Journal of Chromatography, 172 (1979) 15–30

CHROM. 11,648

QUANTITATION OF THE INFORMATION CONTENT OF MULTI-DIMENSIONAL GAS CHROMATOGRAPHY AND LOW-RESOLUTION MASS SPECTROMETRY IN THE IDENTIFICATION OF DOPING DRUGS*

J. F. K. HUBER, E. KENNDLER and G. REICH

Institute of Analytical Chemistry, University of Vienna, Waehringerstrasse 38, A-1090 Vienna (Austria)

(Received December 1st, 1978)

SUMMARY

On the basis of information theory, a comparision of the performance of the identification of compounds by multi-dimensional gas chromatography and binary-encoded low-resolution mass spectrometry is made for doping drugs. This group of compounds was not selected at random but according to their pharmocological effects. For these compounds it was observed that the *m/e* values of the peaks in the mass spectra are highly correlated, while a number of chromatographic columns can be found on which the retention data are only slightly correlated. As a result, the information content of two-dimensional gas chromatographic retention data on such columns equals the information content of binary-encoded low-resolution mass spectra. It is shown that the accuracy and precision of retention data in gas–liquid chromatography, and therefore the information content, can be significantly improved if the data are corrected for adsorption effects of the solid support and extrapolated to a limiting value in the case of asymmetric peaks.

INTRODUCTION

The development of an optimal strategy for problem solving in chemical analysis requires the quantification of the performance of the various unit operations applied in analytical chemistry. For identification by chromatography, two performance characteristics have been proposed: the information content and the discriminating power. Both performance characteristics give a measure of the chance of identifying a component. Information theory was first applied to characterize the informing power of chromatography in 1969[1]. The information content, given in bits, was determined on the basis of the resolving power of the chromatographic system[1–3] or the measuring precision of the retention data[4,5]. In the first approach, a

* This paper was presented at the *12th International Symposium on Advances in Chromatography, November 7–10, 1977, Amsterdam (The Netherlands)*. The majority of papers presented at this symposium has been published in *J. Chromatogr.*, Vol. 142 (1977).

constant probability for an occurrence of the retention values was assumed. In the second approach, a large, randomly composed retention data collection was used.

The discriminating power was defined[6] as the probability that two components selected at random from a given amount of compounds would be discriminated on a given chromatographic system or set of systems. The application of this characteristic was demonstrated for the identification of basic drugs[7,8].

The identification potential of mass spectrometry was also quantified on the basis of information theory[9,10]. It is still an open question how many intensity levels per *m/e* value can be discriminated if different instruments or different working conditions are used. The most restrictive approach is to assume only two intensity levels, *i.e.*, to discriminate between peaks present and peaks absent.

In this paper, the identification potential of multi-dimensional gas chromatography and low-resolution mass spectrometry for a real analysis problem, the identification of doping drugs, is investigated. In practice, the amount of compounds to be discriminated is generally not composed at random, but selected according to given criteria. In doping analysis the selection criterion is the pharmacological effect based on certain structural characteristics of the molecules. The comparison of the two methods is performed in terms of the information content. In contrast to the discriminating power concept, which is also based on the probability, the information content concept is applicable to different chemical analytical methods and allows one to deal with multi-dimensional procedures.

THEORETICAL

In information theory, "information" is defined as a measure of the uncertainty of the incidence of an event: not the meaning of the event, but the eliminated uncertainty is the information, which is measured in bits. In analytical chemistry, information therefore means the decrease in the uncertainty of the nature and amount of the components of a random sample of the investigated system.

The mean information content, H, of a measurement, called the entropy of the measurement, depends on the probabilities, p_r, of the incidence of the single possible results, r:

$$H = -\sum_{r=1}^{n} p_r \log_2 p_r \text{ [bit]} \tag{1}$$

with the condition $\sum_{r=1}^{n} p_r = 1$, whereby n equals the total number of possible results. For equal probabilities of the incidence of the single results, $p_r = 1/n$, the entropy reaches a maximum, the so-called decision content, H_0.

If k different measuring values, x_i, which are independent of each other, are determined from a sample, the total entropy, $H(x_1, x_2, \ldots, x_k)$ of this group of values equals the sum of the entropies, $H(x_r)_i$, of the single values:

$$H(x_1, x_2, \ldots, x_k) = \sum_{i=1}^{k} H(x_r)_i \tag{2}$$

In general, however, the results are not independent of each other, so that the total entropy is less than the sum of the single entropies. In qualitative analysis by mass

spectrometry or gas chromatography, for example, the measuring values are intensities of *m/e* values or retention values. These data are used for the identification of substances by comparison with a data library. Data produced by one of these methods are not independent, but partially correlated, which must be taken into account when calculating the information content.

Information content of mass spectra

Mass spectrometric peaks are characterized by the intensities of their *m/e* values. As shown[9,10], the information content of binary-encoded mass spectra without correlation of the peaks is in the range of 100–200 bits. Taking into consideration the correlation of peaks, it amounts to about 40 bits[10]. Binary encoding of the *m/e* values means that only two intensity levels per *m/e* value are considered: if the intensity of a peak exceeds a specified value, it is considered to be present, otherwise it is denoted as absent. The mean information content has been calculated from the data file of a library of spectra of several thousand compounds. The substances considered were distributed over the whole range of organic compounds.

In real analytical problems, such as the control of prohibited doping drugs, not all possible types of compounds can occur, but only certain species, which are more or less correlated. In this instance, the mean information content of a mass spectrum will be reduced because of a strong interdependence of the probability of the occurrence of peaks.

To calculate the entropy of a mass spectrum, the single *m/e* values are binary encoded. For two possibilities ($n = 2$), eqn. 1 reads

$$H_i = -p_r \log_2 p_r - (1 - p_r) \log_2 (1 - p_r) \tag{3}$$

where p_r is the probability of the occurrence of a certain peak, r, in the spectrum.

Each *m/e* value gives a maximum information of 1 bit for equal probability, $p_r = 1/2$, of the occurrence and non-occurrence of a peak. For probabilities of 0 or 1, the entropy will be zero. In general, the entropy will be between the limiting values 0 and 1 bit per peak depending on the probability of the occurrence of a peak.

For k different *m/e* values, which are totally independent of each other, the total entropy, H, of a mass spectrum is given according to eqn. 2 by

$$H = \sum_{i=1}^{k} H_i \tag{4}$$

The assumption of the independent occurrence of peaks at different *m/e* values is not justified in general. If, for example, a peak at *m/e* 91 for $C_7H_7^+$ is found in a mass spectrum, a peak at *m/e* 65 normally is also found in the same spectrum, being generated by abstraction of C_2H_2 from $C_7H_7^+$. The probability of the occurrence of *m/e* 65 under the condition *m/e* 91 is nearly 1, which means that this peak will not eliminate uncertainty: it does not contribute to the information content.

In order to obtain the entropy of mass spectra considering the correlation of the occurrence of ions with different *m/e* values, the discontinuous probability distribution of the *m/e* values is replaced by a continuous Gaussian distribution, so that an approximate value of the entropy can be calculated[10]:

$$H_{\text{Gauss}} = \tfrac{1}{2} \log_2 \left(\frac{2\pi e}{\Delta x^2} \right)^k |\text{cov}| \tag{5}$$

where Δx is the class width in the corresponding histogram, |cov| is the determinant of the covariance matrix, (cov), which is defined as

$$(\mathrm{cov}) = \begin{pmatrix} \sigma_{11}\sigma_{12} \cdots \sigma_{1k} \\ \sigma_{21}\sigma_{22} \cdots \sigma_{2k} \\ \vdots \quad \vdots \quad \vdots \\ \sigma_{k1}\sigma_{k2} \cdots \sigma_{kk} \end{pmatrix}$$

where k represents the number of m/e values, σ_{ii} represent the variance of the distribution of the intensities for peak i and σ_{ij} the covariance of the intensities for peaks i and j.

Information content of gas chromatographic retention data

For the determination of the entropy of gas chromatographic (GC) data, there are two possibilities. The first approach connects the theory of the chromatographic process with information theory[1–3]. The starting point is the question of the maximal number of peaks, $n_{R\max}$, that can be differentiated in a chromatogram. As can be shown, this number of peaks depends on the desired chromatographic resolution, R, and the mean theoretical plate number, $\bar{N}$:

$$n_{R\max} = \frac{\log_{10}(t_{Rn}/t_{R0})}{\log_{10}(1 + R/\sqrt{\bar{N}})} \tag{6}$$

where t_{Rn} is the retention time of the last component, n, and t_{R0} is the hold-up time of the mobile phase.

The maximal entropy, the so-called decision content, H_0, can be obtained from

$$H_0 = \log_2 n_{R\max} = \log_2\left[\frac{\log_{10}(t_R/t_{R0})}{\log_{10}(1 + R/\sqrt{\bar{N}})}\right] \tag{7}$$

For the usual type of packed GC column, with $t_R/t_{R0} \leq 50$, $R \leq 6$ and $\bar{N} = 2500$, we have $H_0 \leq 5.1$ bits and for a capillary column with $t_R/t_{R0} \leq 50$, $R = 6$ and $\bar{N} = 100{,}000$, we have $H_0 \leq 6.1$ bits.

In the second approach, the entropy is calculated from a collection of retention data[4,5]. For a continuous function with the probability density partition $w(x)$, the entropy, $H(x)$, can be expressed in an analogous manner to eqn. 1 by

$$H(x) = -\int_{-\infty}^{+\infty} w(x) \log_2 w(x)\, dx \tag{8}$$

In the identification of substances by comparing the measured retention indices with retention indices in a data library, an uncertainty remains because of the inevitable statistical error of the measurement. The entropy, H_i, of retention indices determined on a GC column with a particular stationary phase, i, at a certain temperature is given by the difference in the entropies of the measured values, and the measuring error:

$$H_i = -\int_{I_{R1}}^{I_{R2}} w_m(x) \log_2 w_m(x)\, dx + \int_{I_{R1}}^{I_{R2}} w_e(x) \log_2 w_e(x)\, dx \tag{9}$$

where I_{R1} and I_{R2} are the lower and upper limiting values of the retention indices in the data library, $w_m(x)$ is the measured probability density partition function of the retention indices and $w_e(x)$ the probability density partition function of the statistical error of the I_R measurement. Assuming the probability density partition functions to be Gaussian, eqn. 9 reduces to

$$H_{\text{Gauss}} = \tfrac{1}{2}\log_2\left(\frac{\sigma_m^2}{\sigma_e^2}\right) \tag{10}$$

where σ_m^2 and σ_e^2 are the variances of the partition functions of the I_R values and the error, respectively.

The entropies of the retention indices of 248 compounds from a number of different chemical classes have been calculated on 10 stationary phases, and were found to be in the range 6.5–7 bits for a given stationary phase.

The entropy of the retention indices of several columns with different stationary phases is the sum of the entropies of the data for the single columns, if correlation is not taken into account. Considering the correlations of the probabilities of the retention indices, the entropy, $H_{1,2,\ldots,k}^{\text{Gauss}}$, of the retention indices on k different stationary phases can be expressed[4] by

$$H_{1,2,\ldots,k}^{\text{Gauss}} = \tfrac{1}{2}\log_2\left(\frac{|\text{cov}|_m}{|\text{cov}|_e}\right) \tag{11}$$

where $|\text{cov}|_m$ and $|\text{cov}|_e$ are the determinants of the covariance matrices of the I_R values and the measuring error, respectively, whereby the convariance matrix is defined by

$$(\text{cov}) = \begin{pmatrix} \sigma_{11}\sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21}\sigma_{22} & \cdots & \sigma_{2k} \\ \cdots & \cdots & \cdots \\ \sigma_{k1}\sigma_{k2} & \cdots & \sigma_{kk} \end{pmatrix}$$

where σ_{ii} is the variance of the retention indices on the stationary phase i, and σ_{ij} is the covariance of the retention indices on the stationary phase i and j[4]. If no correlation exists, σ_{ij} becomes zero and the matrix is reduced to the diagonal. In this instance, the total entropy equals the sum of the single entropies of the retention indices on the different stationary phases.

In the identification of components in real analytical problems, one can expect a specific pre-selection of the compounds. Not all of the compounds can be found in the sample. In this event, the information content attainable by GC is reduced. The entropy of the retention indices of 248 compounds on 10 stationary phases was found[4] to be 43.3 instead of the 67.2 bits that would result as sum of the entropies on the single columns.

EXPERIMENTAL

Doping drugs

The doping drugs used were obtained from Theta Corp., Media, Pa., U.S.A., kit tk-4000 FP. The amphetamine derivatives were kindly donated by M. Donike,

Institute of Biochemistry, University of Cologne, G.F.R., and all other doping drugs by G. Heinisch, Institute of Pharmaceutical Chemistry, University of Vienna, Austria.

Gas chromatography

GC retention data were determined by the use of a doublecolumn gas chromatograph (Siemens, Karlsruhe, G.F.R.; Model L402), equipped with an alkali flame-ionization detector. The columns were made from silylated glass tubes (2 m × 2 mm I.D.) and the carrier gas was nitrogen (99.995%; Messer-Griessheim, Düsseldorf, G.F.R.). Retention indices for the characterization of doping drugs and for calculation of the information content were determined on two stationary phases of different polarity: methylsilicone OV-101 and nitrilesilicone OV-225, 3% (w/w) on the solid silica support Volaspher A2, 125–150 μm (all E. Merck, Darmstadt, G.F.R.). To investigate the influence of the solid support on the retention indices, a column packing of the solid silica support Volaspher A4 (Merck), 125–150 μm, coated with OV-101 was used. The compounds were applied as free bases, dissolved in ethyl acetate; the volumes injected were usually 1 μl, corresponding to 1 μg of doping drug.

Mass spectrometry

Low-resolution mass spectra were measured by use of a double-focusing mass spectrometer with an electron-impact source (Varian-MAT, Model 112). The mass spectrometer was connected to a gas chromatograph (Varian Aerograph, Model 1400) by means of a slit separator. Helium (99.996%; Messer-Griessheim) was used as the carrier gas. The operating conditions are given in Table I.

Data processing was carried out off-line by means of a process computer (PDP 15, Digital Equipment) coupled with a large central computer (CDC CYBER 73, Control Data) via a MODEM.

TABLE I

CONDITIONS FOR MASS SPECTROMETRIC MEASUREMENT

Parameter	*Value*
EI source pressure	$2 \cdot 10^{-6}$–$5 \cdot 10^{-6}$ torr
Source temperature	200 °C
GC–MS interface temperature	250 °C
Electron energy	70 eV
Electron current	300 μA
Resolution (10% valley)	1000

RESULTS AND DISCUSSION

Retention data for doping drugs

For the characterization and identification of substances, chromatographic retention data can be applied. In GC the retention index, I_{Rl}, which is defined with the aid of the homologous series of *n*-alkanes, is usually used:

$$I_{Rl} = 100\left[n + \frac{\log_{10} r_{ln}}{\log_{10} r_{n(n+1)}}\right] \qquad (12)$$

where r_{in} and $r_{n(n+1)}$ are the relative retentions of the sample component, i, and the n-alkanes n and $n+1$ between which the component i is eluted.

In principle one must take into account that the stationary bed acts not only by dissolution in the stationary liquid, but also by adsorption on the solid surface of the support. The limiting value of the capacity factor, κ_{i0}, at infinite dilution can be expressed[11] as the sum of two increments, κ_{i0}^{s} and κ_{i0}^{σ}, corresponding to solution and adsorption:

$$\kappa_{i0} = \frac{t_{Ri} - t_{R0}}{t_{R0}} = \frac{\varepsilon_s}{\varepsilon_m} \cdot K_{i0}^{s} + \frac{a_\sigma}{\varepsilon_m} \cdot K_{i0}^{\sigma} = \kappa_{i0}^{s} + \kappa_{i0}^{\sigma} \tag{13}$$

where K_{i0}^{s} and K_{i0}^{σ} are the distribution coefficients at infinite dilution of component i between the stationary liquid phase s and the mobile fluid phase m, and between the liquid–solid interface σ and the mobile phase m, respectively; ε_m and ε_s are the volume fractions of the mobile phase m and the stationary liquid phase s in the column; a_σ is the ratio between the area of the solid surface, σ, and the total volume of the column; and t_{Ri} and t_{R0} are the retention time of the component i and the hold-up time of the mobile phase, respectively, the latter being determined by means of an inert tracer ($\kappa = 0$). From eqn. 13 it can be seen that in the determination of retention indices, the influence of the solid support on the distribution process must be considered.

In Table II, retention indices of 43 substances used as doping drugs are given; 21 of them are β-phenylethylamine derivatives. Retention indices for Apiezon L and polyethylene glycol stationary phases are cited from the literature[12] and the others were determined experimentally under standardized conditions with respect to sample size, solid support and stationary liquid loading.

Limiting value of the retention index for asymmetric peaks

If the concentration of a component i in the stationary bed is a linear function of its concentration in the mobile phase (linear distribution isotherm) and the theoretical plate number is not too low, the difference between the retention time (defined as mean residence time of the component, given by the first moment of the residence time distribution function) and the residence time of the maximum of the residence time distribution function is negligible[11]. In this instance, the residence time of the peak maximum, which is easier to determine, can be used as an approximation for the retention time. With non-linear shapes of the distribution isotherms for solution and adsorption, the residence time, t_{ci}, of concentration c_i^m of the component, i, in the mobile phase, m, can be described to a first approximation by the expression[11]

$$t_{ci} = t_{R0}\left(1 + \frac{\varepsilon_s}{\varepsilon_m} \cdot \frac{dc_i^s}{dc_i^m} + \frac{a_\sigma}{\varepsilon_m} \cdot \frac{dc_i^\sigma}{dc_i^m}\right) \tag{14}$$

where c_i^m, c_i^s and c_i^σ are the equilibrium concentrations of component i in the mobile phase, m, the stationary liquid phase, s, and the solid–liquid interface, σ, respectively.

From eqn. 14, one can see that the residence time of the peak maximum of a component depends on the first derivatives of the partial distribution isotherms at the concentration of the peak maximum. If the distribution isotherms approach a linear

TABLE II

RETENTION INDICES FOR 42 STIMULANT DRUGS ON FOUR STATIONARY PHASES

Data on Apezion L and PEG 20M from the literature[12]. Standardized conditions for OV-101 and OV-225: sample size, 1 μg; solid support, Volaspher AII; stationary liquid loading, 3% (w/w).

Compound	*Abbreviation*	I_{RI}							
		OV-101	°C	*OV-225*	°C	*Apiezon L*	°C	*PEG 20M*	°C
Isometheptene	ISOMET	1033	100	1145	150	1016	100	1218	70
Amphetamine	AM	1117	120	1430	120	1132	120	1581	120
Norfenfluramine	NORFEN	1136	120	1435	120	1092	120	1560	115
Perhydrophentermine	PERPHEN	1167	150	1341	150	1152	120	1359	100
Phentermine	PHEN	1173	150	1473	150	1168	120	1573	120
Propylhexedrine	PRHEX	1185	150	1348	150	1179	120	1354	100
Methamphetamine	MEAM	1186	150	1473	150	1187	120	1562	120
Tranylcypromine	TRAN	1206	120	1657	150	1225	120	1834	120
Fenfluramine	FEN	1226	120	1437	120	1183	120	1531	120
N-Ethylamphetamine	ETAM	1232	120	1504	120	1233	120	1571	120
Pargyline	PAR	1232	150	1510	150	1216	120	1648	130
N,N-Dimethylamphetamine	DIMEAM	1243	120	1508	120	1251	120	1568	120
Mephentermine	MEPHEN	1250	120	1566	140	1275	120	1611	120
Pentorex	PENT	1256	150	1563	150	1261	140	1648	130
N-Isopropylamphetamine	IPRAM	1257	120	1442	120	1253	120	1552	120
Octamylamine	OCTAM	1303	150	1358	150	1281	140	1373	100
Norpseudoephedrine	NORPE	1319	150	1771	150	1342	150	2176	180
Alfetamine	ALF	1320	140	1697	150	1311	140	1814	140
N-*n*-Propylamphetamine	PRAM	1325	150	1599	140	1326	140	1634	120
N-2-Butylamphetamine	2 BUTAM	1365	150	1604	140	1356	150	1635	120
Chlorophentermine	CLPHEN	1369	150	1781	160	1394	160	1871	150
N,N-Diethylamphetamine	DIETAM	1371	150	1592	140	1370	150	1609	120
Nicotine	NIC	1377	200	1714	150	1355	150	1848	160
Methoxyphenamine	MOXPHEN	1386	150	1768	150	1365	150	1880	130
N-Methyl-N-*n*-propyl-amphetamine	MEPRAM	1396	150	1628	140	1405	160	1673	140
N-Methylephedrine	MEEPH	1405	150	1808	150	1426	160	2042	160
N-*n*-Butylamphetamine	nBUTAM	1422	150	1687	150	1428	160	1732	130
N-Ethyl-*p*-chloro-amphetamine	ETCLAM	1434	150	1788	160	1456	160	1860	150
Phenmetrazine	PHENMET	1468	200	1930	200	1468	160	2065	150
N-Methyl-N-*n*-butyl-amphetamine	MEBUTAM	1486	150	1717	150	1488	160	1743	130
Phendimetrazine	PHENDIM	1492	200	1878	200	1485	160	1963	140
3,4-Methylenedioxy-amphetamine	NDOXAM	1495	150	2019	180	1484	160	2204	180
N,N-Di-*n*-propyl-amphetamine	DIPRAM	1526	150	1735	150	1534	180	1745	130
Amfepramon	AMPR	1527	200	1903	200	1490	160	1965	160
Nikethamide	NIKAM	1544	200	2231	200	1487	160	2319	180
Pentetrazol	PENTAZ	1570	200	2585	200	1649	180	2683	200
Diphenylamine (internal standard)	DIPH	1619	200	2213	200	1668	180	2517	190
Prolintane	PROL	1626	170	1905	160	1651	180	1946	150
3,4-Methylenedioxyphen-termine	MDOXPH	1646	200	2149	200	1681	180	2379	180
N,N-Di-*n*-butyl-amphetamine	DIBUTAM	1689	170	1891	160	1565	180	1902	150
Fencamfamine	FENC	1697	200	2041	200	1747	230	2125	160
N-Benzylamphetamine	BENZAM	1784	170	2300	180	1866	230	2390	190
Benzphetamine	BENZPH	1902	250	2254	200	1902	240	2392	190

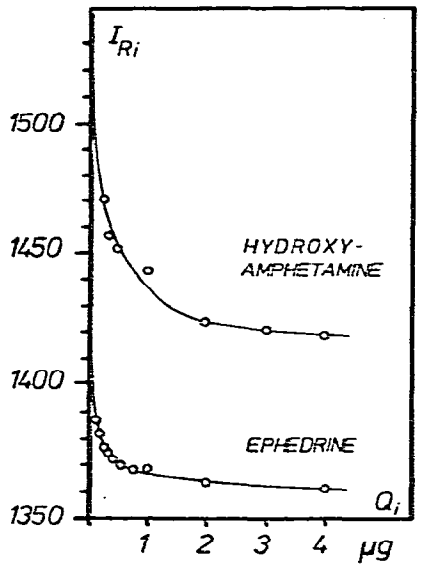
relationship at decreasing concentration, the value of the first derivatives approaches the value of the corresponding distribution coefficients at infinite dilution, and the residence time of the peak maximum approaches the retention time:

$$t_{Ri} = (t_{ci})_{max} = t_{R0}\left(1 + \frac{\varepsilon_s}{\varepsilon_m} \cdot K_{i0}^s + \frac{\alpha_\sigma}{\varepsilon_m} \cdot K_{i0}^\sigma\right) \quad (15)$$

With non-linear distribution isotherms the component has an asymmetric peak and the residence time of the maximum is dependent on the amount of sample injected, as the concentration of the maximum of the elution peak increases with sample size and column efficiency[13]. Consequently, the retention index, I_{Ri}, calculated from the residence time of the peak maximum depends on its concentration, which is determined by the amount injected and the efficiency of the column.

In order to decrease the dependence of the retention index on sample size and column efficiency, the limiting value, I_{Ri0}, at infinite dilution should be extrapolated. We propose two methods for the determination of a fixed value of the retention index.

In the first method, retention indices are determined from the residence times of the peak maxima for different amounts of sample and plotted against the amounts injected. The limiting value, I_{Ri0}, for an infinitely small amount of component is then determined by extrapolation. This method is demonstrated for hydroxyamphetamine and ephedrine in Fig. 1. The curves show a floating approximation to the I_R axis, so that the limiting value cannot be accurately determined. In order to overcome these difficulties, the measured data were fitted with the aid of a computer to the logarith-



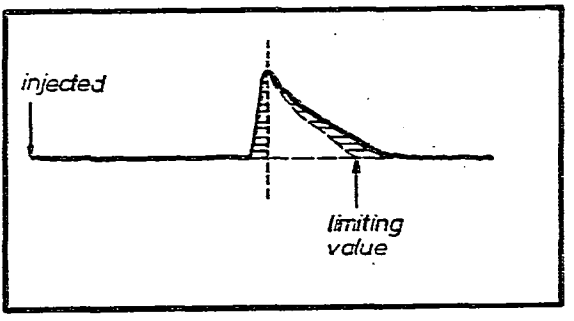


Fig. 1. Dependence of retention indices, I_{Ri}, on the sample amount, Q_i. Column, 1000 × 2 mm; stationary liquid phase, OV-101, 3% (w/w) on Volaspher A4, 125–150 μm; mobile phase, nitrogen; inlet pressure 1.2 bar; temperature, 150°. Sample: ephedrine and amphetamine dissolved in ethyl acetate. Sample volume: 1 μl.

Fig. 2. Illustration of the determination of the limiting value, I_{Ri0}, of the retention index for non-symmetric peaks from the peak flank.

TABLE III

EXTRAPOLATION OF THE RETENTION INDICES, I_{Ri}, TO A LIMITING VALUE, I_{Ri0}, AT A FIXED AMOUNT NEAR TO THE DETECTION LIMIT

Compound	I_{Ri}					
	Calculated from peak maximum				*Extrapolated*	
	Sample size (ng)				*From peak maxima*	*From peak flank*
	4000	*2000*	*500*	*250*		
Ephedrine	1362	1365	1370	1375	1415	1412 ± 6.7 (s = 8.0, n = 8, 95%)
Hydroxyamphetamine	1418	1424	1452	1470	1529	1526 ± 4.2 (s = 7.3, n = 14, 95%)

mic function $I_{Ri} = a + b \ln Q_i$, where Q_i represents the amount of the doping drug injected. With the calculated values of a and b, retention indices for a fixed amount near to the lower detection limit can be computed. Examples of such extrapolations are given in Table III.

In the second method for the determination of the limiting value, I_{Ri0}, for non-symmetric peaks a single peak is used. The peak broadening by kinetic effects is taken into consideration. The best approximation to the thermodynamic residence time is reached in the peak maximum. The flat flank of a non-symmetric peak is related to the distribution isotherm. Both flanks are influenced by the kinetic effects. A partial correction for the kinetic effects can be made by drawing a perpendicular line in the peak maximum and subtracting the distance from the perpendicular line to the steep flank from the flat flank (Fig. 2)[14]. With the aid of this difference method, the limiting values, I_{Ri0}, for hydroxyamphetamine and ephedrine were determined for different peaks corresponding to amounts from 150 to 4000 ng. The results are shown in Table III. The confidence of the peak flank method was calculated for a level of significance of 95% with 8 measurements for ephedrine and 14 for hydroxyamphetamine.

The limiting values obtained by both methods correspond well within the statistical error but differ significantly (*e.g.*, by more than 100 I_R units for hydroxyamphetamine) from the retention indices measured at higher concentrations of the components. A significant improvement in the accuracy has been achieved by the extrapolation to a fixed retention index.

Influence of solid support on retention index

To determine the influence of the solid support on the retention index, measurements at different stationary liquid loadings were carried out. An example of the dependence of the capacity factor κ_i on the stationary liquid loading is shown in Fig. 3. As can be seen, the extension of the linear part of the curve (at higher loading) does not originate from the zero point, but shows a finite intersection on the κ_i axis. With the experimental values a linear regression was performed corresponding to the equation $\kappa_i = a_0 + a_1 x$, where x (%, w/w) is the liquid loading.

Values of $a_0 = 0.80$ and $a_1 = 1.40$ were found. According to eqn. 13, a_0 corresponds to κ_i^{σ} and $a_1 x$ to κ_i^{s}, so that the relationship $\kappa_i = 0.80 + 1.40x$ results. In

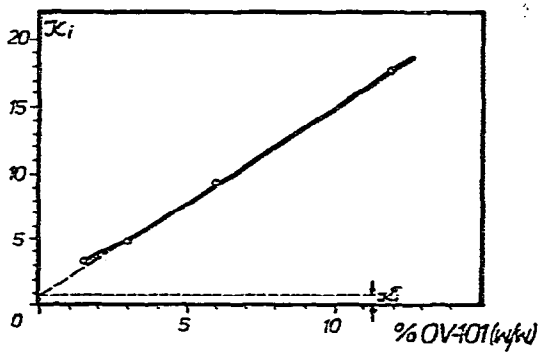


Fig. 3. Dependence of the capacity factor, κ_i, on the stationary liquid loading. $\kappa_i = \kappa_i^s + \kappa_i^a$, where κ_i^s = increment of κ_i due to solution and κ_i^a = increment of κ_i due to adsorption. Stationary liquid phase, OV-101 coated on Volaspher A2, 125–150 μm; mobile phase, nitrogen; inlet pressure, 1.2 bar; temperature, 150°. Sample: 4 μg of ephedrine dissolved in 4 μl of ethyl acetate.

Table IV an example is shown of the error in the determination of retention indices without considering the adsorption effects of the solid support.

The assumption of a negligible contribution to the capacity factor due to adsorption for *n*-alkanes was experimentally verified. The corresponding κ_i^a values are zero.

The dependence of the I_R values on the stationary liquid loading represents a systematic error, which must be considered when retention indices from different sources are compared. The same consideration is necessary if a column has been used over a long period, because of a possible change in loading due to evaporation of the stationary liquid phase. In addition, the chemical nature of the stationary liquid can change, due to decomposition or oxidation for example.

Mass spectra of β-phenylethylamines

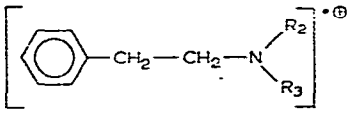
Ionization by electron impact in the ion source of a mass spectrometer takes

TABLE IV

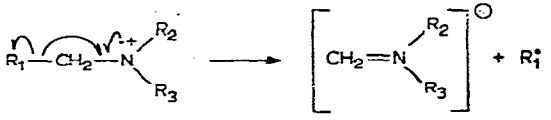
ERROR IN THE DETERMINATION OF RETENTION INDICES DUE TO ADSORPTION ON THE SOLID SUPPORT

Liquid loading: OV-101, % (w/w) on Volaspher A2	*I_{Ri} for 4 μg of ephedrine at 150° calculated by:*		*Systematic error, ΔI_{Ri}*
	$\kappa_i = \kappa_i^s + \kappa_i^a$	$\kappa_i^s = \kappa_i - 0.80$	
1.5	1428	1397	31
3.0	1423	1397	26
6.0	1422	1404	18
12.0	1415	1406	9

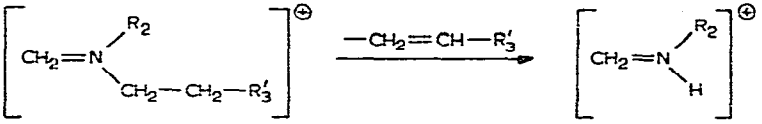
place for N-alkylated β-phenylethylamines at the hetero atom or, less frequently, at the aromatic ring, forming the ion



This molecular ion does not occur in the spectra or, if it does, only with very small intensity (Fig. 4). It is not stable and undergoes fragmentation. The main fragmentation reaction results in a cleavage of the C–C bond in the α-position to the nitrogen atom (α-cleavage):



Depending on the substituents R_2 and R_3 (H or alkyl groups), ions with *m/e* 30, 44, 58, 72, 86, 100, 114, 128, etc., are formed. As shown in the mass spectra in Fig. 4, one of these ions forms the base peak (100% relative abundance) in most instances. The most important secondary fragmentation reaction of the considered compounds, being secondary or tertiary amines, is the abstraction of an olefin by cleavage of the C–N bond and rearrangement of an H atom. This reaction can occur only if the alkyl group on the N atom contains more than one carbon atom:



Ions from the same even *m/e* sequence as given above can also be formed by this reaction. Other ions, formed by β-cleavage or McLafferty rearrangement of the α-cleavage products, occur only with low intensities.

Information content of gas chromatographic and mass spectrometric data

Gas chromatographic retention data. As discussed above, the mean information content (the entropy) of retention indices on one stationary phase can be calculated from eqn. 10, assuming a Gaussian shape of the probability density partition function of the measuring values and the error, respectively. For this calculation, the variance, σ_m^2, of the partition function of the I_{Ri} values and a realistic estimation of the variance σ_e^2 of the statistical error of the measurement of the I_{Ri} values is necessary.

To determine the magnitude of σ_m, the I_{Ri} values on the different stationary phases in Table II were plotted against n (Fig. 5). I_{Ri} values measured on a given stationary phase at different temperatures were extrapolated to the same temperature. Gaussian partition functions were fitted to the line diagrams. From the fitted curves, the standard deviations given in Table V were computed.

To determine the statistical error from the measurement of the I_{Ri} values of

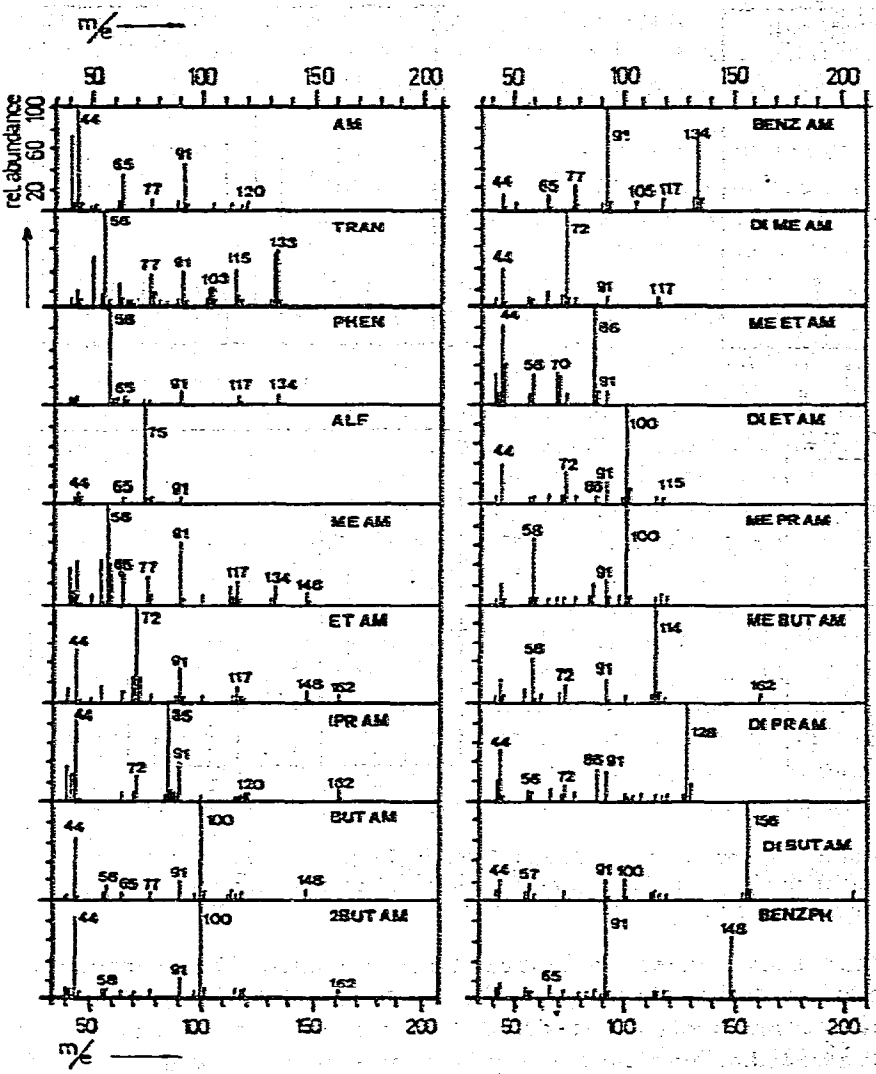


Fig. 4. Mass spectra of β-phenylethylamine derivatives. Conditions as in Table I; abbreviations as in Table II.

TABLE V

STANDARD DEVIATION, σ_m, OF THE GAUSSIAN PARTITION FUNCTIONS FITTED TO THE LINE DIAGRAMS IN FIG. 5

Stationary liquid phase	σ_m (I_R *units*)
OV-101	173
OV-225	284
Apiezon L	174
PEG 20M	326

polar compounds, it must be taken into account that the measuring values depend on the amount of component injected. To reduce this systematic error, the value for a

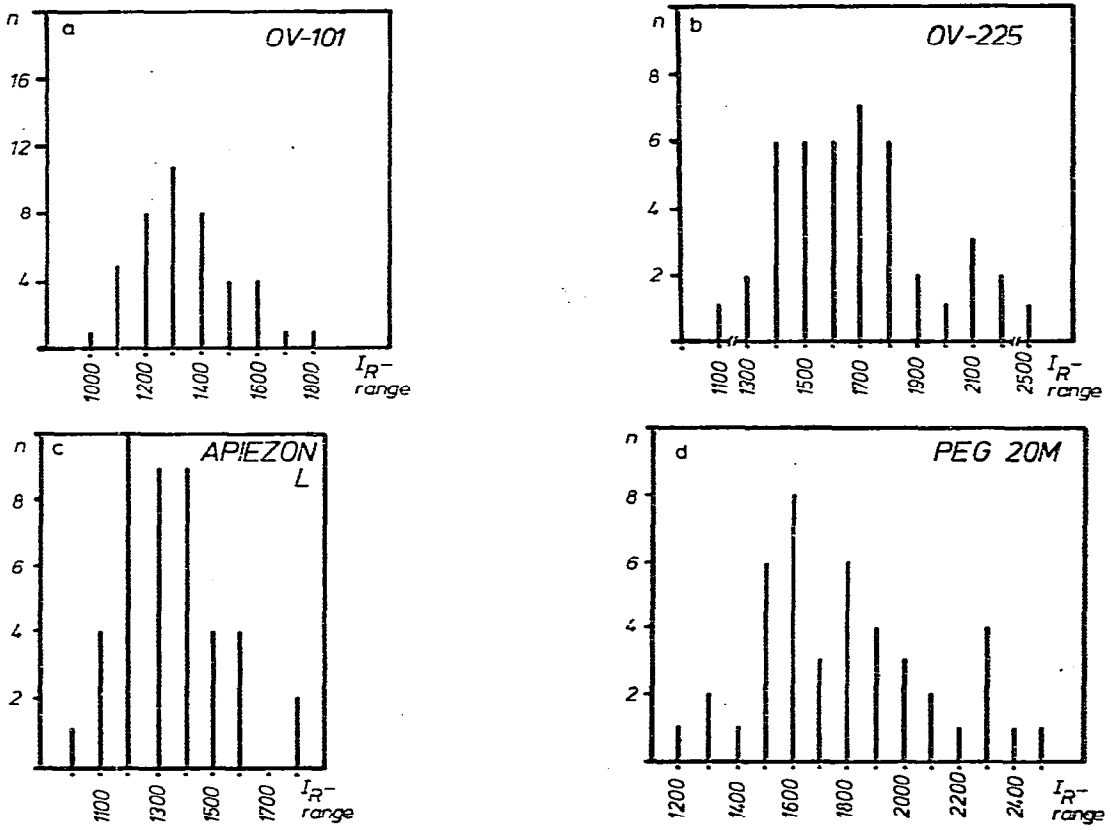


Fig. 5. Line diagrams of the number, n, of compounds listed in Table II within a range of retention indices, I_{Ri}, for four different stationary liquids. I_{Ri} values from Table II extrapolated to the same temperature for each stationary phase.

fixed amount near to the lower detection limit can be extrapolated. The determination of this value has a larger statistical error than the measurement of the I_{Ri} value with a constant amount of component. The standard deviation of I_{Ri} measurements on a given column with a constant amount of component was found to be less than one I_{Ri} unit, whereas the standard deviation of the determination of the extrapolated value by the peak flank method was nearly 7 I_{Ri} units. Even larger errors must be expected if the amount of liquid loading or the type of solid support varies, or if

TABLE VI

MEAN INFORMATION CONTENT (BIT) OF RETENTION INDICES OF 43 DOPING DRUGS ON THE SINGLE STATIONARY LIQUIDS FOR DIFFERENT VALUES OF THE STATISTICAL ERROR, σ_e, OF THE MEASUREMENT

Stationary liquid phase	σ_e (I_R *units*)					
	2	*3*	*4*	*5*	*7*	*10*
OV-101	6.4	5.9	5.4	5.1	4.6	4.1
OV-225	7.2	6.6	6.2	5.8	5.4	4.8
Apiezon L	6.5	5.9	5.5	5.1	4.6	4.1
PEG 20M	7.4	6.8	6.4	6.0	5.6	5.0

packed and capillary columns are used. The results of the calculation of the mean information content of the retention indices in Table II (extrapolated to the same temperature) are given in Table VI for different magnitudes of the statistical measuring error.

Assuming that the retention indices on the different stationary phases are completely independent, the total entropy of the data obtained on columns with different stationary phases is equal to the sum of the entropies of the single columns. In reality, the calculated entropies are lower, owing to the correlation of the data on different stationary phases. In Table VII the entropies, calculated with the aid of the covariance matrices, are given. As can be seen, the average value of the entropy for four different stationary phases for $\sigma_e = 7$ decreases from 5.1 bits per phase (Table VI) to 4.1 bits per phase (Table VII).

Low-resolution mass spectrometric data. To calculate the entropy of low-resolution mass spectra, the following assumptions were made: each integral *m/e* value in the range 1–300 was encoded by 1 bit; hence only the statement "peak present" or "peak absent" is admitted. Additional information obtained by dividing the intensity in various levels will be the subject of future investigations. Peaks with intensities less than 1 % of the base peak were denoted by 0 and all other peaks by 1.

The entropy of the mass spectra of the compounds listed in Table II was computed according to eqn. 4 with $k = 300$. As a result of this calculation, the mean information content, H, is equal to 77.0 bits. As discussed above, the interdependence of the occurrence of the individual peaks on each other must be taken into account. By the aid of the determinant of the covariance matrix, the calculation gave 8.6 bits for the mean information content of the mass spectrum of a doping drug. This value is much less than the 40 bits given in the literature[10] for a large number of different types of compounds.

Owing to the fragmentation scheme, the entropy calculated from the data of compounds with widely different chemical natures must be expected to be higher

TABLE VII

MEAN INFORMATION CONTENT (BIT) OF MULTI-DIMENSIONAL RETENTION INDICES OF 43 DOPING DRUGS CONSIDERING THE CORRELATION

Stationary liquid phase	σ_e (I_R units)					
	2	*3*	*4*	*5*	*7*	*10*
OV-101 + OV-225	12.6	11.5	10.6	10.0	9.0	8.0
OV-101 + Apiezon L	11.1	10.0	9.1	8.5	7.5	6.5
OV-101 + PEG 20M	13.1	11.9	11.1	10.4	9.5	8.4
OV-225 + Apiezon L	12.4	11.3	10.4	9.8	8.8	7.8
OV-225 + PEG 20M	12.6	11.4	10.6	9.9	9.0	7.9
Apiezon L + PEG 20M	12.9	11.7	10.9	10.2	9.3	8.2
OV-101 + OV-225 + Apiezon L	17.1	15.4	14.1	13.2	11.7	10.2
OV-101 + OV-225 + PEG 20M	18.0	16.3	15.0	14.1	12.6	11.1
OV-101 + Apiezon L + PEG 20M	17.6	15.8	14.6	13.6	12.1	10.6
OV-225 + Apiezon L + PEG 20M	17.9	16.1	14.9	13.9	12.5	10.9
OV-101 + OV-225 + Apiezon L + PEG 20M	23.7	21.3	19.7	18.4	16.5	14.4

than that for chemically similar compounds. In this work, data derived from compounds showing special pharmacological effects were used. Nearly half of the compounds are β-phenylethylamine derivatives with similar structural characteristics. These similar compounds show the same fragmentation scheme, so that distinct fragmentation ions frequently occur in the mass spectra of different compounds (Fig. 4). Ions with a large probability of occurrence provide little information for the identification of the compounds. On the other hand, ions that occur very seldom also contribute little to the mean information content. For this reason, the mass spectra of doping drugs were found to have a significantly smaller mean information content than those of an arbitrarily combined group of compounds.

Comparison of gas chromatography and mass spectrometry

In doping drug analysis, the *m/e* values of the peaks in the mass spectra were found to be significantly more correlated than the GC retention indices on columns with chemically greatly different stationary phases. Consequently, the mean information content of the two-dimensional retention indices from two suitable gas chromatographic columns, assuming a standard deviation of 7 retention index units, is about equal to the information content of 8.6 bits of binary-encoded low-resolution mass spectra, as can be seen from Table VII.

This result is confirmed by the observation that all doping drugs investigated can be discriminated by binary-encoded low-resolution mass spectrometry as well as by two-dimensional GC. A computer search proved that there are no two doping drugs in the collection that have all peaks at the same *m/e* values in the mass spectra. From Table II, it can be seen that all doping drugs in the table can be discriminated by at least 20 retention index units by two-dimensional operation on columns with OV-101 and PEG 20M as stationary phases.

REFERENCES

1 J. F. K. Huber and H. C. Smit, *Z. Anal. Chem.*, 245 (1969) 84.
2 H. Vink, *J. Chromatogr.*, 69 (1972) 237.
3 D. L. Massart and R. Smits, *Anal. Chem.*, 46 (1974) 283.
4 F. Dupuis and A. Dijkstra, *Anal. Chem.*, 47 (1975) 379.
5 A. Eskes, F. Dupuis, A. Dijkstra, H. de Clercq and D. L. Massart, *Anal. Chem.*, 47 (1975) 2168.
6 A. C. Moffat, K. W. Smalldon and C. Brown, *J. Chromatogr.*, 90 (1974) 1.
7 A. C. Moffat and K. W. Smalldon, *J. Chromatogr.*, 90 (1974) 9.
8 A. C. Moffat, A. H. Stead and K. W. Smalldon, *J. Chromatogr.*, 90 (1974) 19.
9 S. L. Grotch, *Anal. Chem.*, 42 (1970) 1214.
10 G. van Marlen and A. Dijkstra, *Anal. Chem.*, 48 (1976) 595.
11 J. F. K. Huber and R. G. Gerritse, *J. Chromatogr.*, 80 (1973) 25.
12 M. Donike and D. Stratmann, *Z. Anal. Chem.*, 279 (1976) 129.
13 J. F. K. Huber and A. I. M. Keulemans, in M. van Swaay (Editor), *Gas Chromatography 1962*, Butterworths, London, 1962, p. 26.
14 J. F. K. Huber, *Thesis*, University of Innsbruck, Innsbruck, 1960.