Chem. Soc. Rev., 1986, **15**, 449—474

The Size of Molecules

By A. Y. Meyer
DEPARTMENT OF ORGANIC CHEMISTRY, HEBREW UNIVERSITY,
JERUSALEM 91904, ISRAEL

1 Introduction

> *'Molecules do have size; therefore, they occupy volume. As a result of this fact, molecules collide with one another whenever one molecule touches the sphere of another.'*[1a]

Computer-generated images of molecules have become almost a symbol of our times.[2] But how extended should be imagined a molecule in the bulk? This is not the same as asking about the distance between given atomic nuclei in contiguous molecules.[3] Molecules do not have sharp boundaries, and the topography of demarking surfaces is to a large extent a matter of decision. And yet, a decision has to be made in numerous fields of study.[4] To cite only a few that are of immediate pertinence to the organic chemist, the molecular radius occurs in equations for the energy of solvation,[5] effects of solvent on electronic spectrum,[6,7] effects of pressure on conformation;[8,9] the molecular cross-section was invoked in analysing odour,[10,11] is required in assessing adsorption data,[12–14] and serves the novel field of fractal dimensions;[15] the molecular surface area has to do with solvation,[16]

---

[1] V. Fried, H. F. Hameka, and U. Blukis, 'Physical Chemistry', Macmillan, New York, 1977. (*a*) p. 59; (*b*) p. 12.

[2] See *Chem. Brit.*, January 1985, and reports therein, in particular, C. H. Hassall, *Chem. Brit.*, 1985, **21**, 39, and S. Ramdas and J. M. Thomas, *Chem. Brit.*, 1985, **21**, 49.

[3] This distinction (for atoms, not molecules) is made in E. Schrödinger, 'What is Life?', The University Press, Cambridge, reprinted 1969, p. 7.

[4] Y. Marcus, 'Introduction to Liquid State Theory', Wiley, London, 1977.

[5] R. J. Abraham and E. Bretschneider, in 'Internal Rotation in Molecules', ed. W. J. Orville-Thomas, Wiley, London, 1974.

[6] N. S. Bayliss, *J. Chem. Phys.*, 1950, **18**, 292.

[7] M. F. Nicol, *Appl. Spectrosc. Rev.*, 1974, **8**, 183.

[8] Y. Taniguchi, H. Takaya, P. T. T. Wong, and E. Whalley, *J. Chem. Phys.*, 1981, **75**, 4815.

[9] Y. Taniguchi, *J. Mol. Struct.*, 1985, **126**, 241.

[10] J. E. Amoore, *Ann. N.Y. Acad. Sci.*, 1964, **116**, 457.

[11] J. E. Amoore, G. Palmieri, E. Wanke, and M. S. Blum, *Science*, 1969, **165**, 1266.

[12] A. L. McClellan and H. F. Harnsberger, *J. Colloid Interface Sci.*, 1967, **23**, 577; R. Sh. Mikhail and E. Robens, 'Microstructure and Thermal Analysis of Solid Surfaces', Wiley, Chichester, 1983, pp. 433—453.

[13] J. Koresh and A. J. Soffer, *J. Chem. Soc., Faraday Trans. 1*, 1980, **76**, 2472.

[14] A. Y. Meyer, D. Avnir, and D. Farin, *J. Am. Chem. Soc.*, in press.

[15] See, for example, D. Avnir and D. Farin, *J. Chem. Phys.*, 1983, **79**, 3566, and D. Avnir, D. Farin, and P. Pfeifer, *Nature (London)*, 1984, **308**, 261, and references given therein.

[16] N. Funasaki, S. Hada, S. Neya, and K. Machida, *J. Colloid Interface Sci.*, 1985, **106**, 255.

partition of solutes between solvents,[17] hydrophobicity,[18,19] molecular recognition,[20] and, obviously, is what one inspects in molecular graphics.[21–23] Estimates of molecular size, or, rather, procedures that produce estimates, have to be supplied.

Persual of the literature reveals, interestingly, that authors opt most frequently for either a maximal or a minimal estimate of molecular dimensions.[24] In the maximalistic option, one partitions the volume of a liquid amongst the constituting molecules. In the minimalistic, one computationally builds up a molecule from atomic spheres of van der Waals radii. In both extremes the molecule is conceived as a sharply bounded rigid body, a sort of a non-spherical or an approximately spherical variant of a 'hard sphere without attraction'.[1b] The rare 'non-classical' approach[25] recurs to quantum chemistry to define shape but not to derive size. On the contrary, the notion of size is introduced by fitting the cutoff-density of electron clouds to van der Waals atomic radii.[26]

Numbers furnished by the two approaches have rarely been confronted. One probable reason is the paucity and scatter of data. Another is that only the newly-born technology of molecular graphics brought the problem to the fore. The situation is now changing. More and more frequently does one hear about calculations of molecular size and encounter tabulated results.[27–35] With numbers at hand, assessment, comparison, and extension of approaches is timely and feasible.

This essay is in part an analysis (from the author's perspective) of some procedures to estimate molecular size, and in part an elaboration of procedures. Changes of size that attend conformational transitions are also alluded to. The discussion is illustrated by new numerical results, mainly for alkanes and chlorinated alkanes—two classes of molecules that are referred to in the context of size more often than any other. Our point will be that the maximalistic estimate is an overestimate, the minimalistic an underestimate, and both unsatisfying in certain respects. Also, that a more satisfying mid-path can be outlined by combining aspects of the two extremes.

---

[17] N. Funasaki, S. Hada, and S. Neya, *J. Phys. Chem.*, 1985, **89**, 3046.

[18] R. S. Pearlman, S. H. Yalkowsky, and S. Banerjee, *J. Phys. Chem. Ref. Data*, 1984, **13**, 555.

[19] K. Iwase, K. Komatsu, S. Hirono, S. Nakagawa, and I. Moriguchi, *Chem. Pharm. Bull.*, 1985, **33**, 2114.

[20] M. Santavý and J. Kypr, *J. Mol. Graphics*, 1984, **2**, 47.

[21] M. L. Connolly, *Science*, 1983, **221**, 709.

[22] N. L. Max, *J. Mol. Graphics*, 1984, **2**, 8.

[23] P. A. Bash, N. Pattabiraman, C. Huang, T. E. Ferrin, and R. Langridge, *Science*, 1983, **222**, 1325.

[24] Y. Sicotte, *J. Chim. Phys.*, 1966, **63**, 403.

[25] (*a*) J. L. Rivail, B. Terryn, D. Rinaldi, and M. F. Ruiz-Lopez, *J. Mol. Struct.*, 1985, **120**, 387; (*b*) E. S. Marcos, B. Terryn, and J. L. Rivail, *J. Phys. Chem.*, 1985, **89**, 4695.

[26] M. M. Francl, R. F. Hout, and W. J. Hehre, *J. Am. Chem. Soc.*, 1984, **106**, 563.

[27] R. Pavani and G. Ranghino, *Comp. Chem.*, 1982, **6**, 133.

[28] M. Marsili, P. Floersheim, and A. S. Dreiding, *Comp. Chem.*, 1983, **7**, 175.

[29] A. Gavezzotti, *J. Am. Chem. Soc.*, 1983, **105**, 5220.

[30] A. Gavezzotti, *J. Am. Chem. Soc.*, 1985, **107**, 962.

[31] A. Y. Meyer, *J. Chem. Soc., Perkin Trans. 2*, 1985, 1161.

[32] A. Y. Meyer, *J. Mol. Struct.*, 1985, **124**, 93.

[33] A. Y. Meyer, *J. Comput. Chem.*, 1986, **7**, 144.

[34] A. Y. Meyer, *J. Chem. Soc., Perkin Trans. 2*, 1986, 1567.

[35] T. R. Stouch and P. C. Jurs, *J. Chem. Inf. Comput. Sci.*, 1986, **26**, 4.

2 Maximalistic Estimate

> *'If the usual approximation is made that the chromaphore is a spherical particle . . .'*[7]

The maximalist procedure is based on a datum purely experimental, namely, liquid density (ρ). It assumes that the entire volume of a liquid is made up of material, with no interstitial voids. This assumption may also be construed to imply that it is fitting to append the average void-per-molecule to the molecule proper. Indeed, some authors[36] partition the molar volume ($v = M/\rho$, where M is the molar mass) into cubes, such that each molecule is let occupy a volume $V_m = v/N_A$ (subscript m for molecular, N_A = Avogadro constant) and assigned a length of $(v/N_A)^{1/3}$. More commonly, molecules are assumed spherical or nearly so. Then, $V_m = M/N_A\rho = 4\pi a^3/3$, or

$$4\pi N_A a^3 \rho/3M = 1 \qquad (1)$$

where a stands for the radius. An alternative formulation,[37] in terms of the specific volume ($\bar{v} = 1/\rho$), is

$$4\pi N_A a^3/3M\bar{v} = 1 \qquad (2)$$

Now, in considering the packing of molecules in a condensed phase, one cannot avoid grappling with empty spaces. First, because they are there;[38] second, because they are put there by the models we use, mechanical or conceptual. Hence it is convenient to distinguish two types.

One type of empty space may be considered void 'by human fault'. Molecules do not have sharp boundaries, but man-made space-filling models do. If we cut molecular models too close and still wish a mole of molecules to fill up the molar volume, there is no choice but leave spaces in-between. To give a specific example, models of alkanes, if tailored to van der Waals dimensions, occupy about half the room that an alkane molecule actually reserves in a liquid at ambient temperatures.[39] At the absolute zero, the fraction seems to be 0.63—0.71.[40,41]

The second type of empty space may be considered void 'by molecular fault'. Molecular surfaces do have protrusions and crevices. Despite the tendency 'to pack bumps in hollows',[42] it happens that dents in one molecule are too voluminous to fit cracks in its neighbour. In a different wording, the hole is less attractive than its walls are repulsive, or the pore is wider than its throat.[38] An apposite example will be given when the concept of shape is called into play (next section). Until then,

---

[36] J. J. Fripiat and H. Van Damme, *Bull. Soc. Chim. Belg.*, 1985, **94**, 825.

[37] M. Barón and H. Mechetti, *J. Phys. Chem.*, 1982, **86**, 3464; M. Barón, *J. Phys. Chem.*, 1985, **89**, 4873.

[38] N. Lane, N. Shah, and W. Curtis Conner, *J. Colloid Interface Sci.*, 1986, **109**, 235.

[39] A. Bondi, *J. Phys. Chem.*, 1954, **58**, 929.

[40] A. K. Doolittle, *J. Appl. Phys.*, 1952, **22**, 1471.

[41] B. Terryn and J. Bariol, *J. Chim. Phys.*, 1981, **78**, 207.

[42] A. I. Kitaigorodskii, 'Organic Chemical Crystallography' (translated from the Russian), Consultants Bureau, New York, 1961, in particular pp. viii and 66.

consider the crystalline structure of monoclinic ethane.[43] Distances between hydrogens of adjacent molecules range from 0.246 to 0.303 nm. Well, if hydrogens can come as close as 0.246 nm, what makes some partners keep 0.303 nm apart? If a laboratory technician groomed his hydrogens to a radius as large as 0.246/2 = 0.123 nm (not less!), he would still have to leave holes in a model of the ethane crystal. This time it is not his fault.

Equations 1 and 2 do not allow for holes of any type: in overcoming human fault, they simultaneously disregard molecular fault and overestimate size. If packing-voids are not to be disregarded, equations 1 and 2 have to be reformulated as inequalities, *e.g.*,

$$4\pi N_A a^3 \rho / 3M < 1 \qquad (1a)$$

Furthermore, the two underlying assumptions of the estimate—sphericity and compactness—are contradictory. Closest-packed spherical objects or ellipsoids of revolution[44] can fill no more than 74% of the allotted space; the rest is interstitial sites. If randomly packed, they would fill 60—63%.[45] Indeed, authors have replaced in equation 1 the equality to 1 by an equality to 0.74.[46] Does this improve the estimate? Certainly, because 0.74 is an upper bound while 1 is well beyond the range. Still, no molecule is strictly ellipsoidal, let alone spherical; also, there is no reason that a molecule be closest-packed in the liquid.

At any rate, literature treats the entity of volume V_m as a geometrical body. In line with earlier practice,[14,33] we shall refer to it as the 'm-solid', where 'solid' has the connotation of a body in three dimensions.

The volume V_m, and hence the extension of the m-solid, depends on temperature and pressure. It is usually cited, and will so be cited below, for 20 °C and atmospheric pressure.

3 Minimalistic Estimate

'In the midst of chaos there was shape.'[47]

In the alternative procedure, one attributes van der Waals radii to atoms, and builds up molecules from overlapping atomic spheres.[27] This choice makes molecules as small as conceivably admissible, and leads to the conclusion that the volume of a liquid at room temperature contains occupied and unoccupied zones in about equal lots. Molecular sphericity is not assumed. On the contrary, each nuclear configuration is ascribed a unique geometrical shape.[48] We shall refer to this choice of radii as 'w-radii' and to the outcoming geometrical solid as the 'w-solid' (w for van der Waals). The w-solid is the computational variant of the space-

---

[43] G. J. H. van Nes and A. Vos, *Acta Crystallogr., B*, 1978, **34**, 1947.

[44] G. Godbout and Y. Sicotte, *Can. J. Chem.*, 1968, **46**, 967.

[45] G. D. Scott, *Nature (London)*, 1960, **188**, 908.

[46] W. B. Moniz and S. H. Gutowsky, *J. Chem. Phys.*, 1963, **38**, 1155.

[47] V. Woolf, 'To the Lighthouse', enlarged edition, Hogarth Press, London, 1977, p. 249.

[48] For an analysis of the notion of shape, see J. T. Hougen, *J. Phys. Chem.*, 1986, **90**, 562.

filling mechanical model, improved, of course, by allowing non-bonded atoms to overlap.

To illustrate, Figure 1 shows a pair of overlapping hydrogen and carbon atoms. The w-radii are $r_w(H) = 0.117$ and $r_w(C) = 0.175$ nm,[29] and the internuclear distance *d* is put equal to 0.240 nm. For comparison, the MM2-distances[49] in anti-butane are $d[C^1 \cdots H(C^2)] = 0.217$, $d[C^2 \cdots H(C^1)] = 0.220$ and $d[C^1 \cdots H(C^3)] = 0.278$ nm. So, the overlap-disks for $C^1 \cdots H^2$ and for $C^2 \cdots H^1$ are much more extensive than the disk shown in the Figure, and there is no overlapping in $C^1 \cdots H^3$. Also, shape is defined, unique to the cited combination of w-radii and *d*. Geometrical attributes—volume, surface area, cross-sectional areas, and more—can be derived by analytical or other procedures.[31–34]

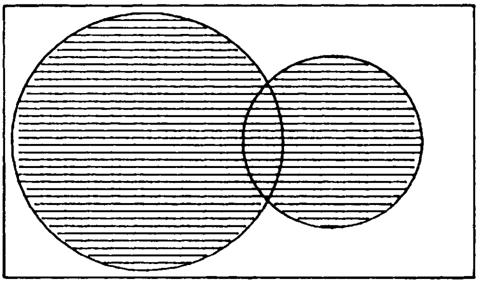


Figure 1 *Overlapping* $C \cdots H$, *non-bonded.* $r(C) = 0.175$, $r(H) = 0.117$, $d = 0.240$ nm

To see why w-radii constitute a minimalistic choice, recall how they are derived.[50] In a molecular crystal, molecules assume equilibrium positions at which the attractive and the repulsive interatomic forces are said to balance. The w-radius of an atom X is defined as half of the smallest distance encountered between atoms X that do not belong to the same molecule. That is, it is half of the lowest observed $d(X \cdots X)$ in $M–X \cdots X–M$, where M stands for variegated residues. As one example, the original estimate of $r_w(Cl)$ was based on intermolecular $Cl \cdots Cl$ distances in the crystal of beta-$C_6H_6Cl_6$. Since these distances fall in the range 0.360—0.382 nm,[51] the value chosen was $r_w(Cl) = 0.360/2 = 0.180$ nm. As another, the closest $H \cdots H$ approach in crystallized alkanes[43,52] is about 0.240 nm. The value of $r_w(H)$, then, was taken as 0.120 nm. A recent tabulation[29] has 0.177 for Cl and 0.117 for H.

Now, the $X \cdots X$ distance in $M^1–X^1 \cdots X^2–M^2$ is not an intrinsic property of X. Rather, it reflects the interplay among properties of both X and M. At the equilibrium distance, X^1 repels X^2 but is attracted to M^2, as is X^2 to M^1. If M^2 and

---

[49] For a description of the MM2 procedure, see N. L. Allinger, *J. Am. Chem. Soc.*, 1977, **99**, 8127.

[50] L. Pauling, 'The Nature of the Chemical Bond', 3rd edition, Cornell University Press, Ithaca, N.Y., 1960, p. 258.

[51] R. G. Dickinson and C. Bilicke, *J. Am. Chem. Soc.*, 1928, **50**, 764.

[52] N. Norman and H. Mathisen, *Acta Chem. Scand.*, 1961, **15**, 1755.

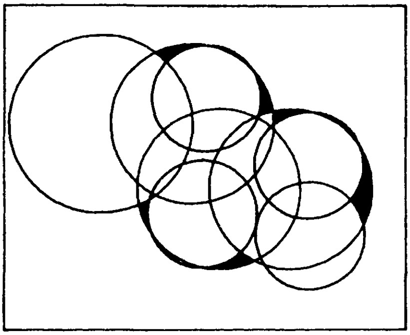


Figure 2 1-*Bromopropane*, anti, *w-solid. Bare carbon-caps shaded*

M^1 were not there, X^2 would be driven farther away from X^1, and half of $d(X^1 \cdots X^2)$ would be larger than the value deduced for $r_w(X)$. By its very definition—outcome of opposing propensities—r_w underestimates the extent of X's range of influence. A numerical demonstration is available for hexane:[53,54] if molecular-mechanical optimization is to reproduce 0.240 nm for intermolecular $d(H \cdots H)$, the hydrogen radius has to be put larger than $r_w(H)$; if $r_w(H)$ is used, d contracts to 0.210 nm. Hydrogen radii used nowadays in molecular mechanics range from 0.143 to 0.182 nm.[55]

The term 'radius' is somewhat misleading, since it makes one visualize a bonded atom as a sphere. Actually, outward-facing extensions of atoms (van der Waals radii) exceed the inward-facing (covalent radii) by factors between 1.5 and 2.2, and diagonally-facing extensions fall in-between.[56,57] Figure 2, where atomic circles are traced explicitly, shows that the assumed shape of atoms affects the external zone of a molecule, but not the internal. If vacuoles happen to occur in simple molecules, they are inaccessible and part of the occupied volume. It is, incidentally, a curious fact that portions of carbon faces come out bare (shaded in the Figure) when r_w-radii are used. The Figure also provides the example we still need of voids 'by molecular fault'. If two molecules of the given shape are to be packed together, the cap of an atomic sphere in one (say, of bromine, Figure 2) is to rest on three caps in the other (say, three hydrogens). The space left inside this arrangement forms an interstitial site. w-Solids, and calculations based on them, sustain both man-made and molecule-made voids.

The volume of w-solids can be estimated by increments,[58] calculated analytically[27,59] or numerically.[29,31,35,41] The numbers cited below were obtained as follows.[33]

---

[53] N. L. Allinger, M. A. Miller, F. A. Van Catledge, and J. A. Hirsch, *J. Am. Chem. Soc.*, 1967, **89**, 4345.
[54] N. L. Allinger, M. T. Tribble, M. A. Miller, and D. H. Wertz, *J. Am. Chem. Soc.*, 1971, **93**, 1637.
[55] A. Y. Meyer and F. R. F. Forrest, *J. Comput. Chem.*, 1985, **6**, 1.
[56] C. Glidewell, *Inorg. Chim. Acta*, 1975, **12**, 219.
[57] S. C. Nyburg, *Acta Crystallogr., A*, 1979, **35**, 641.
[58] A. Bondi, 'Physical Properties of Molecular Crystals, Liquids, and Glasses', Wiley, New York, 1968.
[59] M. L. Connolly, *J. Am. Chem. Soc.*, 1985, **107**, 1118.

The molecule is computationally enclosed in a graduated cuboid. Let the spacing between grid points be g. A search is then conducted, point-by-point (P) and atom-by-atom (X), checking whether $d(\mathrm{P} \cdots \mathrm{X}) < r(\mathrm{X})$ for any X. Whenever the inequality holds, the point is taken to contribute an element g^3 to the volume. Figures 1 and 2 show each a slab (thickness g) of the cuboid. The rectangular frame stands for the sides. The equidistant lines (distance g) in Figure 1 connect streams of points that fulfil the condition.

Numerical estimation of surface area is by necessity more approximate than the estimation of volume.[35] Several suggestions have been made.[30,33,60] In our programme,[33] points are scanned that fulfil $r(\mathrm{X}) < d(\mathrm{P} \cdots \mathrm{X}) < r(\mathrm{X}) + g$, like point Q in Figure 3. This point accounts for g^2 units of area on the surface of a sphere of radius $r(\mathrm{X}) + h$. On the surface that atom X contributes to the molecule, it accounts for a fraction of this quantity, namely, $g^2\{r(\mathrm{X})/[r(\mathrm{X}) + h]\}^2$.

Volumes and surface areas have also been measured on molecular models.[10,17]

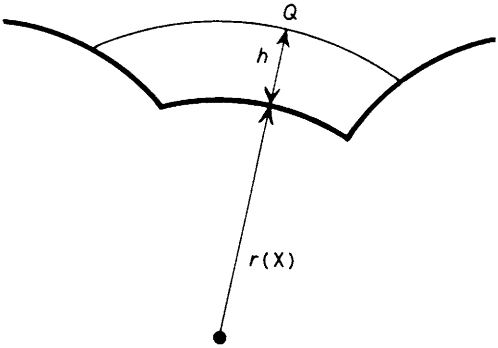


Figure 3 *Estimation of surface area. Thick curves represent the molecular surface*

4 w-Solids—Shape and Size

The trouble with computed attributes of w-solids is not that they do not correlate with bulk properties. They do,[18,31,33,34,41,61,62] and astonishingly well. It is that their magnitudes differ so much from the corresponding measurables. The cross-sections of w-solids are small with respect to established values.[33] For instance,[14] the cross-section of an alkyl chain comes out as 0.13—0.14 nm^2/molecule. In comparison, Langmuir's classical value[63] is 0.22, and the currently accepted range is 0.19—0.25 nm^2.[64]

As an illustration of the nature of w-solids, consider a sample of 20 liquid alkanes,

---

[60] M. L. Connolly, 'MS', Research Institute of Scripps Clinic, La Jolla, California, 1983, programme distributed by QCPE.

[61] S. C. Valvani, S. H. Yalkowsky, and G. L. Amidon, *J. Phys. Chem.*, 1976, **80**, 829.

[62] C. L. de Ligny and N. G. van der Veen, *Chem. Eng. Sci.*, 1972, **27**, 391.

[63] I. J. Langmuir, *J. Am. Chem. Soc.*, 1917, **39**, 1848.

[64] D. H. Wheeler, D. Potente, and H. Witcoff, *J. Am. Oil Chem. Soc.*, 1971, **48**, 125.

comporting unbranched and branched, acyclic and cyclic molecules. The van der Waals volumes (V_w) of these alkanes are listed in Table 1 (column A). In each case, V_w is cited for one conformation only. This does not incur error, since V_w is almost insensitive to stereochemical change[31] (see below, and also items 17—20 in the Table). Column B lists the molar volumes of the neat liquids as measured at 20 °C. The room that a molecule actually reserves in the liquid state is $V_m = v/N_A$, and the fraction of this room that is filled by the w-solid is denoted by f, $f = V_w/V_m = V_wN_A/v$, and is listed in parentheses in column A.

Table 1 *The volume of alkanes*

	A	*B*	*C*	*D*
Alkane	V_w *(and* f_w*)*	v	$v(V_w)$	$v(V_w, G_w)$
(1) pentane	0.0960 (0.50)	115.2	110.0	116.9
(2) hexane	0.1128 (0.52)	130.5	125.9	133.8
(3) heptane	0.1297 (0.53)	146.5	141.8	149.6
(4) octane	0.1466 (0.54)	162.6	157.8	165.2
(5) nonane	0.1634 (0.55)	178.7	173.7	180.0
(6) decane	0.1802 (0.56)	194.9	189.6	194.7
(7) 2-methylbutane	0.0959 (0.50)	116.4	109.9	114.4
(8) 2,2-dimethylbutane	0.1124 (0.51)	132.9	125.5	128.4
(9) 2,3-dimethylbutane	0.1126 (0.52)	130.3	125.7	127.2
(10) 3,3-dimethylpentane	0.1292 (0.54)	144.5	141.4	141.6
(11) 2,2,4-trimethylpentane	0.1459 (0.53)	165.1	157.2	156.9
(12) cyclopentane	0.0855 (0.55)	94.1	100.1	94.3
(13) methylcyclopentane	0.1022 (0.55)	112.4	115.9	112.5
(14) cyclohexane	0.1006 (0.56)	108.1	114.4	109.3
(15) methylcyclohexane	0.1182 (0.56)	127.6	131.0	128.4
(16) cycloheptane	0.1180 (0.59)	121.3	130.8	125.2
(17) *cis*-1,2-dimethylcyclohexane	0.1349 (0.58)	140.9	146.8	142.2
(18) *trans*-1,2-dimethylcyclohexane	0.1349 (0.56)	144.6	146.8	143.3
(19) *cis*-decalin	0.1573 (0.61)	154.2	167.9	156.5
(20) *trans*-decalin	0.1572 (0.60)	158.9	167.8	159.4

Column A: van der Waals volume (V_w) in nm³/molecule, calculated at $g = 0.01$ nm, with $r_w(\mathrm{C}) = 0.175$ and $r_w(\mathrm{H}) = 0.117$ nm. In parentheses: fraction of V_w in molecular volume ($f = V_wN_A/v$). Column B: molar volume (v) in cm³/mol, as derived from reported density at 20 °C ($v = M/\rho$). Column C: molar volume as calculated from one-parameter linear regression (equation 3). Column D: molar volume as calculated from two-parameter linear regression (equation 5)

As can be seen, f ranges from about 0.5 to 0.6, so that w-solids occupy 50—60% of the molar volume at 20 °C. The rest is 'empty space' or 'free volume'. For crystals, incidentally, the packing density f is usually in the range 0.70—0.78.[65] It is noteworthy that f is larger for cyclic than for acyclic alkanes: the former sustain a denser packing.

---

[65] F. M. Richards, *J. Mol. Biol.*, 1974, **82**, 1.

This conclusion is better presented in a different way. Regressing linearly v on V_w, one obtains

$$v = 944.75\ V_w + 6.4590 \qquad (3)$$

(v in cm^3/mol, V_w in $mm^3/molecule$, $r = 0.9469$, $s = 6.4590$). As a predictive tool, equation 3 is not useful. But the sense of deviations is noteworthy. Column C in Table 1 lists the predicted v-values. It is readily seen that the least-squares line underrates the volume of acyclic alkanes and overrates the volume of the cyclic. As an illustration, hexane and cycloheptane have very close V_w; if anything, the w-solid of the latter is somewhat larger (column A). Yet, the more globular cycloheptane packs into a much smaller V_m (column B). Hence, if one is to go beyond equation 3, a measure of globularity is required.

Globularity[66] is usually invoked in a qualitative sense: branching[67] and skewing[68,69] are said—and rightly so—to increase it. Quantitative measures of globularity,[33,34] or of deviation from sphericity,[41] have been proposed. Here it is convenient to refer to the 'equivalent sphere', *i.e.*, a sphere of volume equal to V_w.[70] Its radius is $r_{e,w} = (3V_w/4\pi)^{1/3}$, and its surface area $S_{e,w} = 4\pi r^2_{e,w} \sim 4.836\ V_w^{2/3}$. If there were a perfectly spherical molecule, its surface area S_w would be equal to $S_{e,w}$ and its globularity G_w to $S_{e,w}/S_w = 1$. As one of the molecular axes becomes longer with respect to the others, and/or the surface becomes more convoluted, S_w increases out of proportion with $S_{e,w}$ and the ratio G_w decreases, Table 2 lists computed G_w values for the extended conformation of straight-chain alkanes. These numbers are best fit by an exponential curve,

$$G_w = 0.9715\exp(-0.03764N) \qquad (4)$$

where N is the number of carbon atoms ($r = 0.9975$, $N = 0$ refers to H_2). If one wishes, equation 4 can be turned to ascribe an index N to any given species:

$$N \sim -(\ln G + 0.0289)/0.0376 \qquad (4a)$$

A scale of molecular globularities is thus defined.

When v of the 20 alkanes of Table 1 is regressed against both V_w and G_w, one obtains

$$v = 610.55(49.26)V_w - 221.64(26.91)G_w + 234.69(29.40) \qquad (5)$$

($n = 20$, $R = 0.9931$, $S = 3.1360$). Column D (Table 1) lists v values, as projected from equation 5. A rival formula[71a] uses a five-variable connectivity function, and does not apply to cyclic alkanes. It should be stressed, however, that equation 5 was derived by using V_w and G_w for one sole conformation of each alkane, while v

---

[66] For the origin of the term, see J. Timmermans, *J. Chim. Phys.*, 1938, **35**, 331; J. Timmermans, *Nature*, 1950, **174**, 235.

[67] A. M. Awwad, F. I. Kanbour, and E. I. Allos, *J. Chem. Therm.*, 1984, **16**, 917.

[68] P. E. Schoen, R. G. Priest, J. P. Sheridan, and J. P. Schnur, *J. Chem. Phys.*, 1979, **71**, 317.

[69] D. W. Rebertus, B. J. Berne, and D. Chandler, *J. Chem. Phys.*, 1979, **70**, 3395.

[70] J. T. Edward and P. G. Farrell, *Can. J. Chem.*, 1975, **53**, 2695.

[71] L. B. Kier and L. H. Hall, 'Molecular Connectivity in Chemistry and Drug Design', Academic Press, New York, 1976. (*a*) p. 145; (*b*) p. 108.

Table 2 *Globularity of extended conformations*

N	*Molecule*	G_w
0	H_2	0.977
1	CH_4	0.945
2	C_2H_6	0.903
3	C_3H_8	0.868
4	C_4H_{10}	0.830
5	C_5H_{12}	0.796
6	C_6H_{14}	0.766
7	C_7H_{16}	0.741
8	C_8H_{18}	0.717
9	C_9H_{20}	0.697
10	$C_{10}H_{22}$	0.677

constitutes a weighted mean. In particular, straight-chain alkanes were represented by the all-*anti* conformation, whereas the liquids contain a host of skew-forms,[68] all much more globular than the all-*anti*. Regressing for the branched and cyclic molecules alone, one gets

$$v = 464.45(43.80)V_w - 359.54(31.83)G_w + 363.84(30.02) \quad (6)$$
$$(n = 14, R = 0.9958, S = 2.0520).$$

To obtain an approximation for V_m, volume of the m-solid, we divide equation 5 by N_A;

$$V_m = 0.3897 + 1.0138\ V_w - 0.3680\ G_w \quad (5a)$$

(V_m and V_w in nm^3/molecule). Thus,[70] the volume of alkanes at 20 °C can be said to represent three distinct contributions: Volume of point particle at 20 °C ($\sim$0.39 nm^3/molecule); geometric volume of w-solid; shape-dependent contraction (0.25 $\sim$ 0.32).

From equation 5a, one can approximate the free volume at 20 °C as $V_f = V_m - V_w \sim 0.39 - 0.37\ G_w$. It is a function of globularity alone. Cooling reduces the molecular volume, thereby raising f, which may be taken to mean that the free volume contracts. Table 3 compares V_w with the molecular volume of various liquids, as extrapolated to absolute zero.[40,41] For straight-chain alkanes, f rises from *ca.* 0.5 (Table 1) to ca. 0.7 (Table 3). It is noteworthy that $f \sim 0.7$ for most compounds, and that f attains its upper bound, 0.74, for CH_3Br and CH_3I.

Another measure of molecular volume is the molar refraction[72] (R_M). R_M is easily predicted, by group increments[73] or from connectivity indices.[71b] Regression for the 20 alkanes of Table 1 yields $R_M = 277.72\ V_w - 0.9148$ ($r = 0.9957, s = 0.6518$, R_M in cm^3/mol, V_w in nm^3/molecule). Inclusion of G_w improves the fit:

$$R_M = 311.00(5.23)V_w + 22.07(2.86)G_w - 22.36(2.80) \quad (7)$$

---

[72] C. Hansch, A. Leo, S. Unger, K. Kim, D. Nikaitani, and E. Lien, *J. Med. Chem.*, 1973, **16**, 1207.

[73] A. I. Vogel, 'A Textbook of Practical Organic Chemistry', 3rd edition, Longmans, London, New Impression, 1957, pp. 1035—1036.

Table 3 *Molecular volume of liquids at absolute zero*

Compound	V_m[a]	V_w[b]	*f*
(1) methane	0.0450	0.0284	0.63
(2) ethane	0.0673	0.0455	0.68
(3) propane	0.0911	0.0623	0.68
(4) pentane	0.1373	0.0960	0.70
(5) heptane	0.1829	0.1297	0.71
(6) octane	0.2074	0.1466	0.71
(7) nonane	0.2302	0.1634	0.71
(8) CH_3OH	0.0506	0.0353	0.67
(9) C_2H_5OH	0.0742	0.0522	0.70
(10) CH_3F	0.0467	0.0309	0.66
(11) CH_3Cl	0.0613	0.0431	0.70
(12) CH_3Br	0.0679	0.0504	0.74
(13) CH_3I	0.0789	0.0584	0.74
(14) CH_2Cl_2	0.0800	0.0574	0.72
(15) $CHCl_3$	0.1013	0.0719	0.71
(16) CCl_4	0.1242	0.0867	0.70
(17) C_2H_5Cl	0.0860	0.0602	0.70
(18) $(C_2H_5)_2O$	0.1222	0.0856	0.70

Volumes in nm^3/molecule. [a] Extrapolated to 0 K, Refs. 40, 41. [b] Calculated using radii from ref. 29.

($R = 0.9990, S = 0.3328$). In contrast with its effect on v or V_m, globularity is seen to enhance R_M. For an example, refer again to cycloheptane and hexane: the former has the lower molar volume (Table 1), but a higher molar refraction (32.18 *vs.* 29.98 cm^3/mol). This particular effect of globularity does not seem to be documented. It is obscured in the group-increment pathway to R_M-values and in the multi-variable connectivity functions.

In brief, w-solids are not comparable in magnitude to volumes and cross-sections measured in the bulk, but their attributes correlate fairly well with measurables.

5 The Compromise

> *'Hydrogen exhibits a surprisingly large dimension in complete contrast with its van der Waals radius.'*[13]

How should one visualize a molecule in a liquid, or a solute cavity? w-Solids are too small and too convoluted; m-solids are too big, and —however moulded[25b]—not convoluted enough. Both w-solids[74–76] and m-solids[5,37] have served. But the trend nowadays is towards an intermediate estimate, derived by expanding w-solids beyond w-radii. The expansion smooths over bays and capes in the w-solid's face, thereby making it more globular and akin to the m-solids. Three lines of approach should be cited in sequence.

The first was conceived in studying the dissolution of linear alkanes.[77] The solute

---

[74] S. Miertuš, E. Scrocco, and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117.

[75] J. I. Gersten and A. M. Sapse, *J. Comput. Chem.*, 1985, **6**, 481.

[76] K. Yamamoto, *J. Biochem.*, 1974, **76**, 385.

[77] J. Hylton McCreery, R. E. Christoffersen, and G. G. Hall, *J. Am. Chem. Soc.*, 1976, **98**, 7191.

molecule is still assumed spherical. Its radius is made equal to the distance from the centre of mass to the farthest atom, plus twice the w-radius of that atom.

In the second,[78] one starts with a cuboid. The molecular volume is identified with the volume of a rectangular solid, delineated by the extrema of atomic coordinates in the principal coordinate system, plus 0.2 nm in each direction. The ensuing volume is then equated to the volume of a sphere, and a molecular radius derived.

The third approach is the one initiated by Hermann[79,80] and by Lee and Richards.[81,82] It waives sphericity altogether and, simultaneously, rounds up surface cracks that should be impenetrable to surrounding molecules. A probe sphere of radius r_p, usually taken equal to 0.15 nm, is imagined to roll on the face of the w-solid. The surface traced out by the centre of the sphere is the 'accessible surface area' of the molecule (S_{acc}), defining shape and enclosing volume. S_{acc} can be measured on models[16] or computed. The mathematics has been thoroughly investigated by Connolly,[21,59,83] who also implemented computer programmes. Of these, 'MS' (or 'MSDOT') is in common use,[60] both for computation[84] and for graphics.[85]

MSDOT surface areas have a curious property. When r_p is gradually varied, S_{acc} does not change monotonically. Rather, $S_{acc}(r_p)$ contains sharp peaks, like a spectrum, without a definite trend of elevation or depression in the baseline. For instance, calculations for norbornane[86] (at a surface point density of 5) yield $S_{acc} = 1.34$ nm^2 at $r_p = 0.15$ nm, but only 1.27 at 0.14 and 1.29 nm^2 at 0.17 nm. The intervening peak, 1.38, is at 0.16, and there are four other peaks and one shoulder in the range 0—0.25 nm. S_{acc} in norbornane is as high as 1.39 nm^2 at $r_p = 0.27$, but as low as 1.22 at 0.28 nm. At $r_p = 0$ one does obtain the correct surface area of the w-solid, 1.33 nm^2, which precludes computing bugs. These interesting features are not limited to small molecules.[86] They may turn out to be highly informative, reflecting perhaps packing rearrangements.[87] For the time being, care should be exerted in exploiting numerical results.

The compromise we proposed[14,32,33] is conceptually very simple. One proceeds as in computing w-solids, but uses radii larger than w-radii. Thus, the w-solid improves by annexation of a portion of the surrounding space and by mitigation of surface irregularities. The m-solid improves by reduction of volume and attribution of shape. Concurrently, globularity increases with respect to the w-solid, validating and conserving the notion of molecular radius.

Figure 4 is the expanded counterpart of the w-solid in Figure 2. It stresses the analogy in buildup and the distinction in outcome: the molecule has become larger, smoother, more globular; the bare carbon zones of Figure 2 are now well encased.

---

[78] D. L. Beveridge, M. N. Kelly, and R. J. Radna, *J. Am. Chem. Soc.*, 1974, **96**, 3769.
[79] R. B. Hermann, *J. Phys. Chem.*, 1972, **76**, 2754.
[80] R. B. Hermann, *Proc. Nat. Acad. Sci. USA*, 1977, **74**, 4144.
[81] B. Lee and F. M. Richards, *J. Mol. Biol.*, 1971, **55**, 379.
[82] F. M. Richards, *Annu. Rev. Biophys. Bioeng.*, 1977, **6**, 151.
[83] M. L. Connolly, *J. Appl. Crystallogr.*, 1983, **16**, 548.
[84] M. Lewis and D. C. Rees, *Nature (London)*, 1985, **230**, 1163.
[85] C. Hansch, T. Klein, J. McClarin, R. Langridge, and N. W. Cornell, *J. Med. Chem.*, 1986, **29**, 615.
[86] D. Farin, D. Avnir, and A. Y. Meyer, unpublished results.
[87] M. J. Huron and P. Claverie, *J. Phys. Chem.*, 1974, **78**, 1862.

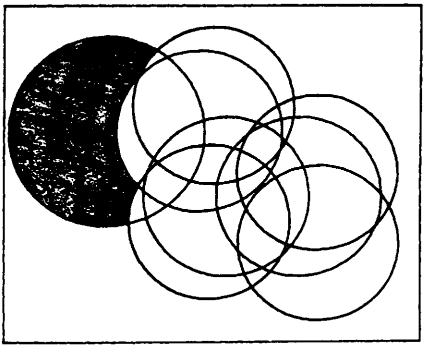


Figure 4 1-*Bromopropane,* anti, *rm-solid* (cf. *Figure* 2). *Hetero-segment shaded*

Smoothing out of surface irregularities is more clearly demonstrated in a branched molecule, far from a principal plane. Figure 5a is a view of 2,2-dimethylbutane, and Figure 5b is a cut in a plane parallel to $C^1C^3C^6$ but distanced from it by 0.15 nm. At this height, only three hydrogen caps (horizontal hatching) remain of the w-solid, separated by an inaccessible moat. In the expanded solid (vertical hatching) this hollow is entirely covered up. A cut farther away from $C^1C^3C^6$ (Figure 5c) shows that the expansion of hydrogens does not replace one unwelcome feature by another.

Actually, only the radii of 'mantle atoms'—atoms outwith the molecular backbone, *e.g.*, hydrogens in an alkane—have to be lengthened.[32] We want the new radii such that the computed volume of a molecule be close to the maximalistic estimate, $V_m = v/N_A$: this should minimize 'man-made voids', by letting each molecule occupy computationally a space of about the size it requires physically. Obviously, and luckily so, a single value for r(H) cannot bring the computed volume of all alkanes close to their respective V_m's. 'Molecule-made' voids are therefore retained. Pilot calculations show that low r(H) values (0.17 nm and less) characterize cyclic alkanes, while high values (0.19 or more) characterize branched alkanes. This is entirely in line with the data in Table 1 and the indications of equation 5a. Hence, if a value for r(H) is selected from the lower end of the range, a tool becomes available to account for the varying amounts of 'molecule-made' voids in alkanes: the computed volume of cyclic alkanes would be very close to V_m, while the volume computed for the branched would be less close.

Cyclohexane was chosen as standard, because its molecules look like hexagonal tiles.[33] By fitting, a hydrogen radius of 0.170 nm brings the computed volume very close to V_m: calculated for cyclohexane, 0.1773; $M/N_A\rho = 0.1795$ nm^3/molecule. To distinguish the new type of solid from its predecessors, we have used the qualifier 'reduced molecular' (subscript rm), and referred to it as the 'rm-solid'.[32] The term 'expanded van der Waals' is equally descriptive. Note that the 'rm-solid' is smaller than Hermann's 'inacessible volume'. Most frequently, the former is $< V_m$, while the

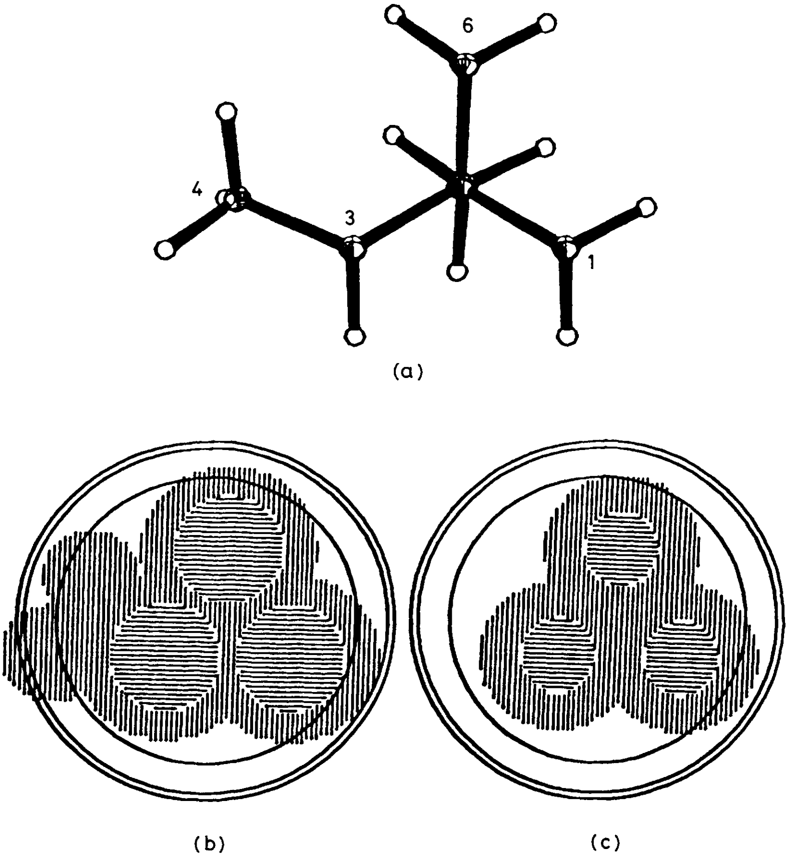


Figure 5 2,2-*Dimethylbutane. (a) ORTEP plot. (b) Cut* 0.15 nm *above centre. (c) Cut* 0.20 nm *above centre. Embedded circular regions are caps of hydrogens on* C-1, C-3, *and* C-6. *The great circles of spheres of volumes equal to* V_w, V_{rm}, *and* V_m *are shown*

latter is $> V_m$. Also, each type of atom in the mantle can be attributed its own r_{rm}, while the radius of the rolling ball is fixed. Hence, the expansion of surface, on going from the w-solid to the rm-solid, is zone-dependent.

Table 4 lists V_{rm}, $f_{rm} = V_{rm}/V_m$ and $G_{rm} = S_{e,rm}/S_{rm}$ for the 20 alkanes enumerated in Table 1. G_w is given for comparison. Several features may be noted. Cyclohexane was chosen as standard, but does not come out as the most compact organic liquid: cycloheptane and the decalins pack even closer, and other cyclics may in time join the list. For cyclic alkanes, $f_{rm} \approx 1$, meaning that $V_{rm} \approx V_m$. This furnishes a direct way to predict their density, on top of multiple correlations that apply to all alkanes (equations 5 and 6 above, 8 and 9 below). For branched alkanes, f_{rm} is lower, in the order of 0.9. With cyclohexane as standard, one might say that 10% of their molar volume is empty 'by molecular fault'. Numbers given

Table 4 *The expanded solids*

Alkane	V_{rm} *(and* f_{rm}*)*	G_{rm}	G_w
(1)	0.1717 (0.89)	0.833	0.796
(2)	0.2028 (0.93)	0.811	0.766
(3)	0.2315 (0.95)	0.787	0.741
(4)	0.2603 (0.96)	0.772	0.717
(5)	0.2891 (0.97)	0.758	0.697
(6)	0.3179 (0.98)	0.740	0.677
(7)	0.1717 (0.89)	0.860	0.807
(8)	0.1969 (0.89)	0.869	0.789
(9)	0.1975 (0.91)	0.857	0.796
(10)	0.2228 (0.93)	0.857	0.776
(11)	0.2501 (0.91)	0.846	0.753
(12)	0.1522 (0.97)	0.867	0.869
(13)	0.1788 (0.96)	0.862	0.833
(14)	0.1773 (0.99)	0.867	0.843
(15)	0.2064 (0.97)	0.855	0.805
(16)	0.2043 (1.01)	0.862	0.819
(17)	0.2309 (0.99)	0.862	0.789
(18)	0.2325 (0.97)	0.850	0.784
(19)	0.2646 (1.03)	0.851	0.786
(20)	0.2677 (1.01)	0.840	0.773

for straight-chain alkanes all refer to the extended conformation. On chain-homologation, G decreases while f increases: as a percentage, less space is wasted as intermolecular gaps. Effects of skewing will be discussed below. For alkanes of all types, the rm-solid is usually more globular than the w-solid, sometimes by two units or more on the scale (N in equation 4a). If the extended conformations of linear alkanes are excluded, G_{rm}-values are quite high, $N \sim 3$. Hence it is not uncalled for to talk of 'molecular radii', $r = (3V_{rm}/4\pi)^{1/3}$.

By cross-sectional areas, rm-solids emulate adsorbed molecules.[33] For example,[14] the calculated section of alkyl chains is about 0.23 nm^2/molecule. As for correlations, V_{rm} does not correlate with v better than V_w,

$$v = 567.25\ V_{rm} + 13.36 \tag{8}$$

(20 alkanes in Table 1, v in cm^3/mol, V_{rm} in nm^3/molecule; $r = 0.9774$, $s = 5.1996$, *cf.* equation 3). Inclusion of G_{rm} actually impairs the correlation, which also becomes worse than with w-descriptors:

$$v = 536.08(40.98)V_{rm} - 48.02(44.88)G_{rm} + 60.38(44.42) \tag{9}$$

($R = 0.9788$, $S = 5.4590$, *cf.* equation 5). Equally, the correlation for R_M is not as good as before:

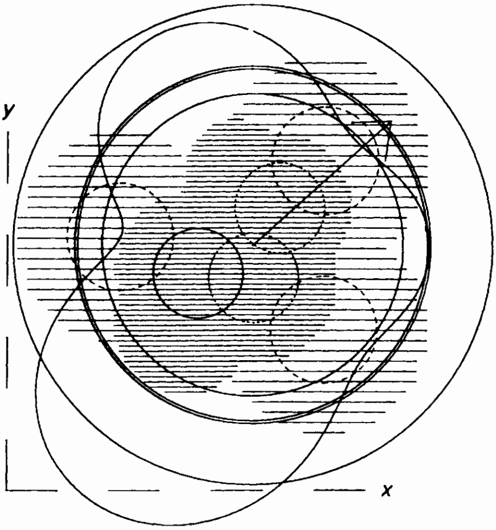
$$R_M = 188.16(4.94)V_{rm} + 39.10(5.41)G_{rm} - 40.07(5.36) \tag{10}$$

($R = 0.9962$, $S = 0.6586$, *cf.* equation 7).

The conclusion is immediate. rm-Solids represent the shape of alkanes infinitely better than w-solids, and computation of their volumes, cross-sections and $r_{e,rm}$-values may replace measurement when required. On the other hand, their attributes do not correlate with molecular and molar properties as well as those of w-solids. It seems that the model of van der Waals core, encased as it is in an envelope of free volume, is more fundamental than the model that appends free volume to the core. The choice of model should therefore be dictated by the end. Coming back to Figure 5, one would say that 2,2-dimethylbutane looks like the rm-solid, and occupies a volume close to V_{rm}. However, if its properties are to be predicted from correlations, regressands such as V_w and G_w are more apposite than V_{rm} and G_{rm}.

6 Heteromolecules

Examination of heteromolecules constitutes, of course, the main thrust of molecular graphics.[2,88] Still, little has been attempted in the quest of size and shape[14,84] beyond the visual inspection of figures. The numbers furnished by the computerized process, if used at all, are not used as such but as regression parameters. Correlations have been reported between calculated volume and measured volume;[41] surface area or volume and solubility;[16,18,19] volume of



Unit: 0.1 nm or 15.0 kJ/mol

Figure 6 1,2,3-*Trichloropropane*, GG(*ga*), rm-*solid, cut through first principal plane. In this and following figures, carbon and chlorine atomic circles are plotted at half their given radius. Dashed circles: atoms in plane of cut; solid: above plane; dotted: below plane; atoms well above or well underneath are not plotted. Axis dashes stand for the unit of measure. See text for other details*

---

[88] T. Liljefors, *J. Mol. Graphics*, 1983, **1**, 111.

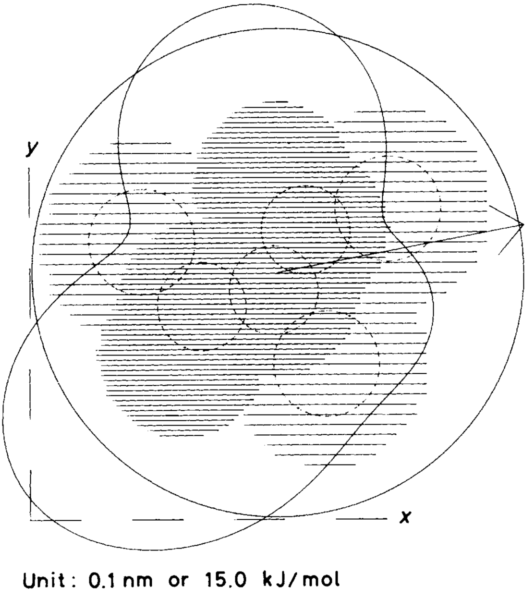


Figure 7 1,2,3-*Trichloropropane*, GA(*ag*), rm-*solid*

delimited portions and steric parameters;[34] surface area and partition coefficients;[17,19] surface and volume and cohesive energy density,[33] and more. The reason, no doubt, is that van der Waals descriptors give small values with respect to the measured cross-sections and volumes. Also, they do not reflect correctly stereochemical changes (see below). In this section we consider descriptors that resemble better the measured attributes.[14,32] Figures 6 and 7 are meant to review some of the points above and lay the background for points to be made below.

Figure 6 is a cut across the first principal plane of the conformer GG(*ga*), the preponderant[89] and most stable[90] nuclear configuration of 1,2,3-trichloropropane (*i.e.*, Cl^1 and Cl^3 *gauche* with respect to CCC, Cl^1 *gauche* and Cl^3 *anti* to Cl^2). Inner circles are great circles of spheres of given volumes: V_w (0.1064 nm^3/molecule), V_{rm} and V_m (almost congruent, 0.1678 and 0.1763). V_{rm} constitutes 95% of V_m, so that the rm-solid represents satisfyingly the average molecular size at 20 °C. The 5% loss in volume is accounted for by surface convolution, which impedes perfect matching of neighbouring molecules. In predicting v, equations 8 and 9—originally derived for unsubstituted alkanes—err by no more than 2—3%. The globularities are $G_w = 0.802$ and $G_{rm} = 0.859$, corresponding to $N = 5$ or 3 by equation 4a. So much for the review.

The outer circle in Figure 6, the curve and the arrow, add information essential to appreciating the space requirements of heteromolecules. The outer circle is a great

---

[89] P. E. Farup and R. Stølevik, *Acta Chem. Scand., A*, 1974, **22**, 871.

[90] Calculated as described in A. Y. Meyer, *J. Mol. Struct.*, 1977, **40**, 127.

circle of the circumscribing sphere, a convenient interfacing[31] of 'inner and outer molecular space'.[91,92] The curve is a plot of the electrostatic energy (W) of a unit of positive charge travelling along the circle. In conformity with common practice, we shall refer to it as the 'molecular electrostatic potential' (MEP). The circle serves as co-ordinate origin for W, so that positive values of the MEP (positive charge repelled) are plotted outwith the circle and negative values (charge attracted) are plotted within. The MEP was traced by the monopole approximation, $W = \Sigma(i)q_i/\bar{r}_i$, using MDR atomic charges.[93] Admittedly, the approximation and graphics are as crude as can be,[94] but they suffice to furnish the information that really counts: a rough distinction between positive and negative zones on the interface, location of extrema, and some idea as to magnitudes.[95] It stands indeed to reason that not only surface topography but also MEP topography affect packing in the liquid. 'Plus avoids plus'[95a] as well as 'bump avoids bump'. The arrow in Figure 6 is the dipole-moment component in the given section (+⟶ convention). The figure also distinguishes the hetero-segment (sparse hatching) from the alkyl segment (dense), the former encompassing those regions of heteroatoms that do not overlap with carbons and hydrogens. Distinction is required since the two types of segment interact with their surroundings in different ways and to different extents,[96] so that they affect measurables separately.[32,33]

To illustrate molecular similarity, Figure 7 shows the second important conformer of 1,2,3-trichloropropane, GA(*ag*). The descriptors of GG(*ga*) and GA(*ag*) differ only slightly, which is not uncommon. For GA(*ag*), one has $V_w = 0.1064$ and $V_{rm} = 0.1683$ nm^3/molecule (not shown in the Figure), $G_w = 0.797$ and $G_{rm} = 0.854$. The MEP is a sensitive indicator of molecular similarity.[97] Denuded, as it is traced here, of almost all detail, it enables the onlooker to compare GG(*ga*) and GA(*ag*) at a glance. So much for the background.

The following procedure to obtain r_{rm}(X) for an heteroatom improves our previous suggestion.[32] Choose as standard a molecule or a set of molecules that contain X, put the radii of mantle atoms other than X equal to their rm-values (previously determined), then fit r_{rm}(X) such that the computed volume be close to the measured V_m. Using equatorial chlorocyclohexane for Cl and equatorial bromocyclohexane for Br, in the same spirit that cyclohexane was used for H, one gets r_{rm}(Cl) = 0.209 and r_{rm}(Br) = 0.219 nm (V_{rm} and V_m 0.1968 and 0.1969 nm^3/molecule for the former, 0.2026 and 0.2027 for the latter). By fitting to $(C_2H_5)_2O$, r_{rm}(L) = 0.228 nm is obtained for the lone electron-pair L on the oxygen. Ether oxygen is not a mantle atom, but carbonylic =O is. Fitting to acetone and to propionic acid yields r_{rm}(=O) of 0.213 nm for ketones and 0.195 for acids,

---

[91] E. Scrocco and J. Tomasi, *Top. Curr. Chem.*, 1973, **42**, 95.

[92] E. Scrocco and J. Tomasi, *Adv. Quantum Chem.*, 1978, **11**, 115.

[93] A. Y. Meyer, *J. Chem. Soc., Perkin Trans. 2*, 1982, 1199.

[94] P. Politzer, K. Jayasuriya, and B. A. Zilles, *J. Am. Chem. Soc.*, 1985, **107**, 121.

[95] For a recent example of using qualitative features of the MEP, see P. R. Andrews, M. Sadek, M. J. Spark, and D. A. Winkler, *J. Med. Chem.*, 1986, **29**, 698.

[95a] G. Náray-Szabó, *J. Mol. Struct.*, 1986, **134**, 401.

[96] M. L. Huggins, *J. Phys. Chem.*, 1970, **74**, 371.

[97] P. E. Bowen-Jenkins, D. L. Cooper, and W. G. Graham, *J. Phys. Chem.*, 1985, **89**, 2195.

both values referring to neat liquids at 20 °C.[14] In this way, one can proceed and obtain rm-radii for other atoms and groupings, provided that molecular geometries are known from measurement or computation.

One pleasant feature of the rm-construction is that it constitutes a step towards the amalgamation of intermolecular and intramolecular non-bonded radii. The intramolecular radii (r^*) are estimated by optimization of molecular-mechanical force fields.[55] It has always been of concern, and a cause for apology, that they come out so much larger than van der Waals radii.[54,98] We now have a set of intermolecular parameters (rm-radii) that also are longer than w-radii, and not discordant at all with r^*. Our r_{rm}-values for H, Cl, and Br (0.170, 0.209, and 0.219) are comparable to the r^*-values of MM2[49,99] (0.150, 0.203, 0.218). Improvements can be envisaged that would bring close other numbers in the two sets. For example, for r^*(=O), MM2 actually incorporates the older MM1 value (0.165), itself raised with respect to an even older estimate[100] (0.154 nm). A detailed examination may yet suggest a further raise and, perhaps, also the distinction of carbonyl types.

There are other satisfying features. In examining partial molar volumes of halogenoalkanes in CCl_4, it has been noticed that halogen atoms behave as if their volumes were smaller than indicated by w-volumes.[101] For one, bromine has about the same V_w as methyl,[58] but its volume increment is considerably smaller. The discord is settled once rm-radii, not w-radii, are adopted as indicators of space requirements. For an isolated methyl group, one calculates V_w of 0.268 nm^3, and hence $r_{e,w} = (3V_w/4\pi)^{1/3} = 0.186$ nm—lower than r_w(Br). By contrast, V_{rm} is 0.487, whence $r_{e,rm} = 0.227$—higher, as it should—than r_{rm}(Br). Also, in the w-solid of bromomethane, the alkyl segment is only slightly larger than the hetero segment: 0.268 *vs.* 0.235 nm^3. In the rm-solid, the difference in bulk is much more pronounced: 0.487 *vs.* 0.322 nm^3.

In Table 5 are assembled data on 20 chlorinated alkanes of variegated types. Included are the volumes of the m-solid (V_m, from measurement), the w-solid, and rm-solid, the latter partitioned into segments. The ratio $f_m = V_{rm}/V_m$ is also listed. Here again, the w-solid is very small, with f_w in the range 0.51—0.62. For cyclic compounds, the rm-solid emulates the m-solid by average volume, and density can be predicted by computation. For other chlorides, f_{rm} varies within the wide range of 0.80—0.93. This is much wider than in unsubstituted alkanes (Table 4).

We believe that the variance is meaningful, reflecting varying amounts of intermolecular gaps 'by molecular fault'. A low f_{rm} does not necessarily imply that computation (V_{rm}) underestimates molecular size. It could well be that the measured number (V_m) includes a high portion of empty space and, hence, is an overestimate. The MEP should be helpful in anticipating cases. If it contains pronounced extrema—negative and positive—close packing is promoted and f_{rm} can be high. This is the case, for example, of 1,2,3-trichloropropane (Figure 6 and 7, $f_{rm} = 0.95$). If extrema are not sharp, molecules are not pulled together as close as

---

[98] N. L. Allinger, *Adv. Phys. Org. Chem.*, 1976, **13**, 1.

[99] A. Y. Meyer, N. L. Allinger, and Y. Yuh, *Israel J. Chem.*, 1980, **20**, 57.

[100] N. L. Allinger, M. T. Tribble, and M. A. Miller, *Tetrahedron*, 1972, **28**, 1173.

[101] F. Shahidi, P. G. Farrell, and J. T. Edward, *J. Phys. Chem.*, 1979, **83**, 419.

Table 5 *Volume of chlorinated alkanes* (nm^3*/molecule*) *averaged over conformers*

Chloride	$V_{\mathrm{rm,alk}}$	$V_{\mathrm{rm,het}}$	f_{rm}	V_{w}	V_{m}
(1) CH_2Cl_2	0.0397	0.0527	0.87	0.0574	0.1063
(2) $CHCl_3$	0.0314	0.0784	0.82	0.0719	0.1336
(3) CCl_4	0.0244	0.1052	0.80	0.0867	0.1602
(4) CH_3CHCl_2	0.0710	0.0518	0.88	0.0748	0.1398
(5) $ClCH_2CH_2Cl$	0.0714	0.0516	0.92	0.0749	0.1330
(6) $CH_3CH_2CH_2Cl$	0.1089	0.0251	0.92	0.0772	0.1464
(7) $CH_3CHClCH_3$	0.1088	0.0252	0.88	0.0770	0.1514
(8) $CH_3CHClCH_2Cl$	0.1009	0.0503	0.93	0.0917	0.1623
(9) $ClCH_2CHClCH_2Cl$	0.0933	0.0748	0.95	0.1064	0.1763
(10) $ClCH_2CH(CH_3)_2$	0.1371	0.0244	0.93	0.0939	0.1745
(11) $CH_3CHClCH_2CH_3$	0.1377	0.0243	0.92	0.0939	0.1760
(12) $(CH_3)_3CCl$	0.1367	0.0245	0.88	0.0937	0.1826
(13) $CH_3CHClCHClCH_3$ (*meso*)	0.1302	0.0482	0.93	0.1085	0.1913
(14) $ClCH_2CCl(CH_3)_2$	0.1293	0.0484	0.92	0.1083	0.1930
(15) C_5H_9Cl	0.1445	0.0248	0.98	0.1002	0.1728
(16) $C_6H_{11}Cl$	0.1721	0.0244	1.00	0.1164	0.1970
(17) $(CH_3)_2CClCH(CH_3)_2$	0.1913	0.0224	0.94	0.1273	0.2281
(18) 1,2-$C_6H_{10}Cl_2$ (*cis*)	0.1646	0.0482	1.01	0.1310	0.2114
(19) 1,2-$C_6H_{10}Cl_2$ (*trans*)	0.1651	0.0475	1.01	0.1310	0.2145
(20) 1,4-$C_6H_{10}Cl_2$ (*cis*)	0.1647	0.0483	1.00	0.1310	0.2136

their shape permits, and, on the average, a comparatively large fraction of V_{m} stays empty. Consider in Figures 8 and 9 cuts through the two chlorides of lowest f_{rm}, CCl_4 and $CHCl_3$. Indeed, the MEP of CCl_4 has only feeble extrema (note the scale), and in $CHCl_3$ the minima are weak and shallow. This contrasts with unfunctionalized alkanes, where MEPs are much less heterogenous.

Despite the variation in f_{rm}, computed attributes of chloroalkanes correlate with molar volumes no worse than for unsubstituted alkanes. Equations 11 and 12, based on the 20 chlorides of Table 5, may be compared with equation 9.

$$v = 513.24\, V_{\mathrm{rm}} + 21.64 \qquad (11)$$

($r = 0.9791$, $s = 3.8190$)

$$v = 523.28(26.33) V_{\mathrm{rm,alk}} + 574.16(58.10) V_{\mathrm{rm,het}} + 17.73(5.31) \qquad (12)$$

($R = 0.9807$, $S = 3.9874$). In equation 12, $V_{\mathrm{rm,alk}}$ and $V_{\mathrm{rm,het}}$ are, respectively, the volumes of the alkyl- and of the hetero-segments.

In brief, it is possible to calculate the size of chloroalkanes, and even obtain an approximation to their densities. Extension of the procedure to other classes of heteromolecules has not been attempted as yet, but is expected to be feasible. Numbers computed for chloroalkanes (V_{rm}) differ a bit from those estimated from densities (V_{m}). It would indeed be of interest to check whether V_{rm} performs better than V_{m} in calculations that require the molecular volume as input, *e.g.*, solvation energies.[5,102]

[102] A. Y. Meyer, *J. Comput. Chem.*, 1981, **2**, 384.

7 Theory-dependent Estimates

Theoretical models for various quantities lead, by mathematical derivation, to equations that contain the molecular radius.[4] As stated in the Introduction, the need to solve such equations constitutes an incentive to estimate molecular size. Computation of solvation energies is a case in point.[5] It also is an illustration of the uncertainties involved when V_m has to be estimated from experimental data for solutes that are unstable or not liquid when pure.[103] Yet, if all other magnitudes in a given equation are determined independently, the equation can be turned to yield the molecular radius. Such procedures are occasionally encountered in the literature, and two instances will be cited here.

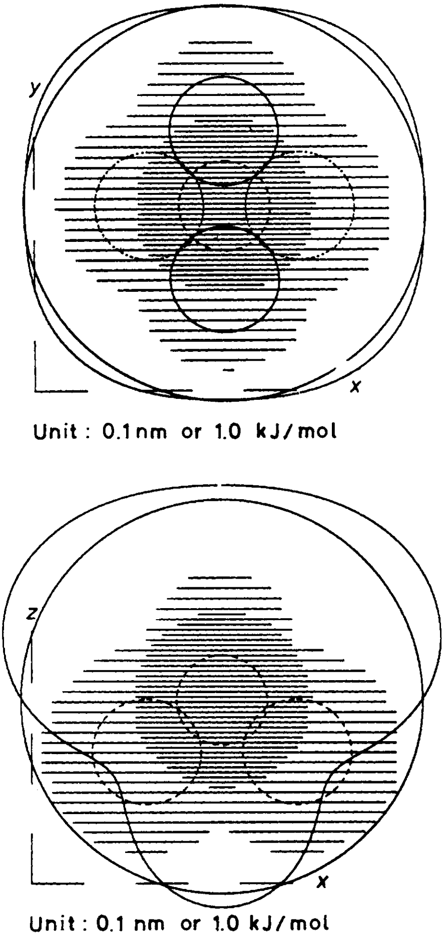


Figure 8 *Two cuts through tetrachloromethane, rm-solid*

[103] A. Y. Meyer, *J. Mol. Struct.*, 1983, **94**, 95.

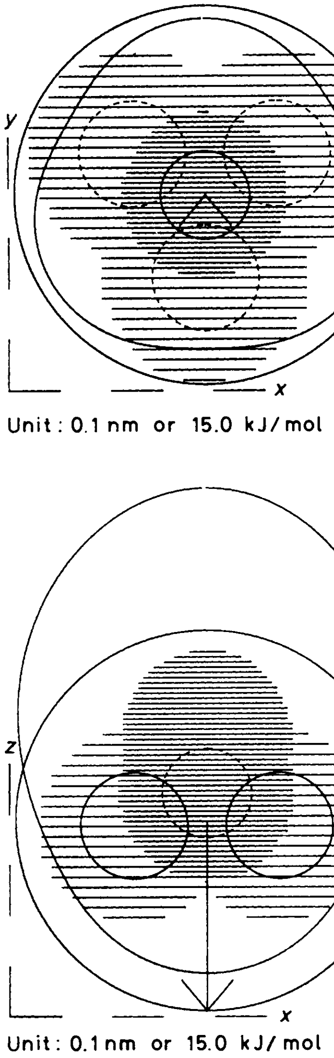


Figure 9 *Two cuts through trichloromethane, rm-solid*

The first dates back to the 1940s. Theoretical derivation[104] relates a function of the refractive index n and the density of a compound to the ratio $\alpha/(1 - f\alpha)$, where α is the molecular polarizability and $f = 2(n^2 - 1)/r^3(2n^2 + 1)$. Measurement of n at various temperatures yields corresponding values for $\alpha/(1 - f\alpha)$. When $(1 - f\alpha)/\alpha$ is plotted against $2(n^2 - 1)/(2n^2 + 1)$, r (as well as α) can be obtained from the plot. For CCl_4, the procedure produced[104] $r = 0.242$ nm, smaller than the

[104] C. J. F. Böttcher, *Rec. Trav. Chim. Pays-Bas*, 1943, **62**, 325.

equivalent radius of the w-solid, $r_{e,w} = (3V_w/4\pi)^{1/3} = 0.275$ nm. For other molecules, the derived r falls between $r_{e,w}$ and $r_{e,rm}$. Thus,[105] a radius of 0.262 nm was found for $CHCl_3$ ($r_{e,w} = 0.258$, $r_{e,rm} = 0.297$), 0.274 for C_2H_5Br (0.252, 0.298), and 0.289 for CH_3COCH_3 (0.250, 0.308).

The other case we cite is more recent. Theoretical derivation[106] equates the halfwidth δ of an infrared band to $\delta_0 + kT/4\pi^2cfr^3\eta$; η is the viscosity coefficient of the medium, and f takes the value 0.163 for a pure liquid. By plotting δ against T/η at various temperatures, the radius of neat CH_2Cl_2 was obtained as 0.220 nm ($r_{e,w} = 0.239$). The technique can be applied to dissolved molecules, but evaluation of f requires then an estimate of the solvent-to-solute radius ratio.

8 Conformational Change

> *'The problem of the failure of van der Waals radii to indicate the effects of atoms or groups on conformational equilibria is an old one.'*[101]

Conformational change is attended by a change in molar volume[9] (Δv). This change has been partitioned in three components:[8] change in the intrinsic volume of the molecules (Δv_{int}); change due to solvent effects and the different packing of conformers in the liquid[107] (Δv_p); and change due to modification in the electrostatic interaction between each molecule and its surroundings (Δv_{es}). In dealing with alkanes, Δv_{es} can be ignored. This is because its formulation contains as multiplicand the diffrence between the squared dipole moments,[8] and dipole moments of alkanes are very small. Δv_{es} can also be ignored in monofunctionalized alkanes, where conformation has little effect on the overall dipole moment, and in centrosymmetric polyfunctionals.

In alkanes, the *anti-to-gauche* transition is usually accompanied by contraction and a rise in energy.[9,108–110] Butane is a prototype.[111,112] When it transforms from *a* to *g*, two *trans*-interactions and two *skew*-interactions disappear (Me/Me-*a*, H/H-*a*, 2Me/H-*g*), and two *trans*-interactions and two *skew*-interactions appear (2Me/H-*a*, Me/Me-*g*, H/H-*g*). Whether computation mimics the contraction depends—strange to realize and seemingly unrealized heretofore—on the choice of atomic radii. If one attributes to hydrogen the comparatively small w-radius, almost no change in volume is computed for *a* → *g*. If anything, *gauche* incorrectly comes out more sizeable than *anti.*[113] This is because non-geminal hydrogens

[105] T. G. Scholte, *Physica*, 1949, **15**, 450.

[106] K. Tanabe, *Spectrochim. Acta, Part A*, 1976, **32**, 1129.

[107] L. R. Pratt, C. S. Hsu, and D. Chandler, *J. Chem. Phys.*, 1978, **68**, 4202.

[108] One tends, though, to label *'gauche'* the putatively less stable conformer. See Y. Morino, *J. Mol. Struct.*, 1985, **126**, 1.

[109] L. S. Bartell, *J. Am. Chem. Soc.*, 1977, **99**, 3279.

[110] J. T. Edward, P. G. Farrell, and F. Shahidi, *J. Phys. Chem.*, 1978, **82**, 2310.

[111] A. L. Verma, W. F. Murphy, and H. J. Bernstein, *J. Chem. Phys.*, 1974, **60**, 1540.

[112] J. Devaure and J. Lascombe, *Nouv. J. Chim.*, 1979, **3**, 579.

[113] A. Raudino, F. Zuccarello, and G. Buemi, *J. Chem. Soc., Faraday Trans. 2*, 1981, **77**, 1331.

overlap neither in *anti* nor in *gauche*, and the C^1/C^4-overlap which occurs only in *gauche* is more than offset by decreased C^1/C^3-, C^2/C^3-, and C^2/C^4-overlaps. That these three are lower in *gauche* than in *anti* has to do with the non-rigidity of internal rotation.[114] When the dihedral $C^1C^2C^3C^4$-angle (ω) closes along the $a \rightarrow g$ pathway, CCC and other bond-angles open up, and the bond C^2–C^3 stretches. Relaxed *g*-butane is therefore not as constricted as one might anticipate.

Opening of the C^1C^4-distance (d) in *g*-butane is worthy of numerical demonstration. The MM2-geometry of *a*-butane ($\omega = 180°$) has $r(C^1C^2) = 0.1534$ nm, $r(C^2C^3) = 0.1536$, $\theta(C^1C^2C^3) = 111.8°$ and $d(C^1 \cdots C^4) = 0.3910$ nm. If the internal rotation were rigid, d would be[115] 0.3641, 0.3033, and 0.2678 nm at $\omega = 120°$, 60° and 0°, respectively, so that

$$d(\text{rigid}) = 0.3323 - 0.0613\cos\omega - 0.0029\cos2\omega - 0.0003\cos3\omega.$$

Actual optimization yields $d = 0.3633$, 0.3144, and 0.2911 at $\omega = 116.0$, 64.7, and 0°, respectively, whence

$$d(\text{relaxed}) = 0.3395 - 0.0519\cos\omega + 0.0015\cos2\omega + 0.0020\cos3\omega$$

and

$$\Delta d = d(\text{relaxed}) - d(\text{rigid}) = 0.0072 + 0.0094\cos\omega + 0.0044\cos2\omega + 0.0022\cos3\omega.$$

Functions of the form[116] $y = A + B\cos\chi + C\cos2\chi + D\cos3\chi$ have extrema at $\sin\chi = 0$ and at $\cos\chi = \{-C \pm [C^2 + 3D(D - B)]^{1/2}\}/6D$. Δd is maximal at the barriers (0°, 115.3°) and minimal at 103.2° and 180°. At $\omega = 64.7°$, it amounts to 0.0062 nm.

The alternative to using w-solids is to assume a large radius for hydrogen, *e.g.*, $r_{rm}(H) = 0.170$ nm. When this is done, more and larger overlaps are generated in *gauche*, which now comes out smaller than *anti*. Thus, for butane ($a \rightarrow g$), $\Delta V_w = +5 \times 10^{-6}$ nm^3/molecule (positive!) but $\Delta V_{rm} = -1.005 \times 10^{-3}$ (negative). As another example,[117] $\Delta V_w = +2.6 \times 10^{-5}$ but $\Delta V_{rm} = -0.74 \times 10^{-3}$ nm^3/molecule for chlorocylohexane ($e \rightarrow a$). In both examples, and many others, w-solids confer on Δv_{int} a sign opposite to that of Δv (positive rather than negative). With rm-solids, Δv_{int} has the same sign as Δv, and also is on the same order of magnitude. A few other examples are given in Table 6, including one (1,2-dichloroethane) where Δv_{es} should contribute to Δv.

The investigator seems thus to have two options. Either lend credit to van der Waals radii, conclude that $\Delta v_{int}(a \rightarrow g) \sim 0$ or >0, and attribute the entire volume change—or, even, more than the entire volume change in butane and related molecules—to packing and solvent effects. Alternatively, admit large radii for mantle atoms, obtain $\Delta v_{int} < 0$ but still $|\Delta v_{int}| < |\Delta v|$, and attribute to Δv_p only some of the volume change. Now, it is true that conformers cannot be examined

---

[114] N. L. Allinger, L. Schäfer, K. Siam, V. J. Klimkowski, and C. Van Alsenoy, *J. Comput. Chem.*, 1985, **6**, 331.

[115] A. K. Ghose and G. M. Crippen, *J. Comput. Chem.*, 1985, **6**, 350.

[116] L. Radom, P. J. Stiles, and M. A. Vincent, *J. Mol. Struct.*, 1978, **48**, 259.

[117] R. A. Ford and N. L. Allinger, *J. Org. Chem.*, 1970, **35**, 3178.

Table 6 *Calculated volume-changes* (cm^3/mol)

Transformation	$N_0\Delta V_w$	*Corr.*[a]	$N_0\Delta V_{rm}$	*Meas.*[b]
(1) Butane ($a \longrightarrow g$)	+0.00	−2.15	−0.61	−1.0
(2) 2-Methylbutane ($a \longrightarrow g$)	+0.01	−2.15	−0.45	−0.9
(3) 2,3-Dimethylbutane ($a \longrightarrow g$)	+0.08	+0.78	−0.36	−1.3
(4) Hexane ($aaa \longrightarrow aag$)	+0.05	−0.32	−0.58	−1.8
(5) Heptane ($aaaa \longrightarrow aaga$)	+0.05	−1.76	−0.71	−1.3
(6) 1-Chloro-2-methylpropane[c] ($P_C \longrightarrow P_H$)	+0.02		−0.21	+0.5
(7) 1-Bromo-2-methylpropane[c] ($P_C \longrightarrow P_H$)	+0.00		−0.33	−1.7
(8) 2-Chlorobutane[c] ($S_{HH} \longrightarrow S_{HH'}$)	−0.03		−0.47	−1.5
($S_{HH} \longrightarrow S_{CH}$)	−0.00		−0.29	−0.8
(9) 2-Bromobutane[c] ($S_{HH} \longrightarrow S_{HH'}$)	−0.01		−0.46	+2.0
($S_{HH} \longrightarrow S_{CH}$)	+0.06		−0.14	+0.6
(10) 1,2-Dichloroethane[d] ($a \longrightarrow g$)	−0.00		−0.16	−3.5
(11) Chlorocyclohexane[e] ($e \longrightarrow a$)	+0.02		−0.45	−1.9
(12) *trans*-1,4-Dichlorocyclohexane[e] ($ee \longrightarrow aa$)	+0.01		−0.93	−3.8

[a] By correlation in equation 6. [b] Data for neat liquids at 25 °C from compilation (Ref. 115) unless marked otherwise. [c] Ref. 129. [d] In hexane, 21 °C. [e] In CS_2, at 50 °C

individually, but configurational stereoisomers can. For pairs of these, calculations based on van der Waals radii have invariably failed to single out the denser species.[31] For example, the volume of *cis* comes out equal or somewhat larger than that of *trans* in 1,2-dimethylcyclohexane, decalin (Table 1) and 1,2-dichlorocyclohexane (Table 5). This reversal—which, incidentally, gainsays the conformational rule[31,118]—does not occur with rm-solids (Tables 4 and 5). Hence, the second choice seems more reasonable.

Notwithstanding, it is not inconceivable that Δv_p outweigh Δv_{int}.[119] Two instances in Table 6 could be of this type. For $XCH_2CH(CH_3)_2$, Δv is reported[120] to be positive for X = Cl and negative for X = Br (conformer P_C has X *gauche* to one Me and *anti* to the other, while P_H has it *gauche* to both). Clearly, whatever set of radii is endorsed, $\Delta v_{int}(P_C \rightarrow P_H)$ cannot have different signs in the chloro-and the bromo-compound. By elimination, Δv_p has to be called upon to differentiate the expansion in the former from the contraction in the latter. The same goes for $CH_3CHXCH_2CH_3$, where pressure experiments[120] identify the conformer S_{HH} as the most voluminous for X = Cl and least voluminous for X = Br (S_{HH} is Me/Me-*a*, $S_{HH'}$ has Me/Me-*g* and Me/X-*g*, S_{CH} has Me/Me-*g* and Me/X-*a*).

9 Conclusion

Different contexts require, or lead to, different estimates of molecular size. The differing estimates correspond to different representations of molecular shape.

---

[118] N. L. Allinger, *Experientia*, 1954, **10**, 328.

[119] T. Bleha, *Polymer*, 1985, **26**, 1638.

[120] H. Nomura, Y. Udagawa, and K. Murasawa, *J. Mol. Struct.*, 1985, **126**, 229.

Judging by available numbers, correlative studies that disregard conformational multiplicity are best based on van der Waals attributes (Section 4). These are obtained by assuming atoms spherical, assigning them van der Waals radii, and mensurating the molecular bodies thus assembled (Section 3). Usage of a set of larger radii for mantle atoms provides a more faithful estimate of the average volume that a molecule occupies in the liquid (Sections 5, 6). It also provides a more acceptable prediction of the sense of volume change in stereochemical transitions (Section 8). Radii in this set are obtained by fitting computed densities to the experimental, and turn out to be similar to the intramolecular non-bonded radii of molecular mechanics. They can serve to estimate non-observables, like the average surface area of a molecular zone in a liquid. By equating the computed volume to the volume of a sphere, a radius can be computed for the molecule itself. These radii are very close to, but somewhat smaller than values currently used as input in calculating liquids (Section 2). They are larger than values calculated on the assumption of molecular sphericity (Section 7).

Acknowledgement. I am greatly indebted to Professor Y. Marcus, of the Hebrew University, Jerusalem, for a critical reading and an in-depth discussion of the draft of this manuscript.