Trans-POLYCHLORINATED AZOBENZENES (PCABs) – NEW POSSIBLY „DIOXIN-LIKE” ENVIRONMENTAL POLLUTANTS. MULTIDIMENSIONAL ANALYSIS OF THE STRUCTURAL FEATURE SPACE OF CONGENER SET

Jerzy Falandysz[1], Agata Wilczyńska[1], Slawomir Piliszek[1], Tomasz Puzyn[1]

[1]University of Gdansk

Introduction

Azoxybenzene is molecule composed of two benzene ring fused to azoxy group (-N(O)=N-). *Trans*-PCAOBs consist at 399 congeners (Table 1), and some of them can be formed as unwanted byproduct in the manufacture of 3,4-dichloroaniline (DCA) and its derivatives, and can be found in herbicides such as Diuron®, Linuron®, Methazole® and Propanil®(1,2). Degradation of chloroaniline herbicides is also a source of environmental pollution with PCAOBs as well as photolysis and biolysis of DCA(3,4). *Trans*-PCAOBs are planar compounds and can be considered as stereoisomers to highly toxic 2,3,7,8-tetrachlorodibenzo-*p*-dioxin.

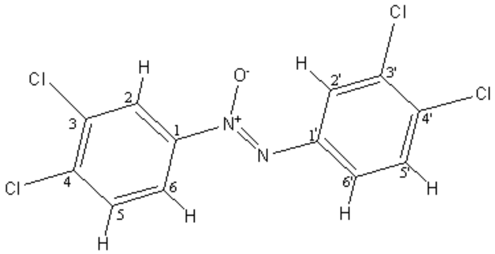

Figure 1. Generalized structure and numbering system for *trans*-chloroazoxybenzene.

They is lack or very limited number of data available regarding to sources, possible mechanism of formation, environmental occurence and fate, analytical chemistry, properties, and effects of *trans*-PCAOBs.

In this work some selected thermodynamic and quantum-mechanical properties of *trans*-PCAOBs were computed using *ab initio* methods. Presented property data of *trans*-PCAOBs can be useful as structural descriptors in further QSAR/QSPR studies of those compounds.

Table 1. Homologue groups, chemical formulas, number of isomers and congeners and molecular weight of chloroazoxybenzenes

Homologue group	Chemical formula	Number of isomers	Molecular weight
Mono-	$C_{12}H_9ClN_2O$	6	232.7
Di-	$C_{12}H_8Cl_2N_2O$	21	267.1
Tri-	$C_{12}H_7Cl_3N_2O$	48	301.6
Tetra-	$C_{12}H_6Cl_4N_2O$	78	336.0
Penta-	$C_{12}H_5Cl_5N_2O$	92	370.4
Hexa-	$C_{12}H_4Cl_6N_2O$	78	404.9
Hepta-	$C_{12}H_3Cl_7N_2O$	48	439.3
Octa-	$C_{12}H_2Cl_8N_2O$	21	473.8
Nona-	$C_{12}HCl_9N_2O$	6	508.2
Deka-	$C_{12}Cl_{10}N_2O$	1	542.7

Materials and Methods

In the first step, input files with internal coordinates were prepared for each of 399 possible congeners of *trans*-PCAOBs. An initial bond lenghts were set to C-C as 1.335 Å, to C-H as 1.089 Å, to C-Cl as 1.750 Å, to C-N as 1.440 Å, to N-O as 1.400 Å, and to N=N as 1.346 Å while, all of the initial valence angles as 120^o and torsion angles as 0^o, 141^o or 180^o.

In this study we used 31 descriptors. The method combines quantum-mechanical and chemometric tools. Qantum-mechanical descriptors, such as: average molecular polarizability, dipole moment, the most positive and negative partial charge on atoms, molecular volume were computed on the level of density functional theory (DFT) using B3LYP hybrid functional in Popple type triple zeta basis 6-311++G** (polarization and diffuse functions were added). In each case, polarizability tensors were diagonalized, and average polarizability was calculated as a mean of eigenvalues from the polarizability tensor.

Thermodynamical descriptors: standard heat of formation in gas phase, standard Gibbs free energy, head capacity in constant volume, entropy were calculated based on frequency analysis on the level of B3LYP/6-311++G** using the algorithm proposed by Ochterski (5). All quantum-mechanical and thermochemical computation were done using GAUSSIAN 03 package(6,7). Topological descriptors, such as: connectivity indices, topological indices, indices of neighborhood symmetry were calculated by means of DRAGON software(8,9). All physico-chemical and quantum-mechanical descriptors are listed in details in Table 2.

Table 2. Descriptors used in multidimensional characterization of *trans*-PCAOBs

Symbol	Descriptor	Units	References
DIPOL	Total dipole moment of the molecule	D	10
DIPOL_X	Dipole vector X	D	10
DIPOL_Y	Dipole vector Y	D	10
DIPOL_Z	Dipole vector Z	D	10
POL	Polarizability	$Å^3$	8,9
POL_X	Polarizability vector X	$Å^3$	8,9
POL_Y	Polarizability vector Y	$Å^3$	8,9
POL_Z	Polarizability vector Z	$Å^3$	8,9
CV	Thermal capacity	J/mol*K	12
S	Entropy	J/mol*K	13
MV	Molar volume	cm^3/mol	8
G	Gibss free energy	kJ/mol	8
H	Standard heat of formation	kJ/mol	8,10
W	Total particle energy	kJ/mol	14
Max Q^+	The most positive partial charge on atoms	-	8,15
Max Q^-	The most negative partial charge on atoms	-	8,15
HOMO	Energy of HOMO	eV	8,10
LUMO	Energy of LUMO	eV	8,10
ET	Thermal energy	kJ/mol	8
MW	Molecular weight	g/mol	8,10
LOGP	Logarithm of the n-octanol/water partition	-	16,17,18
M_REF	Molecular refraction	$Å^3$	13,14
VCl_0	Valence molecular index of zero order, ${}^0\chi^V$	-	8,11,19
VCl_1	Valence molecular index of first order, ${}^1\chi^V$	-	8,11,19
VCl_2	Valence molecular index of second order, ${}^2\chi^V$	-	8,11,19
KAPPA_1	Kappa index of first order, ${}^1\kappa$	-	8,19
KAPPA_2	Kappa index of second order, ${}^2\kappa$	-	8,19
KAPPA_3	Kappa index of third order, ${}^3\kappa$	-	8,19
CI_0	Molecular connectivity index of zero order, ${}^0\chi$	-	11,19
CI_1	Molecular connectivity index of first order, ${}^1\chi$	-	11,19
CI_2	Molecular connectivity index of second order, ${}^2\chi$	-	11,19

Results and discussion

Principal Component Analysis of (31*399) data matrix resulted in five-dimensional model which explains 84,61% (63,79% + 8,87% + 4,65% + 3,84% + 3,45%) of the total variance. The first five significant principal components were chosen applying screen plot criterion(20). Factor matrix obtained after Varimax rotation indicates strong correlations between principal components and congeners.

As it could be noted (Figure 2), the first principal component (PC1) is determined by descriptors connected with molecular shape and size, and what implies on a similar susceptibility to environmental transport and fate (Table 3). There are negatively correlated of HOMO energy, Gibss free energy, standard heat of formation, the total particle energy and thermal energy. PC2 is strongly depended on the energy of LUMO and polarizability vector X, and first of them is negatively correlated. The third principal component is negatively impacted by the most positive partial charge on atoms, but the fourth principal component is negatively determined by dipole vector Y. The last one (PC5) is positively correlated with dipole vector Z.

When analyzing the score plot, it could be noted, that the first principal component distinguishes a different homologue groups of chloroazoxybenzene. Each point represents one congener of chloroazoxybenzene, which is described by the corresponding number. Compounds forming homologue groups are characterized by similar values of indices χ, χ^v, κ, molecular weight, refraction, log K_{ow} , LUMO energy, total particle energy, Gibss free energy, standard heat of formation and thermal energy. The second principal component separates chloroazoxybenzenes owning to values of mainly the most positive and negative partial charge on atoms.

Concluding, the most interesting features of *trans*- PCAOBs set are energy of a highest occupied molecular orbital (which decides about ionization potential of the molecule) and polarity (represented by total dipole moment and dipole vectors). These properties are probably responsible for diverse environmental transport and fate as well as toxicokinetics and effects excerted by each congener of chloroazoxybenzene.

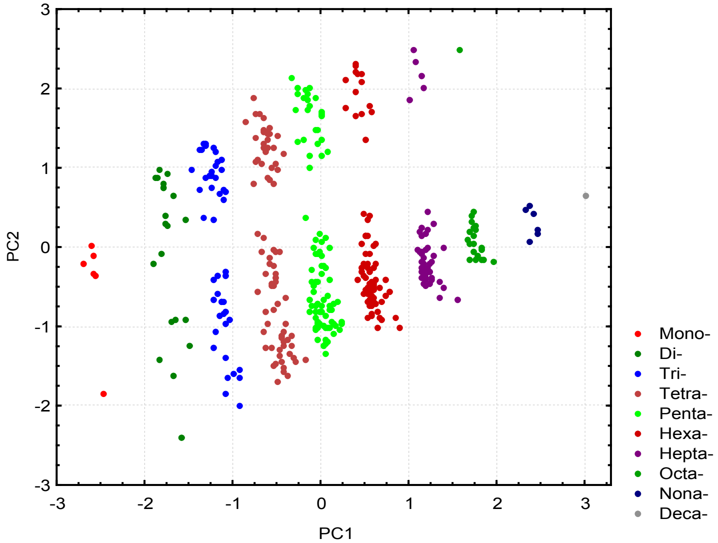

Figure 2. The loading plot of the second principal component (PC2) vs. the first principal component (PC1).

Table 3. Table of loadings

Variable	PC1	PC2	PC3	PC4	PC5
POL	[a]**0,84**	0,52	0,05	0,07	0,00
POL_X	**0,70**	-0,53	0,09	-0,11	0,30
POL_Y	**0,80**	0,33	-0,04	0,09	-0,17
POL_Z	0,32	**0,89**	0,06	0,10	-0,03
maxQ-	-0,02	-0,40	0,67	0,04	-0,24
maxQ+	-0,39	-0,02	**-0,76**	-0,14	0,02
DIPOL	-0,23	0,13	-0,07	-0,54	-0,63
DIPOL_X	-0,02	0,16	**0,57**	-0,19	0,05
DIPOL_Y	0,05	-0,12	0,09	**-0,86**	0,07
DIPOL_Z	-0,11	0,27	-0,15	-0,16	**0,73**
HOMO	**-0,78**	0,41	-0,08	0,15	-0,22
LUMO	-0,09	**-0,86**	0,02	0,03	-0.26
MW	**0,99**	0,07	0,07	0,02	0.01
G	**-0,99**	-0,07	-0,09	0,00	-0.06
H	**-0,99**	-0,05	-0,09	0,00	-0.07
W	**-0,99**	-0,07	-0,07	-0,02	-0.01
logP	**0,99**	0,07	0,07	0,01	0.02
M_REF	**0,99**	0,07	0,07	0,02	0.01
VCI_0	**0,99**	0,07	0,07	0,02	0.01
VCI_1	**0,95**	0,06	0,07	0,01	0.02
VCI_2	**0,95**	0,06	0,06	0,03	0.01
KAPPA_1	**0,98**	0,07	0,07	0,03	0.01
KAPPA_2	**0,92**	0,07	0,01	0,01	0.00
KAPPA_3	0,38	0,14	-0,02	-0,14	0.05
CI_0	**0,97**	0,07	0,06	0,03	0.01
CI_1	**0,93**	0,03	0,09	0,03	-0.02
CI_2	**0,95**	0,10	0,06	0,04	0.03
MV	0,67	0,02	0,04	0,07	-0.03
CV	**0,99**	-0,03	0,09	0,01	0.04
ET	**-0,99**	-0,09	-0,06	-0,03	-0.01
S	**0,98**	-0.11	0.10	-0.01	0.07

[a] **Marked** loadings are > 0.70

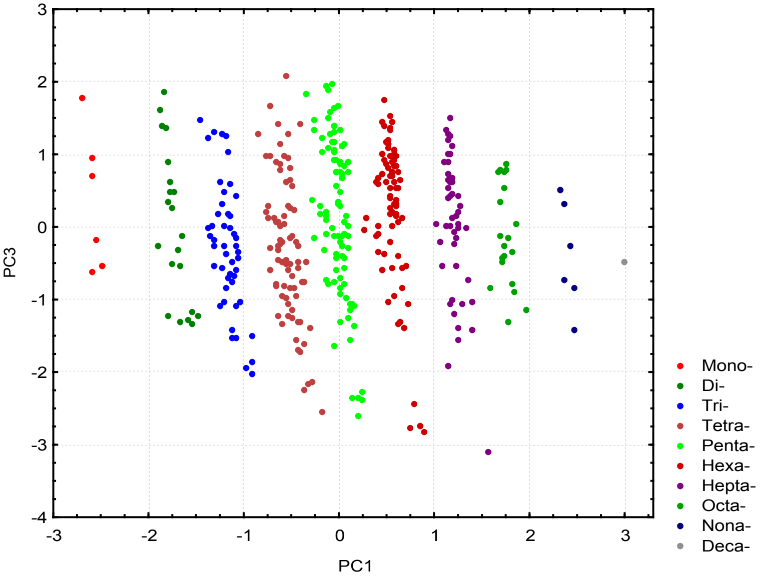

Figure 3. The loading plot of the sixth principal component (PC6) vs. the first principal component (PC1).

Computations were carried out using computers in TASK – Academic Computer Center in Gdańsk.

References

1. Ambrus A., Hamilton D.J., Kuiper H.A., Racke K.D. (2003) Pure Appl. Chem. 75, 937.
2. Taylor J.S., Wuthrich R.C., Loyd K.M., Poland A. (1977) Arch. Dermatol. 113, 616.
3. Kaufman D.D., Plimmer J.R., Iwan J., Klingebiel U.I. (1972) 3,3′,4,4′-Tetrachloroazoxybenzene from 3,4-dichloroaniline in microbial culture. J. Agric. Food Chem. 20, 916.
4. Mansour M., Parlar H., Korte F. (1975) Beiträge zur Ökologischen Chemi Ci: Reaktionsverhalten von 3,4 Dichloranilin und 3,4 Dichlorphenol in Lösung, als Festkörper und in der Gasphase bei UV-Bestrahlung. Chemosphere. 4, 235.
5. Ochterski J.W. (2000) Thermochemistry in Gaussian. Gaussian, Inc.
6. Frisch M.J., Trucks G.W., Schlegel H.B., Scuseria G.E., Robb M.A., Cheeseman J.R., Montgomery J.A., Vreven T., Kudin K.N., Burant J.C., Millam J.M., Iyengar S.S., Tomasi J., Barone V., Mennucci B., Cossi M., Scalmani G., Rega N., Petersson G.A., Nakatsuji H., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Klene M., Knox X. Li, J.E., Hratchian H.P., Cross J.B., Adamo C., Jaramillo J., Gomperts R., Stratmann R.E., Yazyev O., Austin A.J., Cammi R., Pomelli C., Ochterski J.W., Ayala P.Y., Morokuma K., Voth G.A., Salvador P., Dannenberg J.J., Zakrzewski V.G., Dapprich S., Daniels A.D., Strain M.C., Farkas O., Malick D.K., Rabuck A.D., Raghavachari K., Foresman J.B., Ortiz J.V., Cui Q., Baboul A.G., Clifford S., Cioslowski J., Stefanov B.B., Liu G., Liashenko A., Piskorz P., Komaromi I., Martin R.L., Fox D.J., Keith T., Al-Laham M.A., Peng C.Y., Nanayakkara A., Challacombe M., Gill P.M.W., Johnson B., Chen W., Wong M.W., Gonzalez C., and Pople J. A. (2003) Gaussian, Inc., Pittsburgh PA.
7. Curtiss L.A., Raghavachari K., Redfern P.C., Pople J.A. (1997) J. Chem. Phys. 106, 3, 1063.

8. Todeschini R., Consonni V. (2000) Handbook of molecular descriptors, Wiley, Weinheim.
9. Todeschini R., Consonni V., Mauri A., Pavan M. (2003) DRAGON Software version 3. Milano Chemometrics. Milano, Italy.
10. Stewart J.J.P. (1990) in: Review in Computational Chemistry (Lipkowitz K.B., Boyd D.B., Ed.), VCH, New York, ISBN 3-527-27845-1.
11. Hall L.H., Kier L.B. (1991) Rev. Comput. Chem. 2, 367.
12. Jaworski B.M., Dietłaf A.A., (1995) Fizyka poradnik encyklopedyczny (*in Polish*) Wydawnictwo Naukowe PWN, Warszawa, ISBN 83-01-11858-X.
13. Pigoń K. Ruziewicz Z. (1980) Chemia fizyczna (*in Polish*), Państwowe Wydawnictwo Naukowe, Warszawa.
14. Kaliszan R. (1997) Structure and Retention in Chromatography. A Chemometric Approach. Harwood Academic Publishers, Amsterdam, ISBN 90-5702-028-9.
15. Chen J., Peijnenburg W.J.G.M., Quan X., Chen S., Martens D., Schramm K.W., Kettrup A., (2001) Environ. Pollut. 114,137.
16. Hansch C., Leo A. (1995) Exploring QSAR. Fundamentals and applications in chemistry and biology. ACS Professional Reference Book, American Chemical Society, Washington DC, ISBN 0-8412-2992-9.
17. Kubinyi H. (1993) QSAR: Hansch analysis and related approaches, Vol. 1, VCH, New York, ISBN 1-56081-768-2.
18. Seńczuk W. (1999) Toksykologia (*in Polish*), Wydawnictwo Lekarskie PZWL, Warszawa.
19. Kier L.B., Hall L.H. (1986) Molecular connectivity in chemistry and drug design. Academic Press, New York.
20. Mazerski J. (2000) Podstawy chemometrii (*in Polish*). Wydawnictwo Politechniki Gdańskiej, Gdańsk, ISBN 83-88007-37-8.