FLAVOUR AND FRAGRANCE JOURNAL
Flavour Fragr. J., **13**, 400–408 (1998)

Determination of the lipophilicity of aroma compounds by RP–HPLC

Geneviève Piraprez, Marie-France Hérent and Sonia Collin*

Unité de Brasserie et des Industries Alimentaires, Faculté des Sciences Agronomiques, Université Catholique de Louvain, 2/7 Place Croix du Sud, B-1348 Louvain-la-Neuve, Belgium

Received 24 November 1998
Revised 12 March 1998
Accepted 18 March 1998

ABSTRACT: The lipophilicity of aroma compounds was measured by reversed-phase high pressure liquid chromatography (RP–HPLC). The lipophilicity index provided by this technique is the capacity factor, log k_w, obtained by extrapolating series of log k values measured at various proportions of methanol in the methanol–water mobile phase. This article lists the log k_w capacity factors of 96 flavouring compounds belonging to various families (pyrazines, thiazoles, other heterocyclic compounds, alcohols and phenols, ketones, esters, aldehydes and sulphur compounds). © 1998 John Wiley & Sons, Ltd.

KEY WORDS: hydrophobicity; partition coefficient; capacity factor; flavour; pyrazines; thiazoles

Introduction

The lipophilicity of a molecule or moiety is a measure of its affinity for a hydrophobic environment, commonly obtained on the basis of how it distributes between two phases. Thermodynamically, the partition coefficient P is a constant, expressing the ratio of the activities of a solute in two immiscible phases at equilibrium. By convention, the activity in the organic phase is taken as the numerator:

$$P = \gamma_o C_o / \gamma_w C_w \quad (1)$$

In dilute solutions, the partition coefficient can be approximated as:

$$P = C_o / C_w = \text{capacity factor} \times \text{phase ratio} \quad (2)$$

with

$$\text{capacity factor} = k = \left(\frac{\text{weight in the } n\text{-octanol phase}}{\text{weight in the water phase}} \right)$$

and

$$\text{phase ratio} = \left(\frac{\text{water volume}}{n\text{-octanol volume}} \right)$$

A positive log P value thus reflects a preference for the lipid phase, while a negative value indicates a relative affinity for water.[1]

Lipophilicity can be expressed as the sum of two terms, one representing bulk or steric properties, the other electrostatic and polar properties.[1,2] The former, which takes hydrophobic (entropic) and van der Waals interactions into account, can be described by steric parameters such as molar volume or molar refractivity. The polar term is more complex and related to dipolarity–polarizability, H-bond donor acidity and H-bond acceptor basicity.[1–3]

Lipophilicity has been widely used in quantitative structure–activity relationship (QSAR) studies because it is a major determinant of the pharmacokinetic and pharmacodynamic behaviour of many drugs.[1,4,5]

In aroma research, the lipophilicity of a flavouring compound is an indicator of its behaviour in the food matrix, generally composed of water and less hydrophillic constituents such as proteins, lipids, polysaccharides, etc.[6–10] Aroma lipophilicity also plays a determining role in the olfactory response (aroma transport across the aqueous mucus and binding to the olfactory receptors). Hérent *et al.* (1995)[11] have shown that among 32 green- and/or nutty-smelling compounds, most of them pyrazine or thiazole derivatives, only very lipophilic compounds are green odorants and good ligands of bovine and porcine odorant binding properties (OBP).

Unfortunately, although partition coefficients have been determined for numerous compounds of interest

* Correspondence to: S. Collin, Unité de Brasserie et des Industries Alimentaires, Faculté des Sciences Agronomiques, Université Catholique de Louvain, 2/7 Place Croix du Sud, B-1348 Louvain-la-Neuve, Belgium. Tel: +32 10 472113; fax: +32 10 472178; email: collin@inbr.ucl.ac.be

Contract grant sponsor: FRIA, Belgium.

© 1998 John Wiley & Sons, Ltd. CCC 0882–5734/98/060400–09$17.50

to the chemical industry and especially to the pharmaceutical industry, data concerning aroma compounds are scarce. In flavour research, theoretical calculation of partition coefficients[6] and the traditional shake-flask method[8] are still often used to determine the lipophilicity of aroma compounds. These techniques, however, are not reliable for some flavouring compounds, notably very lipophilic aroma or substituted heterocyclic ones (see Results and Discussion).

The aim of this study was to determine, with a reliable experimental technique, the lipophilicity of 96 aroma compounds organoleptically important in many foodstuffs (dairy products, alcoholic beverages, etc.). Among several lipophilicity determination methods reviewed below, we have chosen, for its several advantages, the RP–HPLC method yielding the capacity factor log k_w as a lipophilicity index.

Determination of Lipophilicity

The *n*-octanol/water partition coefficient (log P) measured by the shake-flash technique is commonly used to express lipophilicity, *n*-octanol having several theoretical and practical advantages over other organic solvents.[4]

- A much broader spectrum of compounds are soluble in *n*-octanol than in aliphatic or aromatic hydrocarbons; the hydroxyl group of *n*-octanol can act as both a hydrogen-bond donor and a hydrogen-bond acceptor.
- While *n*-octanol is practically insoluble in water, it dissolves an appreciable amount of water under equilibrium conditions (2.3 M, corresponding to an *n*-octanol : water molar ratio of about 4 : 1). Hence, hydrogen bonds need not be broken during transfer of a solvated molecule from the aqueous phase to the organic phase. *n*-Octanol–water partition coefficients mainly reflect hydrophobic interactions, while hydrocarbon–water partition coefficients are additionally influenced by desolvation energies.
- *n*-Octanol has low vapour pressure at room temperature.
- It is well-suited for direct measurement of concentrations in the ultra-violet region due to its low absorption over a wide range.
- *n*-Octanol–water partition coefficients are available in the literature for a large number of compounds. Several theoretical methods for calculating lipophilicity refer to the *n*-octanol–water system.[12,13]

Unfortunately, the traditional shake-flask method used to measure log P values has some practical drawbacks, notably (a) it is slow, (b) it is limited to log P values between −2 and 4, owing to the required precision and sensitivity of the analytical technique; (c) large errors can be caused by minor impurities. For instance, a strongly UV-absorbing contaminant may seriously interfere with the quantitative determination of the solute; (d) formation of micelles and microemulsions; (e) instability of the solute in aqueous media; (f) dissociation/association of polar solutes; and (g) volatility of the solute.[5,14] Therefore other methods, both theoretical and experimental, have been explored in order to measure lipophilicity.

Theoretical Calculation of Partition Coefficients

In 1964, Fujita and co-workers[12] observed that partition coefficients have an additive–constitutive character. This means that the lipophilicity of a given compound can be calculated by adding the lipophilic increments of its constitutive groups, substituents or fragments. This is possible because these increments are approximately constant, i.e. the value corresponding to a group/substituent/fragment is independent of the molecule to which the entity belongs. They introduced the 'hydrophobic substituent constant' π to express the contribution of each substituent to the overall partition coefficient of a compound. For a substituent X, π_X is defined as:

$$\pi_X = \log P_{RX} - \log P_{RH} \quad (3)$$

where π_X is the hydrophobic constant of substituent X, log P refers to the octanol–water system, and RH and RX are respectively the unsubstituted and X-substituted molecules. In other words, π_X expresses the variation in lipophilicity caused by replacing H with X in RH. An immediate and significant consequence is that $\pi_H = 0$. The first π values were calculated for aromatic substituents, then the system was extended to aliphatic functions. It has been necessary, however, to introduce a number of corrective terms to take into account intramolecular interactions and structural features such as double bonds, branching and ring fusion.[15]

The shortcomings of the hydrophobic substituent constant prompted Rekker and co-workers[13,16] and later Hansch and Leo[17] to develop fragmental methods aimed at overcoming some of these limitations. The two fragmental methods differ in how the hydrophobic increment of each fragment (the hydrophobic constant) is obtained. Using the largest database of measured values available at the time, Rekker (1979)[16] employed statistical methods to determine the average contribution of simple fragments such as C, CH, CH_2, CH_3, OH, NH_2, etc. At first approximation, the log P of a compound can be calculated by adding the f_i values of its fragments[13] plus a certain number of corrective

© 1998 John Wiley & Sons, Ltd.

terms constituting an integral part of Rekker's fragmental system and taking into account various intramolecular interactions. Rekker observed that the corrective terms are all multiples of the so-called 'magic constant' C_M ($C_M = 0.289$). The multiplier is an integer called the 'key number' k_n, whose value depends on the number and type of interactions corrected. The equation is thus:

$$\log P_{RX} = \sum_{i=1}^{n} a_i \cdot f_i + \sum_{j=1}^{m} kn_j \cdot C_M \qquad (4)$$

where a_i is the number of occurrences of fragment f of type i and kn_j is the multiplier of the magic constant for interaction type j.

A fragmental system should meet an important condition: self-consistency, meaning that the calculated log P values must be independent of the extent and mode of fragmentation.[1]

Hansch and Leo (1979)[17] also developed a fragmental system for calculating log P, based partly on the same ideas. Their starting point was a limited set of small molecules (gases such as H_2, N_2 and O_2, lower alkanes, alkenes and alkynes, lower alcohols and alkyl halides, etc.), the log P of which they measured with the best possible precision. Here the fragmental constants are designated as 'f'. As in the previous method, numerous corrective terms of many types were introduced for molecular interactions. The system is thus far from simple to use, and considerable experience is needed to apply all the rules correctly.[1]

Despite the wide use of both systems, the physicochemical nature of all corrective terms is far from understood. These methods have been computerized; the computer-assisted version of Leo's method is called the CLOGP programme and that of Rekker's method is PRO-LOGP.[18]

To avoid corrective terms, some investigators have proposed log P calculation methods based on atomic contributions.[19,20] The authors evaluated the hydrophobicity of individual atoms, taking into account the undeniable intramolecular interactions by employing a large number of atom types. Unfortunately, intramolecular interactions are best expressed by a continuous function, which can only be approximately discretized in terms of atom types.[21] In this method, the formula for estimating the octanol–water partition coefficient is:

$$\log P = \Sigma n_i a_i \qquad (5)$$

where n_i is the number of atoms of type i and a_i is the contribution of the corresponding type.[21] All of these theoretical methods and others based on solvent-accessible surface area (SASA) or molecular properties are reviewed by Leo.[18]

To increase the accuracy of lipophilicity measurements, some chromatographic techniques, especially reversed-phase thin layer chromatography (RP–TLC) and reversed-phase high pressure liquid chromatography (RP–HPLC) have been developed.

Chromatographic Determination of Lipophilicity

RP–TLC

Reversed-phase thin layer chromatography (RP–TLC) provides the lipophilicity index R_m, calculated by means of the expression:

$$R_M = \log\left(\frac{1}{R_F} - 1\right) \qquad (6)$$

where R_F is the ratio of the distances covered by the analyte and solvent.[4]

Usually, silica gel plates are impregnated with liquid paraffin, silicone oil, ethyl oleate, or *n*-octanol to constitute the stationary phase, while the mobile phase is a mixture of a polar organic solvent like methanol, ethanol or acetone with water or an aqueous buffer solution. The lipophilicity index is usually defined as the R_M value extrapolated to 100% water, R_M^0.[4,22] Using R_M^0 values as substitutes for *n*-octanol–water partition coefficients is justified by many close relations between log P and R_M^0.[4,22]

RP–HPLC

The availability of alkyl-bonded phases provides a simple, accurate, and reproducible method for determining the lipophilicity of a wide variety of compounds.[5] The eluent usually consists of water and organic modifiers such as methanol, acetone, tetrahydrofuran or acetonitrile. Methanol/water emerges as the best '*n*-octanol–water-like' eluent, providing an interface with both strong hydrogen-bond donor and strong hydrogen-bond acceptor capacities.

In RP–HPLC, the lipophilicity is derived not from a concentration ratio (log P) but from a weight ratio (capacity factor 'k') (Equation 2). The relation between log P and log k is:

$$\log P = \log k + \log (\text{phase ratio}), \quad \text{with } k = (t_R - t_0)/t_0 \qquad (7)$$

where t_R and t_0 are the retention times of the analyte and a non-retained compound, respectively. The phase ratio is given by the ratio eluent volume : stationary phase volume.

© 1998 John Wiley & Sons, Ltd.

A capacity factor determined at a given percentage of organic modifier 'ϕ' (isocratic capacity factor) is sometimes used as the lipophilicity index. More useful, however, is the capacity factor '$\log k_w$' obtained by extrapolating the retention data obtained with binary eluents to 100% water.

To avoid solute adsorption to the residual silanol sites of the stationary phase (silanophilic interactions), lipophilic amines like *n*-decylamine (0.2%) must be added.[5,14,23] At very low water concentrations, the mobile phase changes from a water-like structure to an organic modifier-determined structure exerting its own solvophobic effect. In this case, adding methanol will change the dielectric constant of the medium and decrease hydrophilicity. At high water concentration, the extended chains shrink to form a rigid surface of interacting hydrocarbon chains that are poorly wetted by the eluent, resulting in a change in the properties of the stationary phase and in the mechanism of retention. In keeping with these observations, El Tayar and co-workers[14] have suggested that extrapolating to $\log k_w$ should best be done by linear extrapolation of the log k values measured:

1. In the range $10 < \phi < 80$ for neutral and/or non-polar compounds.
2. In water-rich ranges of eluent composition for ionogenic polar compounds.

The silica gel matrix is unstable outside the 1.5–7.5 pH range. Therefore, many basic solutes do not elute in their non-ionized state.[5] In this case, a correction involving the pK_a value should be applied as shown by equations (8) and (9).[14,24]

$$\log k_w = \log k_w^{app} + \log(1 + 10^{pH - pK_a}) \text{ for acids} \quad (8)$$

$$\log k_w = \log k_w^{app} + \log(1 + 10^{pK_a - pH}) \text{ for bases} \quad (9)$$

For methanol–water eluents, the statistical significance of $\log k_w$–log P correlations is remarkably high.[5,25,26] As compared to the *n*-octanol–water log P index, $\log k_w$ can, however, be expected to include more steric information, since the stationary phase can discriminate between solutes of different shapes as demonstrated for polycyclic aromatic hydrocarbons, having the same molecular weight, but having rod-like or disc-like shapes.[5]

RP–HPLC has made it possible to determine the hydrophobicity of several important solute groups. These include: (a) complex structures of unknown partition behaviour for which the additivity of hydrophobic substituent constants is inapplicable; (b) permanently charged solutes; (c) hydrophobic compounds with $\log P > 4$, whose partition coefficients cannot be determined with sufficient accuracy by the shake-flask technique.[5]

However, due to the presence of the term log (phase ratio) in Equation (7), the expression of hydrophobicity in terms of log k is relative in nature. This is regarded as a major disadvantage with respect to the classical *n*-octanol–water system which provides a single, continuous hydrophobicity scale. Yet provided one uses methanol–water eluents of comparable ionic strength at near-neutral pH to measure retention, variations in $\log k_w$ due to the stationary phase are minor. This was demonstrated by Braumann[5] for six reversed-phase packings differing in the structure of the bonded phase (monomeric or polymeric), the surface area, the carbon loading and the number of residual silanol groups. An excellent correlation between the mean $\log k_w$ values and the *n*-octanol–water partition coefficient was obtained.

CPC

The most recent chromatographic method used to determine partition coefficients is centrifugal partition chromatography (CPC), a variant of countercurrent chromatography (CCC). As in CCC, two immiscible liquids are used. The first liquid (stationary phase) is held in channels by a centrifugal field while a mobile phase is pumped through it.[27]

The retention of any compound in CPC obeys the classical chromatographic equation:

$$v_r = v_0 + K.v_S \quad (10)$$

where v_r, v_0 and v_S are, respectively, the retention volume, the dead volume and the stationary phase volume. K is the partition coefficient of the compound between the mobile and the stationary phase.[27] The advantage of CPC is that it directly yields partition coefficients, and not capacity factors, in a variety of solvent systems (octanol, hexane, etc.), affording significantly greater precision and taking considerably less time than the shake-flask technique.[2] For many compounds the results correlate very well with published log P values obtained by the shake-flask method.[28] This method does have drawbacks, however: its cost and, as for shake-flask measurements, the restricted lipophilicity range to which it is applicable.[29]

Taking all this into account, RP–HPLC seems to be the best choice for lipophilicity determination in aroma research because of: (a) applicability over a broader range of lipophilicity; (b) decreased perturbations caused by impurities in the analytes; (c) rapidity and minute material consumption; (d) accurate measurement for aroma substances which are, by definition, volatile. This method has been used in this paper for determining the lipophilicity of 96 aroma compounds.

© 1998 John Wiley & Sons, Ltd.

Materials and methods

Aroma Compounds

2-(3-methylpentyl)pyrazine, 1,2-dipyrazinylethane, N,N′-(4-butyl-5-propyl-2-thiazolyl)decanedioic amide, 4-butyl-5-propylthiazole, methylthiomethylpyrazine and 5H-5-methyl-6,7-dihydrocyclopentapyrazine were synthesized and purified along common synthetic routes[30] (Pelosi *et al*., unpublished). All other aroma compounds were commercially available (Aldrich, Fluka, Janssen Chemica or Lancaster; purity at least > 95%).

Lipophilicity Measurements

Lipophilicity was measured by RP–HPLC with a chromatograph equipped with a Waters model 510 isocratic pump and a Waters WISP 710 B autosampler. The UV-detector was a Perkin-Elmer LC 75 operating at different wave-lengths chosen according to the maximum of absorbance of the aroma compound between 500 and 200 nm (spectrophotometer Shimadzu UV-240); i.e. 270 nm for pyrazines and ketones (except for 2,3-pentanedione, 1-carvone, benzophenone, α and β-ionone at 254 nm), 254 nm for aldehydes (except for vanillin at 230 nm and for *trans*-cinnamaldehyde at 270 nm), 245 nm for thiazoles (except for benzothiazole at 220 nm and for 2,4,5-trimethylthiazole at 254 nm), 230 nm for heterocyclic compounds (except for 1,3,5-triazine at 270 nm and for 1,2,4-triazole at 212 nm), 215 nm for sulphur compounds, and 212 nm for esters (except for coumarine at 270 nm and methyl anthranilate at 220 nm), alcohols, and phenolic compounds (except for eugenol and thymol at 230 nm). The Guard-Pak insert was packed with μBondapak™ C18, particle size 10 μm (Waters). The column (25 cm × 4 mm ID) was prepacked with LiChrosorbRP-18, particle size 10 μm (Merck). A digital 380 PC equipped with the Waters 840 acquisition programme (version 6.0) was used as an integrator for peak recording and to calculate retention times. The mobile phase was made up volumetrically from various combinations (30–70%) of methanol (analytical grade, LabScan), and a solution containing MOPS (3-morpholinopropane sulphonic acid, Sigma) buffer (0.01 M) and *n*-decylamine (Sigma) (0.2% v/v). MOPS buffer was chosen to avoid ion-pair formation; *n*-decylamine was used as a masking agent to eliminate silanophilic interactions.[5] The pH of the aqueous solution was adjusted beforehand to 7.4 by addition of HCl. The methanol–aqueous solution mixture was filtered with a Millipore HAWP filter (0.45 μM). Retention times (t_r) were measured at room temperature with a 1.5 ml/min flow rate. The column dead time (t_0) was determined with uracil. The capacity factor was defined as $k = (t_r - t_0)/t_0$. Log k for 100% water ($\log k_w$) was linearly extrapolated from results obtained for different mobile phase compositions.[14,23]

Results and discussion

Reversed-phase high pressure liquid chromatography (RP–HPLC) was used to determine the lipophilicity of 96 aroma compounds: 31 pyrazines, 12 thiazoles, 18 alcohols and phenols, 13 ketones, 8 esters, 5 aldehydes, 7 heterocyclic compounds, and 2 sulphur compounds.

Table 1 shows the isocratic capacity factor $\log k_\phi$, obtained at various percentages of methanol in the mobile phase (ϕ); as advised by El Tayar and co-workers,[14] we linearly extrapolated our measurements from binary eluents to 100% water, to obtain the capacity factor $\log k_w$. Standard deviations for the extrapolated values are reported in Table 1. They were usually below 0.15, except in the case of very lipophilic compounds such as α-ionone ($\log k_w = 3.262$; $S_{\log kw} = 0.243$).

Traditional *n*-octanol–water partition measurements were also performed (in duplicate) for 11 compounds (shake-flask method with UV detection).[31] For log P values below 1.5, the two techniques yielded similar results, but the discrepancy between log P and $\log k_w$ increased with lipophilicity (Figure 1). This may reflect the poor sensitivity of the shake-flask method when most molecules are solubilized in the *n*-octanol phase. In such cases, RP–HPLC is recommendable. As shown below, furthermore, this very fast experimental technique, which takes into account the stereoelectronic properties of the molecules, should be preferred to theoretical calculations.

As expected, the lipophilicity index $\log k_w$ increases in each series with chain length (e.g. methylketones with 4–9 carbon atoms, secondary alcohols with 4–7 carbon atoms), but, contrasting with theoretical methods, no constant increment for a CH_2 group can be used to anticipate all our experimental data.

As shown in the methylketone series, ramification logically results in decreased lipophilicity

$$(\log k_{w\ 3\text{-methyl-2-butanone}} = 0.380 < \log k_{w\ 2\text{-pentanone}} = 0.657).$$

The presence of an additional carbonyl substituent in the molecule

$$(\log k_{w\ \text{diacetyl}} = -0.246 < \log k_{w\ 2\text{-butanone}} = 0.463;$$
$$\log k_{w\ 2,3\text{-pentanedione}} = 0.239 < \log k_{w\ 2\text{-pentanone}} = 0.657)$$

leads to low lipophilicity. However, for other aroma compounds, especially substituted heterocyclic

© 1998 John Wiley & Sons, Ltd.

Table 1. Isocratic capacity factor at different percentages of methanol in the water–methanol mobile phase ($\log k_{\phi}$), extrapolation to 100% water ($\log k_w$), standard deviation of the intercept ($S_{\log k_w}$) and determination coefficient (r^2)

Aroma compound	$\log k_{70}$	$\log k_{60}$	$\log k_{50}$	$\log k_{40}$	$\log k_{30}$	$\log k_w$	$S_{\log k_w}$	r^2
Pyrazine derivatives								
Pyrazine	−0.783	−0.704	−0.631	−0.564	−0.546	**−0.339**	0.035	0.964
2-Methylpyrazine	−0.579	−0.447	−0.311	−0.215	−0.155	**0.199**	0.048	0.978
Acetylpyrazine	−0.552	−0.409	−0.282	−0.178	−0.126	**0.232**	0.053	0.974
1,2-Dipyrazinylethane	−0.614	−0.475	−0.306	−0.123	−0.070	**0.402**	0.064	0.978
2,3-Dimethylpyrazine	−0.409	−0.239	−0.080	0.044	0.141	**0.583**	0.046	0.988
2,6-Dimethylpyrazine	−0.391	−0.223	−0.069	0.056	0.151	**0.586**	0.045	0.988
2-Vinylpyrazine	−0.465	−0.299	−0.211	0.000	0.154	**0.604**	0.050	0.988
2,5-Dimethylpyrazine	−0.379	−0.203	−0.034	0.092	0.182	**0.640**	0.055	0.983
2-Methoxypyrazine	−0.326	−0.156	0.014	0.133	0.204	**0.648**	0.065	0.975
2-Ethylpyrazine	−0.368	−0.196	−0.019	0.119	0.206	**0.680**	0.056	0.984
2-Acetyl-3-methylpyrazine	−0.311	−0.126	0.045	0.189	0.279	**0.763**	0.056	0.984
2,3,5-Trimethylpyrazine	−0.383	−0.131	−0.024	0.164	0.319	**0.838**	0.059	0.987
2-Ethyl-3-methylpyrazine	−0.372	−0.142	0.011	0.203	0.374	**0.933**	0.032	0.997
5H-5-Methyl-6,7-dihydrocyclopentapyrazine	−0.315	−0.068	0.099	0.309	0.487	**1.093**	0.036	0.996
2-Propylpyrazine	−0.340	−0.079	0.095	0.303	0.480	**1.103**	0.043	0.995
3-Methoxy-2-methylpyrazine	−0.130	0.086	0.285	0.455	0.559	**1.124**	0.066	0.984
2-Acetyl-3-ethylpyrazine	−0.310	−0.047	0.126	0.333	0.518	**1.142**	0.042	0.995
2-Ethyl-3,5(6)dimethylpyrazine	−0.232	0.036	0.209	0.428	0.624	**1.265**	0.041	0.996
2,3,5,6-Tetramethylpyrazine	−0.075	0.158	0.365	0.513	0.690	**1.273**	0.048	0.993
2,3-Diethylpyrazine	−0.035	0.216	0.449	0.652	0.789	**1.456**	0.067	0.989
3-Propyl-2-methylpyrazine	−0.022	0.235	0.482	0.690	0.828	**1.520**	0.072	0.988
6-Methylbenzopyrazine	−0.105	0.150	0.382	0.638	0.812	**1.536**	0.043	0.996
Methylthiomethylpyrazine	−0.055	0.236	0.429	0.673	0.890	**1.598**	0.041	0.997
2,3-Diethyl-5-methylpyrazine	−0.082	0.233	0.438	0.683	0.920	**1.665**	0.048	0.996
Ethanethiolpyrazine	−0.201	0.138	0.376	0.695	0.996	**1.876**	0.041	0.998
2-Butyl-3-methylpyrazine	0.187	0.475	0.760	0.999	1.210	**2.011**	0.052	0.996
3-Methoxy-2-isopropylpyrazine	0.228	0.533	0.832	1.099	1.269	**2.116**	0.081	0.990
2-Isopentylpyrazine	0.173	0.501	0.770	1.054	–	**2.226**	0.049	0.998
3-Methoxy-2-isobutylpyrazine	0.373	0.725	1.064	1.367	1.544	**2.507**	0.104	0.987
2-(3-Methylpentyl)pyrazine	0.283	0.645	1.009	1.335	1.551	**2.578**	0.090	0.991
2-Methyl-3-hexyloxypyrazine	0.758	1.248	1.740	2.160	–	**4.060**	0.069	0.999
Thiazole derivatives								
Thiazole	−0.607	−0.434	−0.301	−0.129	−0.077	**0.373**	0.063	0.977
4-Methyl-5-thiazolethanol	−0.813	−0.607	−0.439	−0.251	−0.152	**0.387**	0.055	0.988
2-Acetylthiazole	−0.434	−0.243	−0.077	0.103	0.203	**0.720**	0.048	0.990
4-Methylthiazole	−0.379	−0.170	0.009	0.205	0.299	**0.858**	0.061	0.986
2-Ethoxythiazole	−0.067	0.202	0.446	0.695	0.833	**1.568**	0.072	0.989
Benzothiazole	−0.167	0.091	0.272	0.614	0.906	**1.678**	0.078	0.991
2,4,5-Trimethylthiazole	−0.007	0.263	0.492	0.758	0.992	**1.746**	0.019	0.999
2-Ethyl-4-methylthiazole	0.060	0.351	0.607	0.876	1.014	**1.798**	0.083	0.987
4-Methyl-5-vinylthiazole	−0.005	0.285	0.557	0.844	1.012	**1.835**	0.068	0.992
2-Isopropyl-4-methylthiazole	0.216	0.546	0.842	1.148	1.310	**2.207**	0.092	0.988
N,N′-(4-butyl-5-propyl-2-thiazoyl)-decanedioic amide	0.544	0.926	1.404	1.797	–	**3.498**	0.072	0.998
4-Butyl-5-propylthiazole	0.584	1.141	1.548	1.917	–	**3.721**	0.173	0.990
Other heterocyclic compounds								
1,3,5-Triazine	−0.979	−1.000	−1.060	−1.156	−1.324	**−1.527**	0.084	0.900
Pyridazine	−0.860	−0.783	−0.746	−0.701	−0.677	**−0.529**	0.027	0.961
Pyrimidine	−0.726	−0.670	−0.621	−0.588	−0.563	**−0.430**	0.020	0.974
Pyrazole	−0.896	−0.742	−0.602	−0.470	−0.398	**0.012**	0.046	0.986
1,2,4-Triazole	−2.360	−1.959	−1.663	−1.272	−0.861	**0.219**	0.057	0.997
3-Acetylpyridine	−0.561	−0.362	−0.204	−0.060	0.043	**0.526**	0.053	0.987
Pyridine	−0.366	−0.192	−0.050	0.083	0.181	**0.616**	0.041	0.990
Alcohols and Phenols								
2-Butanol	−0.815	−0.629	−0.522	−0.382	−0.311	**0.096**	0.056	0.979
2-Pentanol	−0.701	−0.413	−0.272	−0.229	−0.058	**0.400**	0.120	0.932
1-Pentanol	−0.542	−0.355	−0.241	−0.123	–	**0.439**	0.068	0.984
3-Methyl-1-butanol	−0.614	−0.383	−0.241	−0.093	0.093	**0.604**	0.044	0.993
1-Hexanol	−0.472	−0.346	−0.115	0.063	0.206	**0.750**	0.047	0.992
2-Hexanol	−0.550	−0.360	−0.230	0.044	0.315	**0.911**	0.089	0.981
Phenol	−0.533	−0.273	−0.009	0.250	0.422	**1.188**	0.054	0.995
2-Phenylethanol	−0.366	−0.099	0.153	0.382	0.500	**1.220**	0.087	0.983
Menthol	–	0.757	0.917	1.087	1.319	**1.855**	0.056	0.992

Table continued on next page

© 1998 John Wiley & Sons, Ltd.

Table 1. Continued

Aroma compound	log k_{70}	log k_{60}	log k_{50}	log k_{40}	log k_{30}	log k_w	$S_{log\ kW}$	r^2
Eugenol	−0.108	0.203	0.495	–	–	**2.006**	0.033	0.999
2-Heptanol	−0.099	0.274	0.515	1.019	1.128	**2.167**	0.158	0.974
1-Octen-3-ol	0.000	0.384	0.660	1.012	1.317	**2.306**	0.044	0.998
Borneol	0.112	0.530	0.798	1.095	–	**2.403**	0.129	0.990
Nerol	0.296	0.718	0.911	1.293	1.563	**2.511**	0.093	0.990
Geraniol	0.185	0.594	0.989	1.327	1.592	**2.711**	0.089	0.993
3,7-Dimethyloctanol	0.490	0.892	1.205	1.477	–	**2.817**	0.117	0.992
Thymol	0.214	0.557	0.910	1.338	1.768	**2.902**	0.065	0.997
Citronellol	0.358	0.792	1.118	1.598	–	**3.192**	0.112	0.995
Ketones								
Diacetyl	−0.957	−0.845	−0.788	−0.636	−0.552	**−0.246**	0.037	0.986
2,3-Pentanedione	−0.776	−0.619	−0.475	−0.344	–	**0.239**	0.023	0.998
3-Methyl-2-butanone	−0.657	–	−0.318	−0.182	−0.085	**0.380**	0.058	0.987
2-Butanone	−0.905	−0.670	−0.518	–	–	**0.463**	0.145	0.985
2-Pentanone	−0.425	−0.230	−0.034	0.036	0.191	**0.657**	0.069	0.977
2-Hexanone	−0.397	−0.071	0.134	0.237	0.439	**1.058**	0.113	0.965
2-Heptanone	−0.082	0.294	0.551	0.749	0.980	**1.788**	0.099	0.984
1-Carvone	0.105	0.437	0.680	0.945	1.171	**1.988**	0.054	0.995
Benzophenone	0.204	0.574	0.900	1.313	–	**2.757**	0.065	0.998
Menthone	0.257	0.711	1.004	1.333	–	**2.763**	0.136	0.991
α-Ionone	0.582	1.031	1.341	–	–	**3.262**	0.243	0.989
β-Ionone	0.597	1.050	1.296	1.826	–	**3.355**	0.206	0.983
2-Nonanone	0.451	0.864	1.280	1.788	–	**3.531**	0.091	0.997
Esters								
Ethyl acetate	−0.645	−0.403	−0.296	−0.223	−0.124	**0.273**	0.095	0.938
Ethyl propionate	−0.413	−0.143	0.034	0.209	–	**1.045**	0.092	0.987
Coumarine	−0.404	−0.155	0.066	0.341	0.604	**1.346**	0.027	0.999
Ethyl butyrate	−0.197	0.111	0.348	0.561	0.711	**1.440**	0.087	0.984
Methyl anthranilate	−0.198	0.081	0.273	0.635	0.942	**1.764**	0.080	0.991
γ-Decalactone	0.172	0.557	0.862	1.406	–	**2.953**	0.189	0.986
Isoamyl butyrate	0.460	0.855	1.193	1.578	–	**3.052**	0.042	0.999
δ-Dodecalactone	0.487	0.943	1.306	1.775	–	**3.453**	0.080	0.998
Aldehydes								
Vanillin	−0.689	−0.457	−0.163	–	–	**1.142**	0.108	0.995
Benzaldehyde	−0.381	−0.148	0.050	0.278	–	**1.146**	0.037	0.998
Phenylacetaldehyde	−0.369	−0.039	0.139	0.479	0.654	**1.455**	0.082	0.989
Phenylpropionaldehyde	−0.308	0.013	0.197	0.474	–	**1.486**	0.099	0.990
trans-Cinnamaldehyde	−0.315	0.038	0.277	0.486	0.786	**1.579**	0.071	0.992
Sulphur compounds								
Dimethylsulphide	−0.054	0.017	0.117	0.178	0.295	**0.540**	0.024	0.991
Dimethyldisulphide	0.044	0.227	0.410	0.505	0.628	**1.085**	0.058	0.982

molecules, theoretical predictions for additional hydrophillic substituents are no more right. In this context, the experimental lipophilicity values of the heterocyclic nitrogen derivatives are particularly instructive.

Unsubstituted pyrazine is hydrophillic ($\log k_W = -0.339$), as are pyrimidine ($\log k_w = -0.430$) and pyridazine ($\log k_w = -0.529$), due to two endocyclic nitrogen lone pairs in the heterocyclic plane which can interact with water molecules. Log k_w is logically higher when only one nitrogen lone pair is present ($\log k_{W\ pyridine} = 0.616$). Among the 5-atom cycles, pyrazole ($\log k_{W\ pyrazole} = 0.012$) is less lipophilic than triazole ($\log k_{W\ 1,2,4\text{-triazole}} = 0.219$) and is unexpectedly lipophilic despite the presence of one more undelocalized nitrogen lone pair in the latter. On the other hand, unsubstituted thiazole ($\log k_{W\ thiazole} = 0.373$) is logically more lipophilic than pyrazole due to the presence of the bulky sulphur atom.

For all the heterocyclic compounds, the presence of an additional lipophilic substituent leads to an increased log k_w. Among alkyl-monosubstituted and -disubstituted pyrazines, for example, the capacity factor increases with the chain length ($\approx 0.5/CH_2$) of the alkyl substituent

$$(\log k_{W\ pyrazine} < \log k_{W\ 2\text{-methylpyrazine}} < \log k_{W\ 2\text{-ethylpyrazine}};$$
$$\log k_{W\ 2,3\text{-dimethylpyrazine}} < \log k_{W\ 2\text{-methyl-3-propylpyrazine}}$$
$$< \log k_{W\ 2\text{-butyl-3-methylpyrazine}} \quad \text{and}$$
$$\log k_{W\ 3\text{-methoxy-2-methylpyrazine}} = 1.124$$
$$\lll \log k_{W\ 3\text{-hexyloxy-2-methylpyrazine}} = 4.060).$$

As regards the methoxy substituent, the various theoretical methods yield inconsistent results. According to Hansch and co-workers, a methoxy moiety is slightly hydrophobic ($f' = 0.28$ in an aromatic

© 1998 John Wiley & Sons, Ltd.

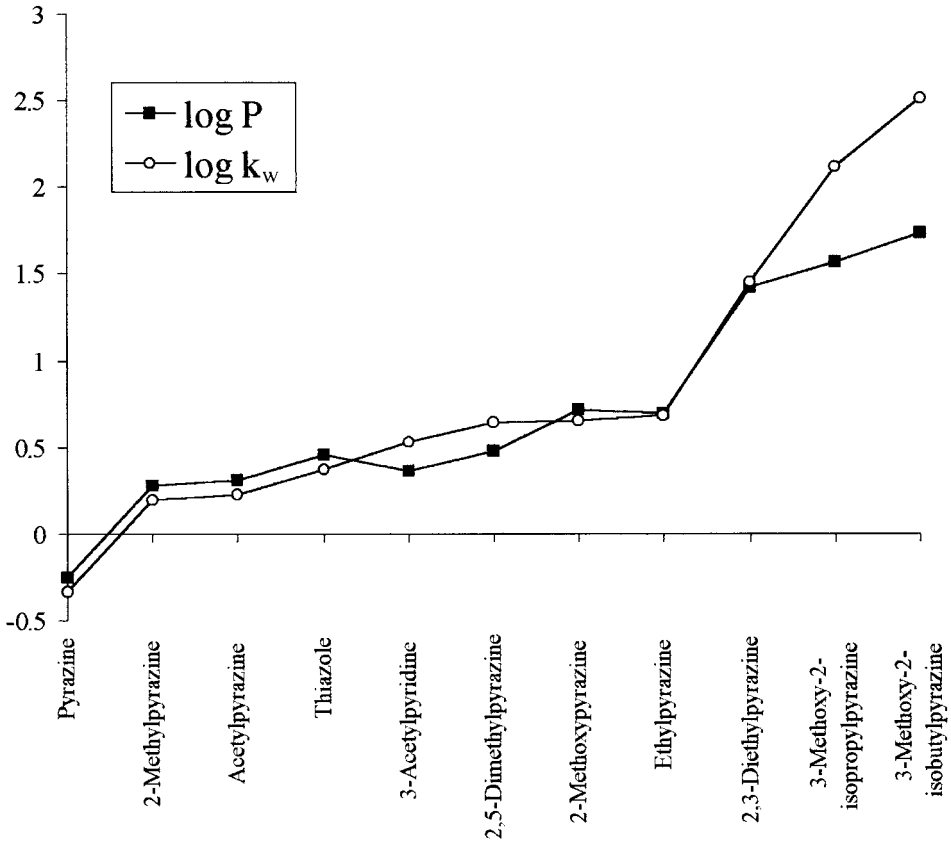

Figure 1. Relationship between the capacity factor, $\log k_w$, determined by RP–HPLC and the partition coefficient, log P, measured by the shake-flask method

structure),[12] while according to Rekker, it is hydrophillic ($f = -0.894$ in an aromatic structure).[13] Our experimental data show that the methoxy substituent induces a great capacity factor increase in pyrazine derivatives

$$(\log k_{W\ 2\text{-methoxypyrazine}} - \log k_{W\ \text{pyrazine}} = 0.987;$$
$$\log k_{W\ 3\text{-methoxy-2-methylpyrazine}} - \log k_{W\ 2\text{-methylpyrazine}} = 0.925).$$

This is due to masking, by the methoxy moiety, of one of the nitrogen lone pairs in such compounds.[11] The same applies to the acetyl substituent, considered hydrophillic by Hansch and Fujita ($\pi = -0.55$ in an aromatic structure),[12] by Rekker ($f = -0.942$ in an aromatic structure),[13] and by Hansch & Léo ($f' = -0.20$ in an aromatic structure).[17] By decreasing the availability of the nitrogen lone pair, an additional acetyl substituent on the pyrazine heterocycle leads to a $\log k_w$ increase of about 0.6

$$(\log k_{W\ \text{acetylpyrazine}} - \log k_{W\ \text{pyrazine}} = 0.57;$$
$$\log k_{W\ 2\text{-acetyl-3-methylpyrazine}} - \log k_{W\ 2\text{-methylpyrazine}} = 0.56).$$

Contrary to the way it affects pyrazine or thiazole derivatives, an acetyl substituent in 3-acetylpyridine ($\log k_{W\ 3\text{-acetylpyridine}} = 0.526$) gives rise to a $\log k_w$ decrease ($\log k_{W\ \text{pyridine}} = 0.616$). In this case, the acetyl moiety cannot hinder the nitrogen lone pair but can participate in electron delocalization leading to hydrophillic mesomeric forms. Likewise, a substituent at position 5 on the thiazole cycle cannot influence the nitrogen lone pair of electrons. On the other hand, the findings about alkoxy and acetyl pyrazines also apply to thiazoles when the substituents are at position 2 or 4

$$(\log k_{W\ 2\text{-acetylthiazole}} - \log k_{W\ \text{thiazole}} = 0.347;$$
$$\log k_{W\ 2\text{-ethoxythiazole}} - \log k_{W\ \text{thiazole}} = 1.195).$$

In the case of aliphatic aldehydes, very low UV-absorbances impeded us in measuring $\log k_w$ values. We assume that reactions with *n*-decylamine occurred in the column, leading to Schiff bases. For such compounds, another kind of masking agent, a lower pH or more sensitive detectors, such as pulsed electrochemical detector, should be tested.

Conclusions

It has recently been shown that there exists, within a same chemical family, a correlation between lipophilicity and aroma retention by triolein.[10] For many

© 1998 John Wiley & Sons, Ltd.

flavouring compounds, accurate lipophilicity measurements are thus needed in order to design pleasant 'light' or 'diet' products. Since RP–HPLC is in several respects more advantageous than the shake-flask technique and theoretical calculation methods, we have chosen this technique, usually applied in pharmacochemistry, to determine the lipophilicity of aroma compounds. The lipophilicity index provided by RP–HPLC is $\log k_w$, the capacity factor extrapolated to 100% water in the mobile phase. Results were obtained for 96 aroma compounds (pyrazines, thiazoles, other heterocyclic compounds, alcohols and phenols, ketones, esters, aromatic aldehydes and sulphur compounds) present in many foodstuffs. As depicted for acetyl- and methoxy-substituted pyrazines and thiazoles, the capacity factor appears to reflect molecular conformation better than do the lipophilicity indices obtained by 'fragmental' techniques. In contrast to the shake-flask experimental method, RP–HPLC was successfully applied to very lipophilic aroma compounds. Currently, the sole restriction concerns linear aldehydes, suspected of reacting with the masking agent *n*-decylamine.

Acknowledgements — Geneviève Piraprez thanks Fonds pour la formation à la recherche dans l'Industrie et dans l'Agriculture (FRIA, Belgium) for financial support.

References

1. H. van de Waterbeemd and B. Testa, *Advances in Drug Research*, **16**, 85 (1987).
2. C. Altomare, R. S. Tsai, N. El Tayar, B. Testa, A. Carotti, S. Cellamare and P. G. De Benedetti, *J. Pharm. Pharmacol.*, **43**, 191 (1991).
3. H. van de Waterbeemd, N. El Tayar, P.-A. Carrupt and B. Testa, *J. Comput.-Aided Molec. Des.*, **3**, 111 (1989).
4. H. Kubinyi, in *Progress in Drug Research*, ed. E. Juker Birkhauser Verlag, Basel, **23**, pp. 97–198 (1980).
5. T. Braumann, *J. Chromatogr.*, **373**, 191 (1986).
6. P. Landy, C. Druaux and A. Voilley, *Food Chem.*, **54**, 387 (1995).
7. G. Piraprez and S. Collin, in '*Proceedings of the Cost Action 96 'Interaction of Food Matrix with Small Ligands Influencing Flavour and Texture*', Dijon, 20–22 November, 58 (1995).
8. D. D. Roberts and T. E. Acree, *J. Agric. Food Chem.*, **43**, 2179 (1995).
9. C. Guyot, C. Bonnafont, I. Lesschaeve, S. Issanchou, A. Voilley and E. Spinnler, *J. Agric. Food Chem.*, **44**, 2341 (1996).
10. G. Piraprez, M.-F. Hérent and S. Collin, *Food Chem.*, **61**, 119–125 (1998).
11. M.-F. Hérent, S. Collin and P. Pelosi, *Chem. Senses*, **20**, 601 (1995).
12. T. Fujita, J. Iwasa and C. Hansch, *J. Am. Chem. Soc.*, **86**, 5175 (1964).
13. R. F. Rekker and H. M. de Kort, *Eur. J. Med. Chem.-Chim. Ther.*, **14**, 479 (1979).
14. N. El Tayar, H. van de Waterbeemd and B. Testa, *J. Chrom.*, **320**, 305 (1985).
15. A. Leo, C. Hansch and D. Elkins, *Chem. Rev.*, **71**(6), 525 (1971).
16. R. F. Rekker, in *The Hydrophobic Fragmental Constant. Its Derivation and Application. A Means to Characterizing Membrane Systems*, Pharmacochemistry Library, eds. W. T. Nauta and R. F. Rekker, Elsevier Scientific, Amsterdam (1979).
17. C. Hansch and A. Leo, in *Substituent Constants for Correlation Analysis in Chemistry and Biology*, ed. C. Hansch and A. Leo, Wiley, Chichester (1979).
18. A. Leo, *Chem. Rev.*, **93**(4), 1281 (1993).
19. P. Broto, G. Moreau and C. Vandycke, *Eur. J. Med. Chem.*, **19**, 71 (1984).
20. A. K. Ghose and G. M. Crippen, *J. Comp. Chem.*, **7**, 565 (1986).
21. A. K. Ghose, A. Pritchett and G. M. Crippen, *J. Comp. Chem.*, **9**(1), 80 (1988).
22. K. P. Dross, R. Mannhold and R. F. Rekker, *Quant. Struct.-Act.Relat.*, **11**, 36 (1992).
23. N. El Tayar, H. van de Waterbeemd and B. Testa, *J. Chromatogr.*, **320**, 293 (1985).
24. S. H. Unger and T. F. Feuerman, *J. Chromatogr.*, **176**, 426 (1979).
25. N. El Tayar, H. van de Waterbeemd and B. Testa, *Quant. Struct.-Act. Relat.*, **4**, 69 (1985).
26. A. Tsantili-Kakoulidou, N. El Tayar, H. van de Waterbeemd and B. Testa, *J. Chromatogr.*, **389**, 33 (1987).
27. A. Berthod and D. W. Armstrong, *J. Liq. Chromatogr.*, **11**, 457 (1988).
28. N. El Tayar, R. S. Tsai, P. Vallat C. Altomare and B. Testa, *J. Chromatogr.*, **556**, 181 (1991).
29. R.-S. Tsai, G. Lisa, P.-A Carrupt and B. Testa, in *Lipophilicity in Drug Action and Toxicology*, eds. V. Pliska, B. Testa and H. van de Waterbeemd, VCH, Weinheim. p. 401 (1996).
30. E. Napolitano and P. Pelosi, *Bioorganic and Mechanical Chemistry Letters*, **2**, 1603 (1992).
31. J. M. Mayer, B. Testa, H. van de Waterbeemd and A. Bornand-Crausaz, *Eur. J. Med. Chem.-Chim. Ther.*, **17**(5), 453 (1982).

© 1998 John Wiley & Sons, Ltd.