Development of a Multiple-Class High-Resolution Gas Chromatographic Relative Retention Time Model for Halogenated Environmental Contaminants

ARTICLE *in* ANALYTICAL CHEMISTRY · APRIL 2003

Impact Factor: 5.64 · DOI: 10.1021/ac020406p · Source: PubMed

2 AUTHORS:

Sierra Rayne
Chemologica Research
270 PUBLICATIONS **2,288** CITATIONS
SEE PROFILE

Michael Ikonomou
Simon Fraser University
98 PUBLICATIONS **4,180** CITATIONS
SEE PROFILE

Anal. Chem. **2003,** *75,* 1049−1057

Articles

Development of a Multiple-Class High-Resolution Gas Chromatographic Relative Retention Time Model for Halogenated Environmental Contaminants

Sierra Rayne

Department of Chemistry, Box 3065, University of Victoria, Victoria, British Columbia, Canada, V8W 3V6

Michael G. Ikonomou*

Contaminants Science Section, Institute of Ocean Sciences, Fisheries and Oceans Canada, 9860 West Saanich Road, Sidney, British Columbia, Canada, V8L 4B2

A predictive model for the relative gas chromatographic retention times (GC-RRTs) of the following nine classes of halogenated environmental contaminants was developed: polybrominated diphenyl ethers (PBDEs); polychlorinated diphenyl ethers (PCDEs); polychlorinated biphenyls (PCBs); polychlorinated naphthalenes (PCNs); polychlorinated dibenzo-*p*-dioxins (PCDDs); polychlorinated dibenzofurans (PCDFs); polybrominated dibenzo-*p*-dioxins (PBDDs); polybrominated dibenzofurans (PBDFs); and organochlorine pesticides. MOPAC calculated physicochemical properties and structural descriptors in the model include molecular weight, square root of the number of halogen substituents, ionization potential, dipole moment, and the number of ortho, meta, and para halogen substituents. Using these variables, individual models for each of the contaminant classes were combined into a multiple class model incorporating the GC-RRTs of the 375 compounds of interest. The individual and multiclass GC-RRT models had acceptable fits between observed and predicted GC-RRTs ($r^2 = 0.9741 - 0.9990$ for PBDEs, PCDEs, PCBs, PCNs, PCDD/Fs, and PBDD/Fs; $r^2 = 0.9250$ for pesticides; and $r^2 = 0.9631$ for the multiclass model) over a wide range of retention times and molecular structures. The combined model was tested on known GC-RRTs of hydroxylated PCBs and chlorinated phenoxyphenols and provided satisfactory results, demonstrating the strength of the model in predicting GC-RRT windows for contaminant classes not used in constructing the model. Such models will be useful in predicting the GC retention characteristics of novel environmental contaminants and their degradation products, for which analytical standards may not be available.

In the period between 1828, when Friedrich Wöhler accidentally performed the first organic synthesis by creating urea from inorganic materials, and the present, chemists have synthesized over 10 million organic compounds. These advances in organic synthesis over the past two centuries have unfortunately created many environmental problems. We now know the problems created by the bioaccumulation of even trace quantities of halogenated contaminants such as DDT and PCBs, among many others. However, analytical chemistry has only touched on the wide diversity of potential contaminants, and little is known of their occurrence and toxicity. Rather than simply determining the total quantities of broad contaminant classes (e.g., total pesticides, Aroclor mixtures, etc.), levels of individual compounds are necessary to properly assess the extent of environmental pollution. Work on the toxicology of polychlorinated dibenzo-*p*-dioxin and dibenzofuran (PCDD/F) congeners has shown some to be several orders of magnitude more toxic than others.[1,2] These studies resulted in the development of the toxic equivalence factor (TEF) concept, which rates all potential acute toxicities to that of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (2378-TeCDD; TEF ≡ 1.0). Long held to be one of the most toxic organic substances known, recent work has demonstrated that similar compounds, namely the 2,3,7,8-substituted brominated dibenzo-*p*-dioxins and dibenzofurans (PBDD/Fs), may be even more toxic.[3] These findings highlight the need to screen environmental samples for compounds of potentially even greater toxicity. In addition, many known contaminants were not intentionally produced and occur

* Corresponding author. Phone: (250) 363-6804. Fax: (250) 363-6807. E-mail: ikonomoum@pac.dfo-mpo.gc.ca.

(1) Kutz, F. W.; Barnes, D. G.; Bretthauer, E. W.; Bottimore, D. P.; Greim, H. *Toxicol. Environ. Chem.* **1990**, *26*, 99−109.
(2) Kutz, F. W.; Barnes, D. G.; Bottimore, D. P.; Greim, H.; Bretthauer, E. W. *Chemosphere* **1990**, *20*, 751−757.
(3) Hornung, M. W.; Zabel, E. W.; Peterson, R. E. *Toxicol. Appl. Pharmacol.* **1996**, *140*, 227−234.

10.1021/ac020406p CCC: $25.00
Published on Web 01/28/2003

as byproducts in technical mixtures of other contaminants or from waste treatment operations (e.g., PCDD/Fs in Agent Orange formulations, from waste combustion, and chlorination of high organic strength wastes), whereas others were both industrial products [e.g., polychlorinated biphenyls (PCBs), polychlorinated naphthalenes (PCNs), polychlorinated and polybrominated diphenyl ethers (PCDEs and PBDEs, respectively), and organochlorine pesticides] and impurities in technical mixtures (e.g., PCNs in PCB mixtures). Thus, it is difficult to predict the nature and distribution of novel contaminants, especially degradation products that may result from a wide range of metabolic, photochemical, and thermal reactions taking place within organisms and various environmental compartments.

The coupling of gas chromatography (GC) with mass spectrometry (MS) is one of the most powerful analytical tools for such compounds. GC/MS allows the separation of these contaminants on the GC column, followed by structural identification and quantitation by the MS. High-resolution GC (HRGC) coupled with high-resolution MS (HRMS; HRGC/HRMS) allows separation and quantitation of nearly all possible environmental contaminants at near or below the parts-per-trillion (ppt) level provided proper sample cleanup procedures are undertaken prior to HRGC/HRMS analysis. However, given the large number of potential contaminants and the lack of analytical standards, predictive tools are needed to assist researchers in screening environmental samples for novel compounds. Several models have been developed to predict relative GC retention times (GC-RRTs) of individual halogenated organic contaminant classes, such as those for PCBs,[4−6] PCDEs,[7] PCDDs,[4,8] PCDFs,[4,9,10] and PBDDs,[11] yet no further attempts have been made to develop a predictive GC-RRT model for more than one contaminant class after such an approach was successfully demonstrated by Ong and Hites.[4]

The objective of this work was to identify a set of easily calculated variables that could be successfully used as descriptors in a predictive GC-RRT model for each of the following individual halogenated contaminant classes: PBDEs, PCDEs, PCBs, PCNs, PCDDs, PCDFs, PBDDs, PBDFs, and organochlorine pesticides. Although in some cases, other published models for a specific individual contaminant class may have a better fit and superior predictive ability, the lack of generalization for the variables used in these published models precludes a ready compilation of published models into a multiclass model. In addition, for several of the classes above, no previously published GC-RRT models are available (PCNs, PBDFs, and organochlorine pesticides). Once a set of successful predictors was found to provide an adequate fit for each class, we sought to combine the GC-RRTs for all compounds into a multiclass model that would calculate retention time windows for new environmental contaminants provided the physicochemical properties and structural parameters could be readily calculated or determined. The utility of the model was then demonstrated on the known and assumed GC-RRTs of some of the chlorinated dihydroxybiphenyl and phenoxyphenol photoproducts of 2378-TeCDD, most of which do not have available analytical standards for conclusive structural identification. Such a model is designed to play a role in screening environmental samples for novel environmental contaminants and their degradation products, as well as provide additional evidence in assigning molecular structure to isomeric compounds for which standards are unavailable.

MATERIALS AND METHODS

Gas Chromatography. Analyses were performed by HRGC/HRMS using a VG-Autospec high-resolution mass spectrometer (Micromass, Manchester, U.K.) equipped with a Hewlett-Packard 5890 series II gas chromatograph and a CTC A200S autosampler (CTC Analytics, Zurich, Switzerland). The HRGC was operated in the splitless injection mode. The volume injected was 1 μL of sample plus 0.5 μL of air. The HRMS was the only on-line detector attached to the HRGC system. For all analyses, ultrahigh-purity He (UHP-He) was the carrier gas at a constant head pressure. All analyses used DB-5 columns from J&W Scientific (Folsom, CA).

For PBDEs, a 30-m DB-5 column (0.25-mm i.d. × 0.25-μm film thickness) was used with UHP-He at 90 kPa and the following temperature program: hold at 100 °C for 1 min, 2 °C·min^{-1} to 150 °C, 4 °C·min^{-1} to 220 °C, 8 °C·min^{-1} to 330 °C, and hold for 1.2 min. The splitless injector port, direct HRGC/HRMS interface, and the HRMS ion source were maintained at 300, 275, and 315 °C, respectively, and the splitless injector purge valve was activated 2 min after sample injection. For PCDEs, a 30-m DB-5 column (0.25-mm i.d. × 0.25-μm film thickness) was used with UHP-He at 40 kPa and the following temperature program: hold at 100 °C for 1 min, 4 °C·min^{-1} to 290 °C, and hold for 2.0 min. The splitless injector port, direct HRGC/HRMS interface, and the HRMS ion source were maintained at 282, 260, and 305 °C, respectively, and the splitless injector purge valve was activated 1 min after sample injection. For mono-ortho and non-ortho PCBs, a 55-m DB-5 column (0.25-mm i.d. × 0.1-μm film thickness) was used with UHP-He at 125 kPa and the following temperature program: hold at 80 °C for 2 min, 8 °C·min^{-1} to 150 °C, and 4 °C·min^{-1} to 285 °C. The splitless injector port, direct HRGC/HRMS interface, and the HRMS ion source were maintained at 282, 260, and 305 °C, respectively, and the splitless injector purge valve was activated 2 min after sample injection. For di-ortho PCBs, a 55-m DB-5 column (0.25-mm i.d. × 0.1-μm film thickness) was used with UHP-He at 120 kPa and the following temperature program: hold at 80 °C for 2 min, 8 °C·min^{-1} to 150 °C, 4 °C·min^{-1} to 300 °C, and hold for 2.0 min. The splitless injector port, direct HRGC/HRMS interface, and the HRMS ion source were maintained at 282, 260, and 305 °C, respectively, and the splitless injector purge valve was activated 2 min after sample injection. PCN analyses were performed by Axys Analytical (Sidney, BC, Canada) using a HP 5890 series II HRGC, a CTC A200S autosampler, and a VG-Autospec VG-70SE HRMS. For PCNs, a 60-m DB-5 column (0.25-mm i.d. × 0.1-μm film thickness) was used with UHP-He at 154 kPa and the following temperature program: hold at 50 °C for 1 min; 1 °C·min^{-1} to 100 °C; and 7 °C·min^{-1} to 300 °C. The splitless

(4) Ong, V. S.; Hites, R. A. *Anal. Chem.* **1991**, *63*, 2829−2834.
(5) Hasan, M. N.; Jurs, P. C. *Anal. Chem.* **1988**, *60*, 978−982.
(6) Robbat, A.; Xyrafas, G.; Marshall, D. *Anal. Chem.* **1988**, *60*, 982−985.
(7) Nevalainen, T.; Koistinen, J.; Nurmela, P. *Environ. Sci. Technol.* **1994**, *28*, 1341−1347.
(8) Liang, X.; Wang, W.; Wu, W.; Schramm, K. W.; Henkelmann, B.; Kettrup, A. *Chemosphere* **2000**, *41*, 923−929.
(9) Liang, X.; Wang, W.; Schramm, K. W.; Zhang, Q.; Oxynos, K.; Henkelmann, B.; Kettrup, A. *Chemosphere* **2000**, *41*, 1889−1895.
(10) Hale, M. D.; Hileman, F. D.; Mazer, T.; Shell, T. L.; Noble, R. W.; Brooks, J. J. *Anal. Chem.* **1985**, *57*, 640−648.
(11) Liang, X.; Wang, W.; Wu, W.; Schramm, K. W.; Henkelmann, B.; Oxynos, K.; Kettrup, A. *Chemosphere* **2000**, *41*, 917−921.

injector port, direct HRGC/HRMS interface, and the HRMS ion source were maintained at 180, 295, and 250 °C, respectively, and the splitless injector purge valve was activated 2 min after sample injection. For PCDDs and PCDFs, a 60-m DB-5 column (0.25-mm i.d. × 0.1-μm film thickness) was used with UHP-He at 170 kPa and the following temperature program: hold at 100 °C for 2 min; 20 °C·min^{-1} to 200 °C; 1 °C·min^{-1} to 215 °C, hold 7 min; 4 °C·min^{-1} to 300 °C; and hold 3 min. The splitless injector port, direct HRGC/HRMS interface, and the HRMS ion source were maintained at 282, 270, and 305 °C, respectively, and the splitless injector purge valve was activated 2 min after sample injection. For the organochlorine pesticides, a 40-m DB-5 column (0.25-mm i.d. × 0.1-μm film thickness) was used with UHP-He at 60 kPa and the following temperature program: hold at 80 °C for 3 min; 15 °C·min^{-1} to 160 °C; 5 °C·min^{-1} to 300 °C, and hold 5 min. The splitless injector port, direct HRGC/HRMS interface, and the HRMS ion source were maintained at 250 , 250, and 300 °C, respectively, and the splitless injector purge valve was activated 0.5 min after sample injection.

Mass Spectrometry. The high-resolution MS was a sector instrument of EBE geometry coupled to the HRGC via a standard Micromass GC/MS interface. For all analyses, the HRMS was operated under positive EI conditions with the filament in the trap stabilization mode at 600 μA, an electron energy of 39 eV, and perfluorokerosene used as the calibrant. The instrument operates at 10 000 resolution, and data were acquired in the selected ion monitoring (SIM) mode for achieving maximum possible sensitivity. Two or more ions of known relative abundance were monitored for each molecular ion cluster representing a group of isomers, as were two for each of the ^{13}C-labeled surrogate standards.

Compounds were identified only when the HRGC/HRMS data satisfied all of the following criteria: (1) two isotopes of the specific compound were detected by their exact masses with the mass spectrometer operating at 10 000 resolving power or higher during the entire chromatographic run, (2) the retention time of the specific peaks was within 3 s of the predicted time obtained from analysis of authentic compounds in the calibration standards, (3) the peak maximums for both characteristic isotopic ions of a specific compound coincided within 2 s, (4) the observed isotope ratio of the two ions monitored per compound were within 15% of the theoretical isotopic ratio, and (5) the signal-to-noise ratio resulting from the peak response of the two corresponding ions was ≥3 for proper quantification of the compound. Concentrations of identified compounds and their method detection limits (MDLs) were calculated by the internal standard isotope-dilution method using mean relative response factors (RRFs) determined from calibration standard runs made before and after each batch of samples was analyzed.

Molecular Modeling and Data Treatment. Physicochemical properties for the analytes of interest were calculated using Chem3D Ultra 6.0 (CambridgeSoft, Cambridge, MA). Molecular structures were optimized using the MM2 energy minimization program. The physicochemical properties were then calculated using the MOPAC2000 MNDO-PM3 program.[12] Data were subsequently treated using Microsoft Excel 2002 and KyPlot v.2.0 beta 9 (32 bit). RRTs were obtained by dividing the RT for the analyte of interest by the RT of 2,2′-PCB4, which had a RT of 17.33 min using the instrument conditions described above. Linking of RRTs between different compound classes was obtained with 2,2′,4,5,5′-PCB101 as the internal standard (i.e., RRTs were adjusted by examining the retention time of PCB101 in respect of the various temperature programs and columns).

RESULTS AND DISCUSSION

Linear predictive equations for the relative GC retention times of the following nine classes of halogenated environmental contaminants were developed on the basis of molecular structures and computer-calculated physicochemical properties (Figure 1) for PBDEs, PCDEs, PCBs, PCNs, PCDDs, PCDFs, PBDD/Fs, and organochlorine pesticides. These individual models were then combined into a single, multiclass predictive model designed to act as a screening tool for novel environmental contaminants and their degradation products covering a wide range of possible molecular structures.

Individual Models for the Contaminant Classes. The following molecular descriptors and computer calculated physical properties were used to develop linear predictive GC-RRT models for each contaminant class: molecular weight (MW), square root of the number of halogen substituents (no. $X^{1/2}$), ionization potential (IP); dipole moment (μ), number of ortho-substituted halogens (no. *o*-X), number of meta-substituted halogens (no. *m*-X), and number of para-substituted halogens (no. *p*-X). Strong correlation was observed between observed and predicted RRTs within each model, ranging from an $r^2 = 0.9215$ for the organochlorine pesticides to $r^2 = 0.9741 - 0.9990$ for the halogenated diaryl systems (e.g., PBDEs, PCBs, etc.). Numerous other potential independent variables were considered in the development of this RRT model (e.g., polarizabilities, ionization potential, Connolly accessible surface area, Connolly solvent-excluded volume, etc.). Correlation analysis was first performed using a matrix of all potential independent variables and the retention time data. Variables strongly correlated with each other were excluded as potential candidates to minimize multicollinearity in the final model. In an ideal multiple linear regression model, none of the independent variables would be correlated other than with the dependent variable. However, in some cases, weakly multicollinear variables are included to minimize curvature of the residuals.[5] In addition, variables that did not have a significant relationship to retention time were also excluded. This process resulted in the reduced set of optimum variables presented above. Selected PCB congeners from each homologue group were examined as potential bases in the PCB model as well as for the multiple-class model. Similar analyses were performed for each contaminant class. In addition, combinations of lower and higher halogenated congeners were examined as potential RRT bases (e.g., PCB15/PCB180) within and between each contaminant class, as has been previously reported for the development of PCB retention time models [13] in which the RRT of a congener is calculated as relative to the average retention time of a low and a higher chlorinated congener. However, there was no increase in the power of the RRT model using either another individual congener or a combination of two congeners as a basis for calculating RRTs.

To our knowledge, linear predictive equations describing the GC behavior of PCNs, PBDFs, and organochlorine pesticides have

(12) Stewart, J. J. P. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1−45.

(13) Frame, G. M. *Fresenius' J. Anal. Chem.* **1997**, *357*, 714−722.

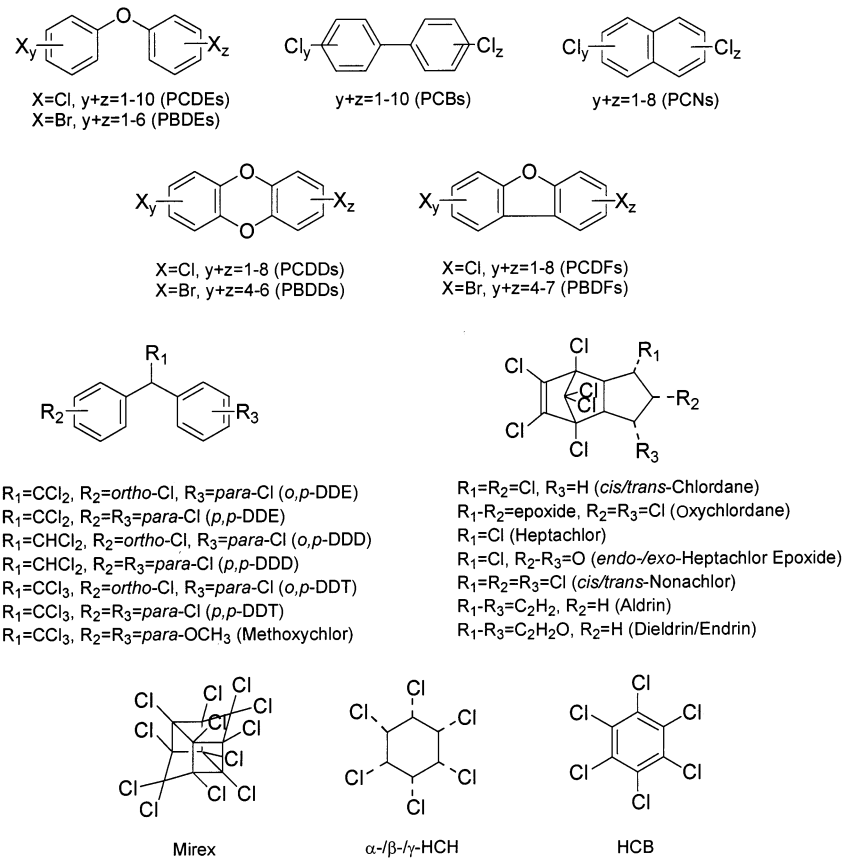

Figure 1. Classes of compounds used in the GC-RRT model.

Table 1. Regression Coefficients and Statistical Descriptors of the Linear Predictive GC-RRT Models: RRT = b_0+b_1(MW)$+b_2$(no. X)$^{1/2}+b_3$(IP)$+b_4(\mu)+b_5$(no. *o*-X)$+b_6$(no. *m*-X)$+b_7$(no. *p*-X)

	PBDEs	PCDEs	PCBs	PCNs	PCDDs	PCDFs	pesticides	PBDD/Fs	multiclass
b_0	−1.415	4.844	−31.05	1287	507.3	6.308	0.2372	−1.629	−1.663
b_1	-4.377×10^{-4}	−0.07118	0.2028	−9.979	−2.6701	0.01343	8.957×10^{-3}	4.841×10^{-3}	0.01225
b_2	3.772	0.4980	0.3078	0.3960	0.002573	0.008741	−0.7841	0.4862	0.1386
b_3	−0.1279	0.8117	6.83×10^{-4}	−0.8799	−1.7399	−0.9338	0.06987	0.04960	0.02708
b_4	0.1199	0.05603	0.03947	0.01967	0.04757	0.1367	-6.101×10^{-3}	0.03854	0.02971
b_5	−0.3321	2.562	−6.863	343.8	92.44	−0.05946	4.138×10^{-3}	n/a	−0.2396
b_6	−0.1302	2.589	−6.849	343.9	92.44	0.1507	−0.1374	n/a	−0.2494
b_7	−0.9156	2.626	−6.823	0	0	0	0.04256	n/a	−0.1951
N	35	39	131	70	33	34	23	10	375
r^2	0.9966	0.9945	0.9741	0.9899	0.9855	0.9815	0.9250	0.9990	0.9631
F	1353	797	660	1033	476	238	26	1207	1369
SE[a]	0.08401	0.04340	0.05567	0.02637	0.08573	0.1116	0.1230	0.02068	0.1924
% CV[b]	2.26	1.88	3.30	2.03	3.98	4.89	5.91	0.72	9.27

[a] Standard error, square root of the variance of the residuals. [b] Coefficient of variation.

not been previously published. The PBDE model was among the strongest [Table 1 and Figure 2] for 35 congeners ranging from mono- through hexa-brominated, with an $r^2 = 0.9966$, an F value of 1353 ($F_{obs} > F_{crit} = 2.4$; $p < 3.8 \times 10^{-33}$), and an even distribution of residuals over the range of predicted RRTs. The standard error (SE; square root of the variance of the residuals) in the PBDE model was 0.084, corresponding to a percent coefficient of variation (%CV) of 2.3%. The observed strong, positive regression coefficient for (no. X)$^{1/2}$ is intuitive, because increasing the number of Br substituents increases the MW and decreases the tendency of a molecule to partition into the vapor phase. Therefore, PBDE congeners with greater numbers of Br substituents are more attracted to the stationary phase and take a longer time to elute than lower brominated congeners. A negative regression coefficient was observed for IP, similar to that previously reported for PCBs.[4] This can be understood in terms of dispersive (London) forces, which have been identified as a form of nonspecific, cohesive solute−solvent interaction force that helps control the attraction of an individual molecule to the stationary phase and, hence, its retention time. Dispersive forces, although also governed by the molecular polarizability, are directly proportional to $(IP_A IP_B)/(IP_A + IP_B)$, where the subscripts A and B refer to the analyte and stationary phase, respectively.[4,14] IPs are generally, to a first approximation, on the order of 10 eV for most

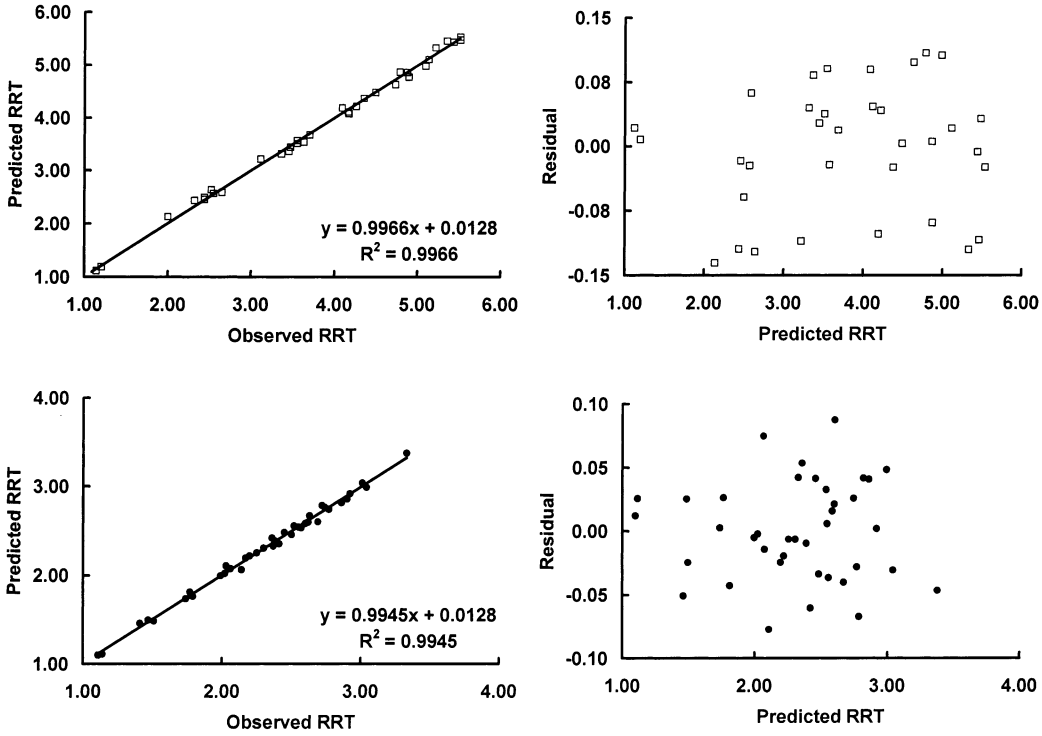

Figure 2. RRT models for PBDEs (□) and PCDEs (●). Graphs on right show distribution of residuals over the range of predicted RRTs.

organic molecules of environmental relevance. Thus, the strength of the dispersion forces (and retention time) should be positively related to the IP of the analyte (as $(IP_AIP_B) \gg (IP_A + IP_B)$). The dipole moment has a positive regression coefficient in the PBDE model. Although the DB-5 stationary phase ((5%-phenyl)-methylpolysiloxane) is relatively nonpolar, it is more polar than the helium mobile phase. Thus, larger dipole moments result in increased PBDE retention times because of the greater attraction [inductive (Debye) and orientation (Keesom) forces] between the analyte and stationary phase than between the analyte and mobile phase.

Negative regression coefficients were found for all three of the halogen substitution variables in the PBDE model. It is unclear why all three of the variables (no. *o*-X, no. *m*-X, and no. *p*-X) would have negative regression coefficients, although similar results have been reported for PCDEs [7] and PCBs.[5,6] For PCBs, the rationalization of a negative *o*-Cl coefficient was an increase in the twist angle[5] between the two biphenyl moieties due to steric and electronic repulsion. For PBDEs, however, molecular modeling using MOPAC shows relatively little effect on three-dimensional structure from Br substitution on PBDEs because of the configuration of the ether linkage (i.e., the diphenyl ether structure is never planar, regardless of substitution). Substitution patterns would determine differences in the dipole moment and ionization potential between congeners sharing the same homologue group, so perhaps what we observe in the differing values of b_5, b_6, and b_7 for PBDEs are the relative influences of Br location on the dipole moment and ionization potential. In other words, *o*-Br are directed approximately toward the orientation of the lone pairs residing on the ether−oxygen linkage, whereas *m*-Br are directed approximately orthogonal to these lone pairs. *p*-Br are, among these three substitution patterns, oriented most strongly with that of the oxygen lone pairs. On the nonsubstituted diphenyl ether (μ = 1.3992 D), it is the O atom that determines the overall dipole moment (Mulliken charge ≈ -0.129), since the benzene rings are nonpolar on their own, although the electronegative O induces small positive charges on the adjacent aryl carbons (Mulliken charge ≈ 0.068). Hence, we observe the magnitude of these three regression coefficients decreases in the order para > meta > ortho, which is similar to their orientation with respect to the oxygen lone pairs. It should be noted that because of space constraints, detailed molecular rationalizations for the observed regression coefficients in each model cannot be presented; however, the qualitative reasoning used above can also be applied to the subsequent models.

A similarly strong model was constructed for the 39 PCDE congeners from mono- through decachlorinated under consideration (Figure 2), with an r^2 of 0.9945, an F value of 797 ($F_{obs} > F_{crit} = 2.3$; $p < 3.7 \times 10^{-33}$), and an even distribution of residuals over the range of predicted RRTs. The SE in the PCDE model was 0.043, corresponding to a %CV of 1.9%. In sharp contrast to the PBDE model, strong positive regression coefficients for no. *o*-Cl; no. *m*-Cl; and no. *p*-Cl were observed in the PCDE model. As with PBDEs, PCDEs are never planar, regardless of substitution pattern, because of the steric and electronic influence of the ether linkage. Despite the increased inductive electron withdrawing ability of a Cl substituent over that of a Br, Cl substituents tend to have an overall positive charge in PCDEs (as do Br substituents on PBDEs). This is a result of the resonance electron donating ability of halogens, which by way of their lone pairs of electrons, can donate electron density into the aromatic rings and thereby accrue a small positive charge. Although Cl is more inductively withdrawing than Br by way of its electronegativity,

(14) Vernon, F. *Dev. Chromatogr.* **1978**, *1*, 1−39.

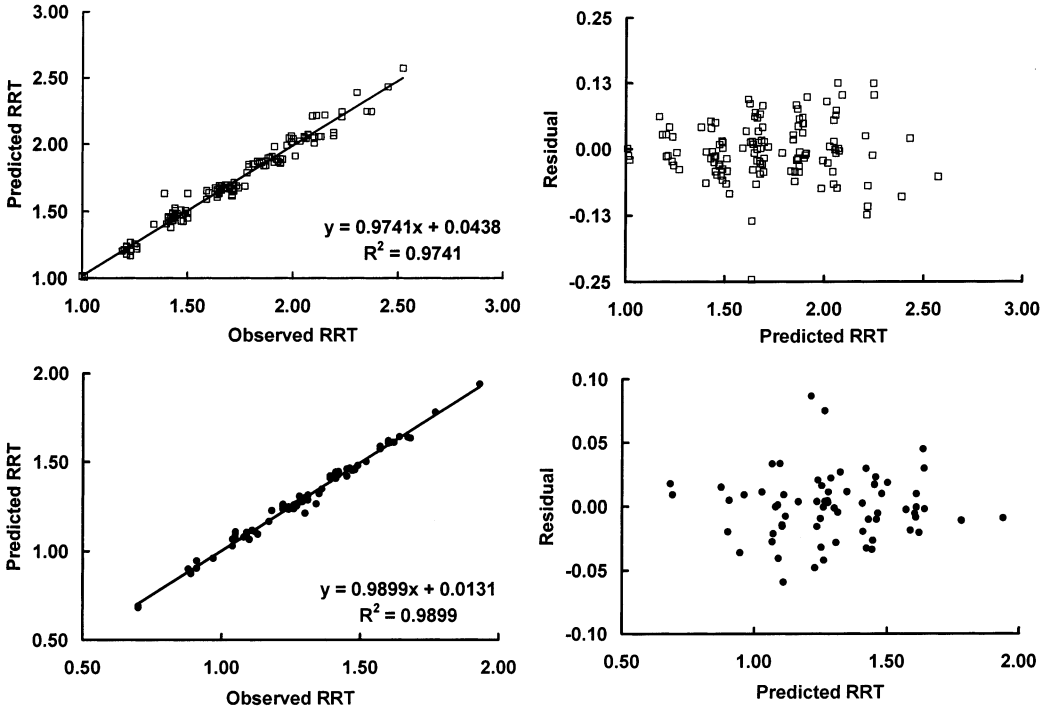

Figure 3. RRT models for PCBs (□) and PCNs (●). Graphs on right show distribution of residuals over the range of predicted RRTs.

Cl is a better resonance electron donor than Br, to such an extent that Cl substituents on PCDEs tend to have a greater positive charge than Br substituents on PBDEs (0.05806 on 2-CDE1 vs 0.04502 on 2-BDE1). For PCDEs, unlike with PBDEs, the values of b_5, b_6, and b_7 are quite similar, suggesting the effect of substitution pattern on retention time is less related to an electronic property, such as the dipole moment (as evidenced by the low value of the regression coefficient for μ; $b_4 = 0.056$), than to a steric property, such as molecular volume.

For PCBs, the model was weaker (Figure 3; $r^2 = 0.9741$) than with PBDEs and PCDEs and certainly poorer than other regression models for PCBs in the literature ($r^2 > 0.99$; see refs 5,6), but comparable with a value reported for PCBs in a similar attempt to link multiple compound classes together in a predictive RRT model ($r^2 = 0.964$, see ref 4). The F value was 660 ($F_{obs} > F_{crit} = 2.1$; $p < 2.6 \times 10^{-94}$), and an even distribution of residuals over the range of predicted RRTs was observed for the 131 congeners used in the model. The SE was 0.055, corresponding to a %CV of 3.3%. In the PCB model, no. *o*-X, no. *m*-X, and no. *p*-X had negative regression coefficients, and the values of these were quite large and did not vary much ($b_5 = -6.86$, $b_6 = -6.85$, and $b_7 = -6.82$). The signs of these values are consistent with PCB−RRT models published previously.[5,6] The small difference between coefficients is not suggestive as to whether the effects of substitution pattern are mainly steric or electronic. This contrasts with PBDEs and PCDEs in that the parent biphenyl system for PCBs are nonpolar, and Cl substitution anywhere on the biphenyl system will have less restricted effects on the dipole moment than with the halogenated diphenyl ether systems, where the halogen may enhance, counter, or have little effect on the dipole induced by the ether linkage. Negative regression coefficients suggest substitution at these positions decreases retention time, possibly by *o*-Cl increasing the twist angle at the biphenyl linkage, as well as all three types of substitution increasing the molecular volume, such as to distribute the molecular mass over a greater region and thereby lower the molecular density. This would result in a higher vapor pressure and greater affinity for the mobile phase. Ionization potential and dipole moment play rather minor roles in determining the RRT of PCBs, as evidenced by their relatively small regression coefficients.

The PCN model also demonstrated a high level of predictability (Figure 3) for 70 congeners ranging from mono- through octachlorinated, with an r^2 value of 0.9899, an F value of 1033 ($F_{obs} > F_{crit} = 2.2$; $p < 6.5 \times 10^{-61}$), and an even distribution of residuals over the range of predicted RRTs. The SE was 0.026, corresponding to a %CV of 2.0%. For PCNs, the four Cl substituents adjacent to the biaryl linkage were modeled as *o*-Cl, whereas the more distant four Cl substituents were considered as *m*-Cl in the model. Because no Cl could be considered para substituents, this field was set to zero for all PCN congeners. PCDDs and PCDFs were modeled in a fashion similar to that for PCNs. Neither PCDDs nor PCDFs have substituents that are accurately considered *p*-Cl; thus, this field was set at zero for the regression analysis. Again, we must emphasize that using a field set to zero has no practical purpose for a single model describing, say, the GC behavior of PCNs and PCDD/Fs, but is important in allowing the linking of GC behavior between compound classes that do and do not have such possible substitution patterns. PCDDs, because of their symmetry about the ether linkages, were modeled as per PCNs. Substituents adjacent to the ether linkage were treated as *o*-Cl, whereas the remaining four Cl were considered as meta substituents. The resulting model had a strong fit (Figure 4; $r^2 = 0.9855$) for 33 congeners from mono- to octachlorinated, with an F value of 476 ($F_{obs} > F_{crit} = 2.5$; $p < 1.7 \times 10^{-16}$) and an even distribution of residuals over the range of predicted RRTs. The SE was 0.086, corresponding to a %CV of 4.0%. That the coefficients for no. *o*-X or no. *m*-X did not differ and the low importance of ionization potential and dipole moment in the PCDD RRT model suggest

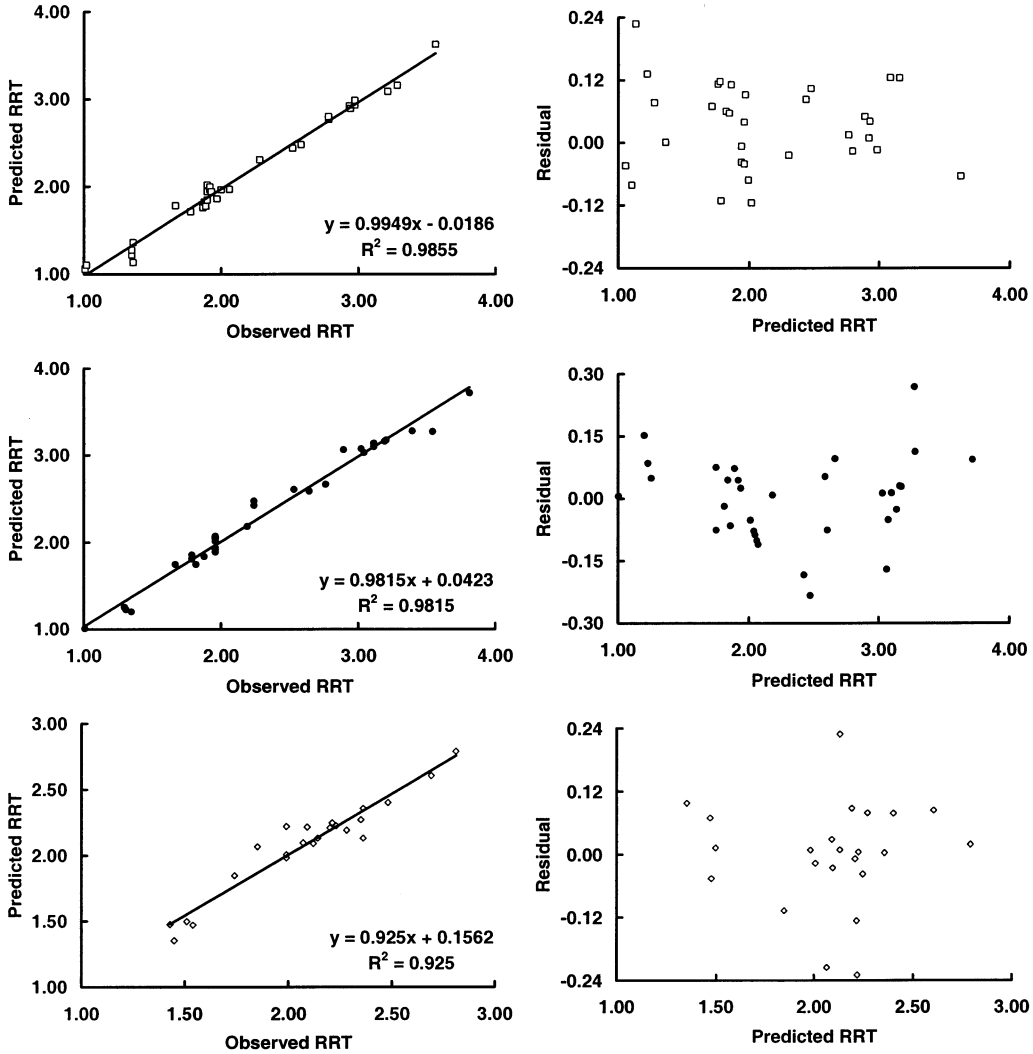

Figure 4. RRT models for PCDDs (□), PCDFs (●), and organochlorine pesticides (◊). Graphs on right show distribution of residuals over the range of predicted RRTs.

that steric considerations are most important in predicting the retention behavior of PCDDs.

The PCDF model had results that were similar to those for PCDDs, with a good, but weaker, fit (Figure 4; $r^2 = 0.9815$) for the 34 congeners from mono- to octachlorinated, an F value of 238 ($F_{obs} > F_{crit} = 2.5$; $p < 4.5 \times 10^{-22}$), and an even distribution of residuals over the range of predicted RRTs. The SE was 0.112, corresponding to a %CV of 4.9%. The slightly poorer predictability of the PCDF model may result from a lack of accounting for the differing diaryl linkages in PCDFs. For PCDDs, both linkages are ether functions, whereas for PCDFs, there is an ether and a direct diaryl linkage. In our model, both linkages were treated equally, such that Cl adjacent to either type of linkage were considered *o*-Cl, whereas the remaining Cl were treated as *m*-Cl. Such an approximation will undoubtedly introduce error, but it was difficult to conceive of another means of dealing with the different linkages, and arbitrarily assigning one of the sets of *o*-Cl (either the two Cl adjacent to the ether linkage, or the two Cl adjacent to the diaryl linkage) as *p*-Cl only acted to decrease the quality of fit. However, the fit we report here is slightly better than that reported for PCDFs in a similar attempt ($r^2 = 0.927$, see ref 4) to link several contaminant classes in a multiclass model.

In contrast to PCDDs and PCNs, the coefficients for no. *o*-X and no. *m*-X in the PCDF model differed in both magnitude and sign (−0.05946 and 0.1507, respectively). Such a difference is expected, because PCDFs have two different aryl linkages, reducing the molecular symmetry and introducing two types of ortho substitution. Molecular modeling using MOPAC indicates the two carbons adjacent to the ether linkage are the only carbons carrying a positive charge (Mulliken population ≈ 0.056); hence, Cl on these carbons adjacent to the ether linkage have the highest positive charges (≈0.125) of any possible Cl substituent. Conversely, the carbon atoms composing the diaryl linkage have negative charges (≈−0.06), and the Cl atoms at this ortho position have the lowest positive charges of any potential Cl substituents (≈0.110). This structure results in the parent dibenzofuran system's having a dipole oriented perpendicularly across the diaryl and ether linkages toward the ether oxygen, unlike the parent dibenzo-*p*-dioxin system, which is nonpolar. Because of this inherent dipole, PCDFs with Cl substituents ortho to the ether linkage increase the dipole moment, but those ortho to the diaryl linkage act to counter the inherent dipole. *m*-Cl are oriented approximately orthogonal to this inherent dipole and, thus, have less overall effect on the dipole moment than *o*-Cl. Because the

regression coefficient for *o*-Cl is negative and lower in magnitude than that for *m*-Cl (−0.05946 vs 0.1507), the effects of substitution on PCDF RRTs are likely steric in nature rather than electronic. Because *o*-Cl have a greater effect on dipole than *m*-Cl, the negative coefficient, if based on electronic grounds, would suggest a more polar molecule has a decreased retention time. Since this does not agree with theoretical expectations, it is likely that *o*-Cl perhaps increase the molecular volume of PCDFs through a decrease in planarity resulting from steric crowding of adjacent *o*-Cl, and between *o*-Cl and the ether and diaryl linkages. This increase in molecular volume would decrease the molecular density over an equivalent planar congener, and this density reduction would increase the affinity of the molecule for the vapor phase and decrease its retention time. On the other hand, *m*-Cl have little effect on the planarity, density, or dipole moment of the molecule; thus, increasing numbers of *m*-Cl would be expected to increase retention time due to the increase in molecular weight.

Perhaps the most robust retention time model is that of the 23 organochlorine pesticides (Figure 4), which although having a poorer fit than the other models ($r^2 = 0.925$), covers a wide range of molecular structures from hexachlorohexane (HCH) through Mirex (see Figure 1 for structures) with a range of 3−12 chlorine substituents. The *F* value was 26 ($F_{obs} > F_{crit} = 2.7$; $p < 1.5 \times 10^{-6}$), and an even distribution of residuals was observed over the range of predicted RRTs. The SE was 0.123, corresponding to a %CV of 5.9%. To allow development of a multiclass model, we chose to approximate the halogen substitution patterns of pesticides to that of an aromatic ring. These approximations clearly suffer from drawbacks, such as how to treat a carbon atom with two halogen substituents (which cannot occur on an aromatic nucleus), which carbon atom to consider the ortho position, and how to deal with two apparent substitution patterns on different regions of the molecule. Rather than attempt a detailed rationalization, we chose a prima facie designation of substitution pattern. In other words, halogen substituents were assigned an ortho, meta, or para position in the model on their subjective similarity to the other aromatic systems under consideration. A more rigorous examination based on molecular symmetry arguments would be desired, but to a first approximation, our model appears to provide a satisfactory screening tool in the determination of reasonable retention time windows for novel organochlorine pesticides and some of their degradation products. The importance of assigned substitution pattern on predicted RRT is less than that of molecular weight, (no. X)$^{1/2}$, and ionization potential, but with the exception of no. *o*-Cl ($b_4 = 4.1 \times 10^{-3}$), which played a very minor role in the model, substitution pattern was a moderate predictor that helped differentiate between pesticides of similar mass, but differing structure (e.g., the isomers of DDD, DDE, and DDT).

There were an insufficient number of identified PBDD/F congeners ($n = 10$) to develop a predictive model with 7 variables, as for the other contaminant classes. Thus, we developed a model utilizing only four variables: molecular weight, (no. X)$^{1/2}$, ionization potential, and dipole moment. However, with four variables and 10 congeners, such a model is not rigorous and serves more as an example that such techniques can also be applied to predicting PBDD/F RRTs. This model is presented in Table 1, and had a strong fit ($r^2 = 0.999$), with an *F* value of 1207 ($F_{obs} >$ $F_{crit} = 5.1$; $p < 1.2 \times 10^{-7}$), a SE of 0.02, and a %CV of 0.7%. All four variables had positive regression coefficients, consistent with the theoretical understanding presented above. Molecular weight was the strongest predictor, followed by (no. X)$^{1/2}$, ionization potential, and dipole moment, in decreasing order. The lack of congeners and combination of both PBDDs and PBDFs into the same regression model precludes a more detailed analysis.

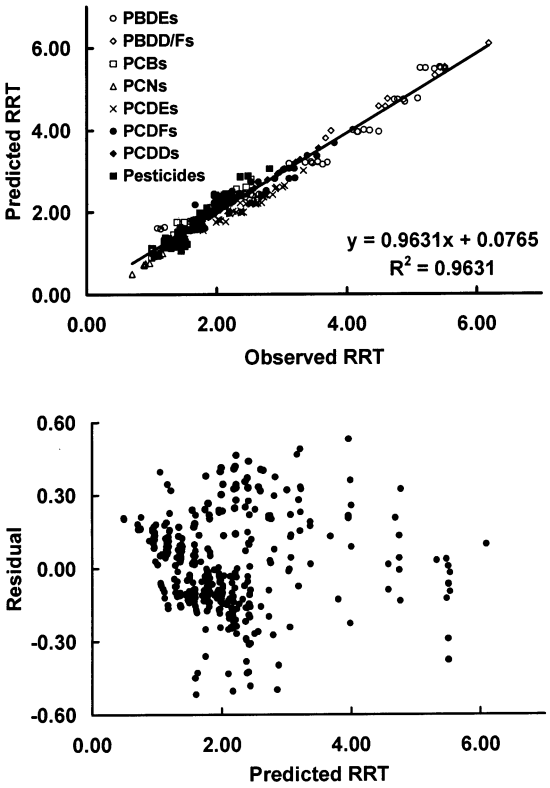

Figure 5. Multiclass RRT model for the nine classes of halogenated organic contaminants under cnsideration and the distribution of residuals over the range of predicted RRTs.

The Multiclass Model. A GC retention time model incorporating the retention behavior of nine major contaminant classes was developed (Figure 5). The multiclass RRT model for the 375 individual compounds had a moderately good fit ($r^2 = 0.9631$), with an *F* value of 1369 ($F_{obs} > F_{crit} = 5.1$; $p < 1.2 \times 10^{-7}$), a SE of 0.19, and a %CV of 9.2%. Positive regression coefficients were obtained for molecular weight, (no. X)$^{1/2}$, ionization potential, and dipole moment, as expected on the theoretical grounds presented above. The strongest predictor among all variables was molecular weight, similar to most other individual models, followed by the three variables representing the halogen substitution pattern (no. *o*-, no. *m*-, and no. *p*-X). All three of the substitution variables had negative regression coefficients ($b_5 = -0.240$, $b_6 = -0.249$, and $b_7 = -0.195$). The wide range of compounds encompassed by these coefficients makes it difficult to determine whether their values result from steric or electronic effects or both, although both effects are likely to be operative.

The development of such a GC-RRT model incorporating a wide range of compound classes and molecular structures is desirable from the standpoint of screening environmental samples for new and emerging contaminants, as well as potential degradation products of well-established contaminants. Our knowledge of the need for detailed analytical work on environmental samples

is increasing, driven by concerns over acute and chronic toxicity to living organisms. Thus, such models may serve an important role for researchers and practitioners in that they allow rapid development of gas chromatographic temperature programs to test whether specific compounds are present in a sample. Our model is not intended to allow the prediction of an analyte's retention time within the window of time (typically 2−3 s) necessary for immediate identification and quantitation. Rather, on the basis of already established GC programs for another class of compounds covered in the model (e.g., PCBs, PCDD/Fs, etc.), the model allows a user to predict a suitable window in which to run an SIM experiment looking for the major isotopic masses of the potential analyte.

For example, in a paper recently published,[15] we examined the aqueous photochemistry of 2378-TeCDD, the most toxic dioxin congener for which the photodegradation pathways were largely unknown. Much of the difficulty in elucidating the photoproducts of such trace environmental contaminants lies in the analysis of the product mixtures. As is the case with biodegradation studies, although the parent compound is typically available in large quantities, analytical standards are usually unavailable for the degradation products and are only possible through complex synthetic work. In the case of 2378-TeCDD, we showed the major, primary photoproduct (>50%) to be 2,2′-dihydroxy-4,4′,5,5′-tetrachlorobiphenyl (4,4′,5,5′-TeCDHBP), which subsequently dechlorinates through a series of tri-, di, and monochlorinated isomers to the parent 2,2′-dihydroxybiphenyl (DHBP). Another photoproduct class from the irradiation of 2378-TeCDD was a suite of mono- through tetrachlorophenoxyphenols, including 4,4′,5,5′-tetrachlorophenoxyphenol (4,4′,5,5′-TeCPP) as a minor primary photoproduct. Since analytical standards were available only for 4,4′,5,5′-TeCDHBP, DHBP, and the parent phenoxyphenol (PP), these compounds only set upper and lower boundaries (∼30 min or ±2 RRTs using the scale presented in this paper) for where the expected photoproducts would elute. At trace levels of these compounds, numerous potential product peaks were evident from our SIM analyses, which monitored the two most abundant expected isotopes of each photoproduct. Each peak had to be rigorously examined as a potential analyte, with little assistance from predicted GC windows, such as the current model would provide. As well, several isomers of each homologue group (i.e., dichlorodihydroxybiphenyls) were possible photoproducts, and in some cases, several were observed but could not be separately identified on the basis of their mass spectra.

To test the multiclass RRT model presented here, we calculated the necessary MOPAC physicochemical properties for some of the 2378-TeCDD photoproducts discussed above and input them into the model to generate their predicted RRTs for comparison with known and assumed RRTs. The purpose of this was not only to test the current model, but also to add further validation to the proposed photoproduct identities reported in our previous paper and to demonstrate the wide utility of such RRT models in environmental monitoring and contaminant degradation studies. The observed and predicted RRTs (adjusted to be relative to 2,2′-PCB4; ± standard error) were as follows: 4,4′,5,5′-TeCDHBP (RRT_{obs} = 2.17, RRT_{pred} = 2.02 ± 0.20); 4,5,5′-TrCDHBP (RRT_{obs} = 1.68, RRT_{pred} = 1.81 ± 0.20); 4,4′,5-TrCDHBP (RRT_{obs} = 1.68, RRT_{pred} = 1.78 ± 0.20); 4,5′-DiCDHBP (RRT_{obs} = 1.39, RRT_{pred} = 1.56 ± 0.20); 4,4′,5,5′-TeCPP (RRT_{obs} = 2.06, RRT_{pred} = 1.97 ± 0.19). For this group of compounds, the model has reasonable predictive ability, and all analytes were within the standard error of the predicted RRT. That such phenolic compounds were not within the classes making up the model demonstrates its robustness. As well, the model is able to predict relative elution orders for isomeric homologues (e.g., 4,5,5′-TrCDHBP and 4,4′,5-TrCDHBP). Thus, where peaks for all possible isomers are observed and mass spectra are unable to provide conclusive structural proof for each isomer, the use of this model to predict retention orders should aid in structural identification.

In addition, if the predicted retention time for an analyte is not practical for the analyst, the temperature program could be adjusted accordingly, and as long as an internal standard (i.e., 2,2′-PCB4, or any other compound suitable as an RRT "anchor" in the model) is used whose retention behavior relative to the potential analyte is known, the analyte's "new" retention time can be calculated. The resulting RRT window calculations presented above are a useful first approximation for a SIM analysis to screen for potential contaminants and to predict relative elution orders of isomers, after which more detailed analytical methods can be used to refine the procedure. Such tools have the potential to greatly reduce the time and guesswork necessary in identifying novel environmental contaminants.

ACKNOWLEDGMENT

M.G.I. thanks DFO for funding and Maike Fischer and Tim He for their help in the HRGC/HRMS work.

SUPPORTING INFORMATION AVAILABLE

Values of each of the descriptors and the relative retention times for the 375 compounds of interest is available as Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review June 21, 2002. Accepted November 26, 2002.

AC020406P

(15) Rayne, S.; Wan, P.; Ikonomou, M. G.; Konstantinov, A. D. *Environ. Sci. Technol.* **2002**, *36*, 1995−2002.