Environmental Mutagens & Carcinogens 21-2 : 113-117 (2001)

Primary Screening of QSAR Molecular Descriptors for Genotoxicity Prediction of Drinking Water Disinfection Byproducts (DBPs), Chlorinated Aliphatic Compounds

Jae-Hyoun Kim*, Jin-nam Jo, Byung-Suk Jin, Dong-Soo Lee, Ki-Tae Kim and Ae-Son Om

**Department of Health Science, School of Natural Science, Dongduk Women's University, Wolgogdong, Sungbuk-gu, Seoul, Korea*

Department of Data Information, School of Information Science, Dongduk Women's University

Department of Applied Chemistry, School of Natural Science, Dongduk Women's University

Graduate School of Environmental Studies, Seoul National University

SeouLin Bioscience Institute, SeouLin Bldg. 452-2, Songnae-Dong, Seoul 134-030, Korea

Department of Food and Nutrition, School of Human Ecology, Hanyang University

(Received October 20, 2001 / Accepted December 10, 2001)

ABSTRACT : **The screening of various molecular descriptors for predicting carcinogenic, mutagenic and teratogenic activities of chlorinated aliphatic compounds as drinking water disinfection byproducts (DBPs) has been investigated for the application of quantitative structure-activity relationships (QSAR). The present work embodies the study of relationship between molecular descriptors and toxicity parameters of the genotoxicity endpoints for the screening of relevant molecular descriptors. The toxicity Indices for 29 compounds constituting the testing set were computed by the PASS program and active values were chosen. We investigate feasibility of screening descriptors and of their applications among different genotoxic endpoints. The correlation to teratogenicity of all 29 compounds was significantly improved when the same analysis was done with 20 alkanes only without alkene compounds. The HOMO (highest occupied molecular orbital) energy and number of Cl parameters were dominantly contributed.**

Keywords : disinfection byproducts (DBPs), genotoxicity, prediction, quantitative structure-activity relationship (QSAR)

Introduction

Agents for drinking water disinfection were used for killing or disable pathogenic microorganisms and reducing the risk of waterborne illnesses (Ford, 1999). An effective disinfectant works quickly to reduce microorganism concentrations below the threshold that is toxic to humans, and provides protection against regrowth throughout the distribution system (National Research Council, 1987; Morris *et al.*, 1992). Common disinfectants include chlorine, chlorine dioxide, and monochloramine. These chemicals may react with organics in the water supply to produce disinfection by-products (DBPs) such as trihalomethanes (THM), haloacetic acids, haloacetonitriles, haloketones, chlorite, and chlorate (Bull, 1991). The public health benefits of disinfection are significant and well-recognized. However, these disinfection practices pose health risks of their own. Although disinfectants such as chlorine, hypochlorites, and chlorine dioxide are effective in controlling many harmful microorganisms, they react with organic and inorganic matter in the water to form disinfection byproducts (DBPs), which pose health risks at certain levels. There is widespread potential for human exposure to disinfection byproducts (DBPs) in drinking water because everyone drinks, bathes, cooks, and cleans with water (U.S. EPA, 1994a; U.S. EPA, 1994b; Boorman *et al.*, 1999; Shin *et al.*, 1999; Symons, 2001).

Quantitative structure-activity relationships (QSAR) represent an attempt to correlate structural or property descriptors of compounds with activities. These physicochemical descriptors, which include parameters to account for hydrophobicity, topology, electronic properties, and steric effects, are determined empirically or, more recently, by computational methods. Activities used in QSAR include chemical measurements and biological assays. QSAR currently are being applied in many disciplines, with many pertaining to drug design and environmental risk assessment. Since a biological effect is the result of interaction between a chemical and a target molecule in an organism, the chemical

*To whom correspondence should be addressed

properties or structure of the chemical can determine the type of effect or degree of toxicity. If this relationship can be illustrated using measured or calculated parameters, a QSAR model can be derived (Klopman and Rosenkranz, 1994; Barratt, 1995; Polloth and Mangelsdorf, 1997).

Health risk assessment will continue to be based upon animal toxicity data, since it is impossible to obtain toxicity data from humans experimentally (Cronin and Schultz, 2001). From this aspect, QSAR methodology is a cost-effective tool for toxicity prediction to allow for hazard identification, setting of testing priorities, and providing scientific support for decisions. This database of DBP properties and QSAR resources provides core information for DBP analysis. Therefore, the DBPs that pose the greatest toxicological risks can be identified, thus allowing prioritization of testing resources to the DBPs that pose the greatest risks. Currently, a variety of databases and commercial software offers toxicity prediction for a single endpoint. However, the water industry is responsible for evaluating effects of various chemicals (i.e. trihalomethanes, haloacetic acids, haloketones, etc.) across a variety of endpoints (Blaha *et al.*, 1998).

The importance of hydrophobicity, frontier orbital (HOMO and LUMO) energies and steric factors as physical descriptors of mutagenicity is emphasized with a variety of compounds. Some possible connections between QSAR models and the general electrophilic theory of genotoxic activity are discussed (Tuppurainen, 1999). Quantitative structure-activity relationships (QSAR) represent an attempt to correlate structural or property descriptors of compounds with genotoxicities (Debnath *et al.*, 1994). QSAR predictions with respect to complex endpoints such as carcinogenicity, chronic toxicity and teratogenicity are still unsatisfactory (Polloth and Mangelsdorf, 1997). QSAR approach has been applied also in prediction of mutagenicity of aromatic (Crebelli *et al.*, 1992) and quinolines (Smith *et al.*, 1997) from structure.

There are no clear findings or even controversial relating to correlationships between toxicity endpoints and molecular descriptors. Many attemps have been made to use QSAR with respect to estimating the carcinogenic or mutagenic risk of chlorinated aliphatic compounds. Quantitative structure-toxicity relationships for 80 chlorinated compounds using quantum chemical descriptors have been throughy examined (Sixt *et al.*, 1995). The QSAR analysis indicated that toxic effects induced by 24 chlorinated aliphatics in *Aspergillus nidulans* are mainly dependent on steric factors, as indicated by the correlation with molar refractivity (MR) and conversely by LUMO (Crebelli *et al.*, 1992). Molecular volume is the key descriptor, with corrections using topological and one electrostatic descriptor to account for features that increase the solubility of the molucules (Huibers and Katritzky, 1998). A preliminary QSBR (quantitative structure biodegradability relationship) model for terminally substituted mono- and dihalogenated alkanes utilized three types of descriptors: hydrophobicity (logP), steric (MW, IX), and electronic (LUMO, total energy, heat of formation) (Damborsky, 1996). A negative linear dependence of the log transformed reaction rate constants on the calculated activation energies was obtained (Verhaar *et al.*, 1996). The linear regression of QSAR using multiple descriptors does not show good predictive properties on the validation set for 10 halogenated aliphatic compounds (Rorije *et al.*, 1997). The toxicity of the tested compounds was influenced by various parameters, such as lipophilicity (logP), electron donor ability (charge) and longest carbon-chlorine (LBC-Cl) bond length. In addition, steric parameters, such as molar refractivity (MR) and LBC-Cl, and electronic parameters, such as LUMO (indicating electrophilicity), were predominant factors discriminating genotoxins from non-genotoxins in the presence but not in the absence of S9 mix (Tafazoli *et al.*, 1998).

There are two objectives to such a screening method, firstly to provide toxicity insight by its molecular structure whether or not there exists a molecular substructure pattern among different genotoxicity endpoints, and secondly to save time and money to predict a toxicity effectively in QSAR study because too many molecular descriptors are used to predict a toxicity of one compound. The substructure pattern is regarded as to constitute a kind of fingerprint of a molecule. Since this screening approach also provides an overall spectrum of genotoxic activities and physicochemical properties, structure-activity relationships can be used quickly through primary screening steps, thus shortening the optimization process.

Materials and Methods

The 29 chlorinated aliphatic compounds in this study were chosen for the screening approach. A total of 14 descriptors was calculated to encompass the relevant physicochemical properties of the compounds. The ISIS/DRAW software (MDL Information Systems, Inc.) is used to draw structures, export to molfiles with structures of 25 compounds and tested with PASS program to get the genotoxic endpoints, such as carcinogenicity, mutagenicity, teratogenicity and alkylation. Second-order valence connectivity index, Kp, LogP (logKow), Energy (Hartree-Fock method), surface area, dipole moment, boiling point, vapor pressure, number of Cl, MW (molecular weight) indices were calculated

Table 1. Physicochemical genotoxic endpoints of descriptors utilized in this study

(1) Physicochemical properties(QSAR descriptors; X)
X1 = LogP (logKow)
X2 = Energy (kJ/mole, Hartfree-Fock theory)
X3 = Valence Connectivity index (χ^2: 2nd order)
X4 = Surface area (cm^2)
X5 = Dipole moment (Modified Del Re)
X6 = HOMO (Winmopec-AM1 calculation)
X7 = LUMO (Winmopac-AM1 calculation)
X8 = Solubility (mol/L)
X9 = Boiling point (BP, degrees °C)
X10 = Vapor pressure (VP, atm, 25 °C)
X11 = Number of chlorine (Cl)
X12 = MW (molecular weight)
(2) Genotoxic Endpoints (Genotoxicity; Y)
Y1 = LD50 (oral-rat toxicity: mg/kg)
Y2 = Logarithm of the Embryotoxicity (logX (%))
Y3 = Logarithm of the Carcinogenicity (logX)
Y4 = Logarithm of the Mutagenicity (logX)
Y5 = Logarithm of the Teratogenicity (logX)
Y6 = Logarithm of the Alkylator (logX)

using the Molpro software (ChemSW Inc.). LD_{50} (oral-rat toxicity: mg/kg) value was taken from the Thomes database (Thomes Plus database, 1997, Micromedex Inc. U.S.A) or MSDS (Material Safety Data Sheet). The Winmopac 3.0 (Fujitsu Limited) software is used for calculation of HOMO and LUMO energy values. A full listing of descriptors is given in Table 1 and their data is described (Table 2). The data set was analysed using the SAS statistical software (SAS Institute Inc.) for Windows.

Statistical Analysis

A QSAR study of genotoxic endpoints as dependent variables of 29 chlorinated aliphatic compounds has been developed using only calculated structural features as independent variables. Multiple linear regression are utilized for model building. A study that performed on both 29 chlorinated aliphatics, and 20 aliphatic alkanes excluding chlorinated aliphatic alkenes among all 29 compounds.

Multiple linear regression analyses of molecular descriptors and the logarithm of the genotoxic endpoints were carriet out. The descriptors are initially screened for significance and correlation to limit the number of descriptors considered.

Results and Discussion

Multiple regression models calculated from genotoxic endpoints of 29 chlorinated aliphatic compounds

In this study, we examined simultaneous use of 12 QSAR descriptors (or independent variables), X_1, X_2, ···. X_{12} to find linear equations to predict each of Genotoxity (Y), that is, Y_1, Y_2, ··· Y_6. The stepwise addition procedure is then applied to derive the best multiparameter linear correlation equations from the set of descriptors to predict

Table 2. Pearson Correlation Coefficient Between 6 dependent variables and 12 independent variables with 29 observations

	Log(1/LD50)	Embryotoxicity	Carcinogenicity	Mutagenicity	Teratogenicity	Alkylator
1. LogP	–0.0078	0.1915	–0.0254	–0.0494	0.3413	–0.0118
	(0.970)	(0.320)	(0.896)	(0.799)	(0.070)	(0.952)
2. TE	–0.0204	–0.0921	–0.2843	–0.1898	0.2262	–0.0671
	(0.916)	(0.635)	(0.135)	(0.324)	(0.238)	(0.729)
3. V-C	0.0167	0.1849	–0.0652	–0.01866	0.4046	0.0315
	(0.932)	(0.337)	(0.737)	(0.333)	(0.030)	(0.871)
4. SA	0.2090	0.2547	0.1380	0.1393	0.3368	0.2265
	(0.277)	(0.182)	(0.475)	(0.471)	(0.074)	(0.237)
5. DM	–0.0583	–0.1613	–0.1901	–0.0979	0.0632	0.0062
	(0.764)	(0.403)	(0.323)	(0.614)	(0.745)	(0.975)
6. HOMO	–0.1000	0.0826	0.2630	0.2462	–0.3358	–0.2098
	(0.660)	(0.670)	(0.168)	(0.198)	(0.075)	(0.275)
7. LUMO	0.0269	–0.0311	–0.0230	0.1938	–0.0230	0.1416
	(0.890)	(0.873)	(0.906)	(0.314)	(0.906)	(0.464)
8. Solubility	–0.1261	0.0290	0.0206	0.0174	0.0975	0.0555
	(0.514)	(0.882)	(0.915)	(0.929)	(0.615)	(0.775)
9. BP	0.2814	0.3020	0.2486	0.1741	0.3474	0.3290
	(0.139)	(0.111)	(0.193)	(0.366)	(0.065)	(0.081)
10. VP	–0.2368	0.0356	0.0739	0.0841	0.0261	–0.2047
	(0.216)	(0.855)	(0.703)	(0.665)	(0.893)	(0.287)
11. N of Cl	0.0749	0.3588	0.1956	0.0034	0.4922	0.1723
	(0.699)	(0.056)	(0.309)	(0.986)	(0.007)	(0.371)
12. MW	0.1164	0.3062	0.1392	–0.0152	0.4566	0.1989
	(0.548)	(0.106)	(0.471)	(0.937)	(0.013)	(0.301)

p-value was indicated in the parenthesis.

genotoxity.

Table 2 shows that Pearson correlation coefficient between 6 dependent variables and 12 independent variables with 29 observations.

QSAR data were obtained for 29 chlorinated aliphatic compounds. Using the data, we obtain the correlation coefficients between 12 descriptors and 6 genotoxity (Y) variables. From the correlation coefficient table (Table 2), log (1/LD50), Carcinogenic (logX), and Mutagenic (logX) are not correlated with any of the 12 descriptors. Alkylator (logX) is positively correlated with BP, whose correlation coefficient value is 0.329 with p-value of 0.081, but not with the others at the significant level of =0.1. Embriotoxic (logX) is correlated with 2 descriptors, Number of CI and MW at the significant level of =0.1, where each value of correlation coefficients is 0.3588 and 0.3062, respectively. Teratogenic (logX) are positively correlated with 6 descriptors among 12 descriptors with the values of correlation coefficients between 0.34 and 0.49, and is negatively correlated with HOMO. It was found that BP was closely associated with log(1/LD50) ($p = 0.139$), Embryotoxic ($p = 0.111$), Teratogenic ($p = 0.065$) and alkylator ($p = 0.081$) with relatively lower p-values. Especially, BP and number of Cl have proven useful for prediction of teratogenicity with p-values of 0.065 and 0.007, respectively.

Linear regression models with R were obtained between molecular descriptors and genotoxic endpoints, at the $\alpha = 0.1$ level.

$$\text{Embryotoxi c(Y2)} = 1.8029 + 0.0193(\text{Cl}) \quad (R = 0.1287)$$
$$(0.0001) \quad (0.0560) \qquad \text{(Eq.1)}$$

$$\text{Carcinogenic (Y3)} = 1.7754 - 0.000017\ (\text{H-F Energy}) \quad (R^2 = 0.0808)$$
$$(0.0001) \quad (0.1350) \qquad \text{(Eq.2)}$$

$$\text{Teratogenic (Y5)} = 1.213 - 1.0487\ (\text{HOMO}) + 0.0539\ (\text{Cl}) \quad (R^2 = 0.3023)$$
$$(0.0002) \quad (0.1465) \quad (0.0133) \qquad \text{(Eq.3)}$$

$$\text{Alkylator (Y6)} = 0.6215 + 0.0040\ (\text{BP}) \quad (R^2 = 0.1083) \quad \text{(Eq.4)}$$
$$(0.0037) \quad (0.0814)$$

None of the significant linear equations are found to predict log (1/LD50), Carcinogenicity (logX), and Mutagenicity (logX), satisfying the condition that the descriptors left in the model are significant at the 0.1 level. Only one descriptor, number of Cl is adopted to predict the Embriotoxicity (logY) among 12 descriptors, satisfying the significant level of $\alpha = 0.1$, and its determination coefficient (R^2) for this model is 0.1287, indicating that Number of Cl explains just about 13% of the variability in Embriotoxicity (logY). BP is also selected to predict Alkylator (logY). The Rvalue of 0.108 for this model is very low. Therefore, the final models for prediction of Embriotoxicity (logY), and Alkylator (logY) are quite poor. The model obtained by the stepwise regression for predictintg Teratogenicity (logX) include two descriptors, that is, Number of Cl and HOMO. Their p-values are 0.013, and 0.146, respectively. Hence, Number of Cl is very significant but HOMO has not much influence on Teratogenic (logX). Therefore, coefficient of determination for this model is 0.302, determining 30% of the total variation in Y that can be accounted by the knowledge of Number of Cl and HOMO. Although this model is not good to predict Teratogenic (logX), it was evident that the model for predicting Teratogenic (logX) is relatively better than any other models for Genotoxity (Y). In conclusion, we can not find the best models to predict 6 genotoxic endpoints (Y).

Multiple regression models calculated from genotoxic endpoints of 20 chlorinated aliphatic alkanes.

The determination coefficient (R^2) for Embryotoxicity, Carcinogenicity, Teratogenicity, and Alkylation activity with all 29 compounds was 0.1287, 0.0808, 0.3023 and 0.1083, respectively (Eq.1~Eq.4). However, it is shown that the determination coefficient of Y2, Y3 and Y5 improves from 0.1287 to 0.1775, from 0.0808 to 0.3750, and from 0.3023 to 0.4649, respectively (Eq. 5~Eq.7) when the same method is run on 20 alkanes only excluding alkene compounds. The Number of Cl was dominantly contributed to the Embryotoxicity, BP and VP to the Carcinogenicity, and VP and Number of Cl to the Teratogenicity, respectively.

$$\text{Embryotoxicity (Y2)} = 1.798 + 0.0195\ (\text{No of Cl}) \quad (R = 0.1775)$$
$$(0.0001) \quad (0.0643) \qquad \text{(Eq.5)}$$

$$\text{Carcinogenicity (Y3)} = 1.641 + 0.0015\ (\text{BP}) + 0.000055\ (\text{VP}) \quad (R = 0.3750)$$
$$(0.0001) \quad (0.0544) \quad (0.0387) \qquad \text{(Eq.6)}$$

$$\text{Teratogenicity (Y5)} = 1.616 - 0.000065\ (\text{VP}) + 0.0667\ (\text{No of Cl}) \quad (R = 0.4649)$$
$$(0.0001) \quad (0.0250) \quad (0.0020) \qquad \text{(Eq.7)}$$

Conclusions

The screening method can allow us to get prioritizing molecular properties before the QSAR analysis, thus, saving time and money. We have applied 14 molecular descriptors to QSAR technique. Steric and electrostatic fields and thermodynamic energy were found to be the relevent descriptors for structure activity relationships. Linear regression models with R were obtained between molecular descriptors and genotoxic endpoints, at the $\alpha = 0.1$ level. While no significant results were found between dependent and independent variables, we observed that the boiling point

(BP) is shown as unique molecular variable with low *p*-values for Teratogenicity and an increasing factor for other endpoints. Hence thermodynamic descriptors may be necessary to develop the new QSAR models.

The statistical treatment of homogeneous group as 20 alkane compounds without 9 alkene compounds allowed us to get higher values of determination coefficient (R) than those tested with all 29 compounds.

From the results obtained from the study, it was concluded that use of additional thermodynamic descriptors might improve the QSAR prediction of genotoxicity of chlorinated aliphatic alkanes in future study.

References

Barratt, M.D., Castell, J.V., Chamberlain, M., Combes, R.D., Dearden, J.C., Fentem, J.H., Gerner, I., Giuliano, A., Gray, T.J.B, Livingstone, D.J., Provan, W.M., Rutten, F.A.J.J.L., Verhaar, H.J.M. and Zbinden, P. (1995): The integrated use of alternative approaches for predicting toxic hazard. ATLA **23**, 410-429.

Blaha, L., Damborsky, J. and Nemec M. (1998): QSAR for acute toxicity of saturated and unsaturated halogenated aliphatic compounds. *Chemosphere*, **36**, 1345-1365.

Boorman, G.A., Dellarco, V., Dunnick, J.K., Chapin,R.E., Hunter, S., Hauchman, F. Gardner, H., Cox, M., and Sills, R.C. (1999): Drinking water disinfection byproducts: review and approach to toxicity evaluation. *Environ. Health Perspect.*, **107**(Suppl. 1), 207-217.

Bull, R.J., and Kopfler, F.C. (1991): Health Effects of Disinfectants and Disinfection By-products. Prepared for the American Water Works Association Research Foundation, Denver, Colo. U.S.A.

Crebelli, R., Andreoli, C., Carere, A., Conti, G., Conti, L., Cotta Ramusino, M. and Benigni, R. (1992): The induction of mitotic chromosome malsegregation in *Aspergillus nidulans*. Quantitative structure activity relationship (OSAR) analysis with chlorinated aliphatic hydrocarbons. *Mutat. Res.*, **266**(2), 117-34.

Cronin, M.T.D. and Schultz, T.W. (2001): Development of quantitative structure-activity relationships for the toxicity of aromatic compounds to *Tetrahymena pyriformis*: comparative assessment of the methodologies. *Chem. Res. Toxicol.*, **14**, 1284-1295.

Damborsky, J. (1996): A Mechanistic Approach to deriving quantitative structure-activity relationship models for microbial degradation of organic compounds. *SAR and QSAR in Environ. Res.*, **5**, 27-36.

Debnath, A.K., Shuterman, A.J., Lopez de Compadre, R.L. and Hansch, C. (1994): The importance of the hydrophobicity interaction in the mutagenicity of organic compounds. *Mutat. Res.*, **305**, 63-72.

Ford, T.E. (1999): Microbiological Safety of Drinking Water: United States and Global Perspectives. *Environ. Health Perspect.*, **107**(Suppl. 1), 191-206.

Huibers, P.D.T. and Katritzky, A.R. (1998): Correlation of the Aqueous Solubility of Hydrocarbons and Halogenated Hydrocarbons with Molecular Structure. *J. Chem. Inf. Comput. Sci.*, **38**, 283-292.

Klopman, G. and Rosenkranz, H.S. (1994): International commission for protection against environmental mutagens and carcinogens. approaches to SAR in carcinogenesis and mutagenesis. Prediction of carcinogenicity/mutagenicity using MULTI-CASE. *Mutat. Res.*, **305**(1), 33-46.

Morris, R.D., Audet, A.M., Angelillo, I.F., Chalmers, T.C. and Mosteller, F. (1992): Chlorination, Chlorination By-products, and Cancer: A Meta-analysis. *Am. J. Pub. Health.*, **82**(7), 955-963.

National Research Council (1987): Drinking water and health (Vol. 7)-chemistry and toxicity of selected disinfectants and by-products. National Academy Press, 2101 Constitution Ave., N.W., Washington, DC 20418, 80-189.

Polloth, C. and Mangelsdorf, I. (1997): Commentary on the application of (Q)SAR to the toxicological evaluation of existing chemicals. *Chemosphere*, **35**(11), 2525-42.

Rorije, E., L., Eriksson, H.J., Verboom, H.J.M., Verhaar, J.L.M., Hermens, and Peijnenburg, W.J.G.M. (1997): Predicting reductive transformation rates of halogenated aliphatic compounds using different QSAR approaches. *Environ. Sci. Poll. Res.*, **4**, 47-54.

Shin, D., Chung, Y., Choi, Y., Kim, J., Park, Y. and Kum, H. (1999): Assessment of disinfection by-products in drinking water in Korea. *J. Expo. Anal. Environ. Epidemiol.*, **9**(3), 192-9.

Sixt, S., Altschuh, J. and Bruggemann, R. (1995): Quantitative structure-toxicity relationships for 80 chlorinated compounds using quantum chemical descriptors. *Chemosphere*, **30**, 2397-2414.

Smith, C.J., Hansch, C. and Morton, M.J (1997): QSAR treatment of multiple toxicities: the mutagenicity and cytotoxicity of quinolines. *Mutat. Res.*, **379**(2), 167-75.

Symons, J.M. (2001): The Early History of Disinfection By-Products: A Personal Chronicle, *Environ. Engineer*, **37**(Part I), 20-26.

Tafazoli, M., Baeten, A., Geerlings, P. and Kirsch-Volders, M. (1998): *In vitro* mutagenicity and genotoxicity study of a number of short-chain chlorinated hydrocarbons using the micronucleus test and the alkaline single cell gel electrophoresis technique (Comet assay) in human lymphocytes: a structure-activity relationship (QSAR) analysis of the genotoxic and cytotoxic potential. *Mutagenesis*, **13**(2), 115-26.

Tomes Plus database (1977): Micromedex Inc. U.S.A.

Tuppurainen, K. (1999): Frontier orbital energies, hydrophobicity and steric factors as physical QSAR descriptors of molecular mutagenicity. A review with a case study: MX compounds. *Chemosphere*, **38**(13), 3015-30.

U.S. EPA. (1994a): National Primary Drinking Water Regulations; Disinfectants and Disinfection Byproducts; Proposed Rule. FR, **59**, 145: 38668.

U.S. EPA. (1994b): Workshop Report and Recommendations for Conducting Epidemiologic Research on Cancer and Exposure to Chlorinated Drinking Water. U.S. EPA, July, 19-21.

Verhaar, H.J.M., Rorije, E., Borkent, H., Seinen, W., and Hermens, J.L.M. (1996): Modelling the nucleophilic reactivity of small organochlorine electrophiles. A mechanistically-based Quantitative Structure-Activity Relationship. *Environ. Toxicol. Chem.*, **15**, 1011-1018.