Journal of Chromatography A, 662 (1994) 269–280
Elsevier Science B.V., Amsterdam

CHROM. 25 688

Estimation of gas–liquid chromatographic retention times from molecular structure

S.H. Hilal and L.A. Carreira*

Department of Chemistry, University of Georgia, Athens, GA 30602 (USA)

S.W. Karickhoff

Environmental Research Laboratory, U.S. Environmental Protection Agency, Athens, GA 30605 (USA)

C.M. Melton

Artificial Intelligence Group, University of Georgia, Athens, GA 30602 (USA)

(First received July 5th, 1993; revised manuscript received October 26th, 1993)

ABSTRACT

A new type of a computer program called SPARC (SPARC Performs Automated Reasoning in Chemistry) was developed to predict chemical reactivity parameters and physical properties of organic molecules from their molecular structures based on fundmental chemical structure theory. SPARC's physical models for vapor pressure and activity coefficient were used to calculate the Henry's constant, which can be related to the Kováts retention index. The Kováts indices for a wide range of compounds at any temperature on a squalane liquid phase were calculated. The Root Mean Square deviation error was found to be less than 7 Kováts units, a value that is close to interlaboratory experimental error.

INTRODUCTION

Despite some limitations, the Kováts index has found much greater usage than all other specialized retention specification schemes. The Kováts index is the only retention value in gas–liquid chromatography (GLC) in which two fundamental quantities, the relative retention and the specific retention volume are united [1]. Moreover, a series of explicit relationships between retention indices and a number of physicochemical quantities related to GLC have been developed. Also many different linear relationships between the Kováts index value for a molecule and other fundamental quantities such as carbon number, boiling point and refractive index have been derived [1,2].

The Kováts index [3] expresses the retention of a compound of interest relative to a homologous series of *n*-alkanes examined under the same isothermal conditions. The Kováts index for a particular compound of interest is defined as the carbon number (c) multiplied by 100 of a hypothetical *n*-alkane having exactly the same net retention volume characteristics of the compound of interest measured under the same conditions:

$$I = 100\left(\frac{\log V_{\mathrm{Nc}} - \log V_{\mathrm{Nx}}}{\log V_{\mathrm{Nc}} - \log V_{\mathrm{N(c+1)}}} + \mathrm{c}\right) \tag{1}$$

where I is the Kováts Index of compound x, x is a compound with a retention between that of the first *n*-alkane and second *n*-alkane standard, c is

---

* Corresponding author.

0021-9673/94/$07.00
SSDI 0021-9673(93)E1160-2

the number of carbon atoms in the first *n*-alkane standard, c + 1 is the number of carbon atoms in the second *n*-alkane standard, V_{Nx} is the net retention volume of compound *x*, V_{Nc} is the net retention volume of the first *n*-alkane standard, and $V_{N(c+1)}$ is the net retention of the second *n*-alkane standard.

Numerous investigators have attempted to calculate or predict *I* using physicochemical descriptors like boiling point, density, dipole moment, etc. Unfortunately, all of the correlations of retention indices and the various physicochemical properties are either relatively limited in scope or their application is restricted to a particular chemical class. Other attempts to predict retention indices for a wide range of molecular structures using molecular bond length, molecular bond angle, topological indices [1,2,4], or other molecular characteristic have not been successful. Most of these studies also were restricted to a particular class of molecules on a specific stationary liquid phase.

Despite all the attempts to predict Kováts indices, no realistic scheme with widespread application for different classes of compounds or for different polarities of stationary liquid phase is available.

Our research goal is to develop mathematical models to calculate the Kováts index at any temperature based on a calculated Henry's constant for a wide range of different classes of compounds on different polar and non-polar stationary liquid phases. In the present study, we report the calculation of Kováts indices of organic compounds on a squalane liquid phase strictly from molecular structure by a new computer program called SPARC.

SPARC

SPARC (SPARC Performs Automated Reasoning in Chemistry) is a prototype computer program being developed by the U.S. Environmental Protection Agency and the University of Georgia for the estimation of chemical reactivities and physical properties from molecular structure for a broad range of compounds using computational algorithms based on fundamental chemical structure theory. This new computer program will cost the user only few minutes of computer time, to produce data that are more accurate and have a broader scope than can be obtained with conventional estimation techniques.

SPARC computational approach

The computational approach in SPARC is based on fundamental chemical structure theory to estimate a variety of reactivity parameters [5] (*e.g.*, ionization pK_a, rate constants, etc.) and physical properties [6] (*e.g.*, vapor pressure, distribution coefficient, heat of vaporization, etc.).

The approach involves primarily deductive reasoning and is theory/mechanism oriented. The SPARC program couples Perturbed Molecular Orbital theory [7], which estimates charge distribution and polarizabilities of π electrons with Linear Free Energy relationships [8] to predict the molecular properties of an almost unlimited range of molecular species.

SPARC presently predicts, for a large number of nonpolymeric organic molecules, ionization pK_a [5] and numerous physical properties [6] such as distribution coefficients between immiscible solvents, solubilities, vapor pressure, etc. The ultimate goal for SPARC is to model the chemical and physical behavior of molecules to predict chemical reactivity parameters and physical properties for the universe of organic and inorganic molecules strictly from molecular structure.

SPARC physical models

For all physical processes (*e.g.*, vapor pressure, activity coefficient, partition coefficient, etc.), SPARC uses one master equation to calculate characteristic process parameters:

$$\Delta G_{\text{process}} = \Delta G_{\text{interaction}} + \Delta G_{\text{monomer}} \quad (2)$$

where $\Delta G_{\text{monomer}}$ describes entropy changes associated with mixing, volume changes, or changes in internal (vibrational, rotational) energies going from the initial state to the final state. $\Delta G_{\text{monomer}}$ depends only on the phase change involved and in the present application is presumed to depend only on solute/solvent volumes in each phase. $\Delta G_{\text{interaction}}$ describes the change in the intermolecular interactions in the initial state and final state. For example, for Henry's constant the interaction term describes the difference in the intermolecular interactions in the gas phase *vs.* those in the liquid phase. The interactions in the liquid phase are modeled explicitly, interactions in the gas phase are ignored, and molecular interactions in the crystalline phase are extrapolated from the subcooled liquid state using the melting point.

The intermolecular interactions in the liquid phase are expressed as a summation over all the intramolecular interaction forces between the molecules:

$$\Delta G_{\text{interaction}} = \Delta G_{\text{dispersion}} + \Delta G_{\text{induction}} + \Delta G_{\text{dipole}} + \Delta G_{\text{H-bonding}} \qquad (3)$$

Each of these interactions is expressed in terms of a limited set of molecular-level descriptors (density-based volume, molecular polarizability, molecular dipole, and H-bonding parameters) which in turn are calculated from molecular structure.

SPARC molecular descriptors

The computational approach for molecular-level descriptors is constitutive with the molecule in question being broken at each essential single bond and the property of interest being expressed as a linear combination of fragment contributions as

$$\chi^0(\text{molecule}) = \sum_i (\chi_i^0 - A_i) \qquad (4)$$

where χ_i^0 are intrinsic fragment contributions (which in most cases are tabulated in SPARC databases) and A_i are adjustments relating to steric or electrometric perturbations from contiguous structural elements for the molecule in question and process model or medium involved. Both χ_i^0 and A_i are empirically trained either on direct measurements of the descriptor in question (*e.g.*, liquid density based molecular volume) or on a directly related property (*e.g.*, index of refraction, which can be related to polarizability) for which large reliable data sets exit.

Average molecular polarizability

In SPARC, fragment polarizability factored into atomic contributions, χ_j, and the polarizability of fragment, i, is expressed as

$$\overline{\alpha_i} = \frac{1}{N_i} \cdot \left[\sum_j \chi_j\right]^2 \qquad (5)$$

where the summation is over all the atoms in fragment i, χ_j is the intrinsic atomic hybrid polarizability contribution, and N_i is the number of electrons in fragment i. The χ_j are empirically determined from measured polarizabilities and stored in the SPARC database (with exception of hydrogen which is calculated from the measured polarizability of H_2).

The average molecular polarizability, α^0, is expressed as

$$\overline{\alpha^0} = 4 \sum_i (\overline{\alpha_i} - A_i) \qquad (6)$$

where α_i is the polarizability of fragment i and A_i are the adjustments for the molecule in question. The only adjustment, A_i, currently implemented in SPARC is a 10% reduction in α_i for hydrocarbon fragments with an attached polar group or atom. The partition of polarizability into atomic contributions enables estimates to be made of molecular polarizabilities for any given molecular structure. The molecular polarizability can be calculated within less than 1% for a wide range of molecules. Fig. 1 shows the observed *vs.* SPARC-calculated refractive index at 25°C for alkane, alkene and aromatic

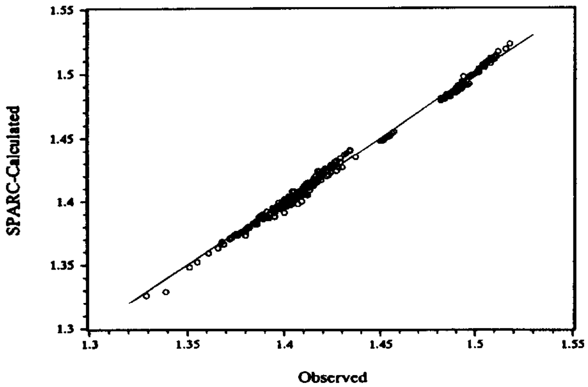

Fig. 1. Observed *vs.* SPARC-calculated refractive indices at 25°C.

systems. The root mean square (RMS) deviation for this set was found to be equal to $3 \cdot 10^{-3}$. Examples of the calculation of molecular polarizability and index of refraction are given in the appendix.

Molecular volume

The zero order density-based molecular volume is expressed as

$$V^0_{25} = \sum_i (V_i^{\text{frag}} - A_i) \tag{7}$$

where V^{frag} is the volume of the fragment and A_i is a correction to that volume based on both the number and size of fragments attached to it. The V^{frag} are determined empirically from the measured volume and then stored in the SPARC database. This zero order volume at 25°C is further adjusted for shrinkage resulting from dipole–dipole and H-bonding interactions:

$$V_{25} = V^0_{25} + A_{\text{d}} \frac{\sum_i D_i^2}{V^0_{25}} + A_{\text{HB}} \frac{\sum_i \alpha_i \sum_i \beta_i}{V^0_{25}} \tag{8}$$

where D_i is the dipole for the molecule, and α and β are the H-bonding parameters of potential proton donor and proton acceptor sites within the molecule, respectively.

The volume at temperature T is then expressed as a polynomial expansion in $(T-25)$ corrected as a function of H-bonding (HB), dipole (D) and polarizability (P) interactions as

$$V_T = V_{25}\left[1 + f(P, D, HB) \sum_n a_n (T-25)^n\right] \tag{9}$$

where a_n are variable parameters. The molecular volumes for a wide range of molecules can be calculated to better than 1%. Fig. 2 shows the observed versus SPARC-calculated density based molecular volumes for alkane, alkene, and aromatic systems. The RMS deviation was found to less than $4 \cdot 10^{-3}$. Examples of the calculation of molecular volume are given in the Appendix.

Solute–solvent interactions

Models for self interactions between like molecules and between solvent–solute molecules have been developed to calculate physical properties.

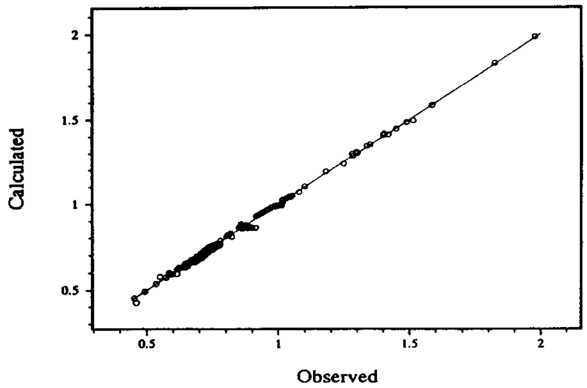

Fig. 2. Observed *vs.* SPARC-calculated densities at 25°C.

These interaction models build on the limited set of molecular-level descriptors (volume, polarizability, molecular dipole and H-bonding parameters) described above. These interaction models are dispersion, induction, dipole–dipole, and H-bonding. Dispersion interactions are present for all the molecules including non-polar molecules. Induction interactions are present between two molecules when at least one of them has a permanent dipole moment. Dipole–dipole interactions exist when both molecules have dipole moments. H-bonding interactions exist when $\alpha_i \cdot \beta_j$ or $\alpha_j \cdot \beta_i$ products are non-zero.

Dispersion interactions

In this paper we present our calculation of the Kováts retention indices for alkanes, alkenes and aromatics on a non-polar stationary phase squalane. For this reason, we shall not discuss the interaction mechanisms for dipoles and H-bonding. As such, this is a definitive test of our dispersion modeling.

Dispersion interactions occur between all molecules as a result of very rapidly varying dipoles formed between nuclei and electrons at zero-point motion of the molecules, acting upon the polarizability of other molecules to produce an induced dipole in the phase. The self interactions are expressed as

$$\Delta G_{ii}(\text{disp}) = \rho_{\text{disp}} (P_i^{\text{d}})^2 V_i \tag{10}$$

whereas solvent–solute interactions are expressed as

$$\Delta G_{ij}(\text{disp}) = \rho_{\text{disp}}\left(P_i^{\text{d}} - P_j^{\text{d}}\right)^2 V_i \qquad (11)$$

$$P_i^{\text{d}}(\text{disp}) = \frac{\alpha_i + A_{\text{disp}}}{V_i} \qquad (12)$$

where i and j designate the solute and squalane molecules respectively; P_i^{d} is the effective polarizability density of molecule i; ρ_{disp} is the susceptibility to dispersion; V_i and α_i are the molar volume and the average molecular polarizability described previously, respectively. $A_{(\text{disp})}$ is the polarizability adjustment for dispersion. $A_{(\text{disp})}$ differentiates the bulk polarizability α_i and the effective or microscopic polarizability that the molecule experiences at a point.

Dispersion is a short range interaction involving surface or near surface atoms and A_{disp} subtracts from the total polarizability, a portion of the contributions of sterically occluded atoms in the molecular lattice. Presently SPARC corrects for access judged to be less than afforded by a linear array of atoms (*i.e.*, for branched structures or rings small enough to prohibit intra penetration of the solvent).

Branched (ternary or quaternary) atoms in an alkane structure will lose a small part of their intrinsic molecular polarizability depending on the size and number of appended groups, and the proximity of other branched carbons. Similarly, carbons in rings may lose their intrinsic polarizability contributions depending on ring sizes and the presence of a ring appendage. Examples of the calculation of effective polarizability are given in the appendix.

Activity coefficient model

For a solute, i, in a liquid phase, j, at infinite dilution, SPARC expresses the activity coefficient as

$$-RT\ \log \gamma_{ij} = \Delta G_{\text{interactions}} + \Delta G_{\text{monomer}} \qquad (13)$$

For the hydrocarbons in this study, the activity coefficient is given as

$$-RT\ \log \gamma_{ij}^{\infty} = \Delta G_{ij_{\text{disp}}} + RT\left(\log\frac{V_i}{V_j} + \frac{\left(1-\frac{V_i}{V_j}\right)}{2.303}\right) \qquad (14)$$

where the last term is the Flory–Huggins [9,10] excess entropy contributions of mixing in the liquid phase [beyond that of $\Delta G_{ij}(\text{disp})$] of placing a solute molecule in the solvent. When the solute and solvent have the same volume, the Flory–Huggins term will go to zero. Table I shows the observed *vs.* SPARC-calculated activity coefficients in squalane. It should be noted that the negative log values are a consequence of the large Flory–Huggins contributions. See the Appendix for sample calculations.

Vapor pressure model

The vapor pressure P_i of a solute, i, is expressed as

$$-2.303\ RT\ \log P_i = \Delta G_{i_{\text{disp}}} - 2.303\ RT(\log T + C) \qquad (15)$$

where $RT(\log(T) + C)$ describes the change in the entropic contributions [14] associated with the volume changes between the liquid and the gas phases. Fig. 3 shows the observed *vs.* the SPARC-calculated values for the vapor pressure

TABLE I

OBSERVED *vs.* SPARC-CALCULATED VALUES FOR THE log ACTIVITY COEFFICIENT IN SQUALANE

Molecule	Activity coefficient	
	Obs.	Calc.
Pentane	−0.24	−0.20
Hexane	−0.19	−0.18
Heptane	−0.15	−0.16
Octane	−0.15	−0.15
Nonane	−0.17	−0.14
2-Methylpentane	−0.19	−0.16
2,4-Dimethylpentane	−0.14	−0.13
2,5-Dimethylhexane	−0.11	−0.12
2,3,4-Trimethylpentane	−0.17	−0.14
Cyclohexane	−0.28	−0.30
Ethylcyclohexane	−0.23	−0.22
Benzene	–	−0.15
Toluene	–	−0.16
1,3-Dimethylbenzene	–	−0.14
1,4-Dimethylbenzene	–	−0.14
1,3,5-Trimethylbenzene	–	−0.11

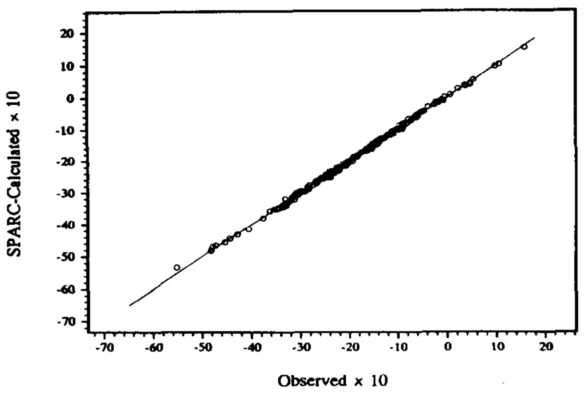

Fig. 3. Observed *vs.* calculated values for the vapor pressures at 25°C.

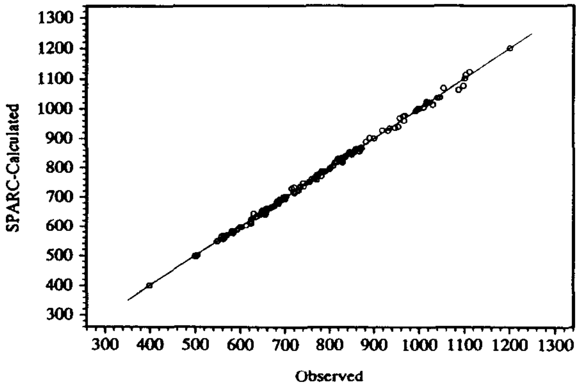

Fig. 4. Observed *vs.* calculated values for the retention indices at 25°C.

for various molecular structures at 25°C. The RMS deviation error for $\log P$ was 0.038.

See the Appendix for sample vapor pressure calculations.

Henry's constant

Henry's constant for a dilute solute, *i*, in a solvent, *j*, may be expressed as

$$H_x = P_i^0 \gamma_{ij}^{\infty} \tag{16}$$

where P_i^0 is the vapor pressure of pure solute *i* and γ_{ij}^{∞} is the activity coefficient of solute *i* in the squalane liquid phase at infinite dilution. SPARC vapor pressure and activity coefficient models are used to calculate Henry's constant for a solute in a squalane liquid phase.

Henry's constant can be related to the net retention volume, V_N, by

$$H_i = \frac{RT}{M} \cdot \frac{V_L}{V_N} \tag{17}$$

where *M* is the molecular weight of the solvent, and V_L is the volume of the stationary phase. Substituting in eqn. 1, we get

$$I = 100 \cdot \left(\frac{\log H_{Nx} - \log H_{Nz}}{\log H_{HN(z+1)} - \log H_{Nz}} + c \right) \tag{18}$$

where H_{Nx}, H_{Nz}, and $H_{N(z+1)}$ are Henry's constant for a compound *x*, first *n*-alkane standard, and second *n*-alkane standard, respectively.

Kováts indices at 25°C

Fig. 4 and Table II show the observed [11,12] versus the SPARC-calculated Kováts indices at 25°C. The RMS deviation was less than 7 Kováts units, a value that is close to the interlaboratory experimental error. We also calculated the Kováts index as a function of temperature using SPARC temperature dependence models discussed below.

SPARC calculates a physical property of interest at 25°C. In addition to the inherent temperature dependence described previously in eqns. 14 and 15 and the temperature dependence built in the volume calculator (eqn. 9), the susceptibility of dispersion at temperature *T* is modeled as a function of the polarizability density and the effective polarizability density. In effect, this describes the small temperature dependence of enthalpy. For temperature, an "activity-driven" process $\rho(t)$ is given by

$$\rho_T = \left[1 + \left(1 - \sum_n^5 a_n \left(\frac{298.15}{T} \right)^n \right) \cdot f(P, P') \right] \rho_{25} \tag{19}$$

where a_n are variable parameters. These parameters were inferred from boiling point measurements at 1, 10, 100, and 760 Torr (1 Torr = 133.322 Pa) for more than 400 compounds spanning a range of over 700°C. Table III and Fig. 5 show the observed *vs.* SPARC-calculated boiling point for wide range of molecules. The RMS deviation for this set was 3.4°C.

Based on these temperature dependent models we calculated the Kováts indices at 80°C as

TABLE II

OBSERVED *vs.* SPARC-CALCULATED RETENTION INDICES AT 25°C ON SQUALANE LIQUID PHASE

No.	Compound	Observed	Calculated	Difference
Kováts Bases				
1	*n*-Dodecane	1200.0	1200.0	0.0
2	*n*-Undecane	1100.0	1100.0	0.0
3	*n*-Decane	1000.0	1000.0	0.0
4	*n*-Nonane	900.0	900.0	0.0
5	*n*-Octane	800.0	800.0	0.0
6	*n*-Heptane	700.0	700.0	0.0
7	*n*-Hexane	600.0	600.0	0.0
8	*n*-Pentane	500.0	500.0	0.0
9	*n*-Butane	400.0	400.0	0.0
10	*n*-Propane	300.0	300.0	0.0
11	*n*-Ethane	200.0	200.0	0.0
12	*n*-Methane	100.0	100.0	0.0
Alkanes				
13	2,2-Dimethylbutane	534.8	525.1	9.7
14	2,3-Dimethylbutane	565.0	561.5	3.5
15	2-Methylpentane	569.4	569.0	0.4
16	3-Methylpentane	583.0	575.6	7.4
17	2,2,3-Trimethylbutane	636.2	634.8	1.4
18	3,3-Dimethylpentane	655.6	641.6	14.0
19	2-Methylhexane	666.2	663.0	3.2
20	2,3-Dimethylpentane	669.9	665.9	4.0
21	3-Methylhexane	675.2	670.2	5.0
22	3-Ethylpentane	684.7	678.4	6.2
23	2,2,4-Trimethylpentane	687.3	693.7	−6.4
24	2,2-Dimethylhexane	718.3	719.2	−0.9
25	2,5-Dimethylhexane	727.6	725.9	1.7
26	2,2,3-Trimethylpentane	733.2	738.4	−5.3
27	2,4-Dimethylhexane	730.8	732.9	−2.0
28	3,3-Dimethylhexane	740.0	737.0	3.0
29	2,3,4-Trimethylpentane	748.6	749.1	−0.5
30	2,3,3-Trimethypentane	753.9	750.8	3.0
31	2-Methyl-3-ethylpentane	757.9	760.1	−2.2
32	2,3-Dimethylhexane	758.5	761.4	−2.9
33	3-Methyl-3-ethylpentane	769.2	759.0	10.3
34	2-Methylheptane	764.6	762.5	2.1
35	3,4-Dimethylhexane	768.0	768.5	−0.4
36	4-Methylheptane	766.5	768.6	−2.1
37	2,2,4,4-Tetramethylpentane	769.7	775.7	−6.0
38	3-Ethylhexane	771.3	772.7	−1.4
39	3-Methylheptane	771.5	772.9	−1.4
40	2,2,5-Trimethylhexane	774.8	778.3	−3.5
41	2,2,4-Trimethylhexane	785.8	788.7	−2.9
42	2,3,5-Trimethylhexane	810.2	817.2	−7.0
43	2,2,3,4-Tetramethylpentane	814.1	821.1	−6.9
44	2,2,-Dimethyl-3-ethylpentane	817.1	830.5	−13.4
45	2,2-Dimethylheptane	814.7	817.5	−2.8
46	2,2,3-Trimethylhexane	817.7	828.7	−11.0

(*Continued on p. 276*)

TABLE II (*continued*)

No.	Compound	Observed	Calculated	Difference
47	2,4-Dimethylheptane	821.0	823.3	−2.3
48	4,4-Dimethylheptane	824.8	832.1	−7.3
49	2-Methyl-4-ethylhexane	823.4	824.2	−0.8
50	2,6-Dimethylheptane	826.6	826.3	0.3
51	2,4-Dimethyl-3-ethylpentane	832.1	833.7	−1.7
52	2,3,3-Trimethylhexane	835.2	840.2	−5.0
53	3,5-Dimethylheptane	832.5	835.8	−3.4
54	3,3-Dimethylheptane	833.7	836.5	−2.8
55	2,3,4-Trimethylhexane	842.6	851.9	−9.3
56	3,3,4-Trimethylhexane	847.8	843.3	4.4
57	3-Methyl-3-ethylhexane	849.4	848.8	0.6
58	2,3,3,4-Tetramethylpentane	852.1	850.6	1.5
59	3-Methyl-4-ethylhexane	851.1	855.6	−4.4
60	2,3-Dimethylheptane	853.6	858.7	−5.1
61	3,4-Dimethylheptane	855.7	861.9	−6.1
62	4-Ethylheptane	856.4	864.7	−8.3
63	2,3-Dimethyl-3-ethylpentane	865.7	856.1	9.6
64	4-Methyloctane	862.3	867.9	−5.6
65	2-Methyloctane	864.2	856.1	8.1
66	3,3-Diethylpentane	870.9	867.9	2.9
67	4-Ethylheptane	866.0	863.6	2.4
68	3-Methyloctane	869.7	870.0	−0.3
Alkenes				
69	*trans*-2-Pentene	501.0	496.3	4.7
70	4-Methyl-*trans*-2-pentene	561.9	555.4	6.5
71	*trans*-3-Hexene	593.4	591.2	2.2
72	3-Methyl-*trans*-2-pentene	612.6	603.9	8.7
73	2-Methyl-*trans*-3-hexene	648.0	655.4	−7.4
74	4-Methyl-*trans*-2-hexene	655.5	656.7	−1.2
75	3,3-Dimethyl-1-butene	505.2	501.4	3.8
76	4-Methyl-1-pentene	548.0	547.6	0.4
77	3-Methyl-1-pentene	549.4	549.9	−0.6
78	2,3-Dimethyl-1-butene	557.4	565.3	−7.9
79	2-Methyl-1-pentene	579.6	580.0	−0.4
80	1-Hexene	581.6	585.0	−3.3
81	2-Ethyl-1-pentene	592.0	586.7	5.4
82	4,4-Dimethyl-1-pentene	602.7	597.5	5.2
83	2-Methyl-2-pentene	598.0	598.5	−0.6
84	3,3-Dimethyl-1-pentene	623.5	617.7	5.8
85	2,3,3-Trimethyl-1-butene	625.6	627.4	−1.8
86	2,3-Dimethyl-2-butene	624.4	611.0	13.4
87	3,4-Dimethyl-1-pentene	634.6	632.6	2.0
88	3-Methyl-1-hexene	643.5	641.5	2.0
89	3-Ethyl-1-pentene	645.0	642.7	2.3
90	2,3-Dimethyl-1-pentene	648.1	652.2	−4.1
91	5-Methyl-1-hexene	648.8	647.2	1.6
92	4-Methyl-1-hexene	656.5	649.2	7.2
93	2-Methyl-1-hexene	677.5	674.7	2.8
94	1-Heptene	681.3	686.0	−4.7
95	2-Ethyl-1-pentene	681.7	679.2	2.6
96	2-Methyl-2-hexene	692.0	691.5	0.6

TABLE II (*continued*)

No.	Compound	Observed	Calculated	Difference
97	3-Ethyl-2-pentene	695.9	701.1	−5.1
98	2,3-Dimethyl-2-pentene	701.9	702.7	−0.7
99	2,4,4-Trimethyl-2-pentene	713.9	729.7	−15.8
100	1-Octene	780.3	788.3	−8.0
101	1-Nonene	881.2	888.9	−7.8
102	3-*trans*-Heptene	687.0	688.9	−1.9
103	2-*trans*-Heptene	698.7	693.5	5.2
104	4-*trans*-Octene	783.0	786.1	−3.1
105	3-*trans*-Octene	789.0	787.8	1.2
106	2-*trans*-Octene	798.6	794.8	3.8
Aromatics				
107	Benzene	629.9	645.7	−15.8
108	Tolune	739.1	748.7	−9.6
109	Ethylbenzene	828.2	835.7	−7.5
110	*p*-Xylene	842.5	854.6	−12.2
111	*m*-Xylene	845.7	846.9	−1.2
112	*o*-Xylene	862.7	856.4	6.3
113	Isopropylbenzene	889.5	902.5	−12.9
114	*n*-Propylbenzene	917.8	927.2	−9.5
115	1-Methyl-3-ethylbenzene	930.2	925.4	4.7
116	1-Methyl-4-ethylbenzene	934.4	934.5	−0.1
117	1-Methyl-2-ethylbenzene	946.1	935.7	10.4
118	1,3,5-Trimethylbenzene	953.1	939.1	14.1
119	*tert*-Butylbenzene	956.5	968.3	−11.8
120	*iso*-Butylbenzene	963.7	974.6	−10.9
121	1,2,4-Trimethylbenzene	965.9	959.6	6.3
122	*sec*-Butylbenzene	967.4	977.3	−9.9
123	1-Methyl-2-isopropylbenzene	994.6	996.2	−1.6
124	1-Methyl-3-isopropylbenzene	991.4	990.9	0.5
125	1-Methyl-4-isopropylbenzene	996.8	1001.6	−4.8
126	1,3-Dimethylbenzene	1009.4	1004.4	5.0
127	1-Methyl-4-*n*-propylbenzene	1015.7	1024.4	−8.7
128	1-Methyl-3-*n*-propylbenzene	1014.7	1015.6	−0.9
129	1,4-Diethylbenzene	1017.8	1020.3	−2.5
130	*n*-Butylbenzene	1018.5	1022.6	−4.1
131	1-Methyl-2-*n*-propylbenzene	1022.8	1023.0	−0.2
132	1,3-Dimethyl-5-ethylbenzene	1029.7	1013.4	16.3
133	1,4-Dimethyl-2-ethylbenzene	1039.3	1035.9	3.5
134	1,2-Dimethyl-4-ethylbenzene	1044.9	1037.6	7.3
135	1,3-Dimethyl-4-ethylbenzene	1044.1	1036.3	7.8
136	1-Methyl-4-*tert*-Butylbenzene	1053.1	1067.7	−14.6
137	1,2,3,5-Tetramethylbenzene	1086.4	1062.2	24.2
138	1,2,3,4-Tetramethylbenzene	1097.0	1074.8	22.2
139	1-Methyl-4-*n*-propylbenzene	1102.1	1111.7	−9.6
140	*n*-Pentylbenzene	1110.6	1121.2	−10.6
Cycloalkanes				
141	Cyclopentane	562.4	568.8	−6.4
142	Methylcyclopentane	624.6	625.7	−1.2

(*Continued on p. 278*)

TABLE II (*continued*)

No.	Compound	Observed	Calculated	Difference
143	1,1-Dimethylcyclopentane	669.3	668.5	0.8
144	1-*trans*-3-Dimethylcyclopentane	682.1	690.0	−7.9
145	1-*trans*-2-Dimethylcyclopentane	685.3	681.0	5.3
146	Cyclohexane	658.3	663.2	−4.9
147	1,1,3-Trimethylcyclopentane	718.9	734.4	−15.5
148	Methylcyclohexane	719.9	714.2	5.7
149	Ethylcyclopentane	729.3	723.2	6.1
150	1,1-Dimethylcyclohexane	779.8	771.3	8.5
151	1-*trans*-4-Dimethylcyclohexane	780.0	771.2	8.8
152	1-Methyl-2-*trans*-ethylcyclopentane	788.4	786.3	2.1
153	Isopropylcyclopentane	806.1	805.3	0.8
154	Ethylcyclohexane	827.5	816.8	10.7
155	*n*-Propylcyclopentane	825.6	818.1	7.5
156	1,1,3-Trimethylcyclohexane	833.8	840.9	−7.1

shown in Fig. 6. The RMS deviation was less than 8 Kováts units.

Models for predicting the Kováts index for polar and non-polar molecules on different polarity liquid phases (*e.g.*, OV-101, SE-30 and PEG-20M) are under development [13].

APPENDIX

Sample calculations

The following calculations will demonstrate the SPARC approach to calculating bulk polarizability, liquid density based volume, refractive

TABLE III

OBSERVED *vs.* SPARC-CALCULATED BOILING POINTS AT DIFFERENT PRESSURES

Molecule	Pressure					
	760 Torr		100 Torr		1 Torr	
	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.
Pentane	36.1	36.6	−12.6	−12.6	−76.6	−76.3
Hexane	68.7	69.0	15.8	15.3	−54.0	−55.3
Heptane	98.4	99.0	41.8	41.8	−33.2	−35.2
Octane	125.7	126.8	65.7	66.8	−14.9	−15.7
Nonane	150.8	152.6	87.9	90.6	3.60	3.00
2-Methylpentane	60.3	55.8	8.10	5.00	−6.00	−6.40
2,4-Dimethylpentane	80.5	76.4	25.4	23.4	−4.70	−5.0
2,5-Dimethylhexane	109.1	106.0	50.5	45.5	−26.7	−31.4
2,3,4-Trimethylhexane	139.0	137.0	76.0	76.0	−7.00	−9.00
Cyclohexane	80.7	81.9	25.5	27.2	−47.0	−44.5
Ethylcyclohexane	131.8	128.7	69.0	67.8	−14.4	−14.7
Benzene	80.1	85.8	26.0	30.2	−45.0	−41.9
Toluene	110.6	113.5	51.9	54.5	−26.1	−23.1
1,3-dimethylbenzene	139.1	138.5	76.8	76.5	−7.2	−5.90
1,4-dimethylbenzene	138.4	140.3	75.9	78.1	−8.1	−5.0
1,3,5-trimethylbenzene	146.7	162.2	99.8	97.7	11.6	10.9

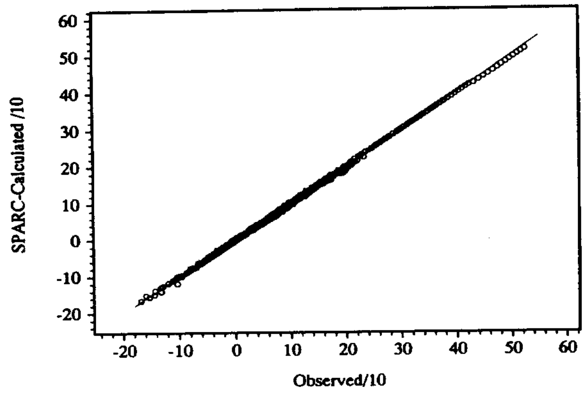

Fig. 5. Observed *vs.* SPARC-calculated values for the boiling points at 1, 10, 100 and 760 Torr.

index, effective molecular polarizability, vapor pressure and activity coefficient. Since the bulk of the molecules in this study are straight-chain and branched alkanes, the two molecules chosen for the sample calculations are *n*-pentane and 2,5-dimethylhexane.

Polarizability

The fragments for polarizability calculations are CH_x units for alkanes. Only two intrinsic atomic polarizabilities are needed, $\chi(\mathrm{sp3{-}C}) = 1.25$ and $\chi(\mathrm{H{-}C}) = 0.314$. From eqn. 5 α_i for the carbon is $(1.25 \cdot 1.25/6) = 0.260$ and α_i for each hydrogen is $(0.314 \cdot 0.314)/1 = 0.0986$. There are no corrections A_i (eqn. 6) for the alkanes since there are no connected polar groups. The bulk molecular polarizability for the alkanes, (C_nH_{2n+2}), from eqn. 6 can be written as $\alpha = 4 \cdot (n \cdot 0.260 + (2n+2) \cdot 0.0986)$ where the units are Å^3/molecule. For *n*-pentane this yields 9.94 and for 2,5-dimethylhexane the value is 15.43.

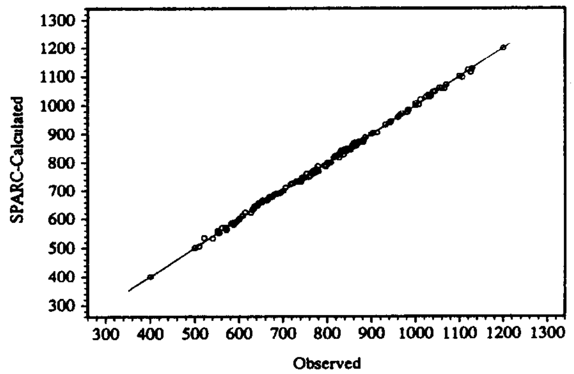

Fig. 6. Observed *vs.* SPARC-calculated values for the retention indices at 80°C.

Volume

For straight-chain and branched alkanes the fragment values V_{frag} in eqn. 7 are all that of $CH_x = 52.945$. The corrections A_i are due to the type, number and size of the substituents. For alkanes the type is always CH_x. The correction for this type is -19.2618. The correction for occluded volume from branching is 0 for branching <2, 3.31 for 2-branching and 8.42 for 3-branching. The size correction is 6.8 times the sum of the cone volumes [5].

n-pentane	V^{frag}	Subtr-type	Branch	Size	Frag-Total
C	52.945	-19.262	0	-6.8(0.113)	32.92
C	52.945	-19.262*2	3.3	-6.8(0.05+0.107)	16.66
C	52.945	-19.262*2	3.3	-6.8(0.091*2)	16.50
C	52.945	-19.262*2	3.3	-6.8(0.05+0.107)	16.66
C	52.945	-19.262	0	-6.8(0.113)	32.92
				Sum	115.6
				Obs	116.1
2,5 DMH					
C	52.945	-19.262	0	-6.8(0.213)	32.23
C	52.945	-19.262*3	8.4	-6.8(0.05*2+0.11)	2.16
C/	52.945	-19.262	0	-6.8(0.213)	32.23
C	52.945	-19.262*2	3.3	-6.8(0.131+0.126)	15.98
C	52.945	-19.262*2	3.3	-6.8(0.131+0.126)	15.98
C	52.945	-19.262*3	8.4	-6.8(0.05*2+0.11)	2.16
C/	52.945	-19.262	0	-6.8(0.213)	32.23
C	52.945	-19.262	0	-6.8(0.213)	32.23
				Sum	165.2
				Obs	166.1

Index of refraction

Index of refraction is a good way to check the polarizability density for the molecule. The polarizability and volume can be related to the index of refraction using the Lorentz–Lorenz equation. For our units of cm^3/mole for volume and Å^3/molecule for polarizability the Lorentz–Lorenz equation can be written as

$$\frac{n^2 - 1}{n^2 + 2} = \frac{4 \cdot \pi \cdot 0.6023\, P}{3\, V} \tag{A1}$$

where n is the index of refraction, P is the molecular polarizability and V is the liquid density based volume. Eqn. A1 leads to a calculated index of refraction of 1.354 (obs 1.358) for *n*-pentane and a calculated index of refraction of 1.388 (obs 1.392) for 2,5-dimethylhexane.

Effective polarizability

Dispersion is a short range interaction involving surface or near-surface atoms and A_{disp} in eqn. 12 subtracts from the total polarizability, a portion of the contributions of sterically occluded atoms in the molecular lattice. Presently SPARC corrects for access judged to be less than afforded by a linear array of atoms. Branched (ternary or quaternary) atoms in an alkane structure will lose a small part of their intrinsic molecular polarizability depending on the size and number of appended groups, and the proximity of other branched carbons. For *n*-pentane there are no corrections so that the effective polarizability equals the calculated molecular polarizability of 9.94. For 2,5-dimethylhexane the polarizability of the 2 and 5 atoms are reduced by 1.72 times the sum of the sizes (see volume above) of the fragments attached to 2 and 5. Each atom is reduced by $1.72 \cdot (0.05 + 0.109 + 0.05) = 0.36$. The effective polarizability for 2,5-dimethylhexane is calculated to be $15.43 - 2 \cdot 1.72 \cdot (0.209) = 14.71$.

Vapor pressure

Once the molecular polarizabilty and volume are known we can use eqn. 10 to calculate the dispersion interactions. ϱ_{disp} is -2.571 (where this number has subsumed in it $-2.303 \cdot RT$) and the polarizability densities for *n*-pentane and 2,5-dimethylhexane are $9.94/115.6 = 0.086$ and $14.71/165.2 = 0.0891$, respectively. The dispersion contribution to the vapor pressure is calculated to be $-2.57 \cdot 0.086 \cdot 0.086 \cdot 115.6 = -2.20$ for *n*-pentane and $-2.57 \cdot 0.0891 \cdot 0.0891 \cdot 165.2 = -3.37$ for 2,5-dimethylhexane. The volume entropy terms $(\log(T) + C)$ are $\log(298) - 0.457 = 2.02$ at room temperature. Log (P) for *n*-pentane is then calculated to be $-2.20 + 2.02 = -0.18$. The observed vapor pressure for *n*-pentane is -0.17. Log(P) for 2,5-dimethylhexane is $-3.37 + 2.02 = -1.35$. The observed vapor pressure is -1.38.

Activity coefficient

In order to calculate the infinite dilution activity coefficient for *n*-pentane and 2,5-dimethylhexane in squalane we need the effective polarizability and volume for squalane. SPARC calculates the molecular polarizability of squalane to be 55.70, the effective polarizability to be 53.22 and the volume to be 530.0. The polarizability density of squalane is then calculated to be $53.22/530 = 0.10$. From eqn. 11 the dispersion contribution to the activity coefficient can be calculated. For *n*-pentane this is $-2.57 \cdot (0.086 - 0.10)^2 \cdot 115.6 = -0.06$ and for 2,5-dimethylhexane the value is $-2.57 \cdot (0.089 - 0.10)^2 \cdot 165.2 = -0.05$. The volume entropy terms from eqn. 15 are 0.26 and 0.17 for pentane and 2,5-dimethylhexane, respectively. The logs of the infinite dilution activity coefficients are $-(-0.06 + 0.26)$ or -0.20 and $-(-0.05 + 0.17)$ or -0.12 for *n*-pentane and 2,5-dimethylhexane, respectively.

REFERENCES

1 G. Tarjan, I. Timar, J.M. Takacs, S.Y. Meszaros, Sz. Nyiredy, M.V. Budahegyl, E.R. Lombosi and T.S. Lombosi, *J. Chromatogr.*, 271 (1982) 213.
2 J.K. Haken and M.B. Evans, *J. Chromatogr.*, 472 (1989) 93.
3 E.sz. Kováts, *Adv. Chrommatogr.*, 1 (1965) 31A.
4 L.S. Anker and P.C. Jurs, *Anal. Chem.*, 62 (1990) 2676.
5 S.W. Karickhoff, V.K. McDaniel, C.M. Melton, A.N. Vellino, D.E. Nute, L.A. Carreira, *Environ. Toxicol. Chem.*, 10 (1991) 1405.
6 S.W. Karickhoff, C.M. Melton, L.A. Carreira and S.H. Hilal, in preparation.
7 L.P. Hammett, *Physical Organic Chemistry*, McGraw Hill, New York, NY, 1970, 2nd ed.
8 M.J.S. Dewar, R.C. Dougherty, *The PMO Theory of Organic Chemistry*, Plenum Press, 1975, New York, NY.
9 P.J. Flory, *J. Chem. Phys.*, 10 (1942) 51.
10 M.L. Huggins, *J. Am. Chem. Soc.*, 64 (1942) 1712.
11 N. Dimov, *J. Chromatogr.*, 347 (1985) 366.
12 D. Papazova and N. Dimov, *J. Chromatogr.*, 356 (1986) 320.
13 S.H. Hilal, L.A. Carreira, S.W. Karickhoff and C.M. Melton, in preparation.
14 K.A. Sharp, A. Nicholls, R. Friedman and B. Honig, *Biochemistry*, 30 (1991) 9686.