Optimal Selection of Gas Chromatographic Columns for the Analytical Control of Chemical Warfare Agents by Application of Information Theory to Retention Data

5 AUTHORS, INCLUDING:

Ernst Kenndler
University of Vienna
258 PUBLICATIONS   **6,364** CITATIONS

SEE PROFILE

Gregor Reich
University of Vienna
12 PUBLICATIONS   **90** CITATIONS

SEE PROFILE

Optimal Selection of Gas Chromatographic Columns for the Analytical Control of Chemical Warfare Agents by Application of Information Theory to Retention Data

Josef F. K. Huber,* Ernst Kenndler, Gregor Reich, Werner Hack, and Josef Wolf
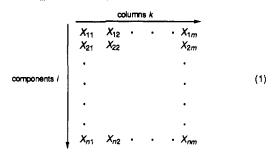
Institute for Analytical Chemistry, University of Vienna, Währingerstrasse 38, A-1090 Vienna, Austria

Retention data can be used to characterize and identify the components of a mixture. The identification power of a single column is rather limited, however. With several columns of different retention characteristics, a high identification power can be achieved. In order to apply chromatographic retention data for the identification of chemical species it is therefore important to select columns with low correlation of their retention characteristics. The retention behavior of five gas chromatographic columns was compared in pairs by linear regression of the retention data of a test set of compounds. The correlation coefficient of the two sets of retention data is an estimate of the degree of correlation of the columns. It was found that five pairs of columns had correlation coefficients of less than 0.900, the lowest value being 0.769. A more exact approach for the selection of columns for the chromatographic identification of compounds is based on the application of information theory. The information content of retention data (in bits) can be calculated from their probability distribution function. The information content was calculated for the same five gas chromatographic columns with the same retention data used in the calculation of the correlation coefficients. The optimal choice with respect to information power and effort is a combination of three columns having an information content of 19.8 bits. The loss of information caused by the correlation of the column retention amounts -2.7 bits in this case. The method of column selection was applied for the development of a gas chromatographic screening method for the identification of chemical warfare agents, precursors, and decomposition products. A test set of 16 chemically different species of such compounds was used.

INTRODUCTION

Chemical compounds are identified by the measurement of properties which depend on their chemical nature. Chromatographic retention is such a property. It depends on the molecular interaction between the solutes and a liquid solvent or a solid surface. Therefore retention data can be used to characterize and identify chemical species. In general, retention data on several chromatographic columns with different retention characteristics have to be determined for a safe identification. In this manner the sample is described by a multidimensional data matrix. Each component is determined by a m-dimensional vector and is depicted by a point in the m-dimensional space. The identification is

retention data matrix x_{ik}

$$\text{(1)}$$

performed by comparing the measured data with those of a library. The certainty of the identification depends on the precision and accuracy and the dimensionality of the data.

High-performance chromatography can be applied in the determination of retention data for identification. In gas chromatography, so-called retention indices are defined in order to reduce the process parameters to those depending on the chemical nature of the components. This aim is achieved by a logarithmic standardization of the retention data by means of a reference scale defined by the n-alkanes.

The number of components which can be differentiated by one or more chromatographic columns is a measure of the column quality for identification. Since the identification can be carried out only with a given probability, it is important to quantify the certainty of identification and to select conditions giving optimal results. This target can be achieved by applying chemometric methods,[1,2] especially information theory.[3,4]

The information content, H, of a continuous signal can be calculated if the probability density function is known. If this function is unknown an approximate calculation of the information content can be carried out by means of a histogram with $z = ld(n + 1)$ intervals[5,6] of an equal width, Δx, where n is the number of data (=components). The information content is calculated by means of the following equation.

$$H = -\sum_{r=1}^{z} p_r \, ld \, p_r + ld \, \Delta x_r - ld \, \sigma_E(2\pi e)^{1/2} \qquad (2)$$

with $\Sigma p_r = 1$, where p_r is the probability of finding a particular retention index within the interval, Δx_r; σ_E is the error defined by the standard deviation of the measurement of the retention

(1) Massart, D. L.; Dijkstra, A.; Kaufman, L. *Evaluation and Optimization of Laboratory Methods and Analytical Procedures*; Elsevier: Amsterdam, 1978.
(2) Massart, D. L.; Vandeginste B. G. M.; Deming, S. N. L. *Chemometrics: A Textbook*, Elsevier: Amsterdam, 1988.
(3) Huber, J. F. K.; Smit, H. C. *Z. Anal. Chem.* **1970**, *254*, 84-88.
(4) Huber, J. F. K.; Kenndler, E.; Reich, G. *J. Chromatogr.* **1979**, *172*, 15-30.
(5) Sturges, H. A. *J. Am. Stat. Assoc.* **1926**, *21*, 65.
(6) Kenndler, E. *Anal. Chim. Acta* **1985**, *173*, 239-251.

Table I. Retention Indexes of Chemical Warfare Agents, Precursors, and Decomposition Products Measured at Two Different Temperatures[a]

analyte	retention index ±SD (temp, °C)				
	MS	MS5	MS50	PEG	CPS
3,3-dimethyl-2-butanol	722.8 ± 1.7 (80)	735.6 ± 0.5 (80)	812.7 ± 0.7 (60)	1118.9 ± 0.6 (90)	1191.4 ± 1.2 (40)
	721.7 ± 1.9 (100)	736.3 ± 0.5 (100)	821.5 ± 1.5 (80)	1114.0 ± 0.4 (110)	1212.5 ± 0.7 (60)
methyl sulfoxide	777.3 ± 0.8 (80)	829.4 ± 0.5 (80)	1140.8 ± 0.3 (100)	1569.1 ± 0.3 (130)	1999.9 ± 0.4 (110)
	780.5 ± 1.4 (100)	832.7 ± 0.5 (100)	1144.1 ± 0.9 (120)	1584.0 ± 0.4 (150)	2084.7 ± 1.1 (130)
methylphosphonic dichloride	802.1 ± 0.4 (100)	844.6 ± 0.4 (110)	e	e	e
	805.0 ± 1.1 (120)	848.5 ± 1.3 (130)			
methylphosphonic acid[b]	838.9 ± 0.1 (80)	875.5 ± 0.1 (80)	1132.7 ± 1.0 (100)	1493.6 ± 0.4 (130)	1784.3 ± 2.1 (90)
	838.8 ± 0.7 (100)	875.7 ± 0.1 (100)	1130.3 ± 0.4 (120)	1501.3 ± 0.6 (150)	1847.9 ± 0.5 (110)
methyl sulfone	849.0 ± 2.1 (80)	914.1 ± 0.4 (80)	1243.6 ± 0.4 (100)	1900.4 ± 0.4 (160)	2538.0 ± 0.8 (150)
	852.2 ± 2.0 (100)	915.5 ± 0.3 (100)	1251.2 ± 0.5 (120)	1913.1 ± 0.5 (180)	2655.8 ± 0.3 (170)
diisopropyl fluorophosphate (DFP)	929.2 ± 0.1 (80)	956.4 ± 0.6 (110)	1123.0 ± 0.5 (100)	1341.6 ± 0.2 (90)	1605.0 ± 0.4 (90)
	927.4 ± 0.6 (100)	954.7 ± 0.3 (130)	1119.4 ± 0.4 (120)	1342.0 ± 0.4 (110)	1646.1 ± 2.1 (110)
1,4-dithiane	1038.0 ± 0.4 (100)	1083.9 ± 0.4 (110)	1341.1 ± 0.6 (130)	1618.9 ± 0.2 (130)	1780.0 ± 0.6 (90)
	1049.7 ± 1.6 (120)	1097.6 ± 0.8 (130)	1364.6 ± 0.5 (150)	1641.3 ± 0.2 (150)	1851.1 ± 1.7 (110)
N-methylbis(2-chloroethyl)amine (HN2)	1052.5 ± 0.3 (100)	1090.7 ± 0.3 (110)	1283.4 ± 0.1 (130)	1566.3 ± 0.5 (130)	1784.3 ± 0.5 (90)
	1058.5 ± 0.2 (120)	1096.8 ± 0.3 (130)	1292.9 ± 0.3 (150)	1574.5 ± 0.7 (150)	1841.2 ± 0.7 (110)
triethyl phosphate	1081.4 ± 0.7 (120)	1118.5 ± 0.4 (140)	1362.1 ± 0.7 (130)	1656.1 ± 0.7 (130)	2011.2 ± 0.3 (110)
	1079.4 ± 1.0 (140)	1118.0 ± 0.4 (160)	1360.2 ± 0.6 (150)	1659.6 ± 0.5 (150)	2064.4 ± 1.8 (130)
thiodiglycol	1132.6 ± 1.4 (120)	1190.4 ± 0.6 (140)	1514.3 ± 0.2 (160)	2382.8 ± 1.0 (190)	3064.3 ± 2.1 (180)[d]
	1140.4 ± 0.5 (140)	1199.5 ± 0.5 (160)	1534.0 ± 1.2 (180)	2393.1 ± 0.4 (210)	3190.9 ± 1.7 (200)[d]
3-quinuclidinol	1151.9 ± 0.9 (120)	1200.7 ± 0.7 (140)	1466.4 ± 1.5 (130)	1985.4 ± 0.7 (190)	2326.8 ± 0.4 (130)
	1162.5 ± 0.5 (140)	1213.3 ± 0.2 (160)	1478.8 ± 0.5 (150)	2001.8 ± 0.5 (210)	2417.5 ± 0.7 (150)
methyl salicylate	1174.7 ± 1.1 (120)	1216.4 ± 0.3 (140)	1426.7 ± 0.6 (130)	1808.6 ± 0.4 (160)	1998.3 ± 0.4 (110)
	1184.6 ± 0.4 (140)	1226.7 ± 0.3 (160)	1444.0 ± 0.3 (150)	1826.5 ± 0.5 (180)	2061.6 ± 1.5 (130)
2-chloracetophenone (CN)	1231.2 ± 0.4 (100)	1301.5 ± 0.2 (140)	1609.6 ± 0.3 (160)	2101.7 ± 0.4 (160)	2501.9 ± 0.4 (130)
	1240.9 ± 0.3 (120)	1312.1 ± 0.2 (160)	1628.4 ± 0.1 (180)	2113.9 ± 0.6 (180)	2581.8 ± 0.7 (150)
tris(2-chloroethyl)amine (HN3)	1351.0 ± 0.3 (120)	1414.1 ± 0.1 (140)	1682.4 ± 1.1 (160)	2089.1 ± 0.3 (160)	2479.7 ± 0.3 (130)
	1361.4 ± 0.2 (140)	1423.7 ± 0.2 (160)	1698.8 ± 0.1 (180)	2101.6 ± 0.3 (180)	2560.7 ± 0.7 (150)
tributyl phosphate	1612.5 ± 1.1 (180)	1650.9 ± 0.3 (200)	1882.5 ± 0.6 (190)	2117.2 ± 0.4 (190)	2539.7 ± 0.6 (150)
	1612.2 ± 1.2 (200)	1651.0 ± 0.3 (220)	1883.8 ± 0.4 (210)	2118.2 ± 0.7 (210)	2587.5 ± 2.1 (170)[d]
bencilic acid[b]	1791.8 ± 0.2 (180)	1865.4 ± 0.3 (200)	2291.3 ± 1.0 (220)	2852.2 ± 1.4 (190)	3367.1 ± 1.6 (180)[c,d]
	1805.2 ± 1.4 (200)	1880.8 ± 0.7 (220)	2236.1 ± 2.1 (240)	2833.6 ± 2.5 (210)	3502.5 ± 9.4 (200)[c,d]

[a] For column code see Experimental Section. [b] Measured as methyl ester derivative. [c] Extrapolation of the retention index to infinite dilution because of an asymmetrical peak of the analyte. [d] Retention index calculated from retention data of the standards extrapolated to infinite dilution because of asymmetrical peaks of the n-alkanes. [e] Compound decomposes on polar columns.

index; $2\pi e$ is a constant factor (with e, the base of the natural logarithm).

If retention indices are measured on m columns with different retention characteristics an m-dimensional histogram results and the information content is calculated as follows.

$$H(1,2,...,m) = -\sum_{r=1}^{z,m} p_r(x_1,x_2,...,x_m) \, \text{ld} \, p_r(x_1,x_2,...x_m) +$$

$$\text{ld}(\Delta x_1 \cdot \Delta x_2 ... \Delta x_m) - \text{ld}(2\pi e)^{m/2} \sigma_E^m \quad (3)$$

The function $p_r(x_1,x_2,...x_m)$ describes the probability of finding the retention indices of a compound in given intervals of the m-dimensional histogram.

On November 30, 1992, the General Assembly of the United Nations adopted the *Convention for the Control of Chemical Weapons* after more than 20 years of negotiations. In practise an optimal compromise between expenditure and safety has to be found for the control procedure. From the analytical viewpoint this goal can be achieved by starting the investigation with a screening by GC followed by GC/MS for final confirmation in the case of suspicious samples.[7-9]

In this work the smallest number of columns with different stationary phases necessary for the safe identification of chemical warfare agents, precursors, and decomposition products by gas chromatographic retention data is evaluated.

<del id="del">(7) D'Agostino, P. A.; Provost, L. R. *J. Chromatogr.* 1985, *331*, 47–54.
(8) *Standard Operating Procedures for the Verification of Chemical Disarmament*; Vol. D.2., *Second Proposal for Procedures Supporting the Reference Database*; Ministry of Foreign Affairs of Finland: Helsinki, 1989.
(9) *Report of the ad hoc Committee on Chemical Weapons to the Conference on Disarmament*; United Nations: Geneva, August 1990.</del>

In a previous paper[10] it was shown that a relatively small test set of properly chosen retention data can represent the retention characteristics of a large number of components without significant loss of information. Therefore 16 species of chemical warfare agents, precursors, and decomposition products were selected as test set representing all types of molecular interactions occurring in this group of compounds.

EXPERIMENTAL SECTION

Apparatus. The experiments were carried out with a gas chromatograph (Model 5890 series II, Hewlett Packard, Palo Alto, CA) equipped with flame ionization detector (FID), split/splitless injector and autosampler (Model 7673, Hewlett Packard). The gas chromatographic columns (all Hewlett Packard, except CPS (Supelco Technologies, Bellefonte, PA); all 530-μm inner diameter) are specified as following: methyl silicone (code MS), 5-m length, film thickness 2.65 μm; methyl silicone with 5% phenyl (code MS5), 10-m length, film thickness 2.65 μm; methyl silicone with 50% phenyl (code MS50), 10-m length, film thickness 2.0 μm; poly(ethylene glycol) (code PEG), mean molecular weight 20 000, 10-m length, film thickness 1.33 μm; methylcyanopropyl silicone (code CPS), 15-m length, 0.2 μm thickness. The hydrogen for the FID was supplied by a hydrogen generator (Model 7525, Packard, Downers Grove, IL).

The signal processing and the control of the chromatograph was made with a computer-based Chromatographic Worksystem (Model HP 3359, Hewlett Packard).

Chemicals. The following solvents and reagents were used in the sample pretreatment and the preparation of sample solutions: n-pentane, diethyl ether, methanol, KOH (all analytical grade, E. Merck, Darmstadt, Germany). The compounds selected as test set for the optimization of identification of chemical

<del id="del">(10) Huber, J. F. K.; Reich, G. *J. Chromatogr.* 1984, *294*, 15.</del>

Table II. Retention Indexes of Chemical Warfare Agents, Precursors, and Decomposition Products Extrapolated for a Constant Temperature on Each Column[a]

	retention index				
analyte	MS (120 °C)	MS5 (120 °C)	MS50 (130 °C)	PEG (130 °C)	CPS (110 °C)
3,3-dimethyl-2-butanol	720.6	737.0	843.5	1109.1	1265.3
methyl sulfoxide	783.7	836.0	1145.8	1569.1	1999.9
methylphosphonic acid[b]	838.7	875.9	1129.1	1493.6	1847.9
methyl sulfone	855.4	916.9	1255.0	1881.4	2304.0
diisopropyl fluorophosphate (DFP)	925.6	955.6	1117.6	1342.4	1646.1
1,4-dithiane	1049.7	1090.7	1341.1	1618.9	1851.1
N-methylbis(2-chloroethyl)amine (HN2)	1058.5	1093.8	1283.4	1566.3	1841.2
triethyl phosphate	1081.4	1119.0	1362.1	1656.1	2011.2
thiodiglycol	1132.6	1181.3	1484.8	2351.9	2621.2[d]
3-quinuclidinol	1151.9	1181.1	1466.4	1936.2	2236.1
methyl salicylate	1174.7	1206.1	1426.7	1781.8	1998.3
2-chloracetophenone (CN)	1240.9	1290.9	1581.4	2083.4	2422.0
tris(2-chloroethyl)amine (HN3)	1351.0	1404.5	1657.8	2070.4	2398.7
tributyl phosphate	1613.4	1650.5	1878.6	2114.2	2444.1[d]
bencilic acid[b]	1751.6	1803.8	2539.7	2908.0	2893.2[c,d]

[a] For column code see Experimental Section. [b] Measured as methyl ester derivative. [c] Retention index extrapolated to infinite dilution because of peak asymmetry of the analyte. [d] Retention index calculated from the retention times of n-alkanes extrapolated to infinite dilution because of peak asymmetry of the standards.

warfare agents, precursors, and decomposition products by means of gas chromatographic retention data are presented in Table I. They were of analytical grade (Aldrich-Chemie, Steinheim, Germany; except 3-quinuclidinol, Sigma, Deisenhofen, Germany). For the determination of the retention indices n-alkanes, GC-grade (Chrompack Middelburg, The Netherlands) were used as reference components.

Nitrogen of 99.99% (v/v) purity (Air Liquide, Mitry-Mory, France) was used as carrier gas for the columns and make-up gas for the detector. For the FID hydrogen supplied by a hydrogen generator and synthetic air of 99.995% (v/v) purity (Air Liquide) were used as burning gases.

Procedures. *Gas Chromatography.* The gas chromatographic columns were operated at constant temperature, depending on the type of sample with a carrier gas flow rate of 10 mL/min. The temperature of the injector was 250 °C, that of the detector 350 °C.

Sample Pretreatment. Free benzilic acid and methylphosphonic acid cannot be run directly on a GC and were derivatized to methyl esters by means of diazomethane in the usual way.

The chemical warfare agents of the 2-chloroamine type (HN2 and HN3) are supplied as hydrochlorides. They have to be extracted as amines prior to gas chromatography. For this purpose 6.2 mg of HN2 hydrochloride or 5.9 mg of HN3 hydrochloride are dissolved in 2 and 1.5 mL, respectively, of 0.1% aqueous solution of KOH and extracted with 5 mL of diethyl ether. In both cases a solution of 1 mg/mL base is obtained. Because of the toxicity of the agents, a glovebox is used for this handling as well as for the handling of diisopropyl fluorophosphate (DFP).

Safety Considerations. For the manipulation of the following compounds special care must be taken. 2-Chloroacetophenone (CN): toxic by inhalation, in contact with skin, and if swallowed. N-Methylbis(2-chloroethyl)amine hydrochloride: very toxic by inhalation, in contact with skin, and if swallowed. May cause cancer. Tris(2-chloroethyl)amine hydrochloride: very toxic by inhalation, in contact with skin, and if swallowed. May cause cancer. Diisopropyl fluorophosphate (DFP): nerve gas; very toxic by inhalation, in contact with skin, and if swallowed.

RESULTS AND DISCUSSION

Estimation of Limiting Retention Indices at a Given Temperature. The retention indices of all test compounds were measured at two temperatures on each column. The results are presented in Table I together with the measuring error, given by the standard deviation obtained from five measurements.

It must be pointed out that, for an asymmetric peak, the retention index is not constant but depends on the analyte

Table III. Similarity Matrix Based on the Correlation Coefficients for the Linear Regression of the Retention Indexes of the Test Set of Chemical Warfare Agents, Precursors, and Decomposition Products[a]

	correlation coefficient, r			
column	MS	MS5	MS50	PEG
MS5	0.999			
MS50	0.953	0.960		
PEG	0.836	0.850	0.923	
CPS	0.769	0.789	0.852	0.969

[a] For column code see Experimental Section.

concentration. This is the consequence of the nonlinear distribution isotherm of the solute. By extrapolation of the retention index to infinite dilution, the constancy of the retention data can be significantly improved. This extrapolation of retention data for nonsymetric peaks is described in a previous paper[4] and was also carried out for the measurements presented here.

Retention indices measured at different temperatures were inter- or extrapolated to a particular temperature on each column. The following approximate linear approach was used to calculate the retention index, $I_R(T_x)$, from the retention indices $I_R(T_2)$ and $I_R(T_1)$, respectively, measured at the higher temperature T_2 and the lower temperature T_1: $I_R(T_x) = I_R(T_1) + [I_R(T_2) - I_R(T_1)](T_x - T_1)/(T_2 - T_1)$. The error resulting from this simplified calculation is included in the standard deviation, σ_E, of the total error in eqs 2 and 3.

The results obtained after both extrapolations are given in Table II. The total error due to measurement, extrapolation to infinite dilution, and conversion to the particular temperature is estimated to be two retention index units.

Linear Regression of the Retention Index Data on Different Columns in Pairs. The retention indices on different columns are compared in pairs by linear regression. The correlation coefficients are given in Table III. Five pairs of columns have a correlation coefficient of less than 0.900. The columns MS and CPS have the most dissimilar retention characteristics for the compounds under consideration ($r = 0.769$), whereas MS and MS5 exhibit the highest linear correlation ($r = 0.999$). From Table II it can be seen that (a) the overall selectivity of the columns relative to n-alkanes indicated by the retention index range of the test species changes. The corresponding retention index ranges are 1031 (MS), 1067 (MS5), 1696 (MS50), 1628 (CPS), and 1799

Table IV. Information Content, H_i, of Gas Chromatographic Columns for the Identification of Chemical Warfare Agents, Precursors, and Decomposition Products[a]

column	H_i (bit)	column	H_i (bit)
MS	7.2	MS50	7.6
MS5	7.2	PEG	7.8
CPS	7.7		

column combination	H_i (bit)	ΣH_i	ΔH
Two Columns			
MS/MS5	9.7	14.4	-4.7
MS5/MS50	13.0	14.8	-1.8
MS/MS50	13.1	14.8	-1.8
MS5/PEG	14.1	15.0	-0.9
MS/PEG	14.1	15.0	-0.9
MS50/PEG	14.0	15.4	-1.4
MS5/CPS	14.2	14.9	-0.7
MS50/CPS	14.4	15.3	-0.9
PEG/CPS	13.5	15.5	-2.0
MS/CPS	14.2	14.9	-0.7
Three Columns			
MS/MS5/MS50	15.3	22.0	-6.7
MS/MS5/PEG	16.1	22.2	-6.1
MS/MS5/CPS	15.9	22.1	-6.2
MS/MS50/PEG	19.4	22.6	-3.2
MS5/MS50/PEG	19.3	22.6	-3.3
MS/MS50/CPS	19.8	22.5	-2.7
MS5/MS50/CPS	19.7	22.5	-2.8
MS/PEG/CPS	19.7	22.7	-3.0
MS5/PEG/CPS	19.7	22.7	-3.0
MS50/PEG/CPS	19.5	23.1	-3.6
Four Columns			
MS/MS5/MS50/PEG	21.4	29.8	-8.4
MS/MS5/MS50/CPS	21.4	29.7	-8.3
MS/MS5/PEG/CPS	21.4	29.9	-8.5
MS5/MS50/PEG/CPS	24.8	30.3	-5.5
MS/MS50/PEG/CPS	24.9	30.3	-5.4
Five Columns			
MS/MS5/MS50/PEG/CPS	26.3	37.5	-11.2

[a] It is assumed that the retention data of a given component on different columns can be clearly assigned to each other with a precision of two index units. H_i, experimentally obtained information content; ΣH_i, sum of the information content of the individual columns; ΔH = H_i - ΣH_i. For column code, see Experimental Section.

(PEG); (b) all test compounds can be separated on all five types of stationary phases with 10-m capillary columns; and (c) a reversal of the retention order occurs on the MS50 column at five positions, on the PEG and CPS columns at six positions compared to MS and MS5.

Characterization of the Identification Power by the Information Content. The information content of the retention indices was calculated according to eq 3 by assuming a standard deviation, σ_E, of two index units for the total error.

The calculation was performed for the single columns as well as for all possible combinations of columns. For the multiple columns it is assumed that the retention data of the same compound on different columns can be definitely attached to each other. This can be achieved without any doubt only by column switching. A less safe approach is the combination of retention data on the basis of the peak areas. In this case false attachments can be caused by the appearance of peak clusters or in the case of too small differences in peak areas. The simple combination of retention data from different columns can lead to artifacts. Furthermore, one has to be alert to the appearance of additional peaks since the analytes have to be determined in unknown matrices such as soil, waste water, etc.

The results of the calculation on the information content are shown in Table IV. The single columns have information contents between 7.2 and 7.8 bits, the combination of two columns between 9.7 and 14.4 bits, the combination of three columns between 15.3 and 19.8 bits, the combination of four columns between 21.4 and 24.9 bits, and the combination of five columns 26.3 bits. The difference, ΔH, between the experimental information content of several columns and the sum of the information contents of single columns indicates the loss of information due to the similarity of the retention characteristics of the columns. This loss due the correlation of data becomes very significant for the combination of five columns, where the gain compared to the combination of four columns with the highest information content is only 1.4 bits. The loss of information compared to the sum of the single column amounts -11.2 bits, which correspond to an amount higher than that of a single column. The loss of information in the combination of columns compared to the sum of the information of the single columns is from -5.4 to -8.5 bits for four columns, from -2.7 to -6.7 bits for three columns, and from -0.7 to -4.7 bits for two columns.

Summarizing, it can be concluded that the optimal choice considering identification power and effort is a combination of three columns with the stationary phases methyl silicone, 50% phenylmethyl silicone, and methylcyanopropyl silicone. This column combination has an information content of 19.8 bits, which corresponds to forensic safety. It can differentiate a maximum of about one million species. The loss of information caused by the correlation of column retention amounts to -2.7 bits in this case.

ACKNOWLEDGMENT