Prediction of Density in Organic Crystals

By Attilio Immirzi and Bruno Perini

Istituto di Chimica delle Macromolecole del CNR, Via A. Corti 12, 20133 Milano, Italy

The possibility is explored of expressing the specific volume of organic substances in the crystal state as the sum of increments, including corrective terms. The unitary volumes listed herein allow an estimate of density with a confidence limit of 2–3%.

Introduction

Kitaigorodsky (1961, 1973) showed how the van der Waals molecular volume (hereafter termed 'molecular volume') can be calculated if the shape of the molecule and the intermolecular atomic radii are known. He also proved that these volumes can be expressed as sums of increments each depending on the nature of individual atoms and on their coordination numbers.

In molecular crystals, the actual volume available to each molecule ('crystal volume'), which is the sum of the molecular volume and the empty spaces, will not necessarily be formed by additive quantities. If the ratio molecular-volume/crystal-volume (the 'packing coefficient' according Kitaigorodsky) were constant, additivity would certainly take place. Actually the packing coefficients assume scattered values in the range 0·65–0·77 (Kitaigorodsky, 1973, p. 18); yet that does not necessarily imply non-additivity of the crystal volume, since different atomic groups might have different packing efficiencies. Moreover, introduction into the sum of proper corrective terms might improve the additivity. Our aim is to verify just this hypothesis, not on theoretical but on empirical grounds, from experimental crystallographic densities, thousands of which are today available in the literature and also in the form of 'crystal data files' (Kennard, Watson, Allen, Motherwell, Town & Rogers, 1975) recorded on magnetic tapes and manageable by computers.

Before undertaking a work based on many thousands of data, which implies also resolving non-trivial programming problems, we decided to carry out some preliminary calculations based upon a moderate number of selected data and without the use of tape-recorded files.

Evaluation of unitary volumes

The additivity of crystal volume is based on the relation:

$$Vc_i = \sum_{j=1}^{p} m_{ij} v_j$$

where Vc_i is the calculated crystal volume for a single molecule, to be compared with the experimental value Ve_i obtained from the crystal density, and v_j are p unit volumes (or increments) associated with atomic species, coordination number and ionic charge (or possibly with groups of atoms) and m_{ij} are the relative stoichiometric multiplicities.

We have also introduced some corrective terms using again the proper stoichiometric coefficient. For the time being we have considered: (i) a correction for hydrogen-bond formation, distinguishing hydrogen-bonds of type O–H $\cdots$ O, N–H $\cdots$ O and N–H $\cdots$ N; (ii) a correction for ring formation; (iii) a correction for ring fusion.

Data for 500 organic crystalline compounds have been selected from the collection of Wyckoff (1971) with molecular weight ranging from 50 to 1000 (see histogram, Fig. 1) and with the following restrictions: (i) compounds not solid at room temperature, or having structural disorder were excluded; (ii) compounds containing molecules of solvent were excluded, except for water for which a special increment was introduced; (iii) only the elements H, C, O, N, S, F, Cl, Br, I, Na, K and Rb were considered; (iv) cyclic compounds have been restricted to benzene and naphthalene derivatives in order to consider a single case of ring formation and a single case of ring fusion; (v) the O $\cdots$ O, N $\cdots$ O and N $\cdots$ O distances taken as indicative of H-bond formation are 2·7–2·8 Å.

Coordination numbers 4, 3 and 2 were considered for C, 2 and 1 for O, and 3, 2 and 1 for N; for S a unique value has been used owing to the scanty observations available. For Cl, Br and I neutral atoms and anions were distinguished, whilst for Na, K and Rb only the corresponding cations were considered.

The unitary volumes were evaluated from the condition:

$$\sum_{i=1}^{N} (Ve_i - Vc_i)^2 = \text{minimum} .$$

Considering for Vc_i the above expression, the p v_j's were obtained as terms of a p-order vector $\mathbf{v}$ given by $\mathbf{v} = \mathbf{A}^{-1} . \mathbf{b}$ where $\mathbf{A}$ and $\mathbf{b}$ are respectively a $p \times p$ matrix and a p-order vector defined as:

$$A_{kl} = \sum_i m_{ik} m_{kl} \quad b_k = \sum_i m_{ik} Ve_i .$$

The resulting unitary volumes are reported in Table 1, while Fig. 1 shows the distribution of the relative deviations $\Delta\% = 100(Ve_i - Vc_i)/Ve_i$. The cumulative frequencies, f, are plotted as $erf^{-1}(2f-1)$ to check for discrepancy compared with the Gaussian distribution. The plot, except at the boundaries, is almost linear, indicating that the distribution is just a Gaussian.
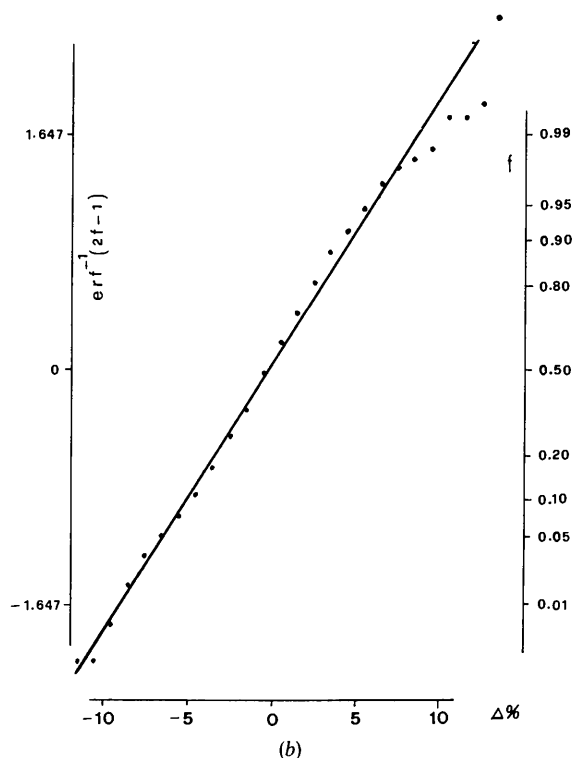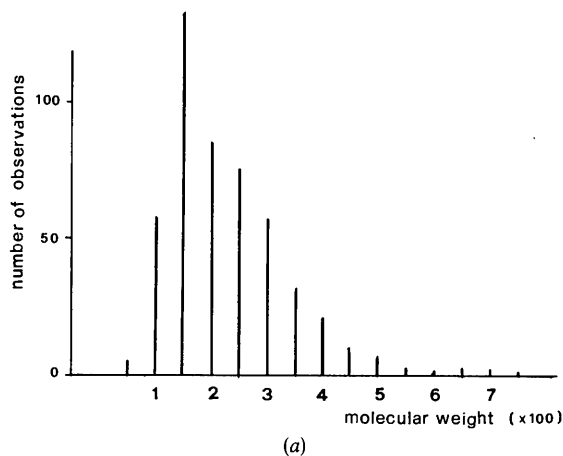

(a)

(b)

Fig. 1. (a) Histogram of the observation vs molecular weight; (b) plot of the distribution of the relative deviations $\Delta\% = 100(Ve_i - Vc_i)/Ve_i$. The cumulative frequencies f are plotted as $erf^{-1}(2f-1)$; on these scales a Gaussian distribution corresponds to a straight line.

Table 1. Volume increments (v) for common elements and ions $(Å^3)$

For C, N and O different coordination numbers are considered. Standard errors (σ) and numbers of contributors (N) are also given.

	v	σ	N		v	σ	N
—H	6·9	0·4	5228	—F	12·8	1·5	14
=C= }	15·3	0·7	74	—Cl	26·7	0·5	134
—C≡ }				—Br	33·0	0·5	120
—C≤	13·7	0·6	453	—I	45·0	1·3	26
>C<	11·0	0·9	1165	Cl⁻	28·9	1·5	39
=O	14·0	0·5	649	Br⁻	39·3	1·5	20
—O—	9·2	0·5	468	I⁻	56·6	2·5	11
N≡	16·0	1·3	30	Na⁺	13·6	2·2	16
N≤	12·8	0·8	68	K⁺	27·3	1·6	32
N≤	7·2	0·8	354	Rb⁺	34·1	2·2	15
S	23·8	0·9	92	H₂O	21·5	0·8	68

Benzene frame (carbons only)	75·2	2·5	443
O–H···O hydrogen-bond	−2·6	0·7	206
N–H···O hydrogen-bond	−2·8	0·5	152
N–H···N hydrogen-bond	−0·3	1·7	11

Non-aromatic rings (rough estimation)	−3·0

Comments

These preliminary results, in spite of the scanty number of data processed, enable us to make some comments. (i) The increments for C, N and O decrease as the coordination number increases, as expected; accordingly the increments for the halogen anions are larger than for the corresponding neutral atoms. (ii) The volume of the benzene frame is less than six times the tri-coordinated C by 7·0 $Å^3$. For saturated rings the correction should be much lower since the ring frame and branched bonds are anything but planar, with intermediate values for more or less flat systems. As shown below we were able to estimate a mean correction of $-3·0$ $Å^3$ for non-aromatic rings. (iii) The volume of the naphthalene frame is less than ten times the tri-coordinated C by 13·3 $Å^3$; the correction of 2×7 $Å^3$ for the presence of two aromatic rings seems thus to be more than sufficient and the ring fusion seems to have a negligible effect. (iv) The corrections for hydrogen-bond formation are, as expected, substractive; the strong N–H···O and O–H···O bonds cause a large contraction, while the weak N–H···N bond gives a negligible correction. (v) Caution must be used in attributing physical meaning to the calculated increments owing to the

Table 2. Comparison between unitary crystal volumes $(c.v.)$ and Kitaigorodsky's (1970, p. 20) molecular volumes $(m.v.)$ $(Å^3)$

	m.v.	c.v.	Ratio		m.v.	c.v.	Ratio
—H	2·0	6·9	0·29	—N	10·0	16·0	0·62
—C≡	5·9	15·3	0·38	—F	9·6	12·8	0·75
—C≤	8·4	13·7	0·61	—Cl	19·9	26·7	0·75
>C<	5·0	11·0	0·45	—Br	26·6	33·0	0·80
=O	13·1	14·0	0·94	≧CH	11·1	17·9	0·62
—O—	11·8	9·2	1·28	≫CH	14·7	22·2	0·67
—NH₂	19·7	21·0	0·94	>CH₂	17·1	24·8	0·69
—NO₂	23·0	25·6	0·90	—CH₃	23·5	31·7	0·74

Table 3. *Observed and calculated crystal volumes* ($\text{\AA}^3$)
for 53 crystalline compounds

A correction of $-7\cdot0$ for aromatic and of $-3\cdot0$ $\text{\AA}^3$ for non-aromatic five or six-membered rings was introduced. The hydrogen-bond correction was introduced only for H atoms belonging to $-CONH_2$ or $-CONH-$ groups ($-2\cdot8$) and to $-COOH$ groups ($-2\cdot6$) [from *Acta Crystallographica*, (1975), B31, indicated page numbers].

No.	Page	Formula	Ve	Ve-Vc	Δ%
1	335	$C_{20}H_{26}O_2$	440·0	−4·2	0·95
2	338	$C_6H_6O_2.CON_2H_4$	207·5	8·7	4·32
3	359	$C_{20}H_{12}O_5.C_3H_6O$	467·6	−0·5	0·11
4	379	$C_8H_2O_4N_2.6H_2O$	317·5	−4·3	1·40
5	384	$C_6H_2O_4F_2$	157·4	1·6	0·99
6	387	$C_6H_3N_2O_5Cl$	206·8	4·5	2·20
7	391	$C_6H_4NO_3^-.K^+.\frac{1}{2}H_2O$	185·9	−4·4	2·20
8	394	$C_5H_6O_3$	129·5	−4·0	3·10
9	418	$C_{17}H_{16}$	319·0	1·8	0·56
10	450	$C_{21}H_{24}O_2$	413·2	−19·7	4·78
11	454	$C_{13}H_{12}N_4O_2.2C_{10}H_8O_2.3H_2O$	753·3	−0·5	0·07
12	462	$C_{10}H_{20}N_2O_4S_2.2HCl$	423·4	−3·0	0·70
13	472	$C_8H_{15}N_2O_2$	244·2	11·2	4·60
14	474	$C_8H_{15}NO_6$	253·6	−0·6	0·24
15	489	$C_6H_6N_4O_2.HCl.H_2O$	228·0	5·0	2·19
16	493	$C_{18}H_{22}O_2$	364·1	−9·6	2·64
17	496	$C_{17}H_{22}N_2O_3$	383·5	−2·9	0·75
18	522	$C_{27}H_{46}$	557·6	−36·5	6·55
19	526	$C_{29}H_{50}$	614·1	−29·2	4·76
20	548	$C_{16}H_{20}N_4O_6$	447·5	29·0	6·48
21	569	$C_{16}H_{27}O_5N$	431·9	6·5	1·51
22	587	$C_{16}H_{16}N_4$	333·4	7·0	2·10
23	601	$C_6H_{13}NO_2$	182·8	0·7	0·40
24	637	$C_{21}H_{30}O_2.C_6H_6O_2$	590·6	−0·1	0·02
25	653	$C_{30}H_{22}O_2N_2Br_2$	609·1	2·0	0·33
26	658	$C_9H_{17}N_3O_4.\frac{1}{2}H_2O$	302·1	3·6	1·21
27	685	$C_{16}H_{22}O_3N_2$	395·4	−3·0	0·76
28	697	$C_6H_{10}O_7.H_2O$	220·1	−1·9	0·84
29	708	$C_{17}H_{10}N_2Cl_2$	375·0	1·6	0·43
30	713	$C_{23}H_{35}O_4N_5S$	619·4	9·8	1·59
31	738	$(C_6H_6N)_2.C_6H_6N_3O_6.\frac{1}{2}C_6H_6$	515·9	9·0	1·76
32	750	$C_{16}H_{10}N_3O_3Cl$	354·0	−8·8	2·49
33	790	$C_{22}H_{31}O_2N$	469·7	−22·4	4·28
34	796	$C_{24}H_{31}O_6F$	571·4	7·4	1·30
35	816	$C_6H_{12}N_2O_4S_2.2HBr.2H_2O$	385·0	−2·5	0·65
36	819	$C_6H_{12}Cl_2$	333·3	−7·6	2·30
37	826	$C_{18}H_{21}N_4O_2Cl$	436·1	0·7	0·16
38	831	$C_{20}H_{14}O_4$	404·4	1·0	0·25
39	831	$C_{20}H_{14}O_4$	412·9	9·5	2·31
40	835	$C_7H_{15}NO$	215·1	14·9	6·90
41	837	$C_9H_{19}NO$	262·9	13·5	5·10
42	857	$C_{10}H_{20}N_2O_3S_2$	337·0	−3·0	0·88
43	862	$C_8H_{15}NO_7$	271·0	−0·0	0·00
44	866	$C_7H_{13}NO_2$	195·9	2·7	1·40
45	882	$C_4H_{13}NO_2^+.C_8H_{11}N_2O_3^-$	385·3	21·8	5·67
46	882	$C_4H_{12}NO^+.C_8H_{11}N_2O_3^-$	364·4	6·0	1·65
47	893	$C_4H_{12}NS_2^+.Br^-.H_2O$	251·0	3·9	1·54
48	903	$C_{10}H_{15}BrO$	256·1	0·6	0·22
49	904	$C_{10}H_{14}Br_2O$	289·8	8·1	2·81
50	919	$C_{14}H_{20}NCl$	322·9	−6·5	2·02
51	921	$C_{15}H_{10}O_5$	299·0	10·3	3·44
52	926	$C_{15}H_{13}O_2F$	315·3	6·8	2·15
53	928	$C_{23}H_{29}ClO_4$	533·3	−1·6	0·31

correlation effects among them. The set of compounds considered should include substances of scattered

composition. As an example of total degeneracy, if only hydrocarbons of compositions C_nH_{2n} were considered, the increments v_C and v_H would be completely meaningless and only v_{C+2H} would be meaningful.

In Table 2 the calculated crystal unitary volumes are compared with the Kitaigorodsky molecular volumes. The ratio between them turns out to be rather variable and smaller for the smaller volumes. A possible interpretation is that the excess of volume, *i.e.* the volume of empty spaces, tends to distribute itself equally among all atoms, giving rise to volume increases proportionally larger for the smaller atoms.

In order to check the reliability of the calculated increments, we examined a random sample of organic crystals. We considered the 53 organic compounds having composition compatible with our regression which were published in two recent issues of *Acta Crystallographica* [(1975), B31, February and March]; this list includes compounds containing five and six-membered rings of any kind. The calculation of volumes was first done by extending the aromatic ring correction $-7\cdot0$ $\text{\AA}^3$ to all types of rings; this gave a r.m.s. value for the relative deviation $\Delta\% = 100(Ve_i - Vc_i)/Ve_i$ of $3\cdot6\%$. Nevertheless the mean value of $Ve_i - Vc_i$ was positive, probably as a consequence of an overestimation of the ring correction. A value of $-3\cdot0$ $\text{\AA}^3$ for non-aromatic rings turns out to be more appropriate: it gives the mean of $Ve_i - Vc_i = 0$ and also lowers the r.m.s. to $2\cdot8\%$. Table 3 reports the volume calculation for these 53 compounds with the adjusted ring correction of $-3\cdot0$ $\text{\AA}^3$ for non-aromatic rings.

Although the opportunity of a large-scale analysis based on thousands of data and on a more sophisticated regression formula remains valid, we believe that even the present increment table, albeit preliminary, allows a quick, simple and reliable prediction of density which is of some practical interest.

References

KENNARD, O., WATSON, D., ALLEN, F., MOTHERWELL, W., TOWN, W. & ROGERS, J. (1975). *Chem. Brit.* 11, 213.
KITAIGORODSKY, A. I. (1961). *Organic Chemical Crystallography*. New York: Consultants Bureau.
KITAIGORODSKY, A. I. (1973). *Molecular Crystals and Molecules*. New York: Academic Press.
WYCKOFF, R. W. G. (1971). *Crystal Structures*, Vols. 5 and 6. New York: Interscience.