Accepted Manuscript

Title: Determination of solvation descriptors for terpene hydrocarbons from chromatographic measurements

Author: Michael H Abraham Joelle R.M. Gola Javier Gil-Lostes William E. Acree Jr. J. Enrique Cometto- Muñiz

Please cite this article as: M.H. Abraham, J.R.M. Gola, J. Gil-Lostes, W.E. Acree Jr., J.E.C.- Muñiz, Determination of solvation descriptors for terpene hydrocarbons from chromatographic measurements, *Journal of Chromatography A* (2013), http://dx.doi.org/10.1016/j.chroma.2013.03.068

1

1   Highlights
2
3
4   •   Prediction of chromatographic rentention data for terpene hydrocarbons
5   •   Development of predictive linear free energy relationships
6   •   Calculation of solute descriptors for terpene hydrocarbons
7
8
9

9 Determination of solvation descriptors for terpene hydrocarbons from

10 chromatographic measurements

11 Michael H Abraham, [a] [*] Joelle R. M. Gola, [a] Javier Gil-Lostes, [a] William E. Acree, Jr, [b] and

12 J. Enrique Cometto- Muñiz [c]

13 [a] Department of Chemistry, University College London, 20 Gordon St, London WC1H 0AJ,

14 UK. Email: m.h.abraham@ucl.ac.uk

15 [b] Department of Chemistry, 1155 Union Circle Drive #305070, University of North Texas,

16 Denton, TX 76203-5017, USA

17 [c] Chemosensory Perception Laboratory, Department of Surgery (Otolaryngology), University

18 of California, San Diego, La Jolla, CA 92093-0957

19

20 ABSTRACT

21 Gas chromatographic retention data on 107 terpene hydrocarbons from the literature together

22 with other data have been used to obtain a set of Abraham descriptors for these 107

23 compounds. For 88 aliphatic cyclic terpene hydrocarbons, a fragmentation scheme was

24 constructed that allows key descriptors to be estimated just from structure. The total set of

25 descriptors, including those estimated by the fragmentation schemes, were then used to

26 predict water-octanol partition coefficients for the 88 compounds, there being good

27 agreement with values calculated from a number of well-known programs. For a small

28 number of terpene hydrocarbons, there was good agreement between predicted and

29 experimental values of nasal pungency thresholds, and predicted and experimental gas-blood,

30 gas-oil, and gas-water partition coefficients. It is suggested that the descriptors obtained for

31 the 107 terpene hydrocarbons can be used to predict water-solvent partition coefficients, gas-

32 solvent partition coefficients, and partition coefficients in a number of biological systems.

33

34

35 *Keywords:*

36 Terpenes

37 Gas liquid chromatography

38 Partition coefficient

39 Descriptors

40 Linear free energy relationships

41

1. Introduction

42

43 Terpenes are found in a wide variety of essential oils. Geraniol is the main constituent of

44 geranium oil, limonene the major constituent of lemon oil, menthol the principal constituent

45 of pepermint oil and α-pinene the main constituent of turpentine. Many terpenes are of

46 industrial importance; Sell [1] lists 20 terpenes that are used on an industrial scale, for

47 example 30,000 tons of myrcene per annum, and limonene has widespread use as a fragrance

48 component.

49 In spite of their widespread occurrence and use, there is very little known as to the effect

50 of terpenes on humans, other than skin irritation and sensitization [2]. Cometto-Muñiz et al.

51 [3] determined odor detection thresholds (ODT), eye irritation thresholds (EIT) and nasal

52 pungency thresholds (NPT) for eleven terpenes. Nagata [4] included only three terpenes in

53 his extensive investigation on ODTs, and Rodriguez et al. [5] only seven terpenes in ODT

54 values for 100 compounds. There is an enormous number of terpenes likely to be encountered

55 in everyday life, and some method of predicting their ODT, EIT, and NPT thresholds

56 thresholds is clearly of importance. Since very little data exists on physicochemical properties

57 of terpenes, the ability to predict such properties would also be extremely useful. Our

58 method is based on the two linear free energy relationships, LFERs, Eq. (1) and Eq. (2) [6-8].

59 $$SP = c + e\,E + s\,S + a\,A + b\,B + v\,V \tag{1}$$

60

61 $$SP = c + e\,E + s\,S + a\,A + b\,B + l\,L \tag{2}$$

62 The dependent variable, SP, is some property of a series of solutes in a given system. SP

63 may be partition coefficients as, log P_s, for a series of solutes in a given water-solvent system

64 or may be log K_s for a series of solutes in a given gas-solvent system. The former are mostly

65 used in Eq. (1) and the latter in Eq. (2). Other systems can also be used. For example SP in

66 Eq. (2) can be a set of retention data for a series of solutes on a given gas liquid

67 chromatographic, GLC, stationary phase.

68 The independent variables, or Abraham descriptors, in Eq. (1) and Eq. (2) are properties

69 of solutes as follows [6-8]: E is an excess molar refraction in $cm^3\,mol^{-1}/10$, S is a combined

70 dipolarity/polarizability descriptor, A is the overall solute hydrogen bond acidity, B is the

71 overall solute hydrogen bond basicity, V is McGowan's [9] characteristic molecular volume

72 in $cm^3\,mol^{-1}/100$ and L is the gas to hexadecane partition coefficient at 25^oC, as log K_s.

73   The constant *c* and the set of coefficients, *e, s, a, b, v* and *l* characterise the system and

74   are determined by multiple linear regression analysis. These coefficients are not just fitting

75   coefficients, but represent the complementary properties of the system: *e* gives the

76   contribution in Eq. (1) and Eq. (2) of the solute-system general dispersion interaction, *s* is the

77   solvent dipolarity (plus some polarizability), *a* is the system hydrogen bond basicity (the

78   complimentary property to solute hydrogen bond acidity) and *b* is the system hydrogen bond

79   acidity (the complimentary property to solute hydrogen bond basicity). The coefficients in

80   Eq. (1) and Eq. (2) for a few processes that have been used to obtain descriptors for terpenes

81   are in Table 1 [8, 10-13].

82   Abraham et al. [14] set out an equation for NPT values, based on Eq. (2) and later

83   incorporated a number of terpenes into the algorithm [15]. Algorithms that included a few

84   terpenes were also constructed for EIT [15] and ODT [16] values. These equations required

85   knowledge of the descriptors in Eq. 2 for the terpenes, but it has not been easy to obtain these

86   for any substantial number.

87   Descriptors for 30 terpenes have been obtained [17], mostly from GLC retention data

88   on various stationary phases. It was pointed out [17] that none of the GLC stationary phases

89   possessed any hydrogen bond acidity, and hence that the important *B*-descriptor could not be

90   obtained from GLC data. Abraham et al. [17] managed to obtain the *B*-descriptor for a few

91   terpenes from known water-octanol partition coefficients, as log P_{oct}, and from their own

92   measurements of HPLC capacity factors but for most of the terpenes they were only able to

93   estimate *B*-descriptors. Ahmed and Poole [12] listed descriptors for a large number of

94   compounds, including 23 terpenes. They gave no reference for the origin of the descriptors

95   but by inspection those for 19 of the terpenes were taken from the paper by Abraham et al.

96   [17]. Thus the *B*-descriptor for these 19 terpenes is subject to the same uncertainty as

97   mentioned by Abraham et al.[17]. In a later paper, Karunasekara and Poole [18] obtained

98   descriptors for 15 terpenes using a combination of GLC data with values of log P_{oct} for nine

99   terpenes and hexane-acetonitrile partition coefficients, as log P_{hex-me}. The *b*-coefficient for

100   the water-octanol system is numerically very large (-3.460) and so can lead to reasonably

101   accurate values of the *B*-descriptor; the *b*-coefficient in the hexane-acetonitrile system is

102   numerically much smaller (-0.966) and might be expected to lead to less accurate values. We

103   can roughly estimate the possible error in the *B*-descriptor, through the term SD(reg)/*b* where

104   SD(reg) is the standard deviation of the particular regression equation used, and *b* is the *b*-

105   coefficient. For the octanol-water system the term is 0.116/3.46 = 0.034 and for the hexane-

106   acetonitrile system the term is 0.117/0.966 = 0.121, very much larger than the estimated error

107   for the octanol-water system. Thus, as pointed out before [17], unless a water-octanol

108   partition coefficient is available (or very rarely another water-solvent system with a large *b*-

109   coefficient), it is very difficult to obtain a reliable *B*-descriptor. Both Abraham et al [17] and

110   Karunasekara and Poole [18] used experimental values of the gas-water partition coefficient,

111   K_w, in their descriptor determinations. These experimental values are invariably obtained

112   from the terpene solubility in water and the terpene vapour pressure. Unfortunately, variation

113   in reported experimental values suggest that there is likely to be considerable uncertainty in

114   literature values of log K_w. For example, log K_w for limonene at 25°C is reported as -0.373

115   [19], -0.144 [20] and + 0.233 [21] so that extreme caution is necessary if log K_w values are

116   used.

117   In the present work we concentrate on terpene hydrocarbons, and collect in Table 2 the

118   Abraham descriptors for the few terpene hydrocarbons that have been obtained through the

119   use of GLC retention data coupled with an experimental value of log P_{oct} [12, 16]. Although

120   the terpenes in Table 2 are structurally quite simple, there are sometimes quite large

121   differences in the values of the descriptors for a given compound. Note that for all the terpene

122   hydrocarbons we shall deal with, the *A*-descriptor is zero, and so we omit the zero *A*-values

123   from Table 2.

124   We therefore set out to obtain a coherent set of descriptors for terpene hydrocarbons in

125   the hope that we would have enough reliable values for the *B*-descriptor to be able to

126   construct an algorithm for the prediction of *B*-values.

127

128   **2. Methods**

129   Our method for the determination of descriptors for a given terpene is based on Eq. (1)

130   and Eq. (2). The 'unknown' descriptors are usually *E, S, A, B, V* and *L* but for the terpene

131   hydrocarbons the unknown descriptors are *E, S, B, V* and *L.* Values of *E* were obtained from

132   known experimental refractive indices at 20°C [22-25], exactly as described before [6, 7], and

133   *V* can simply be calculated from the number of atoms and rings in a molecule [6, 26]. If a

134   water-solvent partition coefficient, P_s, is available, this can be converted into the

135   corresponding gas-solvent partition coefficient, K_s, through Eq. (3). This requires a

136   knowledge of K_w; as we have seen, literature values of K_w (or log K_w) are not very reliable,

137   and we prefer to use log K_w as an extra descriptor to be determined.

138

139    $P_s = K_s / K_w$ (3)

140

141   Then the 'unknown' descriptors are S, B, L and log K_w. There is little problem in obtaining

142   equations for GLC retention data, following Eq. (2), and we have some 90 such equations

143   constructed mostly from retention indices on the constituents of natural oils and waxes. In a

144   very useful and important paper, Babushok et al. [27] surveyed the literature on retention

145   indices, I, on three particular stationary phases, viz.: dimethylsilicone (DIMS),

146   dimethylsilicone with 5% phenyl groups (DIMSP) and polyethylene glycol (PEG). Since

147   their data base covers a wide range of compounds, we thought that our equations for I-values

148   might be of use to other workers who wish to obtain descriptors from GLC data.

149

150   I(DIMS) = 39.1 - 39.2 E + 107.5 S + 128.9 A + 205.0 L (4)
151
152   N = 174, SD = 15.9, R^2 = 0.998, F =17338.6, PRESS = 45930, Q^2 = 0.997, PSD = 16.5
153
154
155
156   I(DIMSP) = 45.9 - 42.3 E + 135.9 S + 126.9 A + 204.0 L (5)
157   N = 173, SD = 13.8, R^2 = 0.998, F =22907.9, PRESS = 34605, Q^2 = 0.998, PSD = 14.4

158

159   I(PEG) = -62.2 + 91.8 E + 652.2 S + 1038.5 A + 213.3 L (6)

160   N = 172, SD = 35.3, R^2 = 0.993, F =5723.7, PRESS = 226094, Q^2 = 0.992, PSD = 36.8

161

162   In some instances, the number of compounds for which we had descriptors was very small,

163   the large majority of compounds being terpenes, and so it was not possible to obtain an

164   equation for the I-values. We then waited until we had determined descriptors for enough

165   terpenes (through other equations) to construct an equation. An example is the set of retention

166   indices obtained by Darjazi et al. [28] using a DB-5 fused silica column.

167
168
169   I(DB-5) = 62.7 - 16.0 E + 124.2 S + 62.5 A + 200.5 L (7)
170
171   N = 55, SD = 10.3, R^2 = 0.998, F = 5994.4, PRESS = 8679.4, Q^2 = 0.997, PSD = 13.2
172
173

174 In Eqs. (5)-(7), N is the number of data points, that is compounds, R is the correlation

175 coefficient, F is the Fisher F-statistic, PRESS and Q^2 are the leave-one-out statistics and PSD

176 is the predictive standard deviation [29]. The rather better statistics of Eq. (7) is probably due

177 to the data of Babushok et al. [27] being derived from several different sources. As is

178 invariably the case for GLC retention data, none of the equations has a statistically significant

179 term in *B*.

180   For a few terpene hydrocarbons, experimental values of log P_{oct} are listed in the EPI

181 software program, available through ChemSpider [20] and in the BioLoom software program

182 [30]. We also calculated log P_{oct} using a number of the most widely used programes, viz the

183 EPI and the ACD programes [20], the ClogP program [30], the AlogPS program [31] and

184 SPARC [32]. Griffin et al. [33] have measured HPLC capacity factors on a C-18 stationary

185 phase using a number of water-methanol eluants for a large number of terpenes, including 10

186 terpene hydrocarbons. Zenkevich and Kushakova [13] have determined perfluorodecalin-

187 acetonitrile partition coefficients for a large number of compounds, including some terpenes.

188 We have used their data to obtain equation coefficients as shown in Table 1. The *b*-

189 coefficient is very small, 0.150, and so the equation is not at all useful in the determination of

190 the solute *B*-descriptor.

191   The set of equations on the lines of Eq. (1) and Eq (2) together with the corresponding

192 dependent variables were then solved for the unknown descriptors *S, B, L* and log K_w by the

193 trial-and-error procedure 'Solver' in Microsoft Excel. The equations that we used are in Table

194 1, together with a large number of equations for GLC retention data that are of little intrinsic

195 value.

196

197 **3. Results**

198   By far the main set of terpene hydrocarbons are aliphatic cyclic compounds with a wide

199 variety of number and sizes of rings. For several of the terpene hydrocarbons we had a large

200 number of equations for GLC retention data, 65 in the case of α-pinene. Many of the

201 equations have large values of the s-coefficient, see for example Eq. (6), so that there is little

202 problem in obtaining the corresponding S-descriptor. We also had experimental values of log

203 P_{oct} [20, 30], HPLC capacity factors [33], and hexane-acetonitrile partition coefficients [12]

204 for a number of terpenes that helped to yield the total set of descriptors, including the *B*-

205 descriptor. We then set out a fragmentation scheme for the estimation of *B*, using the number

206 and type of ethylenic double bonds and the number and type of rings in the terpene as the

207 fragments. The fragments were defined manually and an Excel spreadsheet was used for the

208 calculations. As we dealt with more and more terpenes, we updated the scheme until we

209 reached an equation with the coefficients shown in Table 3. Since we had all the data

210 available, we set out similar fragmentation schemes for *E* and *S*, as shown also in Table 3. It

211 is very important to note that these schemes are intended to apply only to aliphatic cyclic

212 terpene hydrocarbons. Furthermore, they are purely empirical schemes. However, the

213 fragment scheme for *B*, especially, should be of some value in cases where a given set of

214 equations cannot yield a reliable value. We also used the same set of fragments to set out a

215 scheme for *L,* although we had to delete values for the saturated cyclic terpenes. Values of *E,*

216 *S, B* and *L* obtained through the fragment schemes given in Table 3 are intended to be used in

217 Eq. (1) and Eq (2).

218

219 The final set of descriptors for the 88 aliphatic cyclic terpene hydrocarbons that we have

220 studied is given in Table 4; in all cases the *A*-descriptor is zero. Also included in Table 4 are

221 the molecular formula of the terpene and the number of equations for GLC retention data

222 used in the analysis, n(GLC). Since our analysis yields log K_w as a matter of course, we

223 include this very difficult to measure parameter as well. Where descriptors could be obtained

224 from experimental data, for example *E* from a refractive index or *B* from a water-octanol

225 partition coefficient these were taken as the preferred values.

226

227 We also studied a much smaller number, twelve, of aliphatic acyclic terpene

228 hydrocarbons. These were easier to deal with than the aliphatic cyclic hydrocarbons because

229 both *S* and *B* could be estimated by analogy from data on other aliphatic acyclic

230 hydrocarbons, and *S* could also be established from equations for GLC retention data. In the

231 event, there was good agreement between the two sets of S-values, and the final set of

232 descriptors is in Table 4.

233 There are a few terpene hydrocarbons that are derivatives of benzene or of 1,2,3,4-

234 tetrahydronaphthalene, and we examined seven aromatic hydrocarbons, given as the final

235 entries in Table 4. In these cases, there was only poor agreement between the *S*-descriptor

236 obtained from the GLC equations, and the *S*-descriptor obtained by analogy with values for

237 known aromatic hydrocarbons. In all cases, the *S*-descriptor from the GLC equations was

238 smaller than the values estimated from known aromatic hydrocarbons. Fortunately,

239 calculated values of log P_{oct} using several methods give reasonably consistent results, see

240   later, and so we selected S and B values that were consistent with results from the GLC

241   equations, with estimations using known values for aromatic hydrocarbons, and with the

242   calculated log P_{oct} values. This is a quite different analysis from that used for the aliphatic

243   terpene hydrocarbons, but we found it to be the only way that we could estimate descriptors.

244   Our selected values are in Table 4.

245

246   **4. Discussion**

247   As explained, above, for the 88 aliphatic cyclic hydrocarbons and the 12 aliphatic acyclic

248   hydrocarbons, the key descriptor S was obtained from the GLC equations or from the

249   fragmentation scheme in Table 3, and for most of the compounds the 'difficult' descriptor B

250   was obtained from the corresponding fragmentation scheme. We can check if our set of

251   descriptors is reasonable or not by a comparison of log P_{oct} calculated from the descriptors in

252   Table 4 and log P_{oct} calculated by standard literature methods. We therefore assemble in

253   Table 5 values of log P_{oct} calculated using the BioLoom ClogP program [30], the Advanced

254   Chemistry Development program (ACD) and the EPI program as implemented in

255   ChemSpider software [20], the AlogPS program [31] and the SPARC program [32]. These

256   five methods are probably the most used calculations for log P_{oct}. We have shown [11] that

257   experimental values of log P_{oct} are well represented through Eq. 8. Both the s-coefficient and

258   especially the b-coefficient in Eq. (8) are numerically large and so estimations of log P_{oct}

259   should provide a quite rigorous test of our descriptors.

260

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

262   Results of the log P_{oct} calculations are in Table 5. It is not our intention to compare the

263   various literature methods for these calculations, but we note that the AlogPS method usually

264   gives log P_{oct} values smaller than the average, and that the SPARC method usually gives

265   values larger than the average. The differences in the various calculations can become very

266   large, so that for α-cubebene they amount to over four log units. The log P_{oct} values obtained

267   from our descriptors, that include the S- and B-descriptors, are invariably within the upper

268   and lower limits of the calculated descriptors, and are usually close to the average value of

269   the ClogP, ACD, and EPI calculations. We therefore conclude that the fragment calculations

270   for S and B do yield reasonable values, so that the total set of descriptors then leads to

271 reasonable values for log P_{oct}. There is then every reason to think that the descriptors set out

272 in Table 4 will yield good estimates of log P and log K values for the very large number of

273 water-solvent and gas-solvent systems for which we have the required equations [8]. In a

274 similar vein, the descriptors can be used to estimate nasal pungency thresholds. The

275 calculations of log P, log K and NPT are completely trivial – all that is needed is to combine

276 the descriptors with the corresponding equation coefficients.

277   There are a few cases in which our predictions can be compared to experiment. The most

278 recent equation for NPT values [34] is Eq. (9), and our predicted values using the descriptors

279 in Table 4 are compared with the observed values [17] in Table 6. The agreement is excellent

280 and although we have only two sets of data, there seems no reason why the descriptors in

281 Table 4 should not yield reasonable predictions of nasal pungency thresholds.

282

283 $\text{Log}(1/NPT) = -7.770 + 1.543\,S + 3.296\,A + 0.876\,B + 0.816\,L$ (9)

284

285   A more extensive set of experimental results are those of Falk et al [35] on gas-blood,

286 gas-water and gas-olive oil partitions at 37°C. We have equations for gas-blood partitions

287 [36], gas-olive oil partitions [37] and gas-water partitions [38], Eq. (10) – Eq. (12), all of

288 which refer to 37°C, and we compare our predictions with experimental values in Table 6.

289 $\text{Log}\,K_{blood}\,(37^{o}\text{C}) = -1.062 + 0.460\,E + 1.067\,S + 3.777\,A + 2.556\,B + 0.375\,L$ (10)

290 $\text{Log}\,K_{olive\,oil}\,(37^{o}\text{C}) = -0.156 - 0.254\,E + 0.859\,S + 1.656\,A + 0.873\,L$ (11)

291 $\text{Log}\,K_{w}\,(37^{o}\text{C}) = -1.035 + 0.575\,E + 2.482\,S + 3.601\,A + 4.561\,B - 0.973\,V$ (12)

292 Agreement is generally very good, even for equations that have a substantial b-coefficient, so

293 for the four terpenes studied by Falk et al. [35] the B-descriptor that is compatible with the

294 coefficients in Table 3 seems to lead to reasonable values of the gas-solvent partition

295 coefficients.

296   We note that the descriptors in Table 4 can be used for the prediction of gas-water

297 partition coefficients from 0°C to 100°C, in combination with the equation coefficients we

298 have already listed [38].

299

300   It is now possible to use the coefficients in Table 3 to predict E, S and B for aliphatic

301 cyclic terpene hydrocarbons. Then if GLC retention data are known for a few systems for

302 which equations are available, the entire set of descriptors E, S, B, V and L can be obtained.

303 We use the retention data of Babushok et al [27] for the systems characterized through Eq.

304   (4), (5) and (6) and calculate *E, S, B, V* as suggested. These are given in Table 7 for three

305   terpenes listed by Babushok et al [27]. As before, we can use the obtained set of descriptors

306   to estimate log P_{oct} through Eq. (8) and can compare our estimated values with literature

307   calculations, as shown in Table 7. There is good agreement between the log P_{oct} values

308   predicted from our descriptors and those calculated by literature methods. Experimental data

309   on GLC retention on just a few characterized phases are all that is needed to determine the *L*-

310   descriptor. Then partition coefficients can be estimated for transfer from the gas phase to

311   numerous solvents, as well as a large number of biological properties, including the important

312   nasal pungency thresholds. We give predictions for the latter in Table 7, with values of NPT

313   in ppm. Both α-bulnesene and β-sesquiphellandrene are quite potent vapours, of about the

314   same potency as that of hexanoic acid, which has log (1/NPT) = -1.30 [14].

315

316   We carried out two final checks on our fragment scheme shown in Table 3. First we

317   calculate *E, S* and *B* from the fragmentation scheme for the terpenes for which experimental

318   values of log P_{oct} are known, see Table 5. Then knowing *V* from simple calculation of atom

319   fragments [9] we have all the fragment data necessary to calculate log P_{oct} using Eq. (8).

320   Details are in Table 8, where the last column gives the average experimental values of log

321   P_{oct}. Comparison of the calculated and average experimental values gives an average error,

322   AE, (calculated − experimental) of -0.08, an average absolute error, AAE, of 0.12 and a

323   standard deviation, SD, of 0.17 log units. The latter compares favorably with the error in the

324   experimental values of log P_{oct}.

325   Second, we calculate the values of *E, S, B* and *L* from the fragmentation scheme for the

326   terpenes and use these to calculate the first ten values in Table 6 (these refer to the equations

327   in *L*). For the ten sets of data the preferred descriptors from Table 4 yield AE = − 0.12, AAE

328   = 0.19 and SD = 0.22 log units. If we use the descriptors as calculated by the fragmentation

329   scheme we find AE = 0.13, AAE = 0.20 and SD = 0.24 log units, not significantly different to

330   the errors found with the preferred descriptors.

331   We suggest that this indicates that the fragmentation scheme in Table 3 can, indeed, be

332   used reliably to estimate *E, S, B* and *L*. Knowledge of these descriptors, plus *V*, is enough to

333   be able to predict log P_{oct} and a large number of other physicochemical and biological values

334   for aliphatic cyclic terpene hydrocarbons.

335

335 **References**

336

337 [1] C. S. Sell, The Chemistry of Fragrances, 2nd Ed., RSC Publishing, Cambridge, 2006.

338 [2] S. Meakins, The safety and toxicology of fragrances, in C. S. Sell, Ed. The Chemistry of

339   Fragrances, 2nd Ed., RSC Publishing, Cambridge, 2006.

340 [3] J. E. Cometto- Muñiz, W. S. Cain, M. H. Abraham, R. Kumarsingh, Pharmacol.

341   Biochem. Behavior, 60 (1998) 765-770

342 [4] Y. Nagata, Measurement of odor threshold by triangle odor bag method. In Odor

343   measurement review. Office of Odor, Noise and Vibration. Environmental Management

344   Bureau, Ministry of Environment, Tokyo, pp. 118-127, 2003..

345 [5] O. Rodriguez, M. A. Teixeira, A. E. Rodrigues, Flavour and Fragr. J. (2011)

346   Wileyonlinelibrary.com, DOI 10.1002/ffj.2076

347 [6] M. H. Abraham, Chem. Soc. Revs., 22 (1993) 73-83.

348 [7] M. H. Abraham, A. Ibrahim, A. M. Zissimos, J. Chromatogr.A, 1037 (2004)

349   29-47.

350 [8] M. H. Abraham, R. E. Smith, R. Luchtefeld, A. J. Boorem, R. Luo, W. E. Acree,

351   Jr., J. Pharm. Sci., 99 (2010) 1500-1515.

352 [9] M. H. Abraham, J. C. McGowan, Chromatographia, 23 (1987) 243-246.

353 [10] M. H. Abraham, J. Andonian-Haftvan, G. S. Whiting, A. Leo, R. W. Taft,

354   J. Chem. Soc. Perkin Trans. 2, (1994) 1777-1791.

355 [11] M. H. Abraham, H. S. Chadha, G. S. Whiting, R. C. Mitchell, J. Pharm. Sci., 83 (1994)

356   1085-1100.

357 [12] H. Ahmed, C. F. Poole, J. Chromatogr. A., 1104 (2006) 82-90.

358 [13] I. G. Zenkevich and A. S. Kushakova, Russ. J. Gen. Chem., 81 (2011) 337-344.

359 [14] M. H. Abraham, R. Kumarsingh, J. E. Cometto- Muñiz, W. S. Cain, Archives Toxicol.,

360   72 (1998) 227-232.

361 [15] M. H. Abraham, J. M. R. Gola, J. E. Cometto- Muñiz, W. S. Cain, Indoor Built Environ.,

362   10 (2001) 252-257.

363 [16] M. H. Abraham, R. Sánchez-Moreno, J. E. Cometto-Muñiz, W. S. Cain, Chem.

364   Senses, 37 (2012) 207-218.

365 [17] M. H. Abraham, R. Kumarsingh, J. E. Cometto- Muñiz, W. S. Cain, M. Roses,

366   E. Bosch, M. L. Diaz, J. Chem. Soc., Perkin Trans.2, (1998) 2405-2411.

367 [18] T. Karunasekara. C. F. Poole, J. Chromatogr. A., 1235 (2012) 159-165.

368 [19] J. Li, E. M. Perdue, S.G.Paviostathis, R. Araujo, Environ. Int., 1998, 24 (1998),

369   353-358.

370   [20] ChemSpider, the EPI suite and the ACD suite.

371   [21] I. Fichan, C. Larroche, J. B. Gros, J. Chem. Eng. Data, 44 (1999) 56-62.

372   [22] J. L. Simonsen, The Terpenes, vol. I, second Ed., Cambridge University Press, 1947.

373   [23] J. Simonsen and L. N. Owen, The Terpenes, vol. II, second Ed., Cambridge University

374   Press, 1949.

375   [24] J. Simonsen and D. H. R. Barton, The Terpenes, vol. III, Cambridge University Press,

376   1952.

377   [25] www.cas.org/products/scifinder

378   [26] M. H. Abraham, J. C. McGowan, Chromatographia, 23 (1987) 243-246.

379   [27] V. I. Babushok, P. J. Linstrom, I. G. Zenkevich, J. Phys. Chem. Ref. Data, 40 (2011)

380   043101-1 – 043101-46.

381   [28] B. B. Darjazi, A. Rustaiyan, A. Talaei, A. Khalighi, K. Larijani, B. Golein, E.

382   Hayatbakhsh, R. Taghizad, Iran J. Chem. Chem. Eng., 28 (2009) 99-111.

383   [29] M. H. Abraham, W. E. Acree, Jr., A. J. Leo, D. Hoekman, New J. Chem., 33 (2009)

384   568-573.

385   [30] BioLoom, BioByte Corp, 201 W. 4th Street, #204 Claremont, CA 91711-4707,

386   USA.

387   [31] www.vcclab.org/lab/alogps/

388   [32] archemcalc.com/sparc/pKa/

389   [33] S. Griffin, S. G. Wyllie, J. Markham, J. Chromatogr.A, 864 (1999) 221-228.

390   [34] M. H. Abraham, R. Sánchez-Moreno, J. Gil-Lostes, J. E. Cometto-Muñiz, W. S.Cain.

391   Physicochemical Modeling of Sensory Irritation in Humans and Experimental Animals,

392   Chapter 25, pp 376-389, in 'Toxicology of the Nose and Upper Airways, Ed. J. B.

393   Morris, D. J. Shusterman. Informa Healthcare USA, New York, 2010.

394   [35] A. Falk, E. Gullstrand, A. Lof, E. Wigaeus-Hjelm, Brit. J. Ind. Med. 47 (1990) 62-64.

395   [36] M. H. Abraham, A. Ibrahim, W. E. Acree, Jr., Chem. Res. Toxicol. 18 (2005) 904-911.

396   [37] M. H. Abraham, A. Ibrahim, J. Chem. Inf. Model. 46 (2006) 1735-1741.

397   [38] M. H. Abraham, W. E. Acree, Jr., Phys. Chem. Chem. Phys., 14 (2012) 7433-7440.

398

399

399 **Table 1**

400 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

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

402

402 **Table 2**

403 Descriptors for some terpene hydrocarbons

Terpene	E	S	B	V	L	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

404

405

405 **Table 3**

406 Fragment schemes for the estimation of *E*, *S, B* and *L* for aliphatic cyclic terpene

407 hydrocarbons.

Fragment	*E*	*S*	*B*	*L*
Constant term	-0.0108	0.2194	-0.1183	-2.6575
Number of internal double bonds, C-C=C-C	0.1619	0.0575	0.1317	2.4783
Number of external double bonds, $C=CH_2$	0.1704	0.0715	0.1136	2.4276
Number of -C=C-C=C- groups	-0.0091	-0.0449	-0.0368	-0.0668
Number of –C=C-C=C-C=C- groups	0.2089	0.0771	-0.0619	-2.0175
Total number of rings	0.2702	0.0001	0.1112	2.4445
Number of 3-membered rings	-0.1409	-0.0159	-0.0060	-0.1166
Number of 4-membered rings	-0.1696	-0.0547	-0.0101	-0.3862
Number of 5-membered rings	-0.1136	-0.0458	-0.0474	-0.1566
Number of 6-membered rings	-0.0486	-0.0430	-0.0080	-0.0158
Number of 7-membered rings	-0.0133	-0.0009	-0.0024	-0.0664
Number of 10-membered rings	0.0183	0.0393	-0.0818	-0.1531
Number of 11-membered rings	0.0228	-0.1219	-0.0680	-0.4200
Number of compounds	88	88	88	80
SD	0.046	0.005	0.024	0.1653

408

409

409 **Table 4**

410 Descriptors for terpene hydrocarbons

Terpene	MF	n(GLC)	*E*	*S*	*B*	*V*	*L*	log *Kw*
Santene	C9H14	5	0.459	0.19	0.17	1.1165	3.898	-0.39
1,3,8-p-Menthatriene	C10H14	3	0.914	0.44	0.30	1.2800	5.138	0.99
Thuja-2,4(10)-diene	C10H14	4	0.638	0.24	0.24	1.2144	4.448	0.09
Tricyclene	C10H16	15	0.500	0.11	0.17	1.1918	4.280	-0.64
α-Pinene	C10H16	65	0.438	0.20	0.14	1.2574	4.256	-0.64
β-Pinene	C10H16	57	0.515	0.19	0.15	1.2574	4.515	-0.57
Limonene	C10H16	84	0.501	0.31	0.23	1.3230	4.688	0.05
α-Terpinene	C10H16	40	0.526	0.25	0.23	1.3230	4.699	-0.09
β-Terpinene	C10H16	48	0.522	0.29	0.22	1.3230	4.869	-0.04
γ-Terpinene	C10H16	53	0.522	0.29	0.22	1.3230	4.840	0.02
Camphene	C10H16	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
y-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
y-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

411

412

412 **Table 5**

413 Comparison of log P_{oct} calculated from the descriptors in Table 4 (taken) and log P_{oct}

414 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-Guaeine, 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	
y-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-Muurolene	6.35	6.30	6.44	6.19	5.00	7.25	
y-Muurolene	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-Zingberene	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	

415

416

416 **Table 6**

417 Comparison of predicted and experimental values of log (1/NPT) and of log K for some gas-

418 solvent partitions at 37°C

Terpene	Pred [a]	Pred [b]	Obs
log (1/NPT)			
Δ-3-Carene	-3.49	-3.50	-3.21
α-Terpinene	-3.35	-3.34	-3.30
log K(gas-blood)			
a-Pinene	1.31	1.52	1.18 ± 0.11
β-Pinene	1.45	1.47	1.36 ± 0.11
Δ-3-Carene	1.51	1.68	1.51 ± 0.09
Limonene	1.85	1.86	1.62 ± 0.08
log K(gas-oil)			
a-Pinene	3.62	3.64	3.46 ± 0.07
β-Pinene	3.82	3.60	3.63 ± 0.07
Δ-3-Carene	3.99	3.90	3.70 ± 0.10
Limonene	4.08	4.05	3.76 ± 0.11
log K(gas-water)			
a-Pinene	-0.87	-0.55	-0.92 ± 1.30
β-Pinene	-0.81	-0.59	-0.92 ± 0.98
Δ-3-Carene	-0.79	-0.42	-0.39 ± 0.61
Limonene	-0.22	-0.20	0.20 ± 0.36

419 [a] Using the descriptors in Table 4. [b] Using descriptors calculated by the fragmentation

420 scheme, Table 3, see later.

421

422

422 **Table 7**

423 Estimation of descriptors through coefficients in Table 3, and derived estimated values of log

424 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

425

426

426 **Table 8**

427 Calculation of log P_{oct} through the fragmentation scheme in Table 3 [a]

Terpene	E	S	B	V	Calc	Exp [b]	Av exp
α-Pinene	0.473	0.18	0.22	1.2547	4.20	4.44, 4.48, 4.83	4.58
Limonene	0.543	0.31	0.23	1.3230	4.32	4.38, 4.50, 4.57	4.48
α-Terpinene	0.526	0.29	0.21	1.3230	4.39	4.25	4.25
β-Terpinene	0.543	0.31	0.23	1.3230	4.32	4.35, 4.50	4.43
γ-Terpinene	0.535	0.29	0.25	1.3230	4.27	4.35, 4.50	4.43
Camphene	0.538	0.20	0.16	1.2574	4.41	4.22, 4.37	4.30
2-Carene	0.502	0.22	0.22	1.2574	4.17	4.44	4.44
D-3-Carene	0.502	0.22	0.22	1.2574	4.17	4.38	4.38
β-Phellandrene	0.534	0.30	0.19	1.3230	4.45	4.16	4.16
Terpinolene	0.535	0.29	0.25	1.3230	4.27	4.24, 4.47	4.35

428 [a] Values of E, S and B obtained from the fragmentation scheme in Table 3. V calculated from

429 atom and bond values [9]. [b] From Table 5.

430

431