | 1 | 5.114 | 6.160 | 5.600 | 136 | Salicylaldehyde | 3 | 4.725 | 4.178 | 3.708 |
| 42 | *p*-Fluorophenyl ether | 1 | 5.261 | 5.038 | 4.858 | 137 | Valeraldehyde | 3 | 3.825 | 3.240 | 3.259 |
| 43 | 2,2,5,5-Tetramethyltetrahydrofuran | 1 | 2.883 | 2.499 | 2.925 | 138 | Vanillin | 3 | 3.228 | 3.659 | 3.634 |
| 44 | 2,3-Benzofuran | 3 | 3.926 | 3.852 | 3.481 | 140 | 2′,3′,4′-Trichloroacetophenone | 1 | 5.048 | 5.120 | 4.922 |
| 47 | 2-(Bromomethyl)tetrahydro-2H-pyran | 1 | 2.941 | 3.499 | 3.333 | 141 | 2′,3′,4′-Trimethoxyacetophenone | 1 | 3.021 | 3.117 | 2.859 |
| 48 | 3-Furanmethanol | 3 | 2.286 | 1.485 | 1.823 | 143 | 2′-Hydroxy-4′-methoxyacetophenone | 1 | 3.367 | 3.649 | 2.683 |
| 49 | Cineole | 1 | 3.180 | 2.604 | 3.193 | 144 | 2-Amino-4′-chlorobenzophenone | 1 | 5.039 | 4.978 | 4.856 |
| 50 | Dibenzofuran | 3 | 4.967 | 4.767 | 4.594 | 145 | 2-Butanone | 1 | 1.350 | 2.493 | 2.696 |
| 51 | Flavone | 3 | 4.803 | 4.611 | 3.998 | 147 | 2-Dodecanone | 1 | 5.194 | 5.199 | 4.786 |
| 52 | Furan | 3 | 3.048 | 2.559 | 2.330 | 148 | 2-Heptanone | 1 | 2.940 | 2.997 | 3.561 |
| 55 | 1,1,1,3,3,3-Hexafluoro-2-propanol | 1 | 2.838 | 2.995 | 3.125 | 150 | 2-Nonanone | 1 | 3.971 | 4.035 | 4.122 |
| 56 | 1,1,1-Trichloro-2-methyl-2-propanol | 1 | 3.119 | 3.309 | 3.370 | 151 | 2-Octanone | 1 | 3.552 | 3.488 | 3.846 |
| 58 | 1-Chloro-2-propanol | 3 | 2.586 | 2.628 | 2.708 | 152 | 2-Pentanone | 1 | 1.842 | 2.512 | 2.986 |
| 59 | 1-Decanol | 1 | 4.189 | 4.385 | 4.226 | 153 | 2-Undecanone | 1 | 5.055 | 4.931 | 4.611 |
| 60 | 1-Heptanol | 1 | 3.527 | 3.492 | 3.397 | 154 | 3,3-Dimethyl-2-butanone | 1 | 3.061 | 3.039 | 3.589 |
| 61 | 1-Hexanol | 1 | 3.019 | 2.967 | 3.110 | 156 | 3-Pentanone | 1 | 1.748 | 2.581 | 3.155 |
| 63 | 1-Octanol | 1 | 3.975 | 3.909 | 3.682 | 157 | 4′-Chloro-3′-nitroacetophenone | 3 | 4.560 | 4.543 | 4.071 |
| 64 | 1-Pentanol | 1 | 2.271 | 2.426 | 2.821 | 158 | 4-Dimethylamino-3-methyl-2-butanone | 3 | 4.182 | 2.495 | 2.684 |
| 65 | 1-Propanol | 1 | 1.120 | 1.687 | 2.244 | 159 | 4-Methyl-2-pentanone | 1 | 2.282 | 2.627 | 3.226 |
| 66 | 2,2,2-Trichloroethanol | 1 | 2.699 | 2.851 | 3.012 | 160 | 5-Methyl-2-hexanone | 1 | 2.856 | 2.929 | 3.513 |
| 67 | 2,2,2-Trifluoroethanol | 3 | 2.925 | 1.791 | 2.363 | 161 | 5-Nonanone | 1 | 3.662 | 4.341 | 4.278 |
| 68 | 2,3-Dibromopropanol | 3 | 3.487 | 3.402 | 3.270 | 162 | 6-Methyl-5-hepten-2-one | 1 | 3.168 | 2.889 | 3.484 |
| 70 | 2-Butanol | 1 | 1.305 | 1.998 | 2.545 | 165 | $\beta$-Ionone | 1 | 4.577 | 4.664 | 4.453 |
| 72 | 2-Ethyl-1-hexanol | 1 | 3.664 | 3.916 | 3.687 | 166 | Benzophenone | 1 | 4.092 | 4.748 | 4.570 |
| 74 | 2-Methyl-2-propanol | 1 | 1.063 | 1.700 | 2.260 | 168 | 4-Aminopropiophenone | 2 | 3.009 | 3.457 | 3.190 |
| 75 | 2-Methyl-3,3,4,4-tetrafluoro-2-butanol | 1 | 2.439 | 2.399 | 2.806 | 169 | 1-Benzoylacetone | 3 | 5.169 | 3.854 | 3.377 |
| 76 | 2-Phenoxyethanol | 1 | 2.604 | 2.850 | 2.961 | 170 | 5,5-Dimethyl-1,3-cyclohexanedione | 1 | 1.086 | 2.734 | 3.351 |
| 77 | 2-Propanol | 1 | 0.832 | 1.657 | 2.204 | 171 | 2-Adamantanone | 1 | 3.393 | 3.120 | 3.640 |
| 78 | 3-Chloro-1-propanol | 1 | 2.072 | 2.542 | 2.445 | 174 | Rotenone | 3 | 5.243 | 5.602 | 4.333 |
| 79 | 3-Methyl-3-pentanol | 1 | 2.182 | 2.649 | 2.944 | 175 | [1(*R*)-endo]-(+)-3-bromocamphor | 1 | 3.528 | 4.349 | 4.770 |
| 80 | Benzyl-tert-butanol | 1 | 3.393 | 4.194 | 3.901 | 176 | (1*S*)-(−)-camphor | 1 | 3.952 | 3.981 | 4.095 |
| 81 | Cyclohexanol | 1 | 2.153 | 2.403 | 2.808 | 177 | 3,3-Dimethylglutaric acid | 1 | 1.054 | 1.860 | 2.301 |
| 82 | Exo-norborneol | 1 | 2.692 | 2.201 | 2.686 | 178 | Benzoic acid | 1 | 2.474 | 3.159 | 3.248 |
| 83 | trans-2-Phenyl-1-cyclohexanol | 1 | 3.599 | 4.483 | 4.102 | 179 | Hexanoic acid | 1 | 2.560 | 2.847 | 3.048 |
| 84 | [(1S)-endo]-(−)-Borneol | 1 | 3.386 | 3.812 | 3.607 | 181 | Propionic acid | 1 | 1.302 | 1.642 | 2.181 |
| 85 | 1,5-Hexadien-3-ol | 3 | 3.411 | 2.832 | 3.040 | 182 | Salicylic acid | 1 | 1.958 | 2.688 | 2.925 |
| 87 | 3-Hydroxy-3,7,11-trimethyl-1,6,10-dodecatriene | 1 | 5.192 | 4.593 | 4.597 | 183 | $\gamma$-Decanolactone | 1 | 3.976 | 4.870 | 4.377 |
| 88 | cis-3-Hexen-1-ol | 1 | 2.420 | 2.607 | 2.921 | 184 | 2-Ethoxyethyl acetate | 3 | 3.497 | 3.387 | 2.886 |
| 90 | 1,1-Diphenyl-2-propyn 1-ol | 1 | 4.273 | 4.882 | 4.517 | 185 | Butyl acetate | 3 | 3.810 | 4.334 | 3.635 |
| 92 | 1-Heptyn-3-ol | 3 | 4.804 | 2.895 | 3.073 | 186 | Dibutyl adipate | 3 | 4.851 | 5.556 | 5.291 |
| 93 | 1-Octyn-3-ol | 3 | 5.485 | 3.429 | 3.360 | 187 | Dibutyl fumarate | 3 | 5.546 | 5.417 | 5.204 |
| 95 | 2-Decyn-1-ol | 3 | 5.159 | 4.224 | 3.995 | 189 | Diethyl adipate | 3 | 3.998 | 4.563 | 4.046 |
| 96 | 2-Methyl-3-butyn-2-ol | 1 | 1.408 | 1.662 | 2.211 | 190 | Diethyl benzylmalonate | 3 | 4.664 | 5.411 | 4.609 |
| 97 | 2-Phenyl-3-butyn-2-ol | 1 | 3.112 | 3.352 | 3.385 | 191 | Diethyl chloromalonate | 3 | 5.311 | 4.582 | 3.664 |

Table 2. (continued)

| No. | Compound Name | Class[a] | Toxicity[b] | NN[c] | MLR[d] | No. | Compound Name | Class[a] | Toxicity[b] | NN[c] | MLR[d] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 192 | Diethyl malonate | 3 | 4.032 | 4.148 | 2.901 | 289 | 3,4-Dichloroaniline | 2 | 4.329 | 4.386 | 4.471 |
| 193 | Diethyl sebacate | 3 | 4.979 | 5.395 | 5.190 | 290 | 3-Benzyloxyaniline | 2 | 4.338 | 4.490 | 4.465 |
| 194 | Ethyl acetate | 3 | 2.583 | 3.469 | 3.005 | 291 | 4-Butylaniline | 2 | 4.165 | 4.681 | 4.427 |
| 196 | Ethyl salicylate | 3 | 3.914 | 5.104 | 3.930 | 292 | 4-Chloroaniline | 2 | 3.620 | 4.023 | 3.854 |
| 197 | Ethyl trifluoroacetate | 1 | 1.153 | 4.253 | 3.540 | 294 | 4-Ethylaniline | 2 | 3.220 | 3.894 | 3.864 |
| 198 | Ethyl *p*-aminobenzoate | 3 | 3.666 | 4.970 | 3.906 | 295 | 4-Fluoroaniline | 2 | 3.818 | 3.501 | 3.337 |
| 199 | Hexyl acetate | 3 | 4.516 | 5.524 | 4.213 | 296 | 4-Hexyloxyaniline | 2 | 4.807 | 4.773 | 4.490 |
| 200 | Methyl 2,5-dichlorobenzoate | 1 | 4.166 | 5.597 | 5.342 | 297 | 4-Toluidine | 2 | 2.826 | 3.628 | 3.576 |
| 201 | Methyl acetate | 3 | 2.314 | 3.410 | 2.866 | 298 | α,α,α-4-Tetrafluoro-*o*-toluidine | 2 | 3.782 | 4.035 | 3.978 |
| 203 | Methyl 4-chloro-2-nitrobenzoate | 1 | 3.891 | 5.737 | 5.103 | 299 | α,α,α-4-Tetrafluoro-*m*-toluidine | 2 | 3.775 | 4.035 | 3.978 |
| 205 | Methyl *p*-nitrobenzoate | 3 | 3.881 | 5.119 | 4.487 | 300 | Aniline | 2 | 2.936 | 3.468 | 3.236 |
| 206 | Phenyl 4-aminosalicylate | 3 | 4.677 | 5.461 | 4.685 | 301 | *p*-Bromoaniline | 2 | 3.559 | 3.918 | 4.036 |
| 207 | Phenyl salicylate | 3 | 5.259 | 6.158 | 5.008 | 303 | 4-Fluoro-*N*-methylaniline | 1 | 3.513 | 3.854 | 3.375 |
| 208 | Propyl acetate | 3 | 3.231 | 3.846 | 3.347 | 304 | 3-Chloro-o-formotoluidide | 1 | 3.561 | 4.299 | 4.314 |
| 209 | Dicumarol | 3 | 4.818 | 5.702 | 4.132 | 305 | *N*-Allylaniline | 1 | 3.569 | 3.933 | 3.810 |
| 211 | Di-*N*- butylorthophthalate | 1 | 5.455 | 6.179 | 6.045 | 307 | *N*-Methylaniline | 1 | 3.030 | 3.864 | 3.274 |
| 212 | Diethyl phthalate | 1 | 3.844 | 5.001 | 4.922 | 308 | Diphenylamine | 1 | 4.650 | 4.762 | 4.588 |
| 213 | Dimethyl aminoterephthalate | 3 | 4.369 | 5.297 | 4.297 | 309 | 2-(*N*-Ethyl-m-toluidino)ethanol | 1 | 3.530 | 3.595 | 3.539 |
| 214 | Dimethyl nitroterephthalate | 3 | 4.564 | 5.782 | 5.018 | 310 | *N,N*-Diethylaniline | 1 | 3.959 | 4.380 | 4.258 |
| 216 | *o*-Methoxybenzamide | 3 | 3.100 | 4.140 | 2.370 | 311 | *N,N*-Dimethylaniline | 1 | 3.268 | 3.916 | 3.774 |
| 217 | 2,2-Dichloroacetamide | 1 | 1.719 | 2.860 | 1.868 | 312 | *N,N*-Dimethyl-*p*-toluidine | 1 | 3.441 | 4.180 | 4.106 |
| 219 | 2,6-Dichlorobenzamide | 1 | 2.608 | 3.656 | 3.680 | 313 | *N*-Phenyldiethanolamine | 1 | 2.392 | 2.629 | 2.899 |
| 222 | *m*-Bromobenzamide | 1 | 3.334 | 4.215 | 3.244 | 314 | 2,9-Dithiadecane | 3 | 4.247 | 4.802 | 4.518 |
| 223 | *N,N*-Dimethyl-*m*-toluamide | 1 | 3.240 | 3.852 | 3.434 | 316 | 3,8-Dithiadecane | 3 | 4.469 | 4.758 | 4.441 |
| 224 | *N,N*-Dibutylformamide | 1 | 3.246 | 3.894 | 3.938 | 317 | 4,7-Dithiadecane | 3 | 4.375 | 4.949 | 4.754 |
| 225 | 4-Nitrobenzamide | 2 | 3.097 | 3.893 | 2.818 | 318 | 4,9-Dithiadodecane | 3 | 4.839 | 5.193 | 5.111 |
| 226 | *p*-(*tert*-Butyl)benzamide | 1 | 3.745 | 4.396 | 3.623 | 319 | *n*-Butyl sulphide | 3 | 4.611 | 5.171 | 4.580 |
| 227 | 2-Ethoxyethyl methacrylate | 3 | 3.757 | 4.224 | 3.361 | 321 | *tert*-Butyl sulphide | 3 | 3.701 | 4.691 | 4.047 |
| 229 | 2-Hydroxypropyl acrylate | 3 | 4.589 | 4.481 | 2.981 | 322 | Isopropyl disulphide | 3 | 4.257 | 4.982 | 4.376 |
| 230 | Allyl methacrylate | 3 | 5.105 | 4.285 | 3.663 | 324 | Propyl disulphide | 3 | 4.759 | 5.054 | 4.455 |
| 232 | Cyclohexyl acrylate | 3 | 5.018 | 4.739 | 4.272 | 325 | *t*-Butyl disulphide | 3 | 5.115 | 5.084 | 4.487 |
| 233 | Hexyl acrylate | 3 | 5.146 | 5.124 | 4.567 | 326 | Methyl sulphoxide | 1 | 0.361 | 2.455 | 0.276 |
| 234 | Isobutyl acrylate | 3 | 4.786 | 4.308 | 3.728 | 328 | Benzyl sulphoxide | 1 | 3.459 | 4.617 | 3.950 |
| 235 | Isopropyl methacrylate | 3 | 3.528 | 4.331 | 3.782 | 330 | 2,6-Dimethoxytoluene | 1 | 3.877 | 4.403 | 3.776 |
| 236 | Tetrahydrofurfuryl methacrylate | 3 | 3.691 | 4.229 | 3.399 | 331 | Acenaphthene | 1 | 4.950 | 4.911 | 4.571 |
| 237 | 1,6-Dicyanohexane | 1 | 2.412 | 3.063 | 2.937 | 333 | Ethylbenzene | 1 | 3.943 | 4.295 | 4.198 |
| 238 | 2-Amino-5-chlorobenzonitrile | 2 | 3.727 | 4.187 | 3.681 | 334 | Isopropylbenzene | 1 | 4.279 | 4.642 | 4.421 |
| 239 | 2-Chloro-6-methylbenzonitrile | 1 | 4.002 | 4.525 | 4.349 | 335 | *m*-Diethylbenzene | 1 | 4.510 | 5.193 | 4.722 |
| 240 | 3,5-Dibromo-4-hydroxybenzonitrile | 1 | 4.340 | 4.724 | 4.644 | 336 | *m*-Nitrotoluene | 2 | 3.729 | 4.151 | 4.291 |
| 242 | α,α,α-Trifluoro-*o*-tolunitrile | 1 | 3.608 | 3.481 | 4.037 | 338 | Nitrobenzene | 2 | 3.015 | 3.761 | 3.955 |
| 243 | Allyl cyanide | 3 | 2.567 | 3.384 | 2.477 | 339 | *p*-Dimethoxybenzene | 1 | 3.072 | 3.863 | 3.437 |
| 244 | Chloroacetonitrile | 3 | 4.748 | 4.373 | 2.560 | 340 | 1,4-Dinitrobenzene | 3 | 5.374 | 4.908 | 4.327 |
| 245 | Methyl 4-cyanobenzoate | 1 | 3.537 | 3.528 | 3.942 | 341 | *p*-Xylene | 1 | 4.078 | 4.364 | 4.247 |
| 247 | *N*-Undecyl cyanide | 1 | 5.625 | 5.518 | 4.648 | 342 | Toluene | 1 | 3.406 | 4.002 | 3.919 |
| 248 | *o*-Tolunitrile | 1 | 3.418 | 3.293 | 3.746 | 343 | *t*-Butylstyrene | 1 | 5.515 | 5.564 | 4.926 |
| 249 | 3,5-Diiodo-4-hydroxybenzonitrile | 1 | 4.737 | 4.702 | 4.646 | 344 | 2,4-Dinitrotoluene | 1 | 3.875 | 5.374 | 4.664 |
| 250 | 3,5-Dichloro-4-hydroxybenzonitrile | 1 | 3.889 | 4.371 | 4.313 | 345 | 1,2-Dibromobenzene | 1 | 4.765 | 5.223 | 5.108 |
| 251 | 1,2-Dimethylpropylamine | 1 | 2.487 | 2.940 | 2.924 | 348 | 1,3-Dichloro-4,6-dinitrobenzene | 3 | 6.712 | 6.980 | 5.547 |
| 252 | 1,8-Diamino-*p*-menthane | 3 | 3.416 | 3.192 | 2.735 | 349 | 1,3-Dichlorobenzene | 1 | 4.263 | 5.103 | 4.800 |
| 253 | 1-Adamantanamine | 3 | 3.782 | 3.000 | 2.991 | 350 | 1,4-dichlorobenzene | 1 | 4.730 | 5.103 | 4.800 |
| 256 | 3,3-Dimethylbutylamine | 1 | 2.226 | 3.208 | 3.175 | 351 | 1-Fluoro-4-nitrobenzene | 1 | 3.696 | 3.857 | 4.056 |
| 257 | Amylamine | 1 | 2.692 | 2.925 | 2.907 | 352 | 3,4-Dichlorotoluene | 1 | 4.743 | 5.595 | 5.097 |
| 258 | Butylamine | 1 | 2.436 | 2.755 | 2.618 | 353 | α,α′-Dichloro-*p*-xylene | 3 | 6.652 | 5.286 | 4.922 |
| 259 | Hexylamine | 1 | 3.252 | 3.235 | 3.195 | 354 | α,α-2,6-Tetrachlorotoluene | 1 | 5.375 | 5.957 | 5.540 |
| 260 | *N*-Decylamine | 1 | 5.184 | 4.914 | 4.333 | 355 | 1-Chloro-3-nitrobenzene | 2 | 3.923 | 4.981 | 4.569 |
| 262 | Nonylamine | 1 | 4.822 | 4.553 | 4.055 | 356 | *p*-Chloromethyl styrene | 3 | 5.692 | 5.198 | 4.705 |
| 263 | Octylamine | 1 | 4.396 | 4.126 | 3.771 | 358 | Chlorobenzene | 1 | 3.823 | 4.520 | 4.197 |
| 265 | tert-Octylamine | 1 | 3.720 | 3.739 | 3.532 | 359 | 1-Naphthol | 2 | 4.493 | 4.314 | 3.980 |
| 266 | (+/−)-*sec*-Butylamine | 1 | 2.425 | 2.759 | 2.630 | 360 | 2,4-Dimethylphenol | 2 | 3.867 | 4.244 | 3.933 |
| 267 | Benzylamine | 1 | 3.021 | 3.451 | 3.164 | 361 | 2,4-Dinitrophenol | 3 | 4.209 | 4.622 | 4.004 |
| 268 | *N*-Undecylamine | 3 | 5.912 | 5.210 | 4.595 | 362 | 2,5-Dinitrophenol | 3 | 4.739 | 4.622 | 4.004 |
| 269 | Di-N-hexylamine | 1 | 5.376 | 4.944 | 4.619 | 364 | 2-Nitrophenol | 2 | 2.939 | 4.031 | 3.633 |
| 270 | Diethylamine | 1 | 1.932 | 2.563 | 2.312 | 365 | 2-Phenylphenol | 2 | 4.442 | 4.943 | 4.566 |
| 271 | *N*-Ethylbenzylamine | 1 | 3.374 | 3.852 | 3.488 | 366 | 2-*sec*-Butyl-4,6-dinitrophenol | 3 | 5.637 | 5.711 | 5.128 |
| 272 | *N*-(3-Methoxypropyl)-3,4,5-Trimethoxybenzylamine | 3 | 3.297 | 3.404 | 3.046 | 367 | 3-Methoxyphenol | 2 | 3.225 | 3.853 | 3.188 |
| 273 | 2-(Diisopropylamino)ethanol | 1 | 2.859 | 2.285 | 2.739 | 370 | 4-Ethylphenol | 2 | 4.070 | 4.166 | 3.883 |
| 274 | 3-Dimethylaminopropyl chloride·HCl(ignored Cl) | 3 | 3.075 | 3.303 | 2.344 | 371 | 4-Methoxyphenol | 2 | 3.053 | 3.853 | 3.188 |
| 275 | 5-Dimethylamino-2-pentanone | 1 | 2.670 | 2.531 | 2.457 | 372 | *p*-Cresol | 2 | 3.817 | 3.694 | 3.599 |
| 276 | *N,N*-Bis(2,2-Diethoxyethyl)methylamine | 1 | 2.617 | 2.547 | 3.087 | 374 | 4-Propylphenol | 2 | 4.093 | 4.553 | 4.159 |
| 278 | *N,N*-Diethylethanolamine | 1 | 1.818 | 1.614 | 2.137 | 375 | *o*-Cresol | 2 | 3.888 | 3.694 | 3.599 |
| 279 | Tripropylamine | 1 | 3.449 | 3.483 | 3.843 | 376 | *p*-Phenoxyphenol | 2 | 4.575 | 4.730 | 4.403 |
| 280 | *N,N*-Dimethylbenzylamine | 1 | 3.554 | 3.853 | 3.502 | 377 | *p*-Phenylazophenol | 3 | 5.229 | 5.264 | 4.887 |
| 282 | 2,3,4-Trichloroaniline | 2 | 4.732 | 4.882 | 5.080 | 378 | *p*-*tert*-Butylphenol | 2 | 4.465 | 4.792 | 4.400 |
| 284 | 2,4-Dinitroaniline | 3 | 4.091 | 4.268 | 3.979 | 379 | *p*-*tert*-Pentylphenol | 2 | 4.802 | 4.944 | 4.628 |
| 285 | 2,6-Diisopropylaniline | 2 | 4.062 | 5.294 | 4.893 | 380 | Phenol | 2 | 3.443 | 3.175 | 3.260 |
| 286 | 2-Chloro-4-methylaniline | 2 | 3.596 | 4.286 | 4.193 | 381 | 2,4,6-Trimethylphenol | 1 | 4.020 | 4.659 | 4.254 |
| 288 | 2-Chloroaniline | 2 | 4.347 | 4.023 | 3.854 | 383 | 2,6-Dinitrophenol | 3 | 3.666 | 4.262 | 3.515 |

Table 2. (continued)

| No. | Compound Name | Class[a)] | Toxicity[b)] | NN[c)] | MLR[d)] |
|---|---|---|---|---|---|
| 384 | 2,2′-Methylenebis(4-chlorophenol) | 2 | 5.939 | 6.389 | 5.413 |
| 385 | 2,3,4,5-Tetrachlorophenol | 3 | 5.753 | 5.491 | 5.618 |
| 386 | 2,4,6-Tribromophenol | 2 | 4.704 | 5.247 | 5.475 |
| 388 | 2,4,6-Triiodophenol | 2 | 5.550 | 5.108 | 5.002 |
| 389 | 2,4-Dichlorophenol | 2 | 4.323 | 4.908 | 4.487 |
| 390 | 2-Chlorophenol | 2 | 4.045 | 4.464 | 3.877 |
| 391 | 4,5-Dichlorocatechol | 3 | 5.303 | 5.014 | 4.169 |
| 392 | 4,5-Dichloroguaiacol | 2 | 4.635 | 4.448 | 4.420 |
| 395 | Pentabromophenol | 3 | 6.721 | 6.538 | 6.128 |
| 396 | Tetrachlorocatechol | 3 | 5.290 | 5.107 | 5.347 |
| 397 | 2,3,4,6-Tetrachlorophenol | 3 | 5.352 | 5.491 | 5.618 |
| 398 | 4-Chlorophenol | 2 | 4.323 | 4.464 | 3.877 |
| 399 | 3-Trifluoromethyl-4-nitrophenol | 2 | 4.355 | 4.774 | 4.269 |
| 400 | 1,4-Diazabicyclo[2,2,2]octane | 3 | 1.812 | 2.654 | 1.802 |
| 401 | 1-Benzylpiperazine | 3 | 3.570 | 3.890 | 3.094 |
| 402 | 2-Amino-4-chloro-6-methylpyrimidine | 2 | 3.007 | 3.402 | 3.061 |
| 403 | 1,2-Bis(4-pyridyl)ethane | 1 | 3.086 | 3.464 | 3.022 |
| 404 | 2,6-Diphenylpyridine | 1 | 6.042 | 5.792 | 5.046 |
| 405 | 2,6-Pyridinedicarboxylic acid | 1 | 2.715 | 2.599 | 2.245 |
| 406 | 2-Amino-5-bromopyridine | 1 | 2.990 | 3.068 | 3.297 |
| 408 | 2-Chloro-3-pyridinol | 1 | 2.319 | 2.052 | 2.901 |
| 409 | 2-Cyanopyridine | 2 | 2.157 | 2.390 | 2.454 |
| 410 | 2-Dimethylaminopyridine | 1 | 2.983 | 2.560 | 3.037 |
| 411 | 2-Picoline | 1 | 2.016 | 1.991 | 2.460 |
| 412 | 3-(3-Pyridyl)-1-propanol | 1 | 2.961 | 2.258 | 2.301 |
| 413 | 3-Cyano-4,6-dimethyl-2-hydroxypyridine | 1 | 2.975 | 2.690 | 2.785 |
| 416 | 4-Acetylpyridine | 2 | 2.858 | 2.195 | 1.831 |
| 417 | 4-Benzoylpyridine | 1 | 3.250 | 3.083 | 3.355 |
| 419 | 4-Phenylpyridine | 1 | 3.984 | 3.602 | 3.689 |
| 420 | 4-Picoline | 1 | 2.364 | 2.111 | 2.676 |
| 421 | 5-Chloro-2-pyridinol | 2 | 2.056 | 2.354 | 3.140 |
| 422 | 5-Ethyl-2-methylpyridine | 1 | 3.174 | 2.642 | 3.089 |
| 423 | 6-Chloro-2-picoline | 1 | 2.740 | 2.039 | 3.349 |
| 424 | 6-Chloro-2-pyridinol | 2 | 2.782 | 2.684 | 3.409 |
| 425 | Pentachloropyridine | 3 | 5.728 | 5.396 | 5.789 |
| 426 | Piperidine | 3 | 4.561 | 2.557 | 2.339 |
| 427 | Pyridine | 2 | 2.899 | 1.974 | 2.336 |
| 428 | 3-Picoline | 1 | 2.810 | 2.111 | 2.676 |
| 429 | 3-Ethylpyridine | 1 | 2.413 | 2.449 | 2.964 |
| 431 | Hexamethylenetetramine | 1 | 0.449 | 2.644 | 3.249 |
| 432 | 1-(2-Chloroethyl)pyrrolidine·HCl.(Ignored Cl anion) | 3 | 3.046 | 3.398 | 2.675 |
| 434 | 2,4,5-Tribromoimidazole | 3 | 4.683 | 4.950 | 4.646 |
| 435 | 2-Acetyl-1-methylpyrrole | 3 | 2.895 | 2.580 | 2.235 |
| 436 | 2-Methylimidazole | 3 | 2.458 | 2.629 | 1.988 |
| 437 | *N*-Vinylcarbazole | 1 | 4.781 | 4.829 | 4.669 |
| 438 | Pyrrole | 1 | 2.504 | 2.491 | 2.731 |
| 440 | 2-Butanone oxime | 1 | 2.014 | 2.870 | 3.059 |
| 441 | 2,4,5-Trimethyloxazole | 3 | 2.394 | 2.642 | 1.903 |
| 442 | 2,6-Dimethylmorpholine | 3 | 2.474 | 2.595 | 2.166 |
| 444 | 5-Chloro-2-mercaptobenzothiazole | 3 | 4.798 | 4.690 | 4.380 |
| 445 | Saccharin | 1 | 1.050 | 1.114 | 1.050 |
| 446 | 3-Bromothiophene | 3 | 4.421 | 3.732 | 3.908 |
| 447 | 2'-(Octyloxy)-acetanilide | 3 | 5.767 | 5.001 | 4.963 |
| 448 | 2-Acetamidophenol | 3 | 3.739 | 2.075 | 2.855 |
| 449 | 3-Acetamidophenol | 1 | 2.126 | 2.075 | 2.855 |
| 450 | 4-Acetamidophenol | 1 | 2.269 | 2.075 | 2.855 |
| 452 | Diuron | 1 | 4.215 | 3.904 | 4.577 |
| 453 | Propanil | 1 | 4.404 | 4.074 | 4.869 |
| 454 | Salicylanilide | 1 | 4.732 | 4.707 | 4.374 |
| 455 | Diethyl benzylphosphonate | 1 | 2.832 | 3.883 | 3.686 |
| 457 | Triphenylphosphate | 1 | 5.574 | 5.038 | 5.382 |
| 458 | Triphenylphospine oxide | 1 | 3.715 | 4.785 | 4.403 |
| 459 | Tris(2-butoxyethyl)phosphate | 1 | 4.551 | 4.845 | 4.559 |
| 460 | *O*-ethyl *O*-*p*-nitrophenyl phenylphosphonothioate | 3 | 6.613 | 5.763 | 5.514 |
| 461 | 2-Chloroethyl-*N*-cyclohexyl carbamate | 3 | 3.769 | 2.983 | 3.715 |
| 462 | Baygon | 3 | 4.376 | 4.518 | 4.618 |
| 464 | Lannate | 3 | 4.886 | 4.768 | 3.809 |
| 465 | *p*-(*tert*-Butyl)-phenyl-*N*-methylcarbamate | 1 | 4.317 | 4.727 | 5.411 |
| 466 | Carbofuran | 3 | 5.419 | 4.522 | 4.412 |
| 467 | Aminocarb | 3 | 4.976 | 4.542 | 4.782 |
| 468 | Bromacil | 3 | 3.143 | 2.948 | 2.889 |
| 469 | Diazinon | 1 | 4.513 | 5.334 | 4.882 |
| 470 | Disulphoton | 1 | 4.839 | 5.216 | 4.626 |
| 471 | Dursban | 1 | 6.002 | 6.132 | 5.517 |
| 473 | Strychnine | 3 | 5.490 | 3.950 | 2.232 |
| 475 | Pentobarbital | 1 | 3.700 | 3.539 | 3.753 |
| 476 | Phenobarbital | 1 | 2.681 | 3.169 | 3.664 |
| 477 | Secobarbital | 1 | 4.043 | 3.754 | 3.884 |
| 478 | Thiopental | 1 | 4.004 | 4.208 | 4.282 |
| | <Testing Set> | | | | |
| 18 | 1-Bromobutane | 1 | 3.572 | 3.634 | 3.753 |
| 19 | 1-Bromoheptane | 1 | 5.086 | 4.245 | 4.580 |
| 23 | Dichloromethane | 1 | 2.411 | 2.959 | 3.043 |
| 25 | Pentachloroethane | 1 | 4.429 | 3.889 | 4.371 |
| 28 | 3,4-Dichloro-1-butene | 3 | 4.180 | 3.606 | 4.187 |
| 31 | Tetrachloroethylene | 1 | 3.993 | 3.736 | 4.278 |
| 33 | Butyl ether | 1 | 3.606 | 3.269 | 3.728 |
| 37 | *tert*-Butyl methyl ether | 1 | 2.118 | 2.231 | 2.690 |
| 39 | Phenyl ether | 1 | 4.629 | 4.858 | 4.700 |
| 45 | 2,3-Dihydrobenzofuran | 1 | 3.168 | 3.853 | 3.387 |
| 46 | 2,5-Dimethyl furan | 3 | 3.131 | 2.509 | 2.578 |
| 53 | Tetrahydrofuran | 1 | 1.524 | 2.586 | 2.210 |
| 54 | (1*R*,2*S*,5*R*)-(−)-menthol | 1 | 3.917 | 3.746 | 3.978 |
| 57 | 1-Butanol | 1 | 1.632 | 1.982 | 2.533 |
| 62 | 1-Nonanol | 1 | 4.403 | 4.190 | 3.954 |
| 69 | 2,4-Dimethyl-3-pentanol | 1 | 2.853 | 3.612 | 3.471 |
| 71 | 2-Chloroethanol | 3 | 3.152 | 2.536 | 2.407 |
| 73 | 2-Methyl-propanol | 1 | 1.715 | 1.998 | 2.545 |
| 86 | 1-Hexen-3-ol | 3 | 3.518 | 2.992 | 3.123 |
| 89 | *trans*-3-Hexen-1-ol | 1 | 2.568 | 2.607 | 2.921 |
| 91 | 1-Ethynyl-cyclohexanol | 1 | 2.686 | 2.415 | 2.815 |
| 94 | 2-Butyn-1-ol | 3 | 3.841 | 1.717 | 2.281 |
| 100 | 3-Methyl-1-pentyn-3-ol | 1 | 1.906 | 2.007 | 2.552 |
| 103 | 2,4-Dichlorobenzaldehyde | 3 | 4.988 | 5.670 | 5.259 |
| 106 | 2-Chloro-5-nitrobenzaldehyde | 3 | 4.680 | 4.373 | 5.020 |
| 111 | 3-Ethoxy-4-hydroxybenzaldehyde | 3 | 3.278 | 4.225 | 3.773 |
| 118 | 5-Bromosalicylaldehyde | 3 | 5.189 | 4.985 | 4.507 |
| 126 | Isovaleraldehyde | 3 | 4.423 | 3.177 | 3.211 |
| 139 | 4-Nitrobenzaldehyde | 3 | 4.175 | 3.055 | 4.403 |
| 142 | 2′,4′-Dichloroacetophenone | 1 | 4.208 | 4.659 | 4.315 |
| 146 | 2-Decanone | 1 | 4.504 | 4.538 | 4.383 |
| 149 | 2-Hexanone | 1 | 2.369 | 2.664 | 3.274 |
| 155 | 3-Methyl-2-butanone | 1 | 1.999 | 2.556 | 3.107 |
| 163 | α-Bromo-2′,5′-dimethoxyacetophenone | 3 | 6.581 | 4.212 | 3.634 |
| 164 | Acetophenone | 1 | 2.870 | 3.893 | 3.080 |
| 167 | *m*-Aminoacetophenone | 2 | 2.549 | 3.404 | 2.732 |
| 172 | 2-Methyl-1,4-naphthoquinone | 3 | 6.194 | 3.882 | 3.146 |
| 173 | Cyclohexanone | 1 | 2.193 | 2.542 | 3.076 |
| 180 | Nonanoic acid | 1 | 3.182 | 4.144 | 3.901 |
| 188 | Dibutyl succinate | 3 | 4.713 | 4.783 | 4.726 |
| 195 | Ethyl hexanoate | 3 | 4.210 | 4.512 | 4.053 |
| 202 | Methyl 2,4-dihydroxybenzoate | 3 | 3.565 | 5.154 | 3.468 |
| 204 | Methyl *p*-chlorobenzoate | 1 | 4.191 | 5.218 | 4.732 |
| 210 | *tert*-Butyl acetate | 1 | 2.551 | 3.870 | 3.364 |
| 215 | Diphenyl phthalate | 3 | 6.600 | 6.552 | 6.570 |
| 218 | 2,4-Dichlorobenzamide | 1 | 3.298 | 3.656 | 3.680 |
| 220 | Anthranilamide | 3 | 2.537 | 3.838 | 2.094 |
| 221 | Benzamide | 1 | 2.263 | 3.586 | 2.443 |
| 228 | 2-Hydroxyethyl methacrylate | 3 | 2.758 | 4.287 | 2.792 |
| 231 | Benzyl methacrylate | 3 | 4.577 | 5.440 | 4.648 |
| 241 | α,α,α-Trifluoro-*m*-tolunitrile | 1 | 3.555 | 3.481 | 4.037 |
| 246 | *N*-Octyl cyanide | 1 | 4.423 | 4.465 | 4.049 |
| 254 | 1-Methyl heptylamine | 3 | 4.396 | 4.146 | 3.784 |
| 255 | 2,2-Dimethy-1-propylamine | 1 | 2.264 | 3.001 | 2.991 |
| 261 | *N*-Heptylamine | 1 | 3.723 | 3.660 | 3.483 |
| 264 | Propylamine | 1 | 2.283 | 2.694 | 2.329 |
| 277 | *N*,*N*-Diethylcyclohexylamine | 1 | 3.861 | 3.268 | 3.727 |
| 281 | 2,3,4,5,6-Pentafluoroaniline | 2 | 3.693 | 3.767 | 3.743 |
| 283 | 2,3,5,6-Tetrachloroaniline | 3 | 5.932 | 5.483 | 5.670 |
| 287 | 2-Chloro-4-nitroaniline | 2 | 3.932 | 4.674 | 4.226 |
| 293 | 4-Ethoxy-2-nitroaniline | 2 | 3.846 | 4.129 | 3.673 |
| 302 | *p*-Nitroaniline | 2 | 3.043 | 4.083 | 3.608 |
| 306 | *N*-Ethyl-*m*-toluidine | 1 | 3.436 | 3.961 | 3.860 |
| 315 | 3,6-Dithiaoctane | 3 | 3.397 | 4.582 | 4.074 |
| 320 | *n*-Propyl sulphide | 3 | 3.736 | 4.663 | 4.015 |
| 323 | Phenyl disulphide | 3 | 5.298 | 4.714 | 5.665 |
| 327 | Phenyl sulphoxide | 1 | 3.365 | 4.609 | 4.027 |
| 329 | 1,2,4-Trimethylbenzene | 1 | 4.192 | 4.876 | 4.552 |
| 332 | Amylbenzene | 1 | 4.938 | 5.494 | 4.887 |
| 337 | Naphthalene | 1 | 4.320 | 4.433 | 4.293 |
| 346 | 1,2-Dichlorobenzene | 1 | 4.191 | 5.103 | 4.800 |
| 347 | 1,3,5-Trichloro-2,4-dinitrobenzene | 3 | 6.087 | 7.737 | 6.125 |
| 357 | 2-Fluorotoluene | 1 | 3.753 | 4.087 | 4.019 |

Table 2. (continued)

| No. | Compound Name | Class[a] | Toxicity[b] | NN[c] | MLR[d] | No. | Compound Name | Class[a] | Toxicity[b] | NN[c] | MLR[d] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 363 | 2-Allylphenol | 2 | 3.952 | 4.358 | 4.010 | 418 | 4-Bromophenyl 3-pyridyl 3-pyridyl ketone | 1 | 4.109 | 3.934 | 4.147 |
| 368 | 4,6-Dinitro-*o*-cresol | 3 | 5.055 | 5.021 | 4.343 | 430 | Quinoline | 1 | 3.220 | 3.259 | 3.485 |
| 369 | 4-Amino-2-nitrophenol | 2 | 3.629 | 4.679 | 3.284 | 433 | Tolazoline hydrochloride.(Ignored Cl anion) | 3 | 2.745 | 3.433 | 3.063 |
| 373 | 4-Nitrophenol | 2 | 3.484 | 4.031 | 3.633 | 439 | 3-Methylindoole | 1 | 4.171 | 3.908 | 3.756 |
| 382 | 2,3,6-Trimethylphenol | 1 | 4.220 | 4.659 | 4.254 | 443 | 4-Methyloxazole | 3 | 1.777 | 2.620 | 2.035 |
| 387 | 2,4,6-Trichlorophenol | 2 | 4.555 | 5.140 | 5.078 | 451 | Alachlor | 1 | 4.732 | 5.951 | 5.030 |
| 393 | 4-Chloro-3-methyl phenol | 2 | 4.397 | 4.888 | 4.210 | 456 | Tributylphosphate | 1 | 4.444 | 5.199 | 4.786 |
| 394 | 4-Chlorocatechol | 3 | 4.961 | 4.257 | 3.555 | 463 | Carbaryl | 3 | 4.356 | 4.594 | 4.977 |
| 407 | 2-Bromo-3-pyridinol | 1 | 2.569 | 3.420 | 3.102 | 472 | Malathion | 3 | 4.370 | 4.640 | 4.537 |
| 414 | 3-Hydroxy-2-nitropyridine | 2 | 2.924 | 1.708 | 2.893 | 474 | Amobarbital | 1 | 3.423 | 3.539 | 3.753 |
| 415 | 3-Pyridinecarboxaldehyde | 3 | 3.815 | 3.056 | 2.782 | | | | | | |

a) Class 1: nonpolar narcotics, class 2: polar narcotics; class 3: excess toxicants.
b) Data taken from M. Nendza, C.L. Russom, QSAR modelling of the ERL-D *fathead minnow* acute toxicity database, *Xenobiotica*, **21**, 147–170(1991). 96 hour $LC_{50}$ was reported as mmol/L but converted to mol/L by adding 3 and to log ($1/LC_{50}$)
c) By artificial neural network model using log *P* and 16 structural descriptors.
d) By multiple linear regression model (Eq. 4).

Table 3. Results of Regression Analysis Using log *P* as the Only Molecular Descriptor

| Data Set | Training Set[a] | | | Testing Set | | |
|---|---|---|---|---|---|---|
| | *n* | *r* | *s* | *n* | *r* | *s* |
| All | 384 | 0.687 | 0.829 | 94 | 0.673 | 0.806 |
| Class 1 | 202 | 0.876 | 0.549 | 50 | 0.905 | 0.413 |
| Class 2 | 50 | 0.827 | 0.450 | 11 | 0.851 | 0.327 |
| Class 3 | 132 | 0.773 | 0.640 | 30 | 0.733 | 0.793 |

a) *n*, the number of samples; *r*, correlation coefficient; *s*, standard error.

Table 4. Results of Neural Network Modeling Using log *P* and 16 Structural Descriptors

| Data Set | Training Set[a] | | | Testing Set | | |
|---|---|---|---|---|---|---|
| | *n* | *r* | *s* | *n* | *r* | *s* |
| All | 384 | 0.819 | 0.665 | 94 | 0.737 | 0.832 |
| Class 1 | 202 | 0.920 | 0.450 | 50 | 0.849 | 0.586 |
| Class 2 | 50 | 0.956 | 0.229 | 11 | 0.617 | 0.655 |
| Class 3 | 132 | 0.917 | 0.389 | 30 | 0.647 | 1.013 |

a) *n*, the number of samples; *r*, correlation coefficient; *s*, standard error.

est correlation with aquatic toxicity is interesting from a mechanistic viewpoint, and confirms our ideas about which reactive substructures are responsible for toxicity. These substructures are listed in Table 5. This analysis clearly demonstrates that reactive groups, such as aldehydes and esters, contribute to toxicity, as do hydrophobic groups, such as the benzene ring.

## Conclusions and Future Work

The potential that neural networks have for predicting aquatic toxicants has been demonstrated. For creating structure–activity relationships, similar results can be obtained using both neural networks and multiple linear regression analysis. Using either modeling system, better prediction results were obtained for individual classes than when all of the classes were combined. The inclusion of molecular descriptors related to the molecular structure significantly improved the structure-activity relationships over the standard baseline model, which generally use log *P* as the only variable. Therefore, to create a model for the automated prediction of aquatic toxicity, a neural network system should be used for classification, and either neural networks or linear regression analysis should be used to create individual QSAR's for each class. The success of the classifying the toxicants into 3 groups is indicative that classification into smaller groups is possible. Therefore, the prediction ability of the system could be improved by further dividing class 3 (reactive toxicants) into sub-classes based upon the classes of the reactive toxicants identified in this study (Table 1). The structure-activity relationships identified in the analysis of aquatic toxicants will be used to construct a knowledge base for the expert system ChemTox,[27] which is currently under development in the authors's group.

Table 5. The Structural Descriptors Used in the Creation of Equation 4

| Substructures positively contributing to toxicity | Substructures negatively contributing to toxicity |
|---|---|
| SS7: Benzene ring | SS13: -CN |
| SS21: -Br | SS35: $-NHSO_2-$ |
| SS25: -Cl | SS65: Pyridine ring |
| SS48: $-NO_2$ | SS68: -OH |
| SS49: $-OCH_3$ | SS94: $-CONH_2$ |
| SS54: $-OCOCH_3$ | |
| SS72: -S- | |
| SS89: -CHO | |
| SS90: -CONH- | |
| SS91: -COO- | |
| SS96: $-NH_2$ | |

We gratefully acknowledge the European Union, the Japan Society for the Promotion of Science and Japan Chemical Industry Association for funding this research.

## References

1 M. T. D. Cronin and J. C. Dearden, *Quant. Struct. -Act. Relat.*, **14**, 1 (1995).
2 H. Könemann, *Toxicology*, **19**, 209 (1981).
3 C. Hansch and A. Leo, "Exploring QSAR, (1) Fundamentals and Applications in Chemistry and Biology," ACS Professional Reference Book (1995).

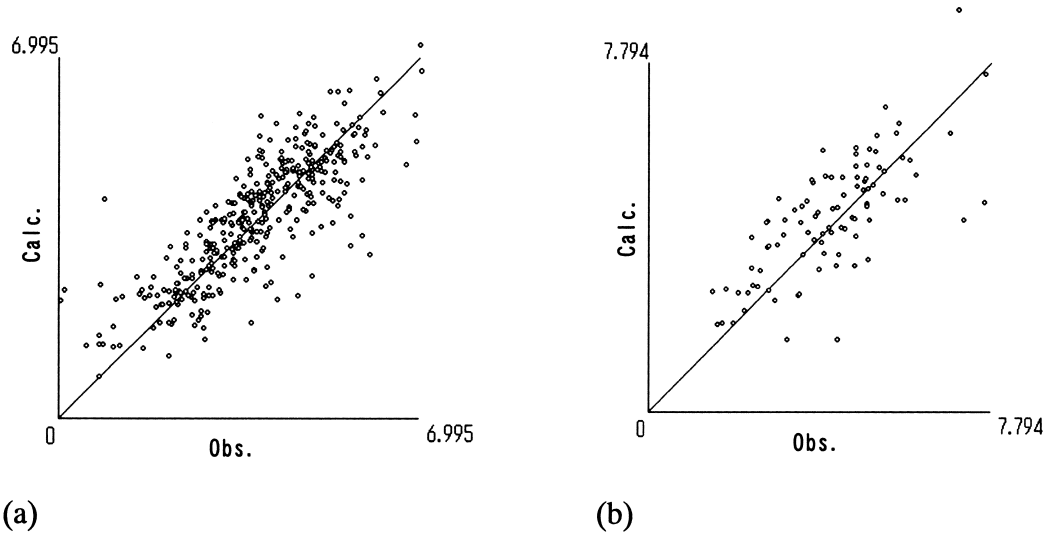


Fig. 2. Correlation plots for observed toxicity versus calculated toxicity by the neural network model based on log *P* and 16 structural descriptors. (a) is for the combined training set of 384 molecules. (b) is for the combined testing set of 94 molecules.

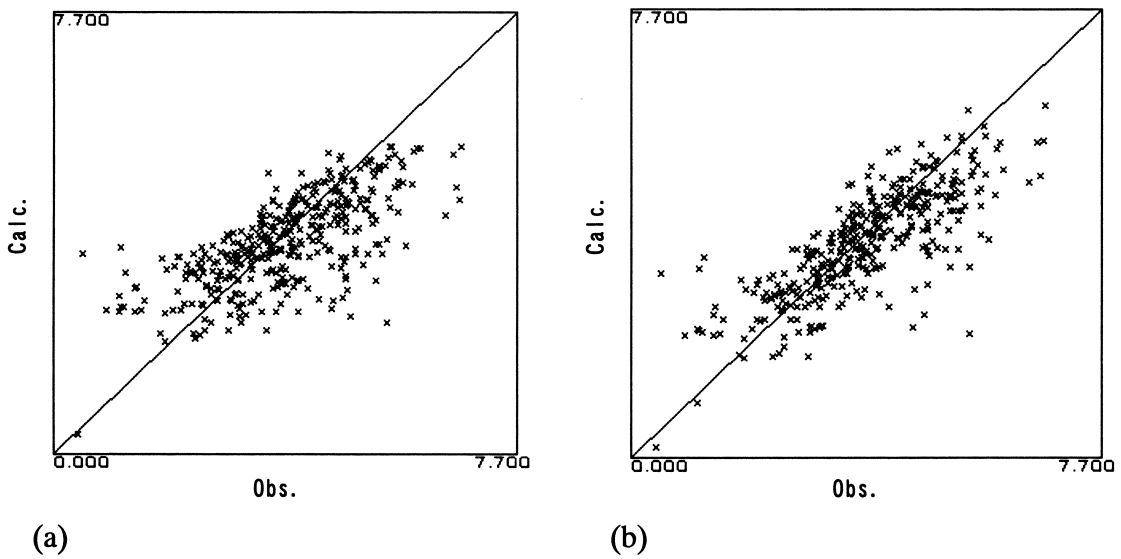


Fig. 3. Correlation plots for QSARs of the combined training set of 384 molecules. (a) is for using only log *P*. (b) is for using log *P* and 16 structural descriptors.

4 G. D. Veith, D. J. Call, and L. T. Brooke, *Can. J. Fish. Aquat. Sci.*, **40**, 743 (1983).

5 R. L. Lipnick, K. R. Watson, and A. K. Strausz, *Xenobiotica*, **17**, 1011 (1987).

6 C. J. van Leeuwen, P. T. J. van der Zandt, T. Aldenberg, H. J. M. Verhaar, and J. L. M. Hermens, *Environ. Tox. Chem.*, **11**, 267 (1992).

7 A. P. van Wezel and A. Opperhuizen, *Crit. Rev. Toxicol.*, **25**, 255 (1995).

8 G. D. Veith and S. J. Broderius, "QSAR in Environmental Toxicology-II," ed by K. L. E. Kaiser, D. Reidel, Dordrecht (1987), p. 385; G. D. Veith and S. J. Broderius, *Environ. Health Persp.*, **87**, 207 (1990).

9 S. P. Bradbury, R. W. Carlson, G. J. Niemi, and T. R. Henry, *Environ. Toxicol. Chem.*, **10**, 115 (1991).

10 J. L. M. Hermens, *Environ. Health. Persp.*, **87**, 219 (1990).

11 J. de Bruijn and J. Hermens, *Aquat. Tox.*, **24**, 257 (1993).

12 R. L. Lipnick, *Sci. Tot. Environ.*, **109/110**, 131 (1991).

13 J. W. Deneer, W. Seinen, and J. L. M. Hermens, *Aquat. Toxicol.*, **12**, 185 (1988).

14 J. Hermens, F. Busser, P. Leeuwanch, and A. Musch, *Toxicol. Environ. Chem.*, **9**, 219 (1985).

15 J. W. Deneer, T. L. Sinnige, W. Seinen, and J. L. M. Hermens, *Aquat. Toxicol.*, **13**, 195 (1988).

16 O. G. Mekenyan and G. D. Veith, *SAR QSAR Environ. Res.*, **1**, 335 (1994).

17 H. J. M. Verhaar, C. J. van Leeuwen, and J. L. M. Hermens, *Chemosphere*, **25**, 471 (1992).

18 J. Hermens, *Sci. Tot. Environ.*, **171**, 235 (1995).

19 H. Jäckel and M. Nendza, *Aquat. Toxicol.*, **29**, 304 (1994).

20 D. V. Eldred, C. L. Weikel, P. C. Jurs and K. L. E. Kaiser, *Chem. Res. Toxicol.*, **12**, 670 (1999).

21 M. Nendza and C.L. Russom, *Xenobiotica*, **21**, 147 (1991).
22 Tutors/Procalc "Laboratory for Molecular Information Systems", Test Version 1.3, Toyohashi University of Technology.
23 SSCheck, "Laboratory for Molecular Information Systems", 1996 version, Toyohashi University of Technology.
24 T. Aoyama, Y. Suzuki, and H. Ichikawa, *J. Med. Chem*., **33**, 2583 (1990).
25 R. D. Brown and Y.C. Martin, J. *Chem. Inf. Comput. Sci*., **37**, 1 (1997).
26 G. D. Veith, R. L. Lipnick, and C. L. Russom, *Xenobiotica*, **19**, 555 (1989).
27 Y. Takahashi, S. Ban, G. Admans, H. Kato, H. Abe, and S. Hanai, Abstracts of Papers in Second Indo-US workshop on Mathematical Chemistry, University of Minnesota Duluth, Minnesota, (2000), p. 65.