

CHROM. 10,538

DETERMINATION OF BOND- AND/OR INTERACTION-INCREMENT CODES AND PRE-CALCULATION OF RETENTION INDICES FOR ALKANES IN GAS-LIQUID CHROMATOGRAPHY BY COMPUTER

E. C. TAKÁCS, J. VÖRÖS and J. M. TAKÁCS

Gas Chromatographic Research Group for Study of Retention Index Systems, Fábián u.27, 1165-Budapest (Hungary)

(First received January 26th, 1977; revised manuscript received July 8th, 1977)

SUMMARY

The determination of bond- and/or interaction-increment codes by computer and the pre-calculation of retention indices are considered. It is shown that computer methods can be applied to any group of compounds (benzene homologues, esters, steroids, etc.), but the general principles have to be adapted to the particular case when preparing the computer program.

INTRODUCTION

The theoretical and practical problems of the retention index system¹, which is used in qualitative evaluations in gas-liquid chromatography (GLC), have been dealt with by many workers in recent years²⁻²⁹.

We have been concerned with the pre-calculation of retention data by means of increments³⁰ based on the principle of the additivity of retention indices. The essence of the method is the coding of the increments to be considered for various compounds, and the determination of the values of the increments on a thermodynamic basis³⁰. The wider application of the method has so far been hindered by the fact that the coding had to be carried out manually, which was difficult and time consuming.

In this work, by choosing appropriate mathematical identifiers, we succeeded in establishing an unambiguous relationship between chemical structure and the increments, which permitted the determination of the codes of the increments to be effected by computer.

THEORETICAL

Utilizing the additivity of retention indices, and in these the contributions, the following equations can be written:

$$I_b = k \cdot i_b(C, \dots) + l \cdot i_b(C, \dots) + \dots + m \cdot i_b(C, \dots) + \dots + n \cdot i_b(H, \dots, \dots) + \dots + s \cdot i_b(H, \dots, \dots) \quad (1)$$

$$I_i = k \cdot i_i(C, \dots) + l \cdot i_i(C, \dots) + \dots + m \cdot i_i(C, \dots) + \dots + n \cdot i_i(H, \dots, \dots) + \dots + s \cdot i_i(H, \dots, \dots) \quad (2)$$

$$I_{ig} = k \cdot i_{ig}(C, \dots) + l \cdot i_{ig}(C, \dots) + \dots + m \cdot i_{ig}(C, \dots) + \dots + n \cdot i_{ig}(H, \dots, \dots) + \dots + s \cdot i_{ig}(H, \dots, \dots) \quad (3)$$

$$I_{ii} = k \cdot i_{ii}(C, \dots) + l \cdot i_{ii}(C, \dots) + \dots + m \cdot i_{ii}(C, \dots) + \dots + n \cdot i_{ii}(H, \dots, \dots) + \dots + s \cdot i_{ii}(H, \dots, \dots) \quad (4)$$

where

I_b = bond index contribution (index units, i.u.);

I_i = interaction index contribution (i.u.);

I_{ig} = general interaction index contribution (i.u.);

I_{ii} = individual interaction index contribution (i.u.);

$$I_i = I_{ig} + I_{ii} \quad (5)$$

i_b = bond-increment value (i.u.);

i_i = interaction-increment value (i.u.);

i_{ig} = general interaction-increment value (i.u.);

i_{ii} = individual interaction-increment value (i.u.);

k, l, m, n and s = number of the same increments in the molecule studied;

C = symbol of the carbon–carbon saturated bond;

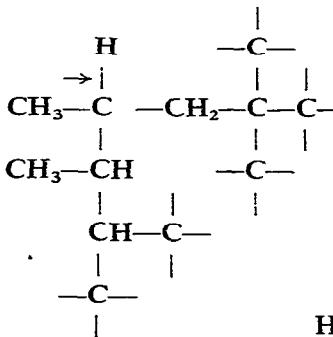
H = symbol of the carbon–hydrogen bond;

(C, ...) = symbol of the code for a carbon–carbon saturated bond;

(H, ..., ...) = symbol of the code for a carbon–hydrogen bond.

The treatment of eqns. 1–4 by a computer required the “linearization” of the block system of codes used earlier³⁰. The rules of the new procedure for coding are summarized in Table I.

It should be noted that in Table I the code, in addition to permitting a direct connection to be established between the chemical structure of a particular chemical bond and computer-calculated codes, contains chemical information that can be used for other purposes. To support this statement, the information content of the simple increment code given in paragraph 15 in Table I is shown below:



H, 3,123,000,004,013

TABLE I
RULES FOR CODING OF ALKANES BY COMPUTER

No.	Rule
1	Carbon-carbon saturated bonds are designated with a "C".
2	"C", and each independent information field, should be followed by a comma.
3	An increment code is ended by a full point.
4	The basic code of a carbon-carbon saturated bond is C,0,000,000.
5	After the comma following "C", the orders of the two carbon atoms between which the bond has been formed should be given in order of decreasing values. For example, the code fraction of a saturated bond between a quaternary and a tertiary carbon atoms will be C,43,...
6	Following the orders of the carbon atoms between which the bond is formed, the orders of the neighbouring carbon atoms have to be given in increasing order. For example, if in the example given in paragraph 5 a quaternary carbon atom is bound to one tertiary, one secondary and one primary carbon atom, whereas the tertiary carbon atom is bound to one secondary and one primary carbon atom, then the corresponding increment code will be C,43,123,012.
7	As can be seen from the increment code given in paragraph 6, if not all the places of the basic code (paragraph 4) are engaged by order values, then the "0" values must remain on the non-engaged places of the basic code. For example, the increment code of the carbon-carbon bond between the second and third carbon atoms of <i>n</i> -butane will be C,22,001,001.
8	If two carbon atoms of the same order are bound to each other, then the code of the neighbour of higher order will be given first. For example, if two tertiary carbon atoms are bound to each other, and two tertiary carbon atoms are bound to the first carbon atom and two primary carbon atoms to the second, the code will be C,33,033,011.
	The rule is similar if the order of the carbon atoms differ by only one unit. For example, if three primary carbon atoms are bound to a quaternary carbon atom and two primary and one secondary carbon atoms are bound to another quaternary carbon atom, bound directly to the other quaternary carbon atom, then the increment code will be C,44,112,111.
9	Carbon-hydrogen bonds are designated with an "H".
10	"H", and every independent information field, must be followed by a comma.
11	The increment code is ended by a full point.
12	The basic code of a carbon-hydrogen bond is H,0,000,000,000,000.
13	After the comma following "H", the order of the carbon atom to which the hydrogen is bound is given. For example, the bond increment code part of a hydrogen atom bound to a tertiary carbon atom is H,3,...
14	Following the order of the carbon atom bound to the hydrogen atom, the orders of the carbon atoms next to the carbon atom to which the hydrogen in question is bound (primary environment) are given. For example, if in the example in paragraph 13 one primary, one secondary and one tertiary carbon atom are bound to the tertiary carbon atom, then the code will continue as follows: H,3,123,...
15	After coding the order of the carbon atoms next to the carbon atom forming a bond with hydrogen (primary environment), the orders of the carbon atoms bound directly to carbon atoms bound to the former are given (secondary environment), in increasing order. For example, if in the example given in paragraphs 13 and 14 one quaternary carbon atom is bound to the secondary carbon atom, and one primary and one tertiary carbon atom are bound to the tertiary carbon atom, the increment code will be H,3,123,000,004,013. as no carbon atom is bound to the primary carbon atom in the primary environment.
16	The increment values must be given to two places of decimals, in index units (i.u.).

Fig. 1. Increment coding and retention index pre-calculation for phytane ($2,6,10,14$ -tetramethylhexadecane) on Apiezon L as stationary phase.

```

1C DEL ISOTPLISIL-MYLLOCATIJS (FILE-4)
2C DISP "TELETYPE(15), TELETYPE(9)"; 
3C IJPUT V
4C V=L
5C LIM FIC[4L],CIC[4G],ASC[5G],VSC[25G],ES[3G]
6C LISP "STATIONARY PHASE";
7C IJPUT ES
8C LISP "COLUMN-TELETYPE (K)";
9C INPUT V3
10C PI=3.67165-(98.34/(V3-1005.16))
11C FOF I=1 TO 250
12C IF I>16 THEN 160
13C FEAL A
14C VCIJ=A
15C GOTO 170
16C VCIJ=L
17C NEXT I
18C DISP "COMPOUND NAME";
19C IJPUT AS
20C IF AS="PFISTAJE" THEN 260
21C IF AS="PRYTANE" THEN 330
22C IF AS="PHAPJESAJE" THEN 400
23C DISP AS" I DO NOT KNOW!!! "
24C WAIT 2E00
25C GOTO 180
26C F[1]=F[2]=R[18]=T[19]=0
27C V1=266
28C H=17
29C C[1]=C[2]=C[3]=C[5]=C[6]=C[7]=C[9]=C[10]=C[11]=C[13]=C[14]=C[15]=C
30C C[17]=C[18]=C[19]=C
31C C[4]=C[6]=C[12]=C[16]=P[17]=F[3]=1
32C GOTO 450
33C R[1]=F[2]=P[19]=F[26]=0
34C V1=282
35C H=18
36C C[1]=C[2]=C[3]=C[5]=C[6]=C[7]=C[9]=C[10]=C[11]=C[13]=C[14]=C[15]=C
37C C[17]=C[18]=C[19]=0
38C C[4]=C[8]=C[12]=C[16]=P[18]=F[3]=1
39C GOTO 450
40C P[1]=P[2]=R[15]=C[1]=C[2]=C[3]=C[5]=C[6]=C[7]=C[9]=C[10]=C[11]=C
41C V1=212
42C C[3]=C[14]=F[16]=C[15]=0
43C F[3]=R[14]=C[4]=C[8]=C[12]=1
44C H=14
45C WRITE (T,*) 
46C WRITE (V,*) 
47C WRITE (W,*) TAB16;AS
48C FOR Z=4 TO H-1
49C IF C[Z]<1 THEN 520
50C R[Z]=3
51C GOTO 530
52C R[Z]=2
53C NEXT Z
54C FOR Z=3 TO H-1
55C IF C[Z]<1 THEN 770
56C GOSUB 580
57C GOTO 650
58C IF R[Z-1]>R[Z+1] THEN 620
59C A=R[Z-1]
60C B=R[Z+1]
61C GOTO 640
62C A=R[Z+1]
63C R=R[Z-1]
64C RETURN
65C IF A*10+B=12 THEN 680
66C J=3
67C GOTO 690
68C J=1

```

Fig. 2.

```

690 WRITE (W,*)"C, 31, 0";A;B;" , 000. "TAB65;V[J]
700 V=V+V[J]
710 FOR I=1 TO 3
720 GOSUB 580
730 J=2
740 WRITE (W,*)"H, 1, 003, 000, 000, 0";A;B"- "TAB65;V[J]
750 V=V+V[J]
760 NEXT I
770 NEXT Z
780 FOR Z=3 TO H-1
790 IF R[Z]<R[Z+1] THEN 870
800 B=R[Z+1]
810 A=P[Z]
820 D=R[Z-1]
830 C=C[Z]
840 F=R[Z+2]
850 E=C[Z+1]
860 GOTO 930
870 E=R[Z]
880 A=P[Z+1]
890 D=R[Z+2]
900 C=C[Z+1]
910 F=R[Z-1]
920 E=C[Z]
930 IF A=B AND D<F THEN 950
940 GOTO 970
950 D=R[Z+2]
960 F=R[Z-1]
970 J=A*10^5+B*10^4+C*10^3+D*10^2+E*10+F*10^0
980 IF J=311200 THEN 1110
990 IF J=321102 THEN 1130
1000 IF J=220302 THEN 1150
1010 IF J=321202 THEN 1090
1020 IF J=210300 THEN 1070
1030 IF J=321201 THEN 1050
1040 STOP
1050 J=13
1060 GOTO 1160
1070 J=12
1080 GOTO 1160
1090 J=6
1100 GOTC 1160
1110 J=1
1120 GOTO 1160
1130 J=4
1140 GOTO 1160
1150 J=5
1160 WRITE (W,*)"C, ";A;B;" , 0";C;D;"0";E;F"- "TAB65;V[J]
1170 V=V+V[J]
1180 NEXT Z
1190 FOR Z=3 TO H
1200 FOR I=1 TO 4-P[Z]
1210 IF R[Z-1]<R[Z+1] THEN 1300
1220 A=C[Z]
1230 B=R[Z+1]
1240 C=R[Z-1]
1250 D=C[Z+1]
1260 E=R[Z+2]
1270 F=C[Z-1]
1280 G=R[Z-2]
1290 GOTO 1370
1300 A=C[Z]
1310 E=P[Z-1]
1320 C=R[Z+1]
1330 D=C[Z-1]
1340 E=R[Z-2]
1350 F=C[Z+1]
1360 G=P[Z+2]

```

```

1370 J=F[Z]*10+7+A*10+6+B*10+5+C*10+4+D*10+3+E*10+2+F*10+G*10+0
1380 IF J=100030012 THEN 1560
1390 IF J=31120002 THEN 1580
1400 IF J=20230211 THEN 1600
1410 IF J=20220303 THEN 1620
1420 IF J=20230212 THEN 1640
1430 IF J=31220202 THEJ 1540
1440 IF J=31220102 THEN 1480
1450 IF J=20130012 THEN 1520
1460 IF J=10020003 THEN 1520
1470 STOP
1480 J=14
1490 GOTO 1650
1500 J=15
1510 GOTO 1650
1520 J=16
1530 GOTO 1650
1540 J=11
1550 GOTO 1650
1560 J=2
1570 GOTO 1650
1580 J=7
1590 GOTO 1650
1600 J=8
1610 GOTO 1650
1620 J=9
1630 GOTO 1650
1640 J=10
1650 VPITE (V,*)"H,";F[Z];",",;A;B;C;";,000,0";D;E;
1660 VFITE (V,*)"Z";F;G;"."TAB65;V[J]
1670 V=V+V[J]
1680 NEXT I
1690 NEXT Z
1700 VPITE (V,*)TAE62"-----"
1710 VPITE (V,*)TAE60"IE=";V
1720 VPITE (V,*)TAE25"AV="V1
1730 VPITE (V,*)TAE25"IA="V1/10
1740 VPITE (V,*)TAE25;ES
1750 VPITE (V,*)TAE25"II=IA+IE="V+(V1/10)
1760 VPITE (V,*)TAE27;VG
1770 WRITE (V,*)TAE25"F(T)="F1
1780 I=(V+V1/10)*F1
1790 VPITE (V,*)
1800 WRITE (V,*)
1810 VPITE (V,1830)" I="I;" INDEX UNITS AT"V3" K COLUMN-TEMPERATURE"
1820 VFITE (V,*)"0.J" ES
1830 FORMAT F8.1,F8.2,2SE
1840 VFITE (V,*)
1850 VPITE (V,*)TAB10;AS,"END"
1860 VFITE (V,*)
1870 WRITE (V,*)
1880 DISP "TABLE";
1890 INPUT ES
1900 IF EI="JO" THEJ 2020
1910 DISP "TEMPERATURE FROM- (K)":
1920 FIXED 2
1930 INPUT I
1940 VPITE (V,*)TAB10"COLUMN-TEMPERATURE K:"TAB40"RETENTION INDEX I.U.:""
1950 FOR I=I TO I+120 STEP 10
1960 F=(3.67165-(98.34/(I-1065.16)))*(V+(V1/10))
1970 WRITE (V,1980)I,F
1980 FORMAT 15X,F8.2,22X,F8.1
1990 NEXT I
2000 DATA 6.62,7,6.97,6.56,7.35,7.39,7.45,7.5,6.67,7.42,6.84,6.78,7.72,7.86
2010 DATA 7.62,7.48
2020 END

```

Fig. 2. Computer program for coding and retention index pre-calculation for isoprenoid hydrocarbons.

The example shows that the total environment (primary = 123 and secondary = 000,004,013) of the carbon-hydrogen bond indicated by the arrow is defined unambiguously.

An example of the coding and pre-calculation of retention indices is given in Fig. 1. The bond-increment values necessary for the pre-calculation of the retention indices of C₁-C₉ alkanes were given earlier³⁰.

The computer program for the coding and pre-calculation of the retention indices of isoprenoid hydrocarbons is given as an example in Fig. 2.

EXPERIMENTAL

Calculated data were compared with measured values (literature data and those measured by this research group). The measurements were carried out on Erba Science Fractovap GI-452 and 2400 and Perkin-Elmer F-11 and 900 gas chromatographs. The chromatograms of model mixtures were obtained under the conditions given in Table II.

Calculations were performed by means of a Hewlett-Packard 9830A calculator fitted with an HP-9866A printer. Programs were written in Basic computer language.

TABLE II

PARAMETERS FOR GAS CHROMATOGRAPHIC MEASUREMENTS CARRIED OUT USING APIEZON L AS STATIONARY PHASE IN ORDER TO CHECK THE CALCULATIONS

Parameter	Condition
Detector	Flame ionization
Sensitivity	Between 4 × 10 and 32 × 100
Sample introduction	0.1–0.3 µl by Hamilton syringes
Columns	(a) Aluminium spiral, 3.0 m × 2.0 mm I.D. (b) Stainless steel capillary, 50.0 m × 0.25 mm I.D.
Stationary phase	(a) 10.0% (w/w) Apiezon L on Chromosorb W, AW (60–80 mesh) (b) Apiezon L
Carrier gas	Nitrogen
Splitting ratio with capillary column	1:40
Auxiliary gases	Air and hydrogen
Column temperature	Varied between 170° and 240° and programmed between 60° and 240° at 1.5, 2.5 and 4.0 °C/min
Recorders	Speedomax W and Hitachi; response time, 1.0 sec; 1.0 or 2.5 mV full scale
Chart speed	1.27 cm/min

RESULTS AND DISCUSSION

The determination of increment codes by computer made possible the wider application of the pre-calculation of retention data. The method has advantages in the planning of gas chromatographic analyses (optimization of temperature, carrier gas flow-rate, etc.) of samples with a wide boiling range. The simple case considered as an example (2,6,10,14-tetramethylhexadecane) demonstrated the potentialities of

the method. Such computer calculations can be applied to any type of compound (benzene homologues, esters, fatty acids, amino acids, steroids, etc.) but, in preparing the computer program, the general principles must be adapted to the particular group of compounds and the appropriate symbols have to be worked out.

For example, with steroids, different symbols must be used for the oxygen in the $-C-OH$ group and that in the $-C=O$ bond. The same is true for the oxygen atoms in the primary and secondary environments.

In Table III, some data are compared in order to show the applicability of the method. Other types of applications (organic structure elucidation, calculations of thermodynamic data, etc.) remain to be explored in the future.

TABLE III

COMPARISON OF MEASURED, LITERATURE²³ AND CALCULATED DATA FOR SOME SATURATED ISOPRENOID HYDROCARBONS ON APIEZON L AT 200°

Compound	Retention index (i.u.)				
	Measured (I)	Literature ²³ (II)	Calculated (III)	I - III	II - III
2,6,10-trimethyltridecane	1446.4	1448.8	1443.9	2.5	4.9
2,6,10-trimethyltetradecane	1538.7	1540.0	1534.2	4.5	5.8
2,6,10,14-tetramethylpentadecane	1681.9	1686.6	1689.0	-7.1	-2.4
2,6,10,14-tetramethylhexadecane	1786.6	1790.9	1794.6	-8.0	-3.7

CONCLUSION

As shown in recent years^{27,28,31}, in the pre-calculation of retention data by means of increments the initial problems have been overcome and a useful approximation method has been developed. By introducing the method of coding by computer in this paper, we consider that the last obstruction to the wider application of the increment method has been overcome.

The coding by computer itself, without the method for the calculation of increment values, can be applied in all instances in which chemical structure is involved. The simplicity and large information content of the method confer several advantages over methods previously applied. Its flexibility renders possible its adaptation to other groups of compounds, and could lead to the setting up of a retention data bank by computer.

REFERENCES

- 1 E. sz. Kováts, *Helv. Chim. Acta*, 41 (1958) 1915.
- 2 J. J. Walraven, A. W. Ladon and A. I. M. Keulemans, *Chromatographia*, 1 (1968) 195.
- 3 J. Jonás, J. Janák and M. Kratochvíl, *J. Gas Chromatogr.*, 4 (1966) 332.
- 4 P. Chovin, *J. Gas Chromatogr.*, 2 (1964) 83.
- 5 C. Landault and G. Guiochon, *J. Chromatogr.*, 13 (1964) 327.
- 6 L. Rohrschneider, *J. Chromatogr.*, 22 (1966) 6.
- 7 M. B. Evans and J. F. Smith, *J. Chromatogr.*, 28 (1967) 277.
- 8 J. C. Loewenguth and D. A. Tourres, *Z. Anal. Chem.*, 236 (1968) 170.
- 9 G. Schornburg, *J. Chromatogr.*, 14 (1964) 157.

- 10 J. Bonastre and P. Grenier, *Bull. Soc. Chim. Fr.*, (1968) 118.
- 11 M. S. Vigdergauz and A. A. Martynov, *Chromatographia*, 4 (1971) 463.
- 12 R. V. Golovnya and V. P. Uraletz, *J. Chromatogr.*, 61 (1971) 65.
- 13 G. Castello, *J. Chromatogr.*, 66 (1972) 213.
- 14 N. Dimov and M. Muchtarova, *J. Chromatogr.*, 91 (1974) 105.
- 15 T. Tóth and J. Takács, *Magy. Kém. Foly.*, 77 (1971) 252.
- 16 L. Soják, I. Hrvnák and P. Majer, *Anal. Chem.*, 45 (1973) 293.
- 17 L. S. Ettre, *Anal. Chem.*, 36 (1964) 31A.
- 18 W. O. McReynolds, *J. Chromatogr. Sci.*, 8 (1970) 685.
- 19 J. Burkhard, J. Vais, L. Vodička and St. Landa, *J. Chromatogr.*, 42 (1969) 207.
- 20 J. K. Haken, *J. Chromatogr.*, 50 (1970) 114.
- 21 W. Engewald, K. Epsch, J. Graefe and Th. Welsch, *J. Chromatogr.*, 91 (1974) 623.
- 22 J. Sanz, J. Calderon and M. V. Dabrio, *An. Quim.*, 71 (1975) 611.
- 23 A. F. Shlyakhov, R. I. Koreshkova and M. S. Telkova, *J. Chromatogr.*, 104 (1975) 337.
- 24 R. A. Hively and R. E. Hinton, *J. Gas Chromatogr.*, 6 (1969) 203.
- 25 C. J. W. Brooks, A. R. Thawley, P. Rocher, B. S. Middleditch, G. M. Anthony and W. G. Stillwell, *J. Chromatogr. Sci.*, 9 (1971) 35.
- 26 G. Tarján, *J. Chromatogr. Sci.*, 14 (1976) 309.
- 27 R. Vanheertum, *J. Chromatogr. Sci.*, 13 (1975) 150.
- 28 Ph. Souter, *J. Chromatogr. Sci.*, 12 (1974) 418.
- 29 F. Riedo, D. Fritz, G. Tarján and E. sz. Kováts, *J. Chromatogr.*, 126 (1976) 63.
- 30 T. S. Lombosi, E. R. Lombosi, I. Bernát, Zs. Sz. Bernát, E. C. Takács and J. M. Takács, *J. Chromatogr.*, 119 (1976) 307.
- 31 J. M. Takács, *J. Chromatogr. Sci.*, 11 (1973) 210.