

Determination of solvation descriptors for terpene hydrocarbons from chromatographic measurements

Michael H Abraham,^{a*} Joelle R. M. Gola,^a Javier Gil-Lostes,^a William E. Acree, Jr,^b and J. Enrique Cometto- Muñiz^c

^a Department of Chemistry, University College London, 20 Gordon St, London WC1H 0AJ, UK. Email: m.h.abraham@ucl.ac.uk

^b Department of Chemistry, 1155 Union Circle Drive #305070, University of North Texas, Denton, TX 76203-5017, USA

^c Chemosensory Perception Laboratory, Department of Surgery (Otolaryngology), University of California, San Diego, La Jolla, CA 92093-0957

ABSTRACT

Gas chromatographic retention data on 107 terpene hydrocarbons from the literature have been used to obtain a set of Abraham descriptors for these 107 compounds. For 88 aliphatic cyclic terpene hydrocarbons, a fragmentation scheme was constructed that allows key descriptors to be estimated just from structure. The total set of descriptors, including those estimated by the fragmentation schemes, were then used to predict water-octanol partition coefficients for the 88 compounds, there being good agreement with experiment. For a small number of terpene hydrocarbons, there was good agreement between predicted and experimental values of nasal pungency thresholds, and predicted and experimental gas-blood, gas-oil, and gas-water partition coefficients. It is suggested that the descriptors obtained for the 107 terpene hydrocarbons can be used to predict water-solvent partition coefficients, gas-solvent partition coefficients, and partition coefficients in a number of biological systems.

Keywords: Terpenes, Gas liquid chromatography, Partition coefficient, Descriptors, Linear free energy relationships

1. Introduction

Terpenes are found in a wide variety of essential oils. Geraniol is the main constituent of geranium oil, limonene the major constituent of lemon oil, menthol the principal constituent

of peppermint oil and α -pinene the main constituent of turpentine. Many terpenes are of industrial importance; Sell [1] lists 20 terpenes that are used on an industrial scale, for example 30,000 tons of myrcene per annum, and limonene has widespread use as a fragrance component.

In spite of their widespread occurrence and use, there is very little known as to the effect of terpenes on humans, other than skin irritation and sensitization [2]. Cometto-Muñiz et al. [3] determined odor detection thresholds (ODT), eye irritation thresholds (EIT) and nasal pungency thresholds (NPT) for eleven terpenes. Nagata [4] included only three terpenes in his extensive investigation on ODTs, and Rodriguez et al. [5] only seven terpenes in ODT values for 100 compounds. There is an enormous number of terpenes likely to be encountered in everyday life, and some method of predicting their ODT, EIT, and NPT thresholds is clearly of importance. Since very little data exists on physicochemical properties of terpenes, the ability to predict such properties would also be extremely useful. Our method is based on the two linear free energy relationships, LFERs, Eq. (1) and Eq. (2) [6-8].

$$SP = c + e E + s S + a A + b B + v V \quad (1)$$

$$SP = c + e E + s S + a A + b B + l L \quad (2)$$

The dependent variable, SP , is some property of a series of solutes in a given system. SP may be partition coefficients as, $\log P_s$, for a series of solutes in a given water-solvent system or may be $\log K_s$ for a series of solutes in a given gas-solvent system. The former are mostly used in Eq. (1) and the latter in Eq. (2). Other systems can also be used. For example SP in Eq. (2) can be a set of retention data for a series of solutes on a given gas liquid chromatographic, GLC, stationary phase.

The independent variables, or Abraham descriptors, in Eq. (1) and Eq. (2) are properties of solutes as follows [6-8]: E is an excess molar refraction in $\text{cm}^3 \text{mol}^{-1}/10$, S is a combined dipolarity/polarizability descriptor, A is the overall solute hydrogen bond acidity, B is the overall solute hydrogen bond basicity, V is McGowan's [9] characteristic molecular volume in $\text{cm}^3 \text{mol}^{-1}/100$ and L is the gas to hexadecane partition coefficient at 25°C , as $\log K_s$.

The constant c and the set of coefficients, e , s , a , b , v and l characterise the system and are determined by multiple linear regression analysis. These coefficients are not just fitting coefficients, but represent the complementary properties of the system: e gives the contribution in Eq. (1) and Eq. (2) of the solute-system general dispersion interaction, s is the solvent dipolarity (plus some polarizability), a is the system hydrogen bond basicity (the

complimentary property to solute hydrogen bond acidity) and b is the system hydrogen bond acidity (the complimentary property to solute hydrogen bond basicity). The coefficients in Eq. (1) and Eq. (2) for a few processes that have been used to obtain descriptors for terpenes are in Table 1 [8, 10-13].

Table 1

Coefficients in Eq. (1) and Eq. (2), at 25°C.

System	c	e	s	a	b	v	Ref
Gas-water	-0.994	0.577	2.549	3.813	4.841	-0.869	10
Water-octan-1-ol ^a	0.088	0.562	-1.054	0.034	-3.460	3.814	11
Water-hexadecane	0.087	0.667	-1.617	-3.587	-4.869	4.433	11
Hexane-acetonitrile	0.097	0.189	-1.332	-1.649	-0.966	0.773	12
Perfluorodecalin-acetonitrile	-0.357	0.477	-2.360	-5.637	0.150	0.167	13 ^b
System	c	e	s	a	b	l	
Gas-water	-1.271	0.822	2.743	3.904	4.814	-0.213	10
Gas-octan-1-ol ^a	-0.198	0.002	0.709	3.519	1.429	0.858	8
Gas-hexadecane	0.000	0.000	0.000	0.000	0.000	1.000	8

^a Wet octan-1-ol. ^b Data from ref. 13, equation coefficients calculated in this work.

Abraham et al. [14] set out an equation for NPT values, based on Eq. (2) and later incorporated a number of terpenes into the algorithm [15]. Algorithms that included a few terpenes were also constructed for EIT [15] and ODT [16] values. These equations required knowledge of the descriptors in Eq. 2 for the terpenes, but it has not been easy to obtain these for any substantial number.

Descriptors for 30 terpenes have been obtained [17], mostly from GLC retention data on various stationary phases. It was pointed out [17] that none of the GLC stationary phases possessed any hydrogen bond acidity, and hence that the important B -descriptor could not be obtained from GLC data. Abraham et al. [17] managed to obtain the B -descriptor for a few terpenes from known water-octanol partition coefficients, as $\log P_{\text{oct}}$, and from their own measurements of HPLC capacity factors but for most of the terpenes they were only able to estimate B -descriptors. Ahmed and Poole [12] listed descriptors for a large number of

compounds, including 23 terpenes. They gave no reference for the origin of the descriptors but by inspection those for 19 of the terpenes were taken from the paper by Abraham et al. [17]. Thus the B -descriptor for these 19 terpenes is subject to the same uncertainty as mentioned by Abraham et al. [17]. In a later paper, Karunasekara and Poole [18] obtained descriptors for 15 terpenes using a combination of GLC data with values of $\log P_{\text{oct}}$ for nine terpenes and hexane-acetonitrile partition coefficients, as $\log P_{\text{hex-me}}$. The b -coefficient for the water-octanol system is numerically very large (-3.460) and so can lead to reasonably accurate values of the B -descriptor; the b -coefficient in the hexane-acetonitrile system is numerically much smaller (-0.966) and might be expected to lead to less accurate values. We can roughly estimate the possible error in the B -descriptor, through the term $\text{SD}(\text{reg})/b$ where $\text{SD}(\text{reg})$ is the standard deviation of the particular regression equation used, and b is the b -coefficient. For the octanol-water system the term is $0.116/3.46 = 0.034$ and for the hexane-acetonitrile system the term is $0.117/0.966 = 0.121$, very much larger than the estimated error for the octanol-water system. Thus, as pointed out before [17], unless a water-octanol partition coefficient is available (or very rarely another water-solvent system with a large b -coefficient), it is very difficult to obtain a reliable B -descriptor. Both Abraham et al [17] and Karunasekara and Poole [18] used experimental values of the gas-water partition coefficient, K_w , in their descriptor determinations. These experimental values are invariably obtained from the terpene solubility in water and the terpene vapour pressure. Unfortunately, variation in reported experimental values suggest that there is likely to be considerable uncertainty in literature values of $\log K_w$. For example, $\log K_w$ for limonene at 25°C is reported as -0.373 [19], -0.144 [20] and + 0.233 [21] so that extreme caution is necessary if $\log K_w$ values are used.

In the present work we concentrate on terpene hydrocarbons, and collect in Table 2 the Abraham descriptors for the few terpene hydrocarbons that have been obtained through the use of GLC retention data coupled with an experimental value of $\log P_{\text{oct}}$ [12, 16]. Although the terpenes in Table 2 are structurally quite simple, there are sometimes quite large differences in the values of the descriptors for a given compound. Note that for all the terpene hydrocarbons we shall deal with, the A -descriptor is zero, and so we omit the zero A -values from Table 2.

We therefore set out to obtain a coherent set of descriptors for terpene hydrocarbons in the hope that we would have enough reliable values for the B -descriptor to be able to construct an algorithm for the prediction of B -values.

Table 2

Descriptors for some terpene hydrocarbons

Terpene	<i>E</i>	<i>S</i>	<i>B</i>	<i>V</i>	<i>L</i>	Ref
α -Pinene	0.446	0.14	0.12	1.2574	4.308	17
α -Pinene	0.441	0.19	0.22	1.2574	4.348	18
β -Pinene	0.530	0.24	0.19	1.2574	4.394	17
β -Pinene	0.525	0.23	0.20	1.2574	4.584	18
Limonene	0.488	0.28	0.21	1.3230	4.725	17
Limonene	0.497	0.34	0.17	1.3230	4.693	18

2. Methods

Our method for the determination of descriptors for a given terpene is based on Eq. (1) and Eq. (2). The ‘unknown’ descriptors are usually *E*, *S*, *A*, *B*, *V* and *L* but for the terpene hydrocarbons the unknown descriptors are *E*, *S*, *B*, *V* and *L*. Values of *E* were obtained from known experimental refractive indices at 20°C [22-25], exactly as described before [6, 7], and *V* can simply be calculated from the number of atoms and rings in a molecule [6, 26]. If a water-solvent partition coefficient, P_s , is available, this can be converted into the corresponding gas-solvent partition coefficient, K_s , through Eq. (3). This requires a knowledge of K_w ; as we have seen, literature values of K_w (or $\log K_w$) are not very reliable, and we prefer to use $\log K_w$ as an extra descriptor to be determined.

$$P_s = K_s / K_w \quad (3)$$

Then the ‘unknown’ descriptors are *S*, *B*, *L* and $\log K_w$. There is little problem in obtaining equations for GLC retention data, following Eq. (2), and we have some 90 such equations constructed mostly from retention indices on the constituents of natural oils and waxes. In a very useful and important paper, Babushok et al. [27] surveyed the literature on retention indices, *I*, on three particular stationary phases, viz.: dimethylsilicone (DIMS), dimethylsilicone with 5% phenyl groups (DIMSP) and polyethylene glycol (PEG). Since their data base covers a wide range of compounds, we thought that our equations for *I*-values might be of use to other workers who wish to obtain descriptors from GLC data.

$$I(\text{DIMS}) = 39.1 - 39.2 E + 107.5 S + 128.9 A + 205.0 L \quad (4)$$

$$N = 174, SD = 15.9, R^2 = 0.998, F = 17338.6, \text{PRESS} = 45930, Q^2 = 0.997, \text{PSD} = 16.5$$

$$I(\text{DIMSP}) = 45.9 - 42.3 E + 135.9 S + 126.9 A + 204.0 L \quad (5)$$

$$N = 173, SD = 13.8, R^2 = 0.998, F = 22907.9, \text{PRESS} = 34605, Q^2 = 0.998, \text{PSD} = 14.4$$

$$I(\text{PEG}) = -62.2 + 91.8 E + 652.2 S + 1038.5 A + 213.3 L \quad (6)$$

$$N = 172, SD = 35.3, R^2 = 0.993, F = 5723.7, \text{PRESS} = 226094, Q^2 = 0.992, \text{PSD} = 36.8$$

In some instances, the number of compounds for which we had descriptors was very small, the large majority of compounds being terpenes, and so it was not possible to obtain an equation for the I-values. We then waited until we had determined descriptors for enough terpenes (through other equations) to construct an equation. An example is the set of retention indices obtained by Darjazi et al. [28] using a DB-5 fused silica column.

$$I(\text{DB-5}) = 62.7 - 16.0 E + 124.2 S + 62.5 A + 200.5 L \quad (7)$$

$$N = 55, SD = 10.3, R^2 = 0.998, F = 5994.4, \text{PRESS} = 8679.4, Q^2 = 0.997, \text{PSD} = 13.2$$

In Eqs. (4)-(7), N is the number of data points, that is compounds, R is the correlation coefficient, F is the Fisher F-statistic, PRESS and Q^2 are the leave-one-out statistics and PSD is the predictive standard deviation [29]. The rather better statistics of Eq. (7) is probably due to the data of Babushok et al. [27] being derived from several different sources. As is invariably the case for GLC retention data, none of the equations has a statistically significant term in B.

For a few terpene hydrocarbons, experimental values of $\log P_{oct}$ are listed in the EPI software program, available through ChemSpider [20] and in the BioLoom software program [30]. We also calculated $\log P_{oct}$ using a number of the most widely used programmes, viz the EPI and the ACD programmes [20], the ClogP program [30], the AlogPS program [31] and SPARC [32]. Griffin et al. [33] have measured HPLC capacity factors on a C-18 stationary phase using a number of water-methanol eluants for a large number of terpenes, including 10

terpene hydrocarbons. Zenkevich and Kushakova [13] have determined perfluorodecalin-acetonitrile partition coefficients for a large number of compounds, including some terpenes. We have used their data to obtain equation coefficients as shown in Table 1. The b -coefficient is very small, 0.150, and so the equation is not at all useful in the determination of the solute B -descriptor.

The set of equations on the lines of Eq. (1) and Eq (2) together with the corresponding dependent variables were then solved for the unknown descriptors S , B , L and $\log K_w$ by the trial-and-error procedure ‘Solver’ in Microsoft Excel. The equations that we used are in Table 1, together with a large number of equations for GLC retention data that are of little intrinsic value.

3. Results

By far the main set of terpene hydrocarbons are aliphatic cyclic compounds with a wide variety of number and sizes of rings. For several of the terpene hydrocarbons we had a large number of equations for GLC retention data, 65 in the case of α -pinene. Many of the equations have large values of the s -coefficient, see for example Eq. (6), so that there is little problem in obtaining the corresponding S -descriptor. We also had experimental values of $\log P_{oct}$ [20, 30], HPLC capacity factors [33], and hexane-acetonitrile partition coefficients [12] for a number of terpenes that helped to yield the total set of descriptors, including the B -descriptor. We then set out a fragmentation scheme for the estimation of B , using the number and type of ethylenic double bonds and the number and type of rings in the terpene as the fragments. The fragments were defined manually and an Excel spreadsheet was used for the calculations. As we dealt with more and more terpenes, we updated the scheme until we reached an equation with the coefficients shown in Table 3. Since we had all the data available, we set out similar fragmentation schemes for E and S , as shown also in Table 3. It is very important to note that these schemes are intended to apply only to aliphatic cyclic terpene hydrocarbons. Furthermore, they are purely empirical schemes. However, the fragment scheme for B , especially, should be of some value in cases where a given set of equations cannot yield a reliable value.

Table 3

Fragment schemes for the estimation of E , S and B for aliphatic cyclic terpene hydrocarbons.

Fragment	E	S	B
----------	-----	-----	-----

Constant term	-0.0108	0.2194	-0.1183
Number of internal double bonds, C-C=C-C	0.1619	0.0575	0.1317
Number of external double bonds, C=CH ₂	0.1704	0.0715	0.1136
Number of -C=C-C=C- groups	-0.0091	-0.0449	-0.0368
Number of -C=C-C=C-C=C- groups	0.2089	0.0771	-0.0619
Total number of rings	0.2702	0.0001	0.1112
Number of 3-membered rings	-0.1409	-0.0159	-0.0060
Number of 4-membered rings	-0.1696	-0.0547	-0.0101
Number of 5-membered rings	-0.1136	-0.0458	-0.0474
Number of 6-membered rings	-0.0486	-0.0430	-0.0080
Number of 7-membered rings	-0.0133	-0.0009	-0.0024
Number of 10-membered rings	0.0183	0.0393	-0.0818
Number of 11-membered rings	0.0228	-0.1219	-0.0680
Number of compounds	88	88	88
SD	0.046	0.005	0.024

The final set of descriptors for the 88 aliphatic cyclic terpene hydrocarbons that we have studied is given in Table 4; in all cases the *A*-descriptor is zero. Also included in Table 4 are the molecular formula of the terpene and the number of equations for GLC retention data used in the analysis, *n*(GLC). Since our analysis yields $\log K_w$ as a matter of course, we include this very difficult to measure parameter as well.

Table 4
Descriptors for terpene hydrocarbons

Terpene	MF	<i>n</i> (GLC)	<i>E</i>	<i>S</i>	<i>B</i>	<i>V</i>	<i>L</i>	$\log K_w$
Santene	C ₉ H ₁₄	5	0.459	0.19	0.17	1.1165	3.898	-0.39
1,3,8-p-Menthatriene	C ₁₀ H ₁₄	3	0.914	0.44	0.30	1.2800	5.138	0.99
Thuja-2,4(10)-diene	C ₁₀ H ₁₄	4	0.638	0.24	0.24	1.2144	4.448	0.09
Tricyclene	C ₁₀ H ₁₆	15	0.500	0.11	0.17	1.1918	4.280	-0.64
α -Pinene	C ₁₀ H ₁₆	65	0.438	0.20	0.14	1.2574	4.256	-0.64
β -Pinene	C ₁₀ H ₁₆	57	0.515	0.19	0.15	1.2574	4.515	-0.57
Limonene	C ₁₀ H ₁₆	84	0.501	0.31	0.23	1.3230	4.688	0.05
α -Terpinene	C ₁₀ H ₁₆	40	0.526	0.25	0.23	1.3230	4.699	-0.09
β -Terpinene	C ₁₀ H ₁₆	48	0.522	0.29	0.22	1.3230	4.869	-0.04
γ -Terpinene	C ₁₀ H ₁₆	53	0.522	0.29	0.22	1.3230	4.840	0.02
Camphene	C ₁₀ H ₁₆	36	0.520	0.20	0.17	1.2574	4.353	-0.45

α -Thujene	C10H16	40	0.376	0.21	0.20	1.2574	4.223	-0.37
Sabinene	C10H16	31	0.457	0.23	0.18	1.2574	4.483	-0.36
2-Carene	C10H16	4	0.489	0.22	0.14	1.2574	4.659	-0.57
Δ -3-Carene	C10H16	50	0.492	0.22	0.14	1.2574	4.679	-0.58
α -Phellandrene	C10H16	40	0.497	0.25	0.22	1.3230	4.643	-0.15
β -Phellandrene	C10H16	39	0.529	0.26	0.20	1.3230	4.759	-0.19
Terpinolene	C10H16	53	0.564	0.29	0.20	1.3230	5.029	-0.12
α -Fenchene	C10H16	24	0.486	0.19	0.19	1.2574	4.330	-0.38
β -Fenchene	C10H16	11	0.451	0.20	0.17	1.2574	4.276	-0.47
Pinane, Z, cis	C10H18	2	0.421	0.12	0.13	1.3004	4.534	-0.95
Pinane, E, trans	C10H18	2	0.421	0.12	0.13	1.3004	4.487	-0.95
Thujane	C10H18	4	0.304	0.16	0.14	1.3004	4.480	-0.83
β -Acoradiene	C15H24	3	0.802	0.26	0.28	1.9189	6.993	-0.18
Aromadendrene	C15H24	16	0.700	0.23	0.29	1.8533	6.839	-0.21
Alloaromadendrene	C15H24	17	0.722	0.23	0.29	1.8533	6.973	-0.20
α -Amorphene	C15H24	4	0.821	0.25	0.36	1.9189	7.074	0.19
α -Bisabolene, E,trans	C15H24	3	0.668	0.35	0.39	1.9845	7.273	0.45
α -Bisabolene, Z, cis	C15H24	5	0.668	0.35	0.39	1.9845	7.069	0.45
β -Bisabolene	C15H24	3	0.656	0.36	0.37	1.9845	7.044	0.37
γ -Bisabolene, E, trans	C15H24	3	0.668	0.35	0.39	1.9845	7.166	0.45
γ -Bisabolene, Z, cis	C15H24	4	0.668	0.35	0.39	1.9845	7.122	0.45
β -Bourbonene	C15H24	19	0.627	0.14	0.20	1.8533	6.645	-0.87
Bicyclogermacrene	C15H24	9	0.800	0.29	0.38	1.9189	7.105	0.38
α -Bergamotene, E, trans	C15H24	13	0.695	0.28	0.38	1.9189	6.711	0.29
α -Bergamotene, Z, cis	C15H24	5	0.695	0.28	0.38	1.9189	6.608	0.29
Cadina-1(2),4-diene, Z, cis	C15H24	7	0.823	0.25	0.36	1.9189	7.358	0.19
α -Cadinene	C15H24	13	0.809	0.29	0.36	1.9189	7.307	0.29
β -Cadinene	C15H24	1	0.791	0.29	0.36	1.9189	6.950	0.28
δ -Cadinene	C15H24	36	0.817	0.29	0.36	1.9189	7.220	0.29
γ -Cadinene	C15H24	23	0.811	0.31	0.34	1.9189	7.162	0.30
α -Caryophyllene	C15H24	13	0.772	0.27	0.32	1.9845	6.708	-0.04
β -Caryophyllene, E, trans	C15H24	31	0.724	0.33	0.26	1.9189	6.613	-0.14
β -Caryophyllene, Z, cis	C15H24	9	0.724	0.33	0.26	1.9189	6.617	-0.14
g-Caryophyllene	C15H24	1	0.699	0.33	0.26	1.9189	6.529	-0.16
α -Cedrene	C15H24	12	0.711	0.18	0.24	1.8533	6.764	-0.57
β -Cedrene	C15H24	7	0.757	0.20	0.22	1.8533	6.785	-0.59
α -Copaene	C15H24	16	0.624	0.14	0.32	1.8533	6.567	-0.34
β -Copaene	C15H24	9	0.624	0.15	0.30	1.8533	6.849	-0.41
α -Cubebene	C15H24	25	0.616	0.17	0.29	1.8533	6.320	-0.41
β -Cubebene	C15H24	19	0.616	0.19	0.27	1.8533	6.624	-0.46
Cyclosativene	C15H24	6	0.613	0.08	0.24	1.7877	6.659	-0.83
Cyperene	C15H24	4	0.720	0.18	0.24	1.8533	6.637	-0.57
β -Elemene	C15H24	20	0.714	0.40	0.33	1.9845	6.411	0.31
δ -Elemene	C15H24	11	0.714	0.38	0.34	1.9845	6.163	0.31
g-Elemene	C15H24	7	0.714	0.38	0.33	1.9845	6.690	0.26
Germacrene A	C15H24	9	0.764	0.45	0.29	1.9845	6.965	0.27
Germacrene B	C15H24	11	0.764	0.45	0.29	1.9845	7.189	0.27
Germacrene D	C15H24	25	0.764	0.40	0.24	1.9845	6.869	-0.10
α -Guaiene	C15H24	6	0.695	0.30	0.31	1.9189	6.821	0.01
β -Guaiene, E, trans	C15H24	3	0.695	0.30	0.31	1.9189	7.047	0.01
β -Guaeine, Z, cis	C15H24	4	0.695	0.30	0.31	1.9189	6.972	0.01

α -Gurjunene	C15H24	10	0.737	0.21	0.28	1.8533	6.659	-0.29
β -Gurjunene	C15H24	7	0.775	0.23	0.26	1.8533	6.792	-0.31
γ -Gurjunene	C15H26	6	0.734	0.30	0.29	1.9619	6.950	-0.11
α -Humulene	C15H24	38	0.764	0.27	0.32	1.9845	6.896	-0.04
Isolongifolene	C15H24	3	0.728	0.14	0.27	1.8533	6.647	-0.52
Longicyclene	C15H24	6	0.656	0.11	0.22	1.7877	6.600	-0.82
Longifolene	C15H24	8	0.757	0.20	0.22	1.8533	6.677	-0.59
α -Longipinene	C15H24	8	0.665	0.18	0.32	1.8533	6.464	-0.21
α -Muurolene	C15H24	19	0.802	0.25	0.36	1.9189	7.138	0.18
γ -Muurolene	C15H24	22	0.802	0.26	0.34	1.9189	7.062	0.11
α -Patchoulene	C15H24	6	0.720	0.18	0.24	1.8533	6.999	-0.57
β -Patchoulene	C15H24	4	0.720	0.18	0.24	1.8533	6.518	-0.57
γ -Patchoulene	C15H24	5	0.720	0.22	0.20	1.8533	7.003	-0.61
α -Santalene	C15H24	5	0.635	0.17	0.29	1.8533	6.801	-0.40
β -Santalene	C15H24	4	0.668	0.26	0.28	1.9189	6.902	-0.26
α -Selinene	C15H24	4	0.823	0.26	0.34	1.9189	6.895	0.12
β -Selinene	C15H24	20	0.765	0.28	0.32	1.9189	7.067	0.04
7-epi- α -Selinene	C15H24	4	0.823	0.26	0.34	1.9189	7.266	0.12
Selena-4,11-diene	C15H24	4	0.823	0.26	0.34	1.9189	6.895	0.12
Selena-3,7(11)-diene	C15H24	6	0.823	0.25	0.36	1.9189	7.416	0.19
Thujopsene	C15H24	10	0.742	0.17	0.29	1.8533	6.898	-0.34
Valencene	C15H24	11	0.807	0.26	0.34	1.9189	7.145	0.11
α -Ylangene	C15H24	16	0.640	0.13	0.33	1.8533	6.557	-0.31
β -Ylangene	C15H24	3	0.640	0.15	0.31	1.8533	6.835	-0.35
α -Zingberene	C15H24	5	0.680	0.30	0.34	1.9845	7.053	0.08
Abietadiene	C20H32	2	0.950	0.20	0.45	2.5148	9.991	0.06
Myrcene	C10H16	48	0.483	0.29	0.32	1.3886	4.513	0.37
α -Ocimene, E, trans	C10H16	42	0.539	0.31	0.31	1.3886	4.797	0.40
β -Ocimene, Z, cis	C10H16	37	0.581	0.29	0.31	1.3886	4.784	0.37
trans-Allo-ocimene	C10H16	9	0.594	0.39	0.26	1.3886	5.179	0.39
cis-Allo-ocimene	C10H16	10	0.600	0.39	0.26	1.3886	5.189	0.40
α -Farnesene, (E, E)	C15H24	18	0.711	0.38	0.45	2.0501	7.039	0.78
α -Farnesene, (Z, E)	C15H24	3	0.711	0.38	0.45	2.0501	6.956	0.78
β -Farnesene, trans, E	C15H24	12	0.711	0.37	0.46	2.0501	6.769	0.80
β -Farnesene, cis, Z	C15H24	4	0.711	0.36	0.46	2.0501	6.756	0.78
Phyt-1-ene	C20H40	1	0.010	0.10	0.10	2.8836	8.517	-2.75
Phyt-2-ene	C20H40	4	0.050	0.10	0.10	2.8836	8.842	-2.73
Phytane	C20H42	1	0.000	0.00	0.00	2.9266	8.727	-3.54
Cadalene	C15H18	4	1.423	0.92	0.25	1.7899	7.674	1.83
α -Calacorene	C15H20	13	0.930	0.65	0.25	1.8329	7.079	0.82
β -Calacorene	C15H20	5	0.930	0.64	0.25	1.8329	7.174	0.79
Calamenene, E, trans	C15H22	16	0.860	0.56	0.20	1.8759	6.932	0.27
Calamenene, Z, cis	C15H22	13	0.860	0.56	0.20	1.8759	6.975	0.27
Cuparene	C15H22	8	0.770	0.54	0.25	1.8759	6.934	0.45
Abietatriene	C20H30	3	0.950	0.65	0.25	2.4718	9.593	0.27

We also studied a much smaller number, twelve, of aliphatic acyclic terpene hydrocarbons. These were easier to deal with than the aliphatic cyclic hydrocarbons because

both S and B could be estimated by analogy from data on other aliphatic acyclic hydrocarbons, and S could also be established from equations for GLC retention data. In the event, there was good agreement between the two sets of S -values, and the final set of descriptors is in Table 4.

There are a few terpene hydrocarbons that are derivatives of benzene or of 1,2,3,4-tetrahydronaphthalene, and we examined seven aromatic hydrocarbons, given as the final entries in Table 4. In these cases, there was only poor agreement between the S -descriptor obtained from the GLC equations, and the S -descriptor obtained by analogy with values for known aromatic hydrocarbons. In all cases, the S -descriptor from the GLC equations was smaller than the values estimated from known aromatic hydrocarbons. Fortunately, calculated values of $\log P_{oct}$ using several methods give reasonably consistent results, see later, and so we selected S and B values that were consistent with results from the GLC equations, with estimations using known values for aromatic hydrocarbons, and with the calculated $\log P_{oct}$ values. This is a quite different analysis from that used for the aliphatic terpene hydrocarbons, but we found it to be the only way that we could estimate descriptors. Our selected values are in Table 4.

4. Discussion

As explained, above, for the 88 aliphatic cyclic hydrocarbons and the 12 aliphatic acyclic hydrocarbons, the key descriptor S was obtained from the GLC equations or from the fragmentation scheme in Table 3, and for most of the compounds the ‘difficult’ descriptor B was obtained from the corresponding fragmentation scheme. We can check if our set of descriptors is reasonable or not by a comparison of $\log P_{oct}$ calculated from the descriptors in Table 4 and $\log P_{oct}$ calculated by standard literature methods. We therefore assemble in Table 5 values of $\log P_{oct}$ calculated using the BioLoom ClogP program [30], the Advanced Chemistry Development program (ACD) and the EPI program as implemented in ChemSpider software [20], the AlogPS program [31] and the SPARC program [32]. These five methods are probably the most used calculations for $\log P_{oct}$. We have shown [11] that experimental values of $\log P_{oct}$ are well represented through Eq. 8. Both the s -coefficient and especially the b -coefficient in Eq. (8) are numerically large and so estimations of $\log P_{oct}$ should provide a quite rigorous test of our descriptors.

$$\text{Log } P_{oct} = 0.088 + 0.562 E - 1.054 S + 0.034 A - 3.460 B + 3.814 V \quad (8)$$

Table 5

Comparison of $\log P_{oct}$ calculated from the descriptors in Table 4 (taken) and $\log P_{oct}$ calculated from literature software.

	Taken	ClogP	ACD	EPI	ALOGPS	SPARC	Experimental
Santene	3.82	4.18	3.82	3.94	3.35	4.03	
1,3,8-p-Menthatriene	3.98	3.96	4.03	4.74	4.15	4.34	
Thuja-2,4(10)-diene	3.99	3.61	3.60	4.47	3.50	5.38	
Tricyclene	4.21	4.38	3.98	4.13	3.43	5.30	
α -Pinene	4.48	4.18	4.31	4.27	3.66	4.72	4.44[20], 4.48 [30], 4.83 [20]
β -Pinene	4.45	4.70	4.24	4.35	3.94	4.84	
Limonene	4.30	4.35	4.55	4.83	4.50	4.71	4.38 [30], 4.50 [20], 4.57 [19]
α -Terpinene	4.35	4.41	4.29	4.75	4.51	4.65	4.25 [30]
β -Terpinene	4.36	4.35	4.25	4.75	4.30	4.70	4.35 [30], 4.50 [20]
γ -Terpinene	4.35	4.35	4.25	4.75	4.36	4.59	4.35 [30], 4.50 [20]
Camphene	4.38	4.70	4.24	4.35	4.56	4.80	4.22 [30], 4.37 [30]
α -Thujene	4.18	4.19	4.02	4.61	4.07	5.11	
Sabinene	4.27	4.03	3.94	4.69	3.04	5.22	
2-Carene	4.44	4.44	4.32	4.61	3.26	5.05	4.44 [20]
D-3-Carene	4.44	4.44	4.32	4.61	3.64	5.05	4.38 [20, 30]
α -Phellandrene	4.39	4.41	4.41	4.62	4.29	4.79	
β -Phellandrene	4.46	4.41	4.39	4.70	3.98	4.91	4.16 [30]
Terpinolene	4.47	4.35	4.21	4.88	3.82	4.69	4.24 [30], 4.47 [20]
α -Fenchene	4.30	4.70	4.24	4.35	4.05	4.88	
β -Fenchene	4.34	4.70	4.19	4.17	4.09	4.84	
Pinane, Z, cis	4.71	5.19	4.69	4.35	4.04	5.11	
Pinane, E, trans	4.73	5.19	4.69	4.35	4.04	5.11	
Thujane	4.57	4.52	4.39	4.70	3.63	5.44	
β -Acoradiene	6.61	6.30	6.56	6.99	5.87	7.10	
Aromadendrene	6.30	6.39	6.41	6.13	3.70	7.55	
Alloaromadendrene	6.32	6.39	6.43	6.13	3.70	7.55	
α -Amorphene	6.36	6.30	6.56	6.19	5.00	7.25	
α -Bisabolene, E, trans	6.31	6.38	6.61	7.05	5.95	7.09	
α -Bisabolene, Z, cis	6.31	6.38	6.61	7.05	5.95	7.09	
β -Bisabolene	6.37	6.38	6.43	7.12	6.01	7.18	
γ -Bisabolene, E, trans	6.31	6.38	6.70	7.18	5.88	7.17	
γ -Bisabolene, Z, cis	6.31	6.38	6.70	7.18	5.88	7.17	
β -Bourbonene	6.67	5.98	6.13	5.44	3.47	7.44	
Bicyclogermacrene	6.24	6.45	6.50	6.22	6.10	7.07	
α -Bergamotene, E, trans	6.19	6.45	6.54	6.57	6.17	7.13	
α -Bergamotene, Z, cis	6.19	6.45	7.03	6.57	6.17	7.13	
Cadina-1(2),4-diene, Z, cis	6.36	6.30	6.34	6.19	5.33	7.13	
α -Cadinene	6.31	6.30	6.56	6.19	5.00	7.25	
β -Cadinene	6.30	6.30	6.43	6.19	4.92	7.16	
δ -Cadinene	6.32	6.30	6.27	6.32	4.92	7.16	
γ -Cadinene	6.36	6.30	6.35	6.27	5.23	7.34	
α -Caryophyllene	6.70	5.74	6.59	6.95	6.07	6.78	
β -Caryophyllene, E, trans	6.57	6.45	6.78	6.30	5.35	6.97	

β -Caryophyllene, Z, cis	6.57	6.45	6.78	6.30	5.35	6.97	
g-Caryophyllene	6.55	6.45	6.42	6.30	5.35	6.97	
α -Cedrene	6.54	6.91	6.38	6.43	5.18	7.64	
β -Cedrene	6.61	6.91	6.23	6.51	4.38	7.74	
α -Copaene	6.25	7.02	6.21	5.36	3.75	7.15	
β -Copaene	6.31	7.02	6.13	5.44	3.38	7.25	
α -Cubebene	6.32	6.50	6.26	6.73	3.49	7.65	
β -Cubebene	6.37	6.50	6.18	6.81	3.49	7.73	
Cyclosativene	5.72	6.70	5.84	5.57	5.33	7.21	
Cyperene	6.54	6.91	6.19	6.21	5.51	7.60	
β -Elemene	6.49	6.36	5.77	7.04	5.40	7.26	
δ -Elemene	6.48	6.36	6.59	6.96	5.39	7.17	
g-Elemene	6.52	6.36	6.49	7.09	5.23	7.27	
Germacrene A	6.61	6.62	6.86	7.12	6.02	6.81	
Germacrene B	6.61	6.62	6.39	7.18	5.87	6.79	
Germacrene D	6.83	6.68	6.57	6.99	5.64	7.00	
α -Guaiene	6.41	6.30	6.60	6.51	5.80	7.04	
β -Guaiene, E, trans	6.41	6.30	6.58	6.51	5.80	7.04	
β -Guaene, Z, cis	6.41	6.30	6.33	6.56	5.52	6.97	
α -Gurjunene	6.38	5.87	6.45	6.18	4.42	7.31	
β -Gurjunene	6.45	6.39	6.43	6.13	3.75	7.59	
γ -Gurjunene	6.66	6.79	6.80	6.40	5.62	7.10	
α -Humulene	6.69	6.77	6.59	6.95	6.07	6.78	
Isolongifolene	6.48	6.91	6.15	6.12	6.04	7.61	
Longicyclene	6.40	6.85	5.91	5.60	5.64	8.09	
Longifolene	6.61	7.17	6.17	5.48	4.65	7.26	
α -Longipinene	6.13	7.17	6.40	5.40	4.45	7.17	
α -Muurolole	6.35	6.30	6.44	6.19	5.00	7.25	
γ -Muurolole	6.41	6.30	6.54	6.27	4.48	7.34	
α -Patchoulene	6.54	6.91	6.31	6.43	5.72	7.63	
β -Patchoulene	6.54	6.65	6.13	5.87	5.68	7.54	
γ -Patchoulene	6.59	6.91	6.37	5.48	5.35	7.72	
α -Santalene	6.33	6.41	6.02	6.43	4.96	8.02	
β -Santalene	6.54	6.73	6.15	6.64	5.56	7.28	
α -Selinene	6.42	6.30	6.41	6.30	5.86	7.25	
β -Selinene	6.43	6.30	6.33	6.38	4.97	7.33	
7-epi- α -Selinene	6.42	6.30	6.41	6.30	5.86	7.25	
Selena-4,11-diene	6.42	6.30	6.63	6.43	5.61	7.18	
Selena-3,7(11)-diene	6.36	6.30	6.73	6.35	5.78	7.24	
Thujopsene	6.39	6.65	6.05	6.12	5.99	7.45	
Valencene	6.41	6.30	6.29	6.30	5.86	7.21	
α -Ylangene	6.24	7.02	6.21	5.36	3.88	7.15	
β -Ylangene	6.29	7.02	6.13	5.44	3.38	7.25	
α -Zingiberene	6.55	6.44	6.60	6.92	5.77	7.12	
Abietadiene	8.44	8.25	8.42		7.28	9.73	
Myrcene	4.17	4.33	4.25	4.88	4.32	4.83	4.17 [30]
α -Ocimene, E, trans	4.17	4.33	4.14	4.88	4.25	4.82	4.17 [20]
β -Ocimene, Z, cis	4.33	4.33	4.20	4.80	4.25	4.82	
trans-Allo-ocimene	4.40	4.36	4.39	4.72	4.36	4.98	4.56 [20]
cis-Allo-ocimene	4.41	4.36	4.39	4.72	4.36	4.98	
α -Farnesene, (E, E)	6.35	6.36	6.30	7.10	5.70	7.18	

α -Farnesene, (Z, E)	6.35	6.36	6.14	7.17	5.70	7.18	
β -Farnesene, trans, E	6.32	6.36	6.14	7.17	5.70	7.27	
β -Farnesene, cis, Z	6.34	6.36	6.14	7.17	5.70	7.27	
Phyt-1-ene	10.64	10.27	9.87	9.73	8.82	11.38	
Phyt-2-ene	10.66	10.27	9.93	9.78	8.76	11.39	
Phytane	11.25	10.75	9.87		9.09	11.59	
Cadalene	5.88	5.74	5.70	5.72	5.70	5.60	
α -Calacorene	6.05	5.88	6.22	6.16	5.47	5.54	
β -Calacorene	6.06	5.88	6.05	6.24	5.45	5.68	
Calamenene, E, trans	6.44	6.18	6.02	6.25	5.60	5.63	
Calamenene, Z, cis	6.44	6.18	6.02	6.25	5.60	5.63	
Cuparene	6.24	6.26	5.89	6.19	6.06	5.74	
Abietatriene	8.50	8.13	8.20	7.76	7.07	7.55	

Results of the $\log P_{oct}$ calculations are in Table 5. It is not our intention to compare the various literature methods for these calculations, but we note that the AlogPS method usually gives $\log P_{oct}$ values smaller than the average, and that the SPARC method usually gives values larger than the average. The differences in the various calculations can become very large, so that for α -cubebene they amount to over four log units. The $\log P_{oct}$ values obtained from our descriptors, that include the *S*- and *B*-descriptors, are invariably within the upper and lower limits of the calculated descriptors, and are usually close to the average value of the ClogP, ACD, and EPI calculations. We therefore conclude that the fragment calculations for *S* and *B* do yield reasonable values, so that the total set of descriptors then leads to reasonable values for $\log P_{oct}$. There is then every reason to think that the descriptors set out in Table 4 will yield good estimates of $\log P$ and $\log K$ values for the very large number of water-solvent and gas-solvent systems for which we have the required equations [8]. In a similar vein, the descriptors can be used to estimate nasal pungency thresholds. The calculations of $\log P$, $\log K$ and NPT are completely trivial – all that is needed is to combine the descriptors with the corresponding equation coefficients.

There are a few cases in which our predictions can be compared to experiment. The most recent equation for NPT values [34] is Eq. (9), and our predicted values using the descriptors in Table 4 are compared with the observed values [17] in Table 6. The agreement is excellent and although we have only two sets of data, there seems no reason why the descriptors in Table 4 should not yield reasonable predictions of nasal pungency thresholds.

$$\text{Log}(1/\text{NPT}) = -7.770 + 1.543 S + 3.296 A + 0.876 B + 0.816 L \quad (9)$$

A more extensive set of experimental results are those of Falk et al [35] on gas-blood, gas-water and gas-olive oil partitions at 37°C. We have equations for gas-blood partitions [36], gas-olive oil partitions [37] and gas-water partitions [38], Eq. (10) – Eq. (12), all of which refer to 37°C, and we compare our predictions with experimental values in Table 6.

$$\text{Log } K_{\text{blood}} (37^{\circ}\text{C}) = -1.062 + 0.460 E + 1.067 S + 3.777 A + 2.556 B + 0.375 L \quad (10)$$

$$\text{Log } K_{\text{olive oil}} (37^{\circ}\text{C}) = -0.156 - 0.254 E + 0.859 S + 1.656 A + 0.873 L \quad (11)$$

$$\text{Log } K_w (37^{\circ}\text{C}) = -1.035 + 0.575 E + 2.482 S + 3.601 A + 4.561 B - 0.973 V \quad (12)$$

Agreement is generally very good, even for equations that have a substantial *b*-coefficient, so for the four terpenes studied by Falk et al. [35] the *B*-descriptor that is compatible with the coefficients in Table 3 seems to lead to reasonable values of the gas-solvent partition coefficients.

We note that the descriptors in Table 4 can be used for the prediction of gas-water partition coefficients from 0°C to 100°C, in combination with the equation coefficients we have already listed [38].

Table 6

Comparison of predicted and experimental values of log (1/NPT) and of log *K* for some gas-solvent partitions at 37°C

Terpene	Pred	Obs
Log (1/NPT)		
Δ-3-Carene	-3.49	-3.21
α-Terpinene	-3.35	-3.30
Log <i>K</i> (gas-blood)		
α-Pinene	1.31	1.18 ± 0.11
β-Pinene	1.45	1.36 ± 0.11
Δ-3-Carene	1.51	1.51 ± 0.09
Limonene	1.85	1.62 ± 0.08
Log <i>K</i> (gas-oil)		
α-Pinene	3.62	3.46 ± 0.07
β-Pinene	3.82	3.63 ± 0.07

Δ -3-Carene	3.99	3.70 ± 0.10
Limonene	4.08	3.76 ± 0.11
Log K (gas-water)		
α -Pinene	-0.87	-0.92 ± 1.30
β -Pinene	-0.81	-0.92 ± 0.98
Δ -3-Carene	-0.79	-0.39 ± 0.61
Limonene	-0.22	0.20 ± 0.36

It is now possible to use the coefficients in Table 3 to predict E , S and B for aliphatic cyclic terpene hydrocarbons. Then if GLC retention data are known for a few systems for which equations are available, the entire set of descriptors E , S , B , V and L can be obtained. We use the retention data of Babushok et al [27] for the systems characterized through Eq. (4), (5) and (6) and calculate E , S , B , V as suggested. These are given in Table 7 for three terpenes listed by Babushok et al [27]. As before, we can use the obtained set of descriptors to estimate $\log P_{oct}$ through Eq. (8) and can compare our estimated values with literature calculations, as shown in Table 7. There is good agreement between the $\log P_{oct}$ values predicted from our descriptors and those calculated by literature methods. This indicates that, indeed, we are in a position reliably to estimate E , S and B , and of course to calculate V . A knowledge of these descriptors is enough to be able to predict $\log P$ values for partition from water to octanol and to a very large number of other solvents. Experimental data on GLC retention on just a few characterized phases are all that is needed to determine the L -descriptor. Then partition coefficients can be estimated for transfer from the gas phase to numerous solvents, as well as a large number of biological properties, including the important nasal pungency thresholds. We give predictions for the latter in Table 7, with values of NPT in ppm. Both α -bulnesene and β -sesquiphellandrene are quite potent vapours, of about the same potency as that of hexanoic acid, $\log (1/\text{NPT}) = -1.30$ [14].

Table 7

Estimation of descriptors through coefficients in Table 3, and derived estimated values of $\log P_{oct}$

Descriptors	MF	n(L)	E	S	B	V	L	$\log Kw$
Verbenene	C10H14	3	0.635	0.25	0.29	1.2144	4.467	0.36

α -Bulnesene	C15H24	2	0.735	0.30	0.30	1.9189	7.107	-0.02
β -Sesquiphellandrene	C15H24	3	0.696	0.36	0.33	1.9845	7.161	0.20
Log P_{oct}	Pred	ClogP	ACD	EPI	ALOGP	SPARC		log(1/NPT)
Verbenene	3.81	3.76	3.86	4.13	3.64	5.01		-3.49
α -Bulnesene	6.47	6.30	6.73	6.27	n/a	7.11		-1.25
β -Sesquiphellandrene	6.53	6.44	6.52	6.99	6.36	7.33		-1.08

References

- [1] C. S. Sell, *The Chemistry of Fragrances*, 2nd Ed., RSC Publishing, Cambridge, 2006.
- [2] S. Meakins, *The safety and toxicology of fragrances*, in C. S. Sell, Ed. *The Chemistry of Fragrances*, 2nd Ed., RSC Publishing, Cambridge, 2006.
- [3] J. E. Cometto- Muñiz, W. S. Cain, M. H. Abraham, R. Kumarsingh, *Pharmacol. Biochem. Behavior*, 60 (1998) 765-770
- [4] Y. Nagata, *Measurement of odor threshold by triangle odor bag method*. In *Odor measurement review*. Office of Odor, Noise and Vibration. Environmental Management Bureau, Ministry of Environment, Tokyo, pp. 118-127, 2003.
- [5] O. Rodriguez, M. A. Teixeira, A. E. Rodrigues, *Flavour and Fragr. J.* (2011) Wileyonlinelibrary.com, DOI 10.1002/ffj.2076
- [6] M. H. Abraham, *Chem. Soc. Revs.*, 22 (1993) 73-83.
- [7] M. H. Abraham, A. Ibrahim, A. M. Zissimos, *J. Chromatogr. A*, 1037 (2004) 29-47.
- [8] M. H. Abraham, R. E. Smith, R. Luchtefeld, A. J. Boorem, R. Luo, W. E. Acree, Jr., *J. Pharm. Sci.*, 99 (2010) 1500-1515.
- [9] M. H. Abraham, J. C. McGowan, *Chromatographia*, 23 (1987) 243-246.
- [10] M. H. Abraham, J. Andonian-Haftvan, G. S. Whiting, A. Leo, R. W. Taft, *J. Chem. Soc. Perkin Trans. 2*, (1994) 1777-1791.
- [11] M. H. Abraham, H. S. Chadha, G. S. Whiting, R. C. Mitchell, *J. Pharm. Sci.*, 83 (1994) 1085-1100.
- [12] H. Ahmed, C. F. Poole, *J. Chromatogr. A.*, 1104 (2006) 82-90.
- [13] I. G. Zenkevich and A. S. Kushakova, *Russ. J. Gen. Chem.*, 81 (2011) 337-344.
- [14] M. H. Abraham, R. Kumarsingh, J. E. Cometto- Muñiz, W. S. Cain, *Archives Toxicol.*,

- 72 (1998) 227-232.
- [15] M. H. Abraham, J. M. R. Gola, J. E. Cometto- Muñiz, W. S. Cain, *Indoor Built Environ.*, 10 (2001) 252-257.
- [16] M. H. Abraham, R. Sánchez-Moreno, J. E. Cometto-Muñiz, W. S. Cain, *Chem. Senses*, 37 (2012) 207-218.
- [17] M. H. Abraham, R. Kumarsingh, J. E. Cometto- Muñiz, W. S. Cain, M. Roses, E. Bosch, M. L. Diaz, *J. Chem. Soc., Perkin Trans.2*, (1998) 2405-2411.
- [18] T. Karunasekara. C. F. Poole, *J. Chromatogr. A.*, 1235 (2012) 159-165.
- [19] J. Li, E. M. Perdue, S.G.Paviostathis, R. Araujo, *Environ. Int.*, 1998, 24 (1998), 353-358.
- [20] ChemSpider, the EPI suite and the ACD suite.
- [21] I. Fichan, C. Larroche, J. B. Gros, *J. Chem. Eng. Data*, 44 (1999) 56-62.
- [22] J. L. Simonsen, *The Terpenes*, vol. I, second Ed., Cambridge University Press, 1947.
- [23] J. Simonsen and L. N. Owen, *The Terpenes*, vol. II, second Ed., Cambridge University Press, 1949.
- [24] J. Simonsen and D. H. R. Barton, *The Terpenes*, vol. III, Cambridge University Press, 1952.
- [25] www.cas.org/products/scifinder
- [26] M. H. Abraham, J. C. McGowan, *Chromatographia*, 23 (1987) 243-246.
- [27] V. I. Babushok, P. J. Linstrom, I. G. Zenkevich, *J. Phys. Chem. Ref. Data*, 40 (2011) 043101-1 – 043101-46.
- [28] B. B. Darjazi, A. Rustaiyan, A. Talaei, A. Khalighi, K. Larijani, B. Golein, E. Hayatbakhsh, R. Taghizad, *Iran J. Chem. Chem. Eng.*, 28 (2009) 99-111.
- [29] M. H. Abraham, W. E. Acree, Jr., A. J. Leo, D. Hoekman, *New J. Chem.*, 33 (2009) 568-573.
- [30] BioLoom, BioByte Corp, 201 W. 4th Street, #204 Claremont, CA 91711-4707, USA.
- [31] www.vcclab.org/lab/alogps/
- [32] archemcalc.com/sparc/pKa/
- [33] S. Griffin, S. G. Wyllie, J. Markham, *J. Chromatogr.A*, 864 (1999) 221-228.
- [34] M. H. Abraham, R. Sánchez-Moreno, J. Gil-Lostes, J. E. Cometto-Muñiz, W. S.Cain. *Physicochemical Modeling of Sensory Irritation in Humans and Experimental Animals*, Chapter 25, pp 376-389, in ‘*Toxicology of the Nose and Upper Airways*, Ed. J. B. Morris, D. J. Shusterman. Informa Healthcare USA, New York, 2010.

- [35] A. Falk, E. Gullstrand, A. Lof, E. Wigaeus-Hjelm, *Brit. J. Ind. Med.* 47 (1990) 62-64.
- [36] M. H. Abraham, A. Ibrahim, W. E. Acree, Jr., *Chem. Res. Toxicol.* 18 (2005) 904-911.
- [37] M. H. Abraham, A. Ibrahim, *J. Chem. Inf. Model.* 46 (2006) 1735-1741.
- [38] M. H. Abraham, W. E. Acree, Jr., *Phys. Chem. Chem. Phys.*, 14 (2012) 7433-7440.

This is a pre-copyedited, author-produced version of an article accepted for publication in Journal of Chromatography A following peer review. The version of record Journal of Chromatography A **1293**:133-141, 2013 is available online at: <http://www.sciencedirect.com/science/article/pii/S0021967313005517> - doi: 10.1016/j.chroma.2013.03.068