



ELSEVIER

Journal of Chromatography A, 849 (1999) 621–627

JOURNAL OF
CHROMATOGRAPHY A

Short communication

Quantitative structure–retention relationships of polychlorinated naphthalenes in gas chromatography

Jesus Olivero^{a,b,*}, Kurunthachalam Kannan^b

^aGroup of Environmental Chemistry, Universidad de Cartagena, A.A. 6541, Cartagena, Colombia

^bInstitute of Environmental Toxicology, National Food Safety and Toxicology Center, Michigan State University, East Lansing, MI 48824, USA

Received 2 November 1998; received in revised form 3 March 1999; accepted 17 March 1999

Abstract

A quantitative structure–retention relationship (QSRR) model has been developed for the gas chromatographic relative retention indices (RRis) of 62 polychlorinated naphthalene (PCN) congeners in a non-polar column, DB-5 (5% phenyl, methylpolysiloxane). Chemical descriptors were calculated from the molecular structures of PCNs and related to their gas chromatographic RRis by multiple linear regression analysis. The proposed model had a multiple square correlation coefficient $R^2=0.995$, and standard error $SE=16.7$. A QSRR reveals that the gas chromatographic retention of PCNs is associated with the number of chlorine substitutions, influenced by electronic descriptors such as heat of formation, maximum value for atomic valence, and the minimum value for electronic orbital population. © 1999 Published by Elsevier Science B.V. All rights reserved.

Keywords: Quantitative structure–retention relationships; Retention indices; Molecular descriptors; Polychlorinated naphthalenes

1. Introduction

Polychlorinated naphthalenes (PCNs) are environmentally persistent compounds. Technical PCN mixtures with chlorine contents ranging from 22 to 70% (w/w), known as Halowaxes, have been produced in the US since the 1920s. PCN congeners have been detected and quantified in several matrices including sediments [1,2], water [3,4], air [5,6] and biota [7–9]. Isomer specific analysis of PCNs is challenging due to the presence of the 75 theoretically possible isomers, lack of individual purified congeners for characterising the Halowax mixtures, and the simi-

larities in chemical properties among different congeners. PCNs have been analysed by high-performance liquid chromatography (HPLC) [10] and gas chromatography (GC) [11,12], however, isomer-specific analysis requires purified standards of individual isomers. Although gas chromatography–mass spectrometry (GC–MS) is the best analytical technique to identify these compounds at congener level, resolving complex mixtures might require additional tools involving the knowledge of molecular interactions of PCNs with a stationary phase. This can be achieved using quantitative structure–retention relationship (QSRRs) which provides statistical equations that relate molecular structure with the retention phenomena [13–17]. The objective of this

*Corresponding author.

paper is to provide a statistical model to predict the chromatographic retention of individual PCN congeners based on their molecular properties in a non-polar column, DB-5.

2. Experimental

2.1. Data set

Relative retention times were determined for 62 PCN congeners in a Hitachi M-80B GC–MS using fused-silica capillary columns coated with DB-5 (J&W Scientific, Folsom, CA, USA) (30 m×0.25 mm I.D., 0.25 μm film thickness). Chromatographic conditions were: column temperature programmed from 160 to 280°C at a rate of 4°C/min. Injector and detector temperatures were 280 and 180°C, respectively. Helium was used as a carrier gas and injections were made splitless. Relative retention indices (RRIs) were calculated based on relative retention to *n*-alkanes, viz., C₂₀H₄₂ to be 2000 and C₃₀H₆₂ to be 3000. Details regarding the identification of individual PCN congeners are given elsewhere [1].

2.2. Descriptor generation

Molecular descriptors used in this study were derived using the program SPARTAN 5.1 (Wavefunction, Irvine, CA, USA), Calculated descriptors for PCN congeners included physicochemical, geometrical and electronic. Molecules were entered and geometries were fully optimized using the semiempirical quantum chemical method AM1 [18], and single point calculations were performed to generate the descriptors. The physicochemical parameter, octanol–water partition coefficient ($\log P$), was obtained using the Ghose–Crippen (LOGP-GC) [19] and the Alkorta and Villar (LOGP-VI) [20] methods. Electronic descriptors are ranked into two categories, namely global and local. Global descriptors included heat of formation (HOF), energy of the highest occupied molecular orbital (ϵ_{HOMO}), energy of the lowest unoccupied molecular orbital (ϵ_{LUMO}), the difference between ϵ_{HOMO} and ϵ_{LUMO} (GAP), maximum (MAX-OP) and minimum (MIN-OP) electronic orbital population, zero point vibrational

energy (ZPVE), dipolar moment (DIP), and thermodynamic descriptors (taken at 298.15 K and 1.00 atm; 1 atm=101 325 Pa) vibrational enthalpy (VIB-H), rotational entropy (ROT-S) and vibrational entropy (VIB-S). Local descriptors included most negative (MNC) and most positive charge (MPC), sum of atomic charges for chlorine atoms (SACl) and maximum and minimum values for atomic valence, MAX-AV and MIN-AV, respectively. Geometrical descriptors comprise moment of inertia 1 (MI1), 2 (MI2) and 3 (MI3). Number of chlorines (NCl) was included as a topological descriptor. QSRR models developed using quantum-derived, physicochemical and geometrical descriptors are found elsewhere [13,14,21–23].

2.4. Statistical analysis

Descriptors and retention data were submitted to cluster analysis in order to establish preliminary relationships between variables. Pearson correlation and between-groups linkage were used as similarity measure and combining method, respectively [24]. Clusters including more than one variable were subsequently correlated with the RRIs and the descriptor with best correlation was introduced into the pool submitted to regression. The QSRR equation was obtained by forward stepwise multiple regression techniques following the multilinear form: $\text{RRI} = a_1D_1 + a_2D_2 + \dots + a_nD_n$, where D_1 , D_2 and D_n are the descriptors and a_1 , a_2 and a_n the respective regression coefficients [25]. The best model was selected based on the multiple square correlation coefficient (R^2), the standard error of estimation (SE) and the value of *F*-ratio [21]. Dependence between the number of descriptors and prediction capacity of the model was calculated applying the regression coefficient for the degrees of freedom (R_f^2) using the expression $R_f^2 = R^2\{[p - (1 - R^2)]/(n - p - 1)\}$, where R is the regression coefficient, p is the number of descriptors in the QSRR model and n is the total number of molecules [26].

2.5. Model validation

The robustness of the regression equation obtained in this study was evaluated by calculating the cross-

validated correlation coefficient R^2_{crossval} [27]. Regression equations were constructed leaving out a data point from the data set and calculating the value for that particular point with the new regression equation. The set of calculated RRis obtained by this process was correlated with the observed data to obtain R^2_{crossval} . A similar approach was applied leaving out sets of 20, 21 and 21 molecules to calculate an analogue $R_{\text{crossval}(20,21,21)}$ value, which measure the effect of groups of molecules in the predictive power of the model. Higher R_{crossval} and

$R_{\text{crossval}(20,21,21)}$ values indicate that the model possesses robustness and prediction capability.

3. Results and discussion

Cluster analysis depicted in Fig. 1 shows the multiple relationship patterns between the variables used to develop the QSRR model. It reveals that the chromatographic retention of PCNs for the DB-5 column is linked to the extent of chlorination in the

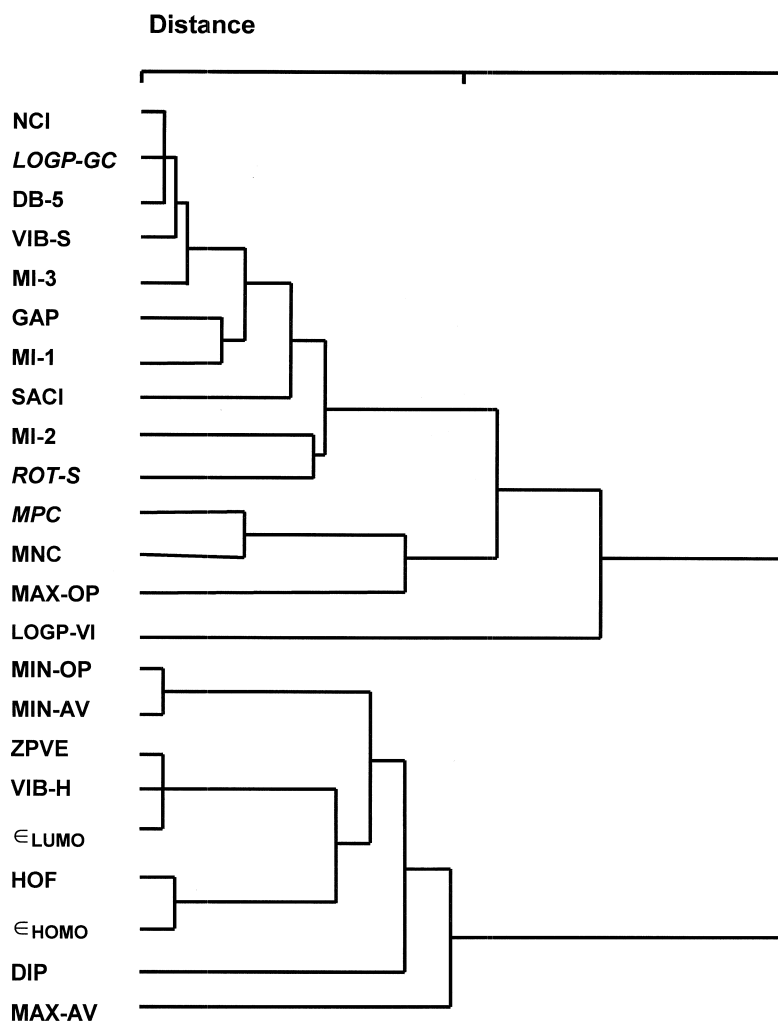


Fig. 1. Dendrogram from cluster analysis of correlation matrix of calculated molecular descriptors and the RRis (DB-5) for polychlorinated naphthalenes.

aromatic rings (NCI) as expected for a homologous series.

Regression equation obtained with selected molecular descriptors, which include NCI, HOF, MAX-AV and MIN-OP is presented in Table 1. The analysis of variance showed a statistical significance at $P < 0.001$ for the calculated model, and the R^2 and SE values for the regression equation were 0.995 and 16.67, respectively. The t -values reveal that the last two descriptors although do not contribute extensively to increase the robustness of the model, they function tuning the prediction power by decreasing the SE value. R^2_{crossval} and $R^2_{\text{crossval}(20,21,21)}$ values were 0.994 and 0.992, which suggest that the developed model is suitable to predict the RRis of PCNs. The high $R^2_{\text{crossval}(20,21,21)}$ value also implies that models developed with structure sets of 20–21 congeners have similar predictive ability to the regression model made with all the congeners. The observed and calculated values for RRis of PCN congeners in the DB-5 column, together with the values for each one of the descriptors involved in the model are shown in Table 2. Fig. 2 illustrates a plot of observed vs. calculated RRi values for all the 62 PCN congeners.

Molecular descriptors included in the regression equation were non-redundant ($R < 0.9$) as deduced from Table 3. Consequently, each descriptor in the model encodes a particular piece of information. Moreover, multicollinearity between the descriptors of the model was checked by calculating their variation inflation factors (VIF) by using the equa-

tion $\text{VIF} = (1 - R^2)^{-1}$, where R is the multiple correlation coefficient obtained after each descriptor has been submitted to linear regression against each other [28]. The average VIF for the four descriptors was 3.10 (range 1.67–4.04), which suggests that multicollinearity was not associated with the selected descriptors.

As expected, the main descriptor for the prediction of the gas chromatographic retention of PCNs in a DB-5 column as deduced from the t -statistics was the number of chlorine atoms, similar to that observed for other homologous series of compounds such as polychlorinated dibenzofurans [29]. NCI is a topological descriptor which encodes information about molecular size, electron number and polarizability [30], and not surprisingly it accounts for most of the variation in the RRis of PCNs on a DB-5 column ($R = 0.981$). Partial correlation controlling for NCI reveals that without considering the degree of chlorination, chromatographic retention is highly correlated with HOF ($R = 0.697$, $P = 0.000$) and MIN-OP ($R = 0.570$, $P = 0.000$). This suggests that the retention of PCNs in the DB-5 column not only involves molecular size but also isomer-specific electronic interactions with the stationary phase.

A model containing only NCI and HOF as molecular descriptors can predict RRis of PCNs with a $R^2 = 0.980$ and $\text{SE} = 33.69$. HOF is a quantum chemical index obtained by subtracting atomic heats of formation from the binding energy. HOF was correlated with ϵ_{HOMO} ($R = 0.949$). ϵ_{HOMO} is a measure of the energy demand of removing an electron from the molecule, which is proportional to the electron releasing abilities of the compounds. From the molecular orbital theory of chemical reactivity, transition states are formed during interaction between LUMO (electron acceptor) and HOMO (electron donor) of reacting species [31]. This orbital-type donor–acceptor interaction can lead to the formation of charge transfer complexes [32] between PCNs and the stationary phase. ϵ_{HOMO} has been found to be an important descriptor for the chromatographic separation of amines [33].

MAX-AV is a valency-derived descriptor encoding information related to spatial electronic distribution in a particular carbon within PCNs. Valency is a measure of the extent of electron sharing between various atomic centres. As a result, molecular orbital

Table 1
Best four descriptor QSRR model^a for gas chromatography retention of PCN congener on a DB-5 column

Descriptor	Regression coefficient \pm standard error
(1) NCI	270.53 \pm 3.86 (70.00) ^b
(2) HOF	15.11 \pm 0.95 (15.87)
(3) MAX-AV	18003.90 \pm 1819.79 (9.89)
(4) MIN-OP	–2525.58 \pm 469.58 (5.38)
Intercept ^c	–72636.75 \pm 7093.14

^a $\text{RRi} = aD_1 + bD_2 + \dots + cD_n$. RRi: relative retention indices taken for a DB-5 column. a , b , c : regression coefficients; D_1 , D_2 , D_n : molecular descriptors.

^b t -Values for the regression terms are given in parenthesis.

^c $R^2 = 0.995$, $R^2_t = 0.925$, $F = 3119.70$, $\text{SE} = 16.67$, $P < 0.001$, $n = 62$.

Table 2
Observed and calculated RRs for PCNs in a DB-5 column and descriptors included in the QSRR model

Molecule	RRis		NCI	HOF	MAX-AV	MIN-OP
	Observed	Calculated				
<i>Trichloronaphthalenes</i>						
1,3,6-Trichloronaphthalene	1759	1780	3	21.46	3.953	0.836
1,3,5-Trichloronaphthalene	1761	1767	3	22.87	3.951	0.836
1,3,7-Trichloronaphthalene	1769	1784	3	21.54	3.953	0.836
1,4,6-Trichloronaphthalene	1772	1763	3	22.58	3.951	0.838
1,2,4-Trichloronaphthalene	1776	1780	3	24.21	3.951	0.834
1,2,5-Trichloronaphthalene	1796	1806	3	23.82	3.951	0.845
1,2,6-Trichloronaphthalene	1802	1826	3	22.48	3.953	0.846
1,2,7-Trichloronaphthalene	1812	1811	3	22.61	3.954	0.838
1,6,7-Trichloronaphthalene	1812	1801	3	22.70	3.953	0.836
2,3,6-Trichloronaphthalene	1819	1821	3	21.46	3.954	0.845
1,2,3-Trichloronaphthalene	1827	1855	3	24.41	3.953	0.845
1,3,8-Trichloronaphthalene	1842	1841	3	26.98	3.952	0.836
1,4,5-Trichloronaphthalene	1852	1846	3	28.52	3.950	0.845
1,2,8-Trichloronaphthalene	1896	1897	3	28.58	3.952	0.846
<i>Tetrachloronaphthalenes</i>						
1,3,5,7-Tetrachloronaphthalene	1911	1941	4	16.82	3.951	0.834
1,2,4,6-Tetrachloronaphthalene	1950	1949	4	17.78	3.951	0.832
1,2,4,7-Tetrachloronaphthalene	1950	1947	4	17.80	3.951	0.832
1,2,5,7-Tetrachloronaphthalene	1950	1956	4	17.77	3.951	0.834
1,3,6,7-Tetrachloronaphthalene	1970	1977	4	16.54	3.953	0.834
1,4,6,7-Tetrachloronaphthalene	1974	1957	4	17.65	3.951	0.834
1,2,5,6-Tetrachloronaphthalene	1993	1981	4	18.67	3.950	0.844
1,3,6,8-Tetrachloronaphthalene	1993	2014	4	20.84	3.952	0.834
1,2,3,5-Tetrachloronaphthalene	2000	1983	4	19.28	3.952	0.833
1,3,5,8-Tetrachloronaphthalene	2000	2003	4	22.46	3.950	0.834
1,2,3,6-Tetrachloronaphthalene	2006	2020	4	17.87	3.953	0.843
1,2,3,7-Tetrachloronaphthalene	2017	2010	4	17.97	3.954	0.837
1,2,3,4-Tetrachloronaphthalene	2018	2026	4	21.04	3.950	0.850
1,2,6,7-Tetrachloronaphthalene	2018	1995	4	17.56	3.954	0.834
1,2,4,5-Tetrachloronaphthalene	2029	1997	4	23.73	3.949	0.832
2,3,6,7-Tetrachloronaphthalene	2034	2015	4	16.45	3.954	0.844
1,2,4,8-Tetrachloronaphthalene	2038	2021	4	24.30	3.950	0.832
1,2,5,8-Tetrachloronaphthalene	2052	2049	4	23.96	3.950	0.843
1,2,6,8-Tetrachloronaphthalene	2052	2043	4	22.38	3.952	0.834
1,4,5,8-Tetrachloronaphthalene	2086	2057	4	29.69	3.946	0.842
1,2,3,8-Tetrachloronaphthalene	2101	2093	4	24.01	3.952	0.844
1,2,7,8-Tetrachloronaphthalene	2114	2100	4	24.07	3.953	0.844
<i>Pentachloronaphthalenes</i>						
1,2,3,5,7-Pentachloronaphthalene	2145	2162	5	13.37	3.952	0.831
1,2,4,6,7-Pentachloronaphthalene	2145	2149	5	12.98	3.951	0.831
1,2,4,5,7-Pentachloronaphthalene	2168	2183	5	17.84	3.949	0.830
1,2,4,6,8-Pentachloronaphthalene	2178	2204	5	18.38	3.950	0.830
1,2,3,4,6-Pentachloronaphthalene	2186	2172	5	14.74	3.951	0.835
1,2,3,5,6-Pentachloronaphthalene	2190	2163	5	14.27	3.951	0.831
1,2,3,6,7-Pentachloronaphthalene	2217	2198	5	13.07	3.954	0.833

Table 2. (Continued)

Molecule	RRis		NCI	HOF	MAX-AV	MIN-OP
	Observed	Calculated				
<i>Pentachloronaphthalenes (Continued)</i>						
1,2,4,5,6-Pentachloronaphthalene	2227	2206	5	19.29	3.949	0.830
1,2,4,7,8-Pentachloronaphthalene	2235	2231	5	19.96	3.950	0.830
1,2,3,5,8-Pentachloronaphthalene	2243	2232	5	19.64	3.950	0.830
1,2,3,6,8-Pentachloronaphthalene	2243	2247	5	17.98	3.952	0.833
1,2,4,5,8-Pentachloronaphthalene	2261	2249	5	25.04	3.947	0.830
1,2,3,4,5-Pentachloronaphthalene	2275	2271	5	21.38	3.949	0.845
1,2,3,7,8-Pentachloronaphthalene	2309	2301	5	19.63	3.953	0.842
<i>Hexachloronaphthalenes</i>						
1,2,3,4,6,7-Hexachloronaphthalene	2378	2370	6	10.02	3.951	0.831
1,2,3,5,6,7-Hexachloronaphthalene	2378	2368	6	9.98	3.951	0.830
1,2,3,4,5,7-Hexachloronaphthalene	2405	2428	6	15.58	3.950	0.831
1,2,3,5,6,8-Hexachloronaphthalene	2405	2420	6	15.09	3.950	0.829
1,2,3,5,7,8-Hexachloronaphthalene	2415	2443	6	15.72	3.951	0.829
1,2,4,5,6,8-Hexachloronaphthalene	2425	2449	6	20.84	3.947	0.829
1,2,4,5,7,8-Hexachloronaphthalene	2425	2453	6	22.03	3.946	0.828
1,2,3,4,5,6-Hexachloronaphthalene	2472	2474	6	17.09	3.950	0.840
1,2,3,4,5,8-Hexachloronaphthalene	2493	2508	6	22.54	3.947	0.840
1,2,3,6,7,8-Hexachloronaphthalene	2505	2504	6	15.32	3.953	0.840
<i>Heptachloronaphthalenes</i>						
1,2,3,4,5,6,7-Heptachloronaphthalene	2694	2663	7	12.93	3.950	0.828
1,2,3,4,5,6,8-Heptachloronaphthalene	2694	2692	7	18.40	3.947	0.828

valency has high values for strongly bonding molecular orbitals [34]. The small *t*-statistics value for MIN-OP shows that it is a less important descriptor

in the model; however, it has reduced the standard error of the estimate from 20.29 to 16.67. MIN-OP is collinear with MIN-AV ($R = 1$), an indicator of the

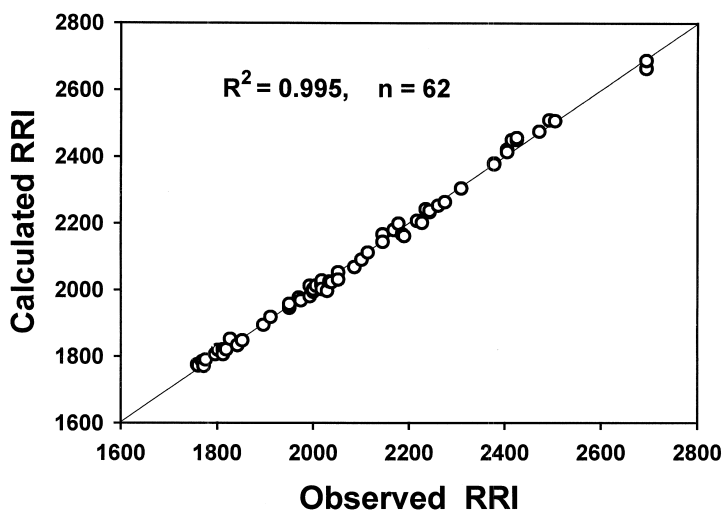


Fig. 2. Plot of calculated RRI vs. observed RRI values for the four-descriptor equation.

Table 3

Correlation matrix for the descriptors used in the QSRR model^a

	NCI	HOF	MAX-AV	MIN-OP
NCI	1.00			
HOF	-0.609	1.000		
MAX-AV	-0.492	-0.188	1.000	
MIN-OP	-0.538	0.473	0.324	1.000

^a Number of data = 62.

valency of hydrogen atoms. Katritzky et al. [17] have reported that MIN-AV was the second most important descriptor in the prediction of GC retention times of different types of organic compounds, and related its presence to the formation of a hydrogen bond between the compound and the stationary phase. In general, both MAX-AV and MP-LOW can account for unspecific weak electronic interactions between the PCN congeners and the stationary phase in which the roles of donor–acceptor are interchangeable on the area of interaction.

In conclusion, the GC retention of PCNs isomers in DB-5 column may be predicted with good accuracy using the chlorine number and three electronic descriptors. The electronic descriptors involved in the retention phenomena suggest that specificity at isomer level discrimination may be due to PCN–stationary phase interactions including the formation of electron charge transfer complexes and weak hydrogen bond-type interactions.

Acknowledgements

J.O. is sponsored by a Colciencias-Fulbright-Laspau scholarship and the University of Cartagena (Cartagena, Colombia). The authors are grateful to the SGI facility (NSF Grant CHE9321436) of the Chemistry Department of Michigan State University, and Paul Reed for his help.

References

- [1] K. Kannan, T. Imagawa, A. Blankenship, J.P. Giesy, Environ. Sci. Technol. 32 (1998) 2507.
- [2] U. Järnberg, L. Asplund, C. deWit, A. Egeback, U. Wideqvist, E. Jakobsson, Arch. Environ. Contam. Toxicol. 32 (1997) 232.
- [3] I. Marti, F. Ventura, J. Chromatogr. A 786 (1997) 135.
- [4] I. Espadaler, E. Eljarrat, J. Caixach, J. Rivera, I. Marti, F. Ventura, Rapid Commun. Mass Spectrom. 11 (1997) 410.
- [5] G. Dörr, M. Hippelein, O. Hutzinger, Chemosphere 33 (1996) 1563.
- [6] T. Harner, T. Bidleman, Atmos. Environ. 31 (1997) 4009.
- [7] J. Falandysz, C. Rappe, Environ. Sci. Technol. 30 (1996) 3362.
- [8] J. Falandysz, B. Strandberg, L. Strandberg, P. Bergqvist, C. Rappe, Sci. Total Environ. 204 (1997) 97.
- [9] A. Lunden, K. Noren, Arch. Environ. Contam. Toxicol. 34 (1998) 414.
- [10] J. Donkerbroek, C. Gooijer, N. Velthorst, R. Frei, Int. J. Environ. Anal. Chem. 15 (1983) 281.
- [11] C. Weistrand, E. Jakobsson, K. Noren, J. Chromatogr. B 669 (1995) 207.
- [12] J. Falandysz, B. Strandberg, L. Strandberg, P. Bergqvist, S. Kulp, C. Rappe, Roczn. Panstw. Zakl. Hig. 47 (1996) 257.
- [13] J. Olivero, T. Gracia, P. Payares, R. Vivas, D. Diaz, E. Daza, P. Geerlings, J. Pharm. Sci. 86 (1997) 625.
- [14] P. Payares, D. Diaz, J. Olivero, R. Vivas, I. Gomez, J. Chromatogr. A 771 (1997) 213.
- [15] J. Kang, C. Cao, Z. Li, J. Chromatogr. A 799 (1998) 361.
- [16] T. Woloszyn, P. Jurs, Anal. Chem. 64 (1992) 3059.
- [17] A. Katritzky, E. Ignatchenko, R. Barcock, V. Lobanov, M. Karelson, Anal. Chem. 66 (1994) 1799.
- [18] M. Dewar, E. Zoebisch, E. Healy, J. Stewart, J. Am. Chem. Soc. 107 (1985) 3902.
- [19] A. Ghose, A. Pritchett, G. Crippen, J. Comp. Chem. 9 (1988) 80.
- [20] I. Alkorta, H. Villar, Int. J. Quantum. Chem. 44 (1992) 203.
- [21] C. Georgakopoulos, J. Kiburis, P. Jurs, Anal. Chem. 63 (1991) 2021.
- [22] D. Stanton, P. Jurs, Anal. Chem. 61 (1989) 1328.
- [23] J. Luco, L. Yamin, H. Ferretti, J. Pharm. Sci. 84 (1995) 903.
- [24] M. Anderberg, Cluster Analysis For Applications, Academic Press, New York, 1973.
- [25] J. Neter, W. Wasserman, M. Kutner, Applied Linear Statistical Models, 2nd ed, Richard D. Irwin, Homewood, IL, 1985.
- [26] W.J. Dixon (Ed.), BMDP Statistical Software Manual, Vol. 1, University of California Press, Los Angeles, CA, 1992, pp. 387–425.
- [27] R. Murugan, M. Grendze, J. Toomey, A. Katritzky, M. Karelson, V. Lobanov, P. Rachwal, Chemtech. 24 (1994) 17.
- [28] G. Small, P. Jurs, Anal. Chem. 24 (1983) 164.
- [29] A. Robbat, C. Kalogeropoulos, Anal. Chem. 62 (1990) 2684.
- [30] A. Katritzky, Y. Wang, S. Sild, T. Tamn, M. Karelson, J. Chem. Inf. Comput. Sci. 38 (1998) 720.
- [31] K. Fukui, in: Theory of Orientation and Stereoselection, Springer-Verlag, New York, 1975, pp. 34–39.
- [32] R. Franke, in: Theoretical Drug Design Methods, Elsevier, New York, 1984, pp. 115–123.
- [33] K. Osmialowski, J. Halkiewicz, A. Radecki, R. Kalisz, J. Chromatogr. 346 (1985) 53.
- [34] M. Gopinathan, P. Siddarth, C. Ravimohan, Theor. Chim. Acta 70 (1986) 303.