Review Commentary

The sizes of molecules—revisited

Yizhak Marcus*

Department of Inorganic and Analytical Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

epoc   ABSTRACT: The size of molecules, be it their volume, surface area or linear extent, is an important quantity in many fields of chemistry. Molecules are often tacitly or explicitly treated as if they were spheres, so that a diameter can be assigned to them. However, the molecules of many kinds of substances are rather cylindrical, being rod- or disk-shaped, so that two linear dimensions are needed for their description. The collision diameters obtainable from experimental data on gases and vapors or diameters obtained from the molar volume or the application of an expression for the interaction potential in liquids, e.g. the Lennard–Jones potential, are explored as ways to yield information on molecular sizes. The ratio of the van der Waals volume and surface area is related to a simple expression derived solely from the composition of the molecules for various types of molecular shapes. These approaches were applied to a database of 350 mainly liquid (but also some gaseous and solid) organic (and a few inorganic) substances. Copyright © 2003 John Wiley & Sons, Ltd.
Additional material for this paper is available from the epoc website at http://www.wiley.com/epoc

KEYWORDS: globular molecules; Lennard–Jones diameter; molecular volume; molecular diameter; oblate molecules; prolate molecules; van der Waals surface; van der Waals volume

INTRODUCTION

Some 17 years ago the late A. Y. Meyer published a review dealing with the sizes of molecules,[1] namely, with their three-dimensional sizes—their volume occupying properties—and also with their one-dimensional extents—their spherical diameters. The sizes of molecules are applicable in many fields of study, such as the Gibbs energy of solvation,[2] adsorption phenomena,[3] absorption in porous media,[4] permeation through polymer membranes,[5] fractal dimensions,[6] dielectric behavior of solvents,[7] evaluation of equations of state,[8] molecular recognition,[9] diffusion rates,[10,11] azeotropic properties of mixtures[12] and guest molecule encapsulation,[13] among many others.[1] They are measures of the maximal proximity of the molecules in either a gaseous or a condensed phase to each other and to solute particles in the latter. Meyer sought reliable estimates of molecular diameters σ that should be situated between a maximal and a minimal estimate.

The maximal estimate, σ_V, is readily obtained from the molar volume, *V*, of the liquid or solid substance of which

the molecules are constituents:

$$\sigma_V = [(6/\pi)V/N_A]^{1/3} \qquad (1)$$

where N_A is Avogadro's number and a spherical shape is tacitly assigned to the molecules. It overestimates the size, since the void space that is inevitably present in condensed phases between the molecules is ascribed to and apportioned among them,[14] yielding the values of σ_V. These values are also temperature dependent (through the values of *V*, its use requiring the experimentally obtained density), although the cube root in Eqn (1) assuages this dependence somewhat.

The minimal estimate, $\sigma_W = [(6/\pi)V_{vdW}/N_A]^{1/3}$, is obtained from the van der Waals volumes, V_{vdW}. These, in turn are obtained from the van der Waals radii of the constituting atoms with account being taken of the overlap of the atomic spheres due to mutual bonding.[15,16] The values of σ_W are underestimates, since depressions in the surfaces of the molecules are inaccessible to other molecules, hence their distances apart must be larger than σ_W. Obtaining the value of σ_W requires a computational program (several such programs are commercially available or can be downloaded from the Internet) and involves an ambiguity in cases where the molecules may have several conformations. The trend nowadays is to rely on packaged quantum-mechanical computational

Correspondence to: Y. Marcus, Department of Inorganic and Analytical Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel.
E-mail: ymarcus@vms.huji.ac.il

programs to provide information on molecular sizes. Such programs even optimize the conformation, pertaining to isolated molecules, however. The derived molecular surface area and volume of isolated molecules may or may not be directly related to the corresponding values for the materials in condensed phases.

In both maximal and minimal estimates one parameter, the diameter, describes the size, it being assumed that the molecules are spherical. Some molecules, called globular, are indeed well approximated by spheres. Examples of globular molecules are tetrachloromethane (CCl_4), sulfur hexafluoride (SF_6), neopentane [2,2-dimethylpropane, $(CH_3)_4C$] and *tert*-butanol [$(CH_3)_3COH$]. However, most relatively small molecules are better described by cylinders: squat ones for oblate (disk-shaped) and elongated ones for prolate (rod-shaped) molecules. Such molecules can be more closely packed in liquids with their major dimensions (diameters for disks, lengths for rods) parallel, as they are in the extreme cases of liquid crystals, but subject to thermal disorder and also to dipole interactions and hydrogen bonding.

Molecules with relatively long chains are folded to some extent in the liquid state and may then be described by a prolate ellipsoid of rotation for their most common conformation. There are, however, molecules that cannot be well described by ellipsoids of rotation if they are significantly bent, e.g. some conformations of dipropylamine [$(C_3H_7)_2NH$] or diethyl ether [$(C_2H_5)_2O$]. However, in some other conformations of the alkyl chains of such molecules the overall elongated shape is retained and in the liquid state these may be dominant so as to ensure the closest packing.

The cavities in which molecules are located in liquids have been said to have diameters that are 120% those of σ_W, as is widely accepted,[17] but this concept of a cavity still ascribes sphericity (or globularity) to the molecules. (Klamt *et al.*,[17] however, did calculate molecular cavities that were not spherical but having a volume corresponding to a spherical cavity.) Ellipsoids of rotation, either oblate or prolate, requiring two diameters, those of the major and minor axes, may be better descriptions of the sizes of the cavities in which molecules can rotate in liquids.

Meyer[1] suggested a compromise between the minimal and maximal estimates of the sizes described above. According to this compromise, the van der Waals radii of the atoms (at the surface of the molecule) are expanded. In particular, the radii of hydrogen atoms are expanded from the non-bonded van der Waals value of 0.120 nm to 0.170 nm. This expedient recognizes the actual shapes of the molecules but increased their apparent globularity, so that the notion of a (single) molecular radius was preserved. This approach was applied to the calculation of the modified σ_W' of 20 alkanes and 20 haloalkanes. Similarly, rolling a sphere with an appropriately chosen diameter over the surface of a molecule, necessarily avoiding crevices in the surface, defines a molecular volume that depends, however, on the diameter of the sphere.[1]

It is attempted in this paper to explore other estimates of the sizes of molecules and by means of data on 350 compounds to emphasize the differences between prolate and oblate molecules. The former include chain-like molecules and the latter, of course, include aromatic, alicyclic and heterocyclic molecules. The data pertain mainly to substances that are liquid at room temperature, although a few solid and gaseous substances are included. Compounds with up to 10 carbon atoms (but also alkanes up to hexadecane) with the most important functional groups are considered, in addition to a few inorganic compounds. The data are derived mainly from the DIPPR compilation[18] and those used include the van der Waals volume, the van der Waals surface area, the molar volume, the isobaric expansivity and the molar enthalpy of vaporization. The temperature-dependent values were those at 298.15 K, where available, or else those closest to room temperature. In the cases of compounds that are gases at room temperature, values of these three properties extrapolated to 298.15 K were employed when available, but if not, values characteristic of gaseous compounds slightly below their normal boiling-points were used. Values for compounds that are solid at room temperature were chosen at as low a temperature above their melting-points as possible.

CATEGORIES OF MOLECULES: ROD-SHAPED, DISK SHAPED AND GLOBULAR

Consider a rod-shaped cylinder of radius r nm and length l nm. Its volume is $v = \pi r^2 l$ nm^3 and its surface area is $a = 2\pi r(l + r)$ nm^2 [Fig. 1(a)]. As the length l becomes longer, the ratio v/a tends to a limit:

$$\lim(l \to \infty)v/a = r/2 \qquad (2)$$

The ratio v/a of rod-shaped molecules is given by the ratio of the listed[18] molar van der Waals volume, V_{vdW}, and surface area, A_{vdW}, each being N_A times the corresponding molecular quantities. Empirically, for chain-like molecules the ratio V_{vdW}/A_{vdW} tends towards a limit, yielding

$$r = 2\lim(l \to \infty)V_{vdW}/A_{vdW} = 0.147 \text{ nm} \qquad (3)$$

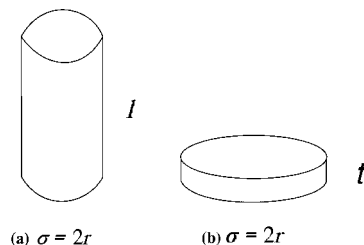(a) $\sigma = 2r$ (b) $\sigma = 2r$

Figure 1. Schematic diagrams of (a) rod-shaped and (b) disk-shaped cylindrical molecules
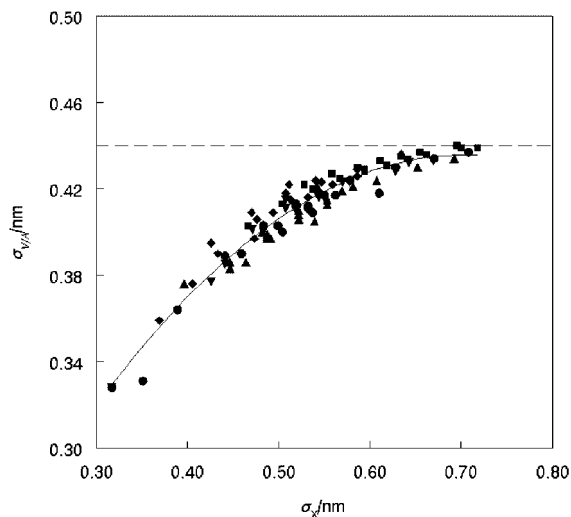
Figure 2. The diameters $\sigma_{V/A}$ [$= \sigma_W$ of Eqn (7)] of prolate molecules derived from van der Waals volumes and areas as a function of the McGowan diameters σ_X: ●, alkanols; ▲, carboxylic acids and esters; ▼, ethers and ketones; ■, hydrocarbons; and ◆, nitriles and amines

This is the case for n-alkanes, 1-alkenes, 1-alkynes, alkyl chlorides and bromides, 1-alkanols, 1-alkanoic acids and their esters, ethers and ketones (irrespective of the positions of the —O— or —C(O)— groups), nitriles and amines, as well as bifunctional compounds, such as α,ω-alkanediols, amides, and 2-substituted ethanols, among others (Fig. 2). These chain-like molecules approximating elongated cylinders should all have the same radius r. Hence the length of the rod-shaped molecule is proportional to its van der Waals volume:

$$l = V_{vsW}/N_A\pi r^2 = 0.0246(V_{vdW}/\text{cm}^3\,\text{mol}^{-1})\,\text{nm} \quad (4)$$

In order to be considered a rod-shaped cylindrical (prolate) molecule, its length l should exceed its diameter, $\sigma = 2r$. This condition is fulfilled according to Eqns (3) and (4) if $V_{vdW} > 120\,\text{cm}^3\,\text{mol}^{-1}$. For chain-like molecules with smaller V_{vdW} values, a more globular than a rod-like cylindrical shape is inferred.

Consider now a squat, disk-shaped molecule, of which the thickness $t = v/\pi r^2$ must also be compatible with the surface area $a = 2\pi r(r + t)$ [Fig. 1(b)]. An iterative calculation using data for such molecules yields the value $t = 0.452\,\text{nm}$. The radius of a disk-shaped molecule is then proportional to the square root of the van der Waals volume:

$$r = 0.0342(V_{vdW}/\text{cm}^3\,\text{mol}^{-1})^{1/2}\,\text{nm} \quad (5)$$

The prototype for such rigid disk-shaped molecules is benzene, for which $v/a = V_{vdW}/A_{vdW} = 0.0807\,\text{nm}$. The diameter of this disk then emerges from Eqn (5) as $\sigma = 2r = 0.476\,\text{nm}$ that is somewhat larger than its thick-

ness. In order for a molecule to be considered disk-shaped, its diameter σ should indeed be larger than its thickness t, and this requires a van der Waals volume $V_{vdW} > 44\,\text{cm}^3\,\text{mol}^{-1}$. It turns out that three- and four-membered cyclic molecules (e.g. cyclopropane, cyclobutane and β-propiolactone) do not fulfil this requirement and should be considered to be globular. Only cyclic compounds with five-membered or larger rings can be considered to be disk-shaped, with the thickness inferred above and a diameter given by Eqn (5).

Globular molecules are approximately spherical and are characterized by a radius

$$r = 0.5[(6/\pi)V_{vdW}/N_A]^{1/3} \quad (6)$$

Van der Waals volumes and areas have been calculated and reported for a very large number of molecules.[18] The ratios of these quantities have a linear dimension and can be translated into molecular diameters for globular (spherical) molecules:

$$\sigma_W = 6V_{vdW}/A_{vdW} \quad (7)$$

Examination of the values of V_{vdW}/A_{vdW} of the 350 substances considered in the present study showed that for oblate (disk-shaped) compounds (aromatic, alicyclic or heterocyclic), $V_{vdW}/A_{vdW}(\text{oblate}) \geq 0.0735\,\text{nm}$ in substantially all cases. On the other hand, for prolate (elongated) compounds, $V_{vdW}/A_{vdW}(\text{prolate}) \leq 0.0735\,\text{nm}$ except for some alkyl halides with molar masses $>170\,\text{g}\,\text{mol}^{-1}$ (e.g. 1,1,2,2-tetrachloroethane, bromoform and butyl iodide) and very long chain alkanes (tetradecane and above). For the latter $V_{vdW}/A_{vdW}(\text{prolate})$ is slightly larger than $0.0735\,\text{nm}$ but does not exceed $0.0745\,\text{nm}$. The value $V_{vdW}/A_{vdW} = 0.0735\,\text{nm}$ can be used as a criterion for distinguishing prolate from oblate molecules.

COLLISION DIAMETERS

Substances in the gaseous (vapor) state can be characterized by their collision diameters (Fig. 3). These are inferred from diverse experimental data, but unfortunately the values are in not very good agreement with each other. One simple way to obtain these diameters is from the exclusion volumes b in the van der Waals equation of state,[19] that are four times the volumes of
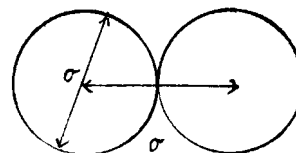

Figure 3. Schematic diagram of the collision of two spherical particles (atoms, molecules)

the molecules. Then, assuming the molecules to be spherical, the diameters are

$$\sigma_b = (3b/2\pi N_A)^{1/3} \qquad (8)$$

These diameters have been used in the estimation of the permeation of gases through polymer films.[5]

More commonly used are collision diameter values obtained from the second virial coefficient B_2 that is obtained from the equation of state, i.e. the P–V–T properties of the substance. The compressibility factor $Z = PV/RT$ yields B_2 as the limit:

$$B_2(T) = \lim(V^{-1} \to 0) \, V(Z - 1) \qquad (9)$$

It is related to the two-body interaction potential energy $u(r)$ that is a function of the distance apart r of the centers of the molecules (or atoms of monatomic substances). It is customary to employ for $u(r)$ either the hard sphere expression, $u^{hs}(r) = \infty$ for $r < \sigma$ and $u^{hs}(r) = 0$ for $r \geq \sigma$, or the Lennard–Jones expression:

$$u^{LJ}(r) = 4\varepsilon\left[(\sigma/r)^{12} - (\sigma/r)^{6}\right] \qquad (10)$$

According to this expression, the mutual interaction energy vanishes when the molecules are at a distance σ apart and increases steeply (infinitely for u^{hs}) for shorter distances. It becomes attractive (i.e. negative) for u^{LJ} (though not for u^{hs}) at $r > \sigma$, reaching a minimum at $r_0 = 2^{1/6}\sigma$, the equilibrium distance, at which $u^{LJ}(r_0) = \varepsilon$, the depth of the potential well, beyond which it becomes less negative, approaching zero asymptotically. The relationship between the experimental quantity $B_2(T)$ and the collision diameter σ via $u(r)$ is far from direct.[20] Some values obtained from B_2 are shown in Table 1. The parameters ε and σ of the Lennard–Jones expression

Table 1. Collision diameters of permanent gases or substances in the vapor state to fit the Lennard–Jones expression (10) and experimental data from various sources

Substance	σ^{LJ}/nm^{10}	σ^{LJ}/nm^{11}	σ^{LJ}/nm^{35}	$\sigma^{LJ}/nm^{34}(B_2)$	$\sigma^{LJ}/nm^{34}(\eta)$
Helium	0.330			0.257	0.263
Neon			0.252	0.275	0.279
Argon	0.333	0.343	0.350	0.340	0.347
Krypton	0.364		0.383	0.360	0.368
Xenon	0.397		0.410	0.396	0.406
Hydrogen	0.316				0.287
Nitrogen	0.353		0.375	0.370	0.375
Oxygen				0.358	0.346
Carbon monoxide	0.358			0.376	0.359
Carbon dioxide	0.359	0.345	0.433	0.449	0.390
Carbon disulfide	0.421	0.425	0.444		0.444
Water			0.302		
Ammonia			0.274		
Methane	0.365	0.364	0.378		0.370
Ethylene	0.400	0.397		0.452	0.407
Ethane	0.418	0.413		0.395	0.438
Propane	0.464			0.564	0.506
n-Pentane	0.530				0.577
Neopentane			0.745		
n-Hexane	0.566	0.557	0.592		0.591
Cyclohexane	0.535	0.548	0.614		0.609
Benzene	0.503	0.505	0.563	0.692	0.527
Toluene	0.537	0.532			0.593
Ethylbenzene	0.568				
Mesitylene	0.618				
Naphthalene	0.587				
Chloromethane		0.403			
Dichloromethane		0.439		0.475	0.476
Chloroform		0.478	0.543		0.543
Tetrafluoromethane	0.446		0.588	0.470	
Tetrachloromethane	0.514	0.523			0.588
Tetramethylsilane		0.561			
Sulfur hexafluoride	0.460	0.472		0.551	0.546
Methanol					0.359–0.367
Ethanol					0.437–0.446
Diethyl ether			0.544		0.554
Acetone	0.468		0.467		0.467
Ethyl acetate	0.520		0.517		0.516
Pyridine		0.486			
Acetonitrile	0.433	0.410			

are not independent for a given $u(r)$ and, furthermore, other valid expressions for $u(r)$, such as the Kihara core-and-periphery potential, yield other values of these parameters, so that the collision diameters obtained according to this route are not unique.

Gas viscosities are another route to the collision diameters. The mean free path l^{fp} between collisions of molecules in the gas at low pressures is given by

$$l^{fp} = 1/\left(2^{1/2}\pi\sigma^2\rho\right) \qquad (11)$$

where $\rho = N_A/V$ is the number density of the substance at the given temperature and pressure. The mean free path is related to the attenuation of the number of molecules $n(x)$ along a distance x in a molecular beam through collisions. The viscosity of a gas of hard sphere molecules is given by:

$$\eta = (5/16)(\pi m k_B T)^{1/2}/\pi\sigma^2 \qquad (12)$$

Its diffusion coefficient is

$$D = (6/5)\eta\rho = (3/8)(\pi m k_B T)^{1/2}\rho/\pi\sigma^2 \qquad (13)$$

and the coefficient of thermal conductivity is

$$\lambda = (5/2)\eta c_V/m = (25/32)(\pi m k_B T)^{1/2}c_V/\pi m\sigma^2 \quad (14)$$

where m is the molecular mass and c_V is the heat capacity at constant volume per molecule. All the expressions of these transport properties depend on the cross-sectional area of the molecules, i.e. on σ^2. However, to obtain the collision diameter of real gases and vapors from experimental values of l, η, D and λ, it is necessary to invoke a realistic interaction potential energy rather than the hard sphere one.

It is common usage to employ a perturbation method, according to which the reference property is the hard sphere quantity whereas the Lennard–Jones potential [Eqn (10)] or some other empirical function is used as a perturbation. For instance, the compressibility factor for hard spheres, $Z^{hs} = f(y,T)$, is a function of the packing fraction $y = \pi\rho d^3/6$ of the molecules, where d is an effective diameter, expressed in units of the hard sphere diameter σ^{hs} (i.e. $d^{hs} = 1$). For the perturbation method, the Verlet–Weiss expression[21] for d is well suited:

$$d^{VW} = (0.3837T + 1.068)/(0.4293T + 1) \qquad (15)$$

which is only a mild function of the temperature: it changes from 0.9069 at 100 K to 0.9035 at 1000 K. The hard sphere diameter, σ^{hs}, itself must be obtained from other sources. Good accuracy was achieved for the equation of state with a relatively simple function, $Z^{hs} = f(y,T)$.[8]

The self-diffusion coefficient can be obtained from Eqn (13) for the hard sphere analog and the Lennard–Jones potential [Eqn (10)], as a perturbation. The nonsphericity of the molecules is taken into account by means of an acentric factor ω. The diameter of the molecules, required for the application of the Lennard–Jones expression, is obtained empirically[10,11] from the critical volume V_C of the substance:

$$\sigma^{LJ} = 0.7889V_C^{1/3} \approx 0.5058(RT_C/P_C)^{1/3} \qquad (16)$$

where the approximation can be used if the critical temperature, T_C, and pressure, P_C, but not the volume, V_C, are known. These values and the appropriate expressions have been used successfully for the fitting of the diffusion of a large number of substances.

An extensive list of thus calculated σ^{LJ} values for gases and vapors, obtained mainly from gas viscosity, η, and the second virial coefficient, B_2, is shown in Table 1. Note, however, the appreciable differences between the values, including those given by the same authors in two papers.[10,11]

Diffraction data (x-ray or neutron) of liquids provide values of the pair correlation function $g(r)$, from which the equilibrium interatomic distance (for monatomic liquids) is obtained as the value of r to the first maximum in this function,[22,23] i.e. $r_0 = 2^{1/6}\sigma$. For molecules the first maximum corresponds to an *intra*molecular distance of mutually bonded atoms and much less clear information is obtained from farther peaks. The values are dependent on the temperature and density. For instance, for argon $r_0/nm = 0.346 + 2.6 \times 10^{-3}(T/K) - 4 \times 10^{-4}$ (ρ/nm^{-3}). At the reduced temperature $T/T_C = 0.56$ and densities $\rho \approx 30\ nm^{-3}$, the values of r_0 are 0.318, 0.379, 0.402 and 0.443 for neon, argon, krypton and xenon, respectively. Values were also reported for nitrogen, oxygen, chlorine, water, methane and tetrachloromethane.[22,23] It should be noted that the value of $r_0/2^{1/6}$ obtained for liquid chlorine from x-ray diffraction, 0.374 nm, is considerably shorter than σ that is obtained from η or B_2 in the gas phase, 0.427 nm. No clear information concerning the sizes of non-globular molecules is obtained from diffraction measurements.

SIZE VALUES BASED ON MOLAR VOLUMES

Substances that are not appreciably volatile at the temperature of interest (generally, room temperature) or polyatomic (so that diffraction does not lead to reliable intermolecular values) require different approaches for the determination of their molecular sizes. Workers to whom sources[18,24] of van der Waals volumes, V_{vdW}, required for the use of Eqns (4)–(6) are not accessible may resort to other means to obtain the linear dimensions of molecules. The most obvious drawback of the use of the molar volumes of liquids for the estimation of the molecular sizes is the already mentioned fact that the void

space inherent in the liquid is ascribed to the molecules, so that σ_V is an overestimate of the diameter even of globular molecules. Several methods have been proposed to overcome this difficulty.

One way to deal with the empty space is to invoke the packing factor, k_p, of close-packed hard spheres representing the molecules in the liquid. The actual volume occupied by close-packed hard spheres is $1/k_p = 1/1.725 = 0.5797$ of the total volume. This value is obtained by computation for hexagonal close packing or empirically by placing ball bearings of a uniform size in a cylinder and filling up with water from a burette to the top layer. Translated into a diameter this approach yields $\sigma'_V = k_p^{-1/3}\sigma_V = 0.834\sigma_V$ or $\sigma_V'/nm = 0.1223\,(V/cm^3\,mol^{-1})^{1/3}$.

Another approach starts from the scaled particle theory (SPT) as expressed by Pierotti[25] that relates thermodynamic quantities to the diameters of molecules of liquids. For instance, inert gas solubilities in liquids were thus related by Pierotti, who equated the hard sphere diameters of solute gas and solvent, required for the calculation of the cavity term, with the Lennard–Jones diameters σ, required for the calculation of the interaction term. The arithmetic mean of the sum of the diameters of the (assumed spherical) molecules of the solute and solvent were taken to calculate the radius of the cavity and the distance apart of neighboring molecules interacting according to the $u(r)^{LJ}$ potentials. Wilhelm and Battino[26] followed this practice and Blum and Fawcett[7] used the values of σ^{LJ} of polar solvents derived by these authors[26] to estimate the relative permittivity of the solvents. They[7] also applied a stickiness parameter and intricate expressions to account for non-sphericity and polarizability as well as other electrostatic features in their calculations.

Inert gas solubilities were also used by Kim,[27] who showed that the derived diameters, σ_K, are linearly correlated with σ_V for 50 liquids (hydrocarbons and halocarbons). The σ_K values are significantly smaller than the σ_V values (cf. the σ_V' values above) and hence more likely to describe the actual sizes of the molecules. Kim's expression[27] can be rewritten in terms of the molar volumes of the solvents (presumably those at 298.15 K):

$$\sigma_K/nm = 0.1363(V/cm^3\,mol^{-1})^{1/3} - 0.085 \quad (17)$$

These empirical diameters, σ_K, although derived from the molar volumes, should be devoid of the empty spaces between the molecules and should describe the diameters of the molecules pictured as hard spheres. The average uncertainty involved by the use of Eqn (17) was estimated[27] as 1.6%. Values of σ obtained from gas solubilities according to various authors, traceable to the concepts of Pierotti,[25] are compared in Table 2.

These approaches, leading to σ_V' and σ_K, consider the molecules as close-packed hard spheres for all liquids. They specify a linear dependence of the molecular diameter on the cube root of the molar volume with constant coefficients and do not recognize differences in the shapes of the molecules. However, the SPT permits the calculation of the packing factor k_p or equivalently the packing fraction $y = k_p^{-1/3}$ from thermodynamic quantities other than inert gas solubilities, leading to a specific dependence of the diameter on the molar volume for each substance. The enthalpy of vaporization, $\Delta_V H$, is related, according to Pierotti,[25] to the packing fraction and the isobaric expansivity, α_P, as follows:

$$\Delta_V H = RT(1 + \alpha_P TY) \quad (18)$$

where

$$Y = (1 + 2y)/(1 - y)^3 \quad (19)$$

y being, as before, the packing fraction. If the value of α_P is not available, it can be approximated by $0.001\,K^{-1}$ for the present purpose. A numerical calculation according to Marcus[28] showed that within 0.2% Eqn (19) can be transformed into

$$y = -0.0520\log(Y^2) + 0.4418\log(Y) - 0.0469 \quad (20)$$

In this manner the packing fraction of the molecules in a liquid is no longer a constant but depends empirically on the thermal properties of the liquid. The derived diameter is then

$$\sigma_T/nm = 0.1469(yV)^{1/3} \quad (21)$$

Again, a single parameter describes the size of the molecules, as if they were spherical, although they are packed to a different extent for different liquids. Values of σ_T have been reported[24] for some 250 solvents.

In all these approaches, experimental (and temperature-dependent) values of certain properties of the substances are required. It would be useful if the molecular size could be estimated solely from the molecular formula of the substance, as, more laboriously, obtaining σ_W does. This can be achieved by using the intrinsic volumes according to McGowan.[29] The volume is calculated additively from values ascribed to the constituent atoms:

$$V_X/cm^3\,mol^{-1} = \Sigma_i N_i V_{Xi} - N_b V_{Xb} \quad (22)$$

where N_i and N_b are respectively the numbers of atoms of kind i and the number of bonds in the molecule, V_{Xi} is the molar volume ascribed to an atom of kind i (see Table 3) and $V_{Xb} = 6.56\,cm^3\,mol^{-1}$ is the molar decrement per bond. No account is taken of the structure of the molecule (whether linear, branched or cyclic) or of the bond order (single, double or triple). Consequently, the same intrinsic volumes, V_X, are assigned to positional isomers belonging to different classes of organic compounds [such as $1\text{-}C_4H_9OH$ and $(C_2H_5)_2O$], structural isomers [such as $1\text{-}C_4H_9OH$ and $(CH_3)_3COH$] and geometric

Table 2. Diameters σ ascribed to molecules in liquid solvents obtained from gas solubilities and the SPT expressions of Pierotti[25] by various workers

Solvent	σ^{26}	σ^8	σ^{25}	σ^{36}	$\sigma^{27}(\sigma_K)$
n-Hexane	0.592		0.594	0.597	0.608
n-Heptane	0.625		0.627		0.635
n-Octane	0.654		0.655	0.658	0.661
3-Methylheptane	0.652			0.654	0.659
2,3-Dimethylhexane	0.650			0.627	0.635
2,4-Dimethylhexane	0.652			0.655	0.661
2,2,4-Trimethylpentane	0.652			0.653	0.664
n-Nonane	0.683			0.685	0.684
n-Decane	0.708			0.711	0.707
n-Dodecane	0.758			0.756	0.749
n-Tetradecane	0.800			0.800	0.787
Cyclohexane	0.563		0.565	0.563	0.566
Methylcyclohexane	0.599			0.600	0.603
Benzene	0.526		0.526	0.525	0.525
Toluene	0.564		0.565	0.565	0.562
m-Xylene	0.597				0.594
Water	0.277	0.274		0.275	0.273
Methanol	0.369	0.371	0.371	0.371	0.384
Ethanol	0.434	0.436	0.436	0.434	0.445
1-Propanol		0.498			
2-Butyl-1-propanol	0.529				0.532
Cyclohexanol	0.575				0.560
Acetone	0.476	0.476	0.479	0.475	0.488
Propylene carbonate		0.536			0.515
Perfluoroheptane	0.704				0.711
Perfluoromethylcyclohexane	0.668			0.678	0.707
Hexafluorobenzene	0.565				0.580
Perfluorotri-n-butylamine	0.854				0.883
Tetrachloromethane	0.537		0.538	0.537	0.542
1,1,2,2-tetrachloroethane	0.559				0.560
1,1,2-trichloro-1,2,2-trifluotoethane	0.559				0.587
Fluorobenzene	0.530		0.531	0.530	0.536
Chlorobenzene	0.561			0.560	0.552
Bromobenzene	0.572			0.571	0.558
Iodobenzene	0.592			0.591	0.572
Hydrazine	0.362		0.363		0.348
Methylhydrazine	0.433				0.427
1,1-Dimethylhydrazine	0.485				0.494
Nitromethane	0.431	0.431		0.431	0.430
Nitrobenzene	0.574	0.574		0.571	0.554
Acetonitrile		0.427			0.427
Propanonitrile		0.477			
Butanonitrile		0.532			
Benzonitrile		0.574			
Formamide		0.345			0.381
N-Methylformamide		0.452			0.447
N,N-Dimethylformamide		0.517			0.496
N-Methylacetamide	0.496				0.497
N,N-Dimethylacetamide		0.548			0.533
Hexamethylphosphoric triamide		0.698			0.679
N-Methylpyrrolidin-2-one		0.569			0.541
N,N,N′,N′-Tetramethylurea		0.544			
Carbon disulfide			0.453		0.451
Dimethyl sulfoxide	0.491	0.491		0.491	0.481

Table 3. Values of the atomic volume increments $V_{Xi}/\text{cm}^3\,\text{mol}^{-1}$ according to McGowan[29,30]

C	16.35	H	8.71	O	12.43	N	14.39
F	10.48	Cl	20.95	Br	26.21	I	34.53
B	18.32	Si	26.83	P	24.87	S	22.91

isomers (such as *cis*-CHCl=CHCl and *trans*-CHCl=CHCl). Values of V_X for 250 solvents have been reported.[24] The corresponding intrinsic diameter is then

$$\sigma_X = [(6/\pi)V_X/N_A]^{1/3} \quad (23)$$

It was shown by Abraham and McGowan[30] that the values of V_X and V_{vdW}, i.e. the van der Waals volume, are linearly related for 209 solid and mainly liquid compounds according to:

$$V_{vdW}/cm^3\,mol^{-1} = (0.597 \pm 0.003) \\ + (0.6823 \pm 0.0023)V_X \quad (24)$$

with a correlation coefficient of 0.9988 and a standard deviation of $1.24\,cm^3\,mol^{-1}$.

The molecular sizes estimated according to Eqns (17), (21) and (23), although improvements on σ_V in the sense that they do not include the void space in a liquid in the sizes of the molecules still consider the molecules as spheres. Since σ_X is easiest to calculate, needing only the formula of the molecule, and is temperature independent, it is used in the following as an independent variable in correlations, in spite of the drawback of its insensitivity to the structure and configuration of the molecule. Good linear correlations exist for σ_K and σ_T with σ_X, as might be expected:

$$\sigma_K/nm = (-0.070 \pm 0.010) + (1.160 \pm 0.020)\sigma_X \quad (25)$$

$$\sigma_T/nm = (-0.076 \pm 0.009) + (1.184 \pm 0.017)\sigma_X \quad (26)$$

with standard errors of 0.0023 nm when all the 350 compounds studied are included. The correlations remain essentially the same but with lower deviations of the coefficients and the fits and higher correlation coefficients when highly branched compounds are excluded.

CORRELATIONS OF VAN DER WAALS VOLUME AND AREA RATIOS

For organic compounds with alkyl chains and a functional group at the end of the chain and for bifunctional compounds with an alkyl chain, the values of V_{vdW}/A_{vdW} increase with increasing σ_X asymptotically to the limit of 0.0735 nm as mentioned above. For the range $0.32 \leq \sigma_X/nm \leq 0.72$ to which the data pertain, the dependence

$$V_{vdW}/A_{vdW} = (0.0137 \pm 0.0020) \\ + (0.1677 \pm 0.0077)\sigma_X \\ - (0.1195 \pm 0.0072)\sigma_X^2 \; nm \quad (27)$$

holds for n-alkanes, 1-alkenes, 1-alkynes and n-alkyl chain-like compounds and also moderately branched chains and bifunctional compounds (see Fig. 2). This correlation has a standard error of the fit of 0.0008 and a correlation coefficient of 0.9765 for 135 items. It is robust, in the sense that the addition of the 30 bifunctional compounds to the 105 monofunctional ones hardly changed the coefficients but decreased their probable errors. The fact that V_{vdW}/A_{vdW} does not exceed 0.0735 nm for these compounds can be attributed to the constant contribution of the increments of V_{vdW} and of A_{vdW} per —CH_2— group with lengthening chains. The lengths of such molecules can be calculated from Eqn (4) and the surface areas A_{vdW} from Eqn (27). If V_{vdW} values are not available then these are estimated with the help of Eqn (24).

For globular molecules, as mentioned, a single diameter is an adequate description of the molecular size. The following linear relationship was established for 27 compounds listed in Table 4 that are considered to be more or less globular:

$$V_{vdW}/A_{vdW} = (0.0365 \pm 0.0045) \\ + (0.0653 \pm 0.092)\sigma_X \; nm \quad (28)$$

but with considerable scatter, so that the standard error of the fit is 0.0037 nm. Expressions (25) for σ_K and (26) for σ_T are valid for this subset of the data, as they are for the prolate molecules with extensive alkyl chains.

Table 4. Van der Waals volume to surface area ratios V_{vdW}/A_{vdW} of fairly globular molecules and their MacGowan diameters σ_X

Compound[a]	V_{vdW}/A_{vdW}/nm	σ_X/nm
Ammonia (g)	0.0563	0.341
Boron trifluoride (g)	0.0703[b]	0.386
Tetrafluoromethane (g)	0.0593	0.394
Chlorotrifluoromethane (g)	0.0637	0.433
Tetrafluorosilane (g)	0.0687	0.433
Sulfur hexafluoride (g)	0.0572[b]	0.446
Carbon disulfide	0.0657	0.454
Dichloromethane	0.0697	0.455
Dichlorodifluoromethane (g)	0.0670	0.466
Dimethyl sulfide (g)	0.0687	0.473
Thionyl chloride	0.0682	0.479
Dibromomethane	0.0713	0.486
Trichloromethane (chloroform)	0.0722	0.490
2-Methylpropane (isobutane) (g)	0.0690	0.504
Phosphoryl chloride	0.0693	0.521
Tetrachloromethane	0.0718	0.521
1,1,1-Trichloroethane	0.0708	0.525
Diiodomethane	0.0792[b]	0.527
Tribromomethane (bromoform)	0.0733	0.529
2-Chloropropane	0.0692	0.531
2,2-Dimethylpropane (neopentane) (g)	0.0683	0.537
2-Methylbutane (isopentane)	0.0700	0.537
2-Bromopropane	0.0703	0.543
Tetrachlorosilane	0.0777[b]	0.544
Tetramethylsilane	0.0725	0.560
2-Iodopropane	0.0722	0.560
3-Ethylpentane	0.0715	0.594

[a] Gaseous at room temperature.
[b] Outliers of the linear correlation in Eqn (28).

For oblate compounds as a set, where invariably as mentioned V_{vdW}/A_{vdW} (oblate) ≥ 0.0735 nm, there exists no simple correlation with σ_X, analogous to Eqn (27). The subset of benzene derivatives, however, shows clear dependences, with σ_K and σ_T conforming to Eqns (25) and (26), but

$$V_{vdW}/A_{vdW} = (0.0892 \pm 0.0022) \\ - (0.0167 \pm 0.0065)\sigma_X \text{ nm} \quad (29)$$

The negative sign of the coefficient of σ_X and its large probable error should be noted. The former feature arises from the loss of the simple disk-like shape of the molecules whenever larger substituents, increasing σ_X, are added to the benzene nucleus. The latter feature and the not very good correlation coefficient arise from the very short range of the dependent variable: $0.0778 \leq (V_{vdW}/A_{vdW})/\text{nm} \leq 0.0838$. Bromobenzene ($V_{vdW}/A_{vdW} = 0.0820$ nm) and iodobenzene ($V_{vdW}/A_{vdW} = 0.0838$ nm), included in the correlation of Eqn (29), are slightly above the line. When a second ring is added to benzene (in tetralin, 1-methylnaphthalene, biphenyl, cyclohexylbenzene, benzophenone, diphenyl ether and dibenzyl ether), still higher values of V_{vdW}/A_{vdW} are obtained, and these substances are not included in the correlation. Simple aromatic molecules would be stacked as parallel disks in the liquid (permitting efficient π–π interactions) if thermal motion did not disrupt this optimal structure.

Non-aromatic ring molecules, including alicyclic hydrocarbons, heterocycles and such molecules as cyclopentanone, γ-butyrolactone, ethylene carbonate and pyrrolidin-2-one, are also considered oblate. They conform to the rule that V_{vdW}/A_{vdW} (oblate) ≥ 0.0735 nm and their σ_K and σ_T values conform to Eqns (25) and (26), respectively. The values of V_{vdW}/A_{vdW} do increase with σ_X, in contract to the benzene derivatives, but only moderately so:

$$V_{vdW}/A_{vdW} = (0.062 \pm 0.005) \\ + (0.0287 \pm 0.008)\sigma_X \text{ nm} \quad (30)$$

The scatter is considerable, as seen in the probable errors of the coefficients, and there are several outliers that are not included in the correlation. For three of these, pyrimidine, pyridine and 2-methylpyridine, either or both V_{vdW} and A_{vdW} may be erroneous, since for the former the ratio is much too high and for the latter two much too low. Other outliers are furan, oxazole, pyrazole, thiophene and tetrahydrothiophene, which have appreciably larger V_{vdW}/A_{vdW} values than expected from Eqn (30).

Substances that are gaseous at room temperature have small molecules and σ_X values in the range $0.28 \leq \sigma_X/\text{nm} \leq 0.52$. The corresponding V_{vdW}/A_{vdW} values conform to Eqn (27) and also to Eqn (28) (for $\sigma_X \geq 0.35$ nm), derived from substances that are liquids
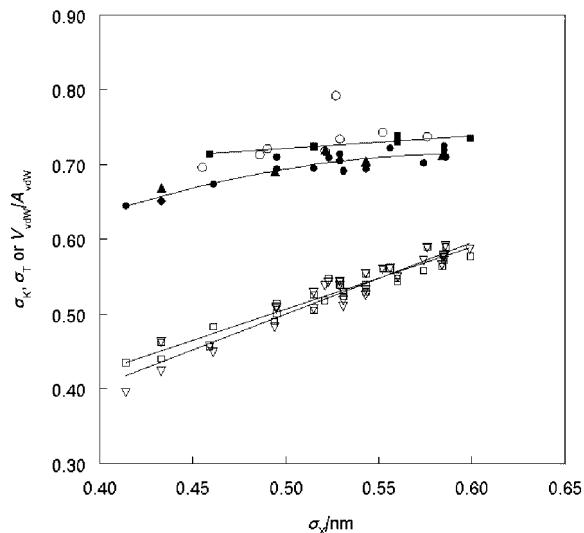

Figure 4. Haloalkane sizes: $\square$ σ_K, $\triangledown$ σ_T and $V_{vdW}/A_{vdW} = (5/3)\,\sigma_{V/A}$ plotted in order to avoid crowding: $\blacklozenge$, V/A(fluorides); $\bullet$, V/A(chlorides); $\blacktriangle$, V/A(bromides); $\blacksquare$, V/A(iodides); and $\bigcirc$ V/A of globular halocarbons

and have globular molecules. Boron trifluoride and sulfur hexafluoride are outliers, the former with too high and the latter with too low V_{vdW}/A_{vdW}, and either or both of their V_{vdW} and A_{vdW} values may be erroneous. The cyclic cyclopropane has the expected larger V_{vdW}/A_{vdW} value than predicted from Eqn (27) but cyclooctafluorobutane does conform.

Halocarbons appear to present a special case, in the sense that iodoalkanes have larger values of V_{vdW}/A_{vdW} than chloroalkanes for a given σ_X, whereas fluoroalkanes and bromoalkanes conform to the curve determined by the chloroalkanes:

$$V_{vdW}/A_{vdW} = -0.005 + 0.258\sigma_X - 0.218\sigma_X^2 \text{ nm} \quad (31)$$

This curve, although with appreciable scatter, is higher than that established in Eqn (27) for chain-like non-halogen-substituted alkanes. Halogen atoms provide larger van der Waals volumes or smaller van der Waals surface areas than other atoms and the iodine atom provides these deviating features more than the other halogen atoms. Globular halocarbons, as a subset, have larger values of V_{vdW}/A_{vdW} than Eqn (31) predicts and mostly lie on the curve representing the iodoalkane values (Fig. 4). On the other hand, the σ_K and σ_T lines have lower slopes (by ca 15%) than Eqns (25) and (26) require. Hense haloalkanes as a subset need special consideration when their sizes are to be estimated.

CONCLUSIONS

For many purposes the molecular volumes, $v_{vdW} = V_{vdW}/N_A$ can be and have been used for the correlation of properties of organic substances. For non-associating

and not very polar organic liquids (hydrocarbons and halocarbons) the solubilities, partition coefficients, vapor pressures, enthalpies of vaporization and boiling-points have been shown[16] to be linear functions of their molecular volumes, v_{vdW}.

However, for other purposes the linear size (one-dimensional extent) of the molecules is relevant. Whether one parameter, the diameter for assumed spherical molecules, or two parameters, diameters and thickness or length of cylindrical molecules or the major and minor axes of ellipsoids of rotation, are to be used must be decided for each application.

The best data available[18,24,31] have been used for obtaining the various estimates of the size (the linear extent) of the 350 molecules considered here. These primary data included the van der Waals surface areas and volumes, the molar volume, the molar enthalpy of vaporization and the isobaric expansivity (the latter three at 298.15 K where available). It is expected that in the cases of compounds that are gases or solids at room temperature σ_K and σ_T would not agree well with σ_X.

Appendix A (Supplementary material) lists the van der Waals size ratios V_{vdW}/A_{vdW} multiplied by 6 as if all the compounds were globular (i.e., molecular $\sigma_{V/A}$ values are listed), in order to make them commensurate with the diameters listed: σ_X, σ_K and σ_T. In very few cases these ratios are so completely out of line ($\sigma_{V/A} > 0.441 = 6 \times 0.0735$ nm for non-cyclic or < 0.441 nm for cyclic compounds) that it is suspected that either their reported[18] van der Waals surface areas or volumes (or both) are incorrect.

One of the quantities σ_X, σ_K, σ_T and σ_W (for globular molecules) or l (for rod-shaped molecules) or $2r$ (for disk-shaped molecules) might have been recommended as approximating best the sizes (diameters, linear dimensions) of the molecules. However, this is not advisable in general. Apart from globular molecules (Table 4), at least two quantities are required to express the sizes: the diameter and the length or thickness of cylindrical molecules or the major and the minor axes if the molecules can be represented as ellipsoids of revolution. These can be estimated for prolate and oblate molecules from V_{vdW} as described above, given the diameter $\sigma = 2r = 0.293$ nm of the former and the thickness $t = 0.452$ nm of the latter and Eqns (3) and (5), respectively.

The σ_X values can be used as first approximations of the sizes of the molecules, but it should be remembered that these values are insensitive to the structure and yield the same values for molecules with diverse structures or positions of the atoms in them. Many such cases can be easily seen in Appendix A. If such pitfalls are to be avoided, the values of σ_K, requiring only the molar volumes, or of σ_T, requiring in addition the molar enthalpies of vaporization and the isobaric expansivities, should be used. Equations (25) and (26) show these estimates to be not too different from σ_X but more sensitive to the individual structures of the molecules.

As seen in Appendix A, they are as a rule up to 10% larger than σ_X, except for the smallest molecules, which, being gaseous, have uncertain σ_K and σ_T values at room temperature.

For some applications the surface fraction that a functional group occupies in a molecule is required. For instance, applications of the UNIQUAC or UNIFAC methods[32,33] for the estimation of the chemical potentials of components in a liquid mixture need this information. Estimation of hydrophobic contact areas also requires this quantity. The surface fractions are often calculated from additive listed values, but can be obtained from the surface area of the functional groups available from Bondi,[15] divided by the van der Waals surface areas A_{vdW} of the molecules. When these A_{vdW} values are not available from compilations, they can be calculated from the appropriate one among Eqns (27)–(31) that yield the ratio V_{vdW}/A_{vdW} and then with the aid of Eqn (24) the desired area A_{vdW} is obtained.

The following cautionary remark is needed as a conclusion. Even if the sizes of molecules constituting the solvent in a solution are known, and assuming them to be independent of the composition of the solution, it is by no means certain that the distances of nearest approach to a solute particle (atom, molecule or ion) are thereby given. Molecules may turn towards a solute particle a particular part of their surface so as to permit hydrogen bonding or donor–acceptor interactions, or to facilitate ion–dipole or dipole–multipole interactions, subject to the constraints imposed by self-interactions of the solvent molecules and thermal agitation. Therefore, the sizes described in this paper can serve as a guide, but each case ought to be studied on its own merits.

Supplementary material

Appendix A, a table giving the data set of molecular sizes, is available at the epoc website at http://www.wiley.com/epoc.

REFERENCES

1. Meyer AY. *Chem. Soc. Rev.* 1986; **15**: 449–474.
2. Abrams RJ, Bretschneider E. In *Internal Rotation of Molecules*, Orville-Thomas WJ (ed). Wiley: Chichester, 1974.
3. Koresh J, Soffer AJ. *J. Chem. Soc., Faraday Trans. 1* 1980; **76**: 2472–2485.
4. Buchwald P. *J. Chem. Phys. B*, 2002; **106**: 6864–6870.
5. Molyneux P. *J. Appl. Polym. Sci.* 2001; **79**: 981–1024.
6. Avnir D, Farin D. *J. Chem. Phys.* 1983; **79**: 3566–3571.
7. Blum L, Fawcett WR. *J. Phys. Chem.* 1996; **100**: 10423–10427.
8. Mulero A, Galan C, Cuadros F. *Phys. Chem. Chem. Phys.* 2001; **3**: 4991–4999.
9. Santavy M, Kypr J. *J. Mol. Graphics* 1984; **2**: 47–52.
10. Liu H, Ruckenstein E. *Ind. Eng. Chem. Res.* 1997; **36**: 5488–5500.
11. Ruckenstein E, Liu H. *Ind. Eng. Chem. Res.* 1997; **36**: 3927–3936.
12. Pandit SP, Kofke DA. *AIChE J.* 1999; **45**: 2237–2244.
13. Mecozzi S, Rebek J Jr. *Chem. Eur. J.* 1998; **4**: 1016–1022.

14. Marcus Y, Jenkins HDB, Glasser L. *J. Chem. Soc., Dalton Trans.* 2002; 3795–3798.
15. Bondi A. *Physical Properties of Molecular Crystals, Liquids and Gases*. Wiley: New York, 1968.
16. Buchwald P, Bodor N. *J. Phys. Chem. B* 1998; **102**: 5715–5726.
17. Klamt A, Jonas V, Bürger T, Lohrenz JCW. *J. Phys. Chem. A* 1998; **102**: 5074–5085.
18. DIPPR, *Design Institute for Physical Property Data Project 801*, Evaluated Process Design Data, American Institute of Chemical Engineers: Brigham Young University, Provo, UT, 2000.
19. Reid RC, Prausnitz JM, Poling BE. *The Properties of Gases and Liquids* (3rd edn). McGraw-Hill: New York, 1977.
20. Sherwood AE, Prausnitz JM. *J. Phys. Chem.* 1964; **41**: 429–437.
21. Verlet L, Weiss JJ. *Phys. Rev. A* 1972; **5**: 939–952.
22. Schmidt PW, Thompson CW. In *Simple Dense Fluids*, Frisch HL, Salsburg ZW (eds). Academic Press: NewYork, 1968: 31.
23. Kruh RF. *Chem. Rev.* 1962; **62**: 319–346.
24. Marcus Y. *The Properties of Solvents*. Wiley: Chichester, 1998.
25. Pierotti RA. *Chem. Rev.* 1976; **76**: 717–726.
26. Wilhelm E, Battino R. *J. Chem. Phys.*, 1971; **55**: 4012–4017.
27. Kim J-I. *Z. Phys. Chem. (NF)* 1978; **113**: 129–150.
28. Marcus Y. *J. Solution Chem.* 1984; **13**: 599–624.
29. McGowan JC. *J. Appl. Chem. Biotechnol.* 1978; **28**: 599–607.
30. Abraham MH, McGowan JC. *Chromatographia* 1987; **23**: 243–246.
31. Riddick JA, Burger WB, Sakano TK. *Organic Solvents* (4th edn). Wiley–Interscience: New York, 1986.
32. Abrams DS, Prausnitz JM. *AIChE J.* 1975; **21**: 116–128.
33. Fredenslund A, Jones RL, Prausnitz JM. *AIChE J* 1975; **21**: 1086–1099.
34. Hirschfelder JO, Curtiss CF, Bird RB. *Tthe Molecular Theory of Gases and Liquids*. Wiley: New York, 1954.
35. Marcus Y. *Introduction to Liquid State Chemistry*. Wiley: Chichester, 1977, and references cited therein.
36. deLigny CL, van der Veen NG. *Chem. Eng. Sci.* 1972; **27**: 391–401.