

## Fast Calculation of van der Waals Volume as a Sum of Atomic and Bond Contributions and Its Application to Drug Compounds

Yuan H. Zhao, Michael H. Abraham,\* and Andreas M. Zissimos†

Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, U.K., and  
ID Business Solution Ltd, 2 Occam Court, Surrey Research Park, Guildford, Surrey, GU2 7QB, UK

m.h.abraham@ucl.ac.uk

Received June 11, 2003

The van der Waals volume is a widely used descriptor in modeling physicochemical properties. However, the calculation of the van der Waals volume ( $V_{\text{vdW}}$ ) is rather time-consuming, from Bondi group contributions, for a large data set. A new method for calculating van der Waals volume has been developed, based on Bondi radii. The method, termed Atomic and Bond Contributions of van der Waals volume (VABC), is very simple and fast. The only information needed for calculating VABC is atomic contributions and the number of atoms, bonds, and rings. Then, the van der Waals volume ( $\text{\AA}^3/\text{molecule}$ ) can be calculated from the following formula:  $V_{\text{vdW}} = \sum \text{all atom contributions} - 5.92N_{\text{B}} - 14.7R_{\text{A}} - 3.8R_{\text{NR}}$  ( $N_{\text{B}}$  is the number of bonds,  $R_{\text{A}}$  is the number of aromatic rings, and  $R_{\text{NR}}$  is the number of nonaromatic rings). The number of bonds present ( $N_{\text{B}}$ ) can be simply calculated by  $N_{\text{B}} = N - 1 + R_{\text{A}} + R_{\text{NR}}$  (where  $N$  is the total number of atoms). A simple Excel spread sheet has been made to calculate van der Waals volumes for a wide range of 677 organic compounds, including 237 drug compounds. The results show that the van der Waals volumes calculated from VABC are equivalent to the computer-calculated van der Waals volumes for organic compounds.

## Introduction

Intrinsic molecular volume has been widely employed as a molecular descriptor in modeling physicochemical properties and biological activity.<sup>1–4</sup> Molecular volume is one of most popular descriptors in QSAR (Quantitative Structure–Activity Relationship) studies.<sup>5–9</sup> Many physicochemical properties and biological processes are related to molecular volume, such as intestinal absorption and blood–brain barrier penetration.<sup>10,11</sup> Various procedures have been proposed to calculate the intrinsic volume. Traditionally, a method developed by Bondi is very popular. The volume of a molecule is calculated as the sum of volumes of mutually intersecting spheres centered on single atoms based on the bond distances,

bond angles, and intermolecular van der Waals radii.<sup>12</sup> van der Waals volume ( $V_{\text{vdW}}$ ) increments reported by Bondi are only valid for groups bonded to carbon, as pointed out by the author. Therefore, the  $V_{\text{vdW}}$  values of any non-hydrocarbon molecule obtained by summation of these latter increments is approximate and may be in error by up to  $2 \text{ cm}^3/\text{mol}$  ( $3.32 \text{ \AA}^3/\text{molecule}$ ).<sup>13</sup>

Nowadays, various computer programs are commonly used to calculate the intrinsic molecular volume. Molecular volume calculated from these programs should be more accurate than the manual group contribution method. However, volumes calculated from different programs are not the same. For example, the van der Waals volume from PcModel (PcModel for windows, Version 7.0, Serena Software) for ethanol is  $76 \text{ \AA}^3$ , from the TSAR program (Oxford Molecular Ltd) it is  $41 \text{ \AA}^3$ , and from MacroModel<sup>14</sup> (version 3.0) it is  $52 \text{ \AA}^3$ . The ACD program (1994–2002, Advanced Chemistry Development Inc.) gives  $98 \text{ \AA}^3$  if we convert  $\text{cm}^3/\text{mol}$  to  $\text{\AA}^3/\text{molecule}$  by dividing by 0.602, but this volume is a calculated molar volume.<sup>14</sup> Different atomic radii used in the van der Waals programs result in the discrepancies among the calculated volumes. For example, the radius of carbon is 1.70 from MicroModel but 1.90–1.94 from PcModel.

Although van der Waals volumes can be easily calculated from many programs, some of the programs are quite expensive and calculation of molecular volume is rather time consuming because in most cases it is

\* Address correspondence to this author at University College London.

† ID Business Solutions Ltd.

(1) Abraham, M. H. *Pure Appl. Chem.* **1993**, *22*, 73–83.

(2) Leahy, D. J. *J. Pharm. Sci.* **1986**, *75*, 629–636.

(3) McGowan, J. C.; Mellors, A. *Molecular Volumes in Chemistry and Biology-Applications Including Partitioning and Toxicity*; Ellis Harwood: Chichester, UK, 1986.

(4) McGowan, J. C. *J. Appl. Chem. Biotechnol.* **1978**, *28*, 599–607.

(5) Agrawal, V. K.; Khadikar, P. V. *Oxid. Commun.* **2002**, *25*, 184–192.

(6) Ericksen, D.; Wilding, W. V.; Oscarson, J. L.; Rowley, R. L. *J. Chem. Eng. Data* **2002**, *47*, 1293–1302.

(7) Consonni, V.; Todeschini, R.; Pavan, M. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 682–692.

(8) Vanyur, R.; Heberger, K.; Kovesdi, I.; Jakus, J. *Photochem. Photobiol.* **2002**, *75*, 471–478.

(9) Rugutt, J. K.; Rugutt, K. J. *Nat. Prod. Lett.* **2002**, *16*, 107–113.

(10) Abraham, M. H.; Chadha, H. S.; Martins, F.; Mitchell, R. C.; Bradbury, M. W.; Gratton, J. A. *Pestic. Sci.* **1999**, *55*, 78–88.

(11) Zhao, Y. H.; Abraham, M. H.; Le, J.; Hersey, A.; Luscombe, C. N.; Beck, G.; Sherborne, B.; Cooper, I. *Pharm. Res.* **2002**, *19*, 1444–1455.

(12) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441–451.

(13) Lepori, L.; Gianni, P. *J. Solution Chem.* **2000**, *29*, 405–447.

(14) Tran, D.; Hunt, J. P.; Wherland, S. *Inorg. Chem.* **1992**, *31*, 2460–2464.

**TABLE 1.** Van der Waals Volumes of Alkanes and Bond Contribution in Å<sup>3</sup>/Molecule<sup>a</sup>

no.	name	formula	$V_{\text{vdW}}(\text{Bondi})$	$V_{\text{SA}}$	$V_{\text{B}}$	$V_{\text{vdW}}(\text{eq 3})$	residue of Bondi – eq 3
1	methane	CH <sub>4</sub>	28.44	49.53	5.27	25.83	2.60
2	ethane	C <sub>2</sub> H <sub>6</sub>	45.42	84.59	5.60	43.12	2.29
3	propane	C <sub>3</sub> H <sub>8</sub>	62.41	119.64	5.72	60.40	1.99
4	butane	C <sub>4</sub> H <sub>10</sub>	79.40	154.70	5.79	77.68	1.70
5	isobutane	C <sub>4</sub> H <sub>10</sub>	79.39	154.70	5.79	77.68	1.68
6	pentane	C <sub>5</sub> H <sub>12</sub>	96.40	189.76	5.84	94.96	1.41
7	neopentane	C <sub>5</sub> H <sub>12</sub>	96.36	189.76	5.84	94.96	1.37
8	hexane	C <sub>6</sub> H <sub>14</sub>	113.39	224.81	5.86	112.24	1.11
9	heptane	C <sub>7</sub> H <sub>16</sub>	130.38	259.87	5.89	129.53	0.82
10	octane	C <sub>8</sub> H <sub>18</sub>	147.38	294.92	5.90	146.81	0.53
11	nonane	C <sub>9</sub> H <sub>20</sub>	164.37	329.98	5.91	164.09	0.23
12	decane	C <sub>10</sub> H <sub>22</sub>	181.36	365.04	5.92	181.37	–0.06
13	undecane	C <sub>11</sub> H <sub>24</sub>	198.36	400.09	5.93	198.66	–0.35
14	dodecane	C <sub>12</sub> H <sub>26</sub>	215.35	435.15	5.94	215.94	–0.65
15	tridecane	C <sub>13</sub> H <sub>28</sub>	232.34	470.20	5.95	233.22	–0.94
16	tetradecane	C <sub>14</sub> H <sub>30</sub>	249.34	505.26	5.95	250.50	–1.24
17	pentadecane	C <sub>15</sub> H <sub>32</sub>	266.33	540.32	5.96	267.78	–1.53
18	hexadecane	C <sub>16</sub> H <sub>34</sub>	283.32	575.37	5.96	285.07	–1.82
19	heptadecane	C <sub>17</sub> H <sub>36</sub>	300.32	610.43	5.96	302.35	–2.12
20	octadecane	C <sub>18</sub> H <sub>38</sub>	317.31	645.48	5.97	319.63	–2.41
21	nonadecane	C <sub>19</sub> H <sub>40</sub>	334.30	680.54	5.97	336.91	–2.70
22	eicosane	C <sub>20</sub> H <sub>42</sub>	351.30	715.60	5.97	354.20	–3.00
	minimization				5.92		
	standard deviation						1.73

<sup>a</sup>  $V_{\text{vdW}}(\text{Bondi})$ : van der Waals volume calculated from the Bondi method.<sup>12</sup>  $V_{\text{SA}}$ : the sum of atomic contributions in a molecule calculated from values in Table 2.  $V_{\text{B}}$ : bond contribution calculated from eq 2.  $V_{\text{vdW}}(\text{eq 3})$ : van der Waals volume calculated from eq 3.

necessary to generate the 3D molecular geometry of the molecule or a SMILES (Simplified Molecular Input Line Entry System).<sup>15</sup> However, the Bondi group contribution method for calculating molecular volumes is still very popular.<sup>16–20</sup> It does not take a long time to use the method to calculate molecular volumes for a small number of compounds. However, it is very time consuming and easy to make a miscalculation if one deals with a large number of compounds. The aim of this paper is to develop a method that can calculate van der Waals volume very rapidly, based on atomic and bond contributions, without using a specialized computer program.

## Results and Discussion

**Theoretical Background of Methodology.** The McGowan characteristic volume is well-known in QSAR studies.<sup>4,5</sup> McGowan volumes ( $V_x$ ) have an advantage in that they are more easily calculated than computer-calculated intrinsic volumes.  $V_x$  is calculated by simply adding up the atomic contribution for the compound and then subtracting 6.56 cm<sup>3</sup>/mol for each bond ( $B$ ) regardless of whether it is a single, double, or triple bond.<sup>21,22</sup>

$$V_x(\text{cm}^3/\text{mol}) = \sum \text{all atom contributions} (\text{cm}^3/\text{mol}) - \sum 6.56B \quad (1)$$

or

(15) Ertl, P.; Rohde, B.; Selzer, P. *J. Med. Chem.* **2000**, *43*, 3714–3717.

(16) Edward, J. T. *Can. J. Chem. Rev. Can. Chim.* **1998**, *76*, 1294–1303.

(17) Simha, R.; Carri, G. *J. Polym. Sci. Part B* **1994**, *32*, 2645–2651.

(18) Beckman, E. J.; Hoefling, T. A.; Vanopstal, L.; Koningsveld, R.; Porter, R. S. *Fluid Phase Equilib.* **1994**, *99*, 121–133.

(19) Hanai, T.; Hatano, H.; Nimura, N.; Kinoshita, T. *J. Liq. Chromatogr.* **1993**, *16*, 1453–1462.

(20) Hanai, T.; Hatano, H.; Nimura, N.; Kinoshita, T. *Anal. Sci.* **1993**, *9*, 43–46.

$$V_x(\text{\AA}^3/\text{molecule}) = \sum \text{all atom contributions} (\text{\AA}^3/\text{molecule}) - \sum 10.90B$$

Abraham and McGowan showed that McGowan volumes are well correlated to Leahy's computer-calculated intrinsic volumes ( $V_i$ ).<sup>2</sup> The related equation for 209 gaseous, liquid, and solid solutes is  $V_i(\text{cm}^3/\text{mol}) = 0.597 + 0.6823 V_x(\text{cm}^3/\text{mol})$  ( $n = 209$ ,  $r = 0.998$ ,  $\text{SD} = 1.24$ ). Because the McGowan volume is well correlated to the intrinsic volume, it was reasonable to believe that the McGowan method could be used to calculate van der Waals volumes ( $V_{\text{vdW}}$ ) provided that the bond contribution of a molecule could be derived.

**1. Bonding Contribution in Alkanes.** We start from alkanes because carbon and hydrogen are the most common elements in organic compounds.

The methodology for the calculation of van der Waals volumes is simply based on the atomic radii and covalent bond distance. Bondi developed a method to calculate the van der Waals volume of a molecule from spheres 1 and 2 (elements 1 and 2) and overlap of the two elements. Bondi then used the method to calculate group contributions to the van der Waals volume. These values are listed in Tables 15 and 16 in ref 12. Because group increments reported by Bondi are correct only for attachment to carbon atoms, we simply add each group contribution to carbon(s) in Tables 15 and 16 of ref 12 and calculate van der Waals volumes ( $V_{\text{vdW}}$ ). The values are listed in Table 1.

To investigate the bond contribution for alkanes, we calculate the sum of atomic volumes ( $V_{\text{SA}}$ ) by using the Bondi volumes in Table 2. Therefore, the average bond contribution to a molecular volume can be obtained from

(21) McGowan, J. C. *Recl Trav. Chim. Pays-Bas* **1956**, *75*, 193–207.

(22) Abraham, M. H.; McGowan, J. C. *Chromatographia* **1987**, *23*, 243–246.

**TABLE 2. Bondi Radii of Atoms and Their Volumes**

atom	$R_{\text{Bondi}} (\text{\AA})$	$V_{\text{vdW}} (\text{\AA}^3)$	atom	$R_{\text{vdW}} (\text{\AA})$	$V_{\text{vdW}} (\text{\AA}^3)$
H	1.20	7.24	P	1.80	24.43
C	1.70	20.58	S	1.80	24.43
N	1.55	15.60	As	1.85	26.52
O	1.52	14.71	B	2.13	40.48
F	1.47	13.31	Si	2.10	38.79
Cl	1.75	22.45	Se	1.90	28.73
Br	1.85	26.52	Te	2.06	36.62
I	1.98	32.52			

the following equation:

$$V_B (\text{\AA}^3/\text{molecule}) = [V_{\text{SA}} - V_{\text{vdW}}(\text{Bondi})]/N_B \quad (2)$$

Here,  $N_B$  is the number of bonds in a molecule. From Table 1 it can be seen that the average bond contribution reaches a constant 5.97 with increasing number of carbon atoms. Branching does not have a large influence on bond contribution (e.g. nos. 4–7 in Table 1). The average value of the bond contribution for alkanes is 5.84. If we

minimize the residue between Bondi volumes and volumes calculated from atomic and bond contributions, we obtain 5.92 for a bond contribution. Table 1 shows the calculated van der Waals volume from eq 3 by using a bond contribution defined as 5.92 and atomic volume contributions in Table 2. Compared with Bondi  $V_{\text{vdW}}$ , eq 3 slightly underestimates the volumes for small molecules and overestimates volumes for large molecules.

$$V_{\text{vdW}} = \sum \text{all atom contributions} - 5.92N_B \quad (3)$$

**2. Bond Contributions in Alkenes, Alkynes, Oxygen, Nitrogen, and Halogen Compounds.** To investigate the bond contribution of double bonds and O, N, and Cl single bonds, van der Waals volumes calculated from Bondi group contributions based on Tables 15 and 16 in ref 12 are listed in Table 3. The van der Waals volumes calculated from eq 3 are also listed in Table 3 for comparison. The results show that double and triple bonds make a large contribution to van der Waals

**TABLE 3. Van der Waals Volumes Calculated from Bondi Group Contributions in  $\text{\AA}^3/\text{Molecule}$** 

no.	name	formula	$V_{\text{vdW}}(\text{Bondi})$	$V_{\text{SA}}$	$V_B$	$V_{\text{vdW}}(\text{eq 3})$	residue of Bondi – eq 3
1	allene	$\text{C}_3\text{H}_4$	51.2	90.7	5.64	55.2	–3.94
2	ethene	$\text{C}_2\text{H}_4$	39.7	70.1	6.09	40.5	–0.84
3	2-methylproene	$\text{C}_4\text{H}_8$	73.6	140.2	6.06	75.1	–1.53
4	2-methyl-2-butene	$\text{C}_5\text{H}_{10}$	90.5	175.3	6.05	92.4	–1.87
5	2,3-dimethyl-2-butene	$\text{C}_6\text{H}_{12}$	107.5	210.3	6.05	109.7	–2.22
6	ethyne	$\text{C}_2\text{H}_2$	38.4	55.6	5.75	37.9	0.50
7	propyne	$\text{C}_3\text{H}_4$	55.3	90.7	5.90	55.2	0.09
	average				5.94		–1.40
	standard deviation						1.50
8	dimethyl ether	$\text{C}_2\text{H}_6\text{O}$	51.6	99.3	5.97	51.9	–0.38
9	methanol	$\text{CH}_4\text{O}$	36.1	64.2	5.64	34.6	1.42
10	acetone	$\text{C}_3\text{H}_6\text{O}$	64.9	119.9	6.11	66.6	–1.75
11	dimethyl sulfide	$\text{C}_2\text{H}_6\text{S}$	63.4	109.0	5.71	61.7	1.70
12	methanethiol	$\text{CH}_4\text{S}$	47.3	74.0	5.33	44.4	2.93
13	methylamine	$\text{CH}_5\text{N}$	40.2	72.4	5.36	36.8	3.37
14	dimethylamine	$\text{C}_2\text{H}_7\text{N}$	58.8	107.4	5.40	54.1	4.69
15	trimethylamine	$\text{C}_3\text{H}_9\text{N}$	75.3	142.5	5.60	71.4	3.87
16	acetonitrile	$\text{C}_2\text{H}_3\text{N}$	47.1	78.5	6.27	48.9	–1.75
17	nitromethane	$\text{CH}_3\text{NO}_2$	50.6	87.3	6.12	51.8	–1.18
18	trimethylphosphine	$\text{C}_3\text{H}_9\text{P}$	85.5	151.3	5.49	80.3	5.19
	average				5.73		1.65
	standard deviation						2.58
19	fluoromethane	$\text{CH}_3\text{F}$	32.2	55.6	5.85	31.9	0.29
20	2-fluoropropane	$\text{C}_3\text{H}_7\text{F}$	67.0	125.7	5.87	66.5	0.46
21	difluoromethane	$\text{CH}_2\text{F}_2$	36.9	61.7	6.19	38.0	–1.06
22	chloromethane	$\text{CH}_3\text{Cl}$	42.0	64.7	5.68	41.1	0.95
23	2-chloropropane	$\text{C}_3\text{H}_7\text{Cl}$	77.0	134.9	5.78	75.7	1.35
24	bromomethane	$\text{CH}_3\text{Br}$	46.6	68.8	5.55	45.1	1.49
25	2-bromopropane	$\text{C}_3\text{H}_7\text{Br}$	80.9	138.9	5.80	79.7	1.20
26	iodomethane	$\text{CH}_3\text{I}$	54.6	74.8	5.06	51.1	3.44
27	2-iodopropane	$\text{C}_3\text{H}_7\text{I}$	90.5	144.9	5.44	85.7	4.76
	average				5.69		1.43
	standard deviation						1.73
28	benzene	$\text{C}_6\text{H}_6$	80.3	166.9	7.21	95.9	–15.5
29	naphthalene	$\text{C}_{10}\text{H}_8$	122.9	263.7	7.41	151.2	–28.4
30	toluene	$\text{C}_7\text{H}_8$	98.9	202.0	6.87	113.2	–14.3
31	hexylbenzene	$\text{C}_{12}\text{H}_{18}$	183.8	377.2	6.45	199.6	–15.8
32	hexylnaphthalene	$\text{C}_{16}\text{H}_{20}$	226.4	474.0	6.69	255.0	–28.6
	average				6.93		–14.66 <sup>a</sup>
	standard deviation						7.30
33	cyclohexanol	$\text{C}_6\text{H}_{12}\text{O}$	109.6	225.0	6.08	112.6	–2.98
34	tetrahydropyran	$\text{C}_5\text{H}_{10}\text{O}$	91.1	190.0	6.18	95.3	–4.16
35	adamantane	$\text{C}_{10}\text{H}_{16}$	147.0	321.6	6.24	155.8	–8.84
36	cyclohexane	$\text{C}_6\text{H}_{12}$	102.0	210.3	6.02	103.8	–1.82
37	dicyclohexane	$\text{C}_{12}\text{H}_{22}$	192.5	406.2	6.11	199.0	–6.54
38	tricyclohexane	$\text{C}_{18}\text{H}_{32}$	283.0	602.1	6.14	294.2	–11.26
39	testosterone	$\text{C}_{19}\text{H}_{28}\text{O}_2$	292.7	623.2	6.24	309.4	–16.73
	average				6.14		–3.84 <sup>a</sup>
	standard deviation						5.26

<sup>a</sup> The average value is based on the number of rings.

volumes because the double bond is shorter than a single bond. The average bonding contribution for a molecule containing a double bond is 5.94, which is slightly higher than the average bonding contribution of a single bond (5.92).

For oxygen, nitrogen, and halogen compounds, the bonding contribution of some compounds is higher than the bond contribution in alkanes and for some compounds it is lower than the bond contribution in alkanes. The average values are lower than the bond contribution in alkane. However, the biggest difference between Bondi van der Waals volume and those values from eq 3 is for the aromatic ring. Also eq 3 overestimates the van der Waals volume for nonaromatic rings as well (average = 3.8). We take the bond contribution as 5.92 for all bonds, no matter whether they are single, double, or triple bonds and no matter what elements compose the bond (i.e. CC, CO, CN, CS, NN, etc.). We make a correction by subtracting 14.7 for each aromatic ring and 3.8 for each nonaromatic ring. Then, a general formula (eq 4) for calculating the van der Waals volume based on Bondi principles can be obtained.

$$V_{\text{vdW}} (\text{\AA}^3/\text{molecule}) = \sum \text{all atom contributions} - 5.92N_{\text{B}} - 14.7R_{\text{A}} - 3.8R_{\text{NR}} \quad (4)$$

Calculating the number of bonds present ( $N_{\text{B}}$ ) is simplified by using an algorithm,  $N_{\text{B}} = N - 1 + R_{\text{g}}$ , where  $N$  is the total number of atoms and  $R_{\text{g}}$  is the total number of ring structures ( $R_{\text{g}} = R_{\text{A}} + R_{\text{NA}}$ ).<sup>22</sup> For example, the volume of naphthalene ( $\text{C}_{10}\text{H}_8$ ) is calculated as follows:  $V_{\text{vdW}} = (10 \times 20.58) + (8 \times 7.24) - (19 \times 5.92) - (2 \times 14.7) = 121.84 \text{ \AA}^3$ .

**Validation of the Methodology.** To test the method developed in this paper, two different sources of data were selected: first, from Tran et al., who used the MacroModel program,<sup>14</sup> and second, calculated values from the computer program of TSAR.

The van der Waals volumes listed in Table 4 are all for organic compounds from Tran et al. The van der Waals volumes for these compounds calculated from eq 4 are also listed in the table. The results show the absolute values between the two sets are very close. The correlation coefficient is 0.999 for 27 organic compounds studied (eq 5).

$$V_{\text{vdW}}(\text{Tran}) = 1.03 V_{\text{vdW}}(\text{eq 4}) + 0.40 \quad (5)$$

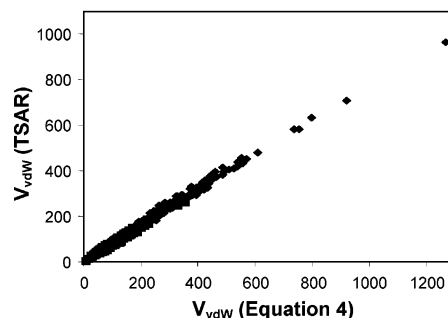
$$n = 27; S = 4.75; r = 0.999$$

Although the relationship between the values calculated from eq 4 and calculated from Tran et al. is excellent, there is a limited number of organic compounds in the paper of Tran et al.<sup>14</sup> Also the range of structures is rather small. Therefore, 440 general organic compounds containing all kinds of substituents (molecular weight from 16 to 334) and 237 drug compounds<sup>11</sup> (molecular weight from 75 to 1692) were chosen for analysis. van der Waals volumes ( $\text{\AA}^3/\text{molecule}$ ) have been calculated from both a computer program TSAR and eq 4 for these compounds. Figure 1 shows the relationship between the two sets for 677 organic compounds. The

**TABLE 4.** Comparison of van der Waals Volumes in  $\text{\AA}^3/\text{Molecule}$  Calculated by Tran et al.<sup>14</sup> and through Eq 4<sup>a</sup>

compd	formula	$V_{\text{vdW}}(\text{Tran})$	$V_{\text{vdW}}(\text{eq 4})$
N((CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub> ) <sub>4</sub> BPh <sub>4</sub>	C <sub>40</sub> H <sub>56</sub> NB	609	592
N((CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub> ) <sub>4</sub> BF <sub>4</sub>	C <sub>16</sub> H <sub>36</sub> NBF <sub>4</sub>	335	327
N((CH <sub>2</sub> ) <sub>2</sub> CH <sub>3</sub> ) <sub>4</sub> BF <sub>4</sub>	C <sub>12</sub> H <sub>28</sub> NBF <sub>4</sub>	270	258
As(C <sub>6</sub> H <sub>5</sub> ) <sub>4</sub> I	C <sub>24</sub> H <sub>20</sub> Ias	365	354
As(C <sub>6</sub> H <sub>5</sub> ) <sub>4</sub> Cl	C <sub>24</sub> H <sub>20</sub> ClAs	351	343
As(C <sub>6</sub> H <sub>5</sub> ) <sub>4</sub> Br	C <sub>24</sub> H <sub>20</sub> BrAs	360	348
HC(C <sub>6</sub> H <sub>5</sub> ) <sub>3</sub>	C <sub>19</sub> H <sub>16</sub>	243	243
N((CH <sub>2</sub> ) <sub>3</sub> CH <sub>3</sub> ) <sub>4</sub> I	C <sub>16</sub> H <sub>36</sub> NI	333	330
N(C <sub>4</sub> H <sub>9</sub> ) <sub>4</sub> Br	C <sub>16</sub> H <sub>36</sub> NBr	327	324
N(C <sub>3</sub> H <sub>7</sub> ) <sub>4</sub> Br	C <sub>12</sub> H <sub>28</sub> NBr	262	255
N(C <sub>2</sub> H <sub>5</sub> ) <sub>4</sub> Br	C <sub>8</sub> H <sub>20</sub> NBr	194	186
N(C <sub>2</sub> H <sub>5</sub> ) <sub>4</sub> Cl	C <sub>8</sub> H <sub>20</sub> NCl	185	182
N(CH <sub>3</sub> ) <sub>4</sub> BF <sub>4</sub>	C <sub>4</sub> H <sub>12</sub> NBF <sub>4</sub>	134	120
N(CH <sub>3</sub> ) <sub>4</sub> Br	C <sub>4</sub> H <sub>12</sub> NBr	126	117
N(CH <sub>3</sub> ) <sub>4</sub> Cl	C <sub>4</sub> H <sub>12</sub> NCl	117	113
benzyl alcohol	C <sub>7</sub> H <sub>8</sub> O	108	107
butan-1-ol	C <sub>4</sub> H <sub>10</sub> O	85	87
butan-2-ol	C <sub>4</sub> H <sub>10</sub> O	83	87
propan-1-ol	C <sub>3</sub> H <sub>8</sub> O	71	69
propan-2-ol	C <sub>3</sub> H <sub>8</sub> O	69	69
ethanol	C <sub>2</sub> H <sub>6</sub> O	52	52
methanol	CH <sub>4</sub> O	37	35
bipyridine	C <sub>10</sub> H <sub>8</sub> N <sub>2</sub>	146	141
dimethylglyoxime	C <sub>4</sub> H <sub>8</sub> N <sub>2</sub> O <sub>2</sub>	106	112
nioxime	C <sub>6</sub> H <sub>10</sub> N <sub>2</sub> O <sub>2</sub>	128	123
phenanthroline	C <sub>12</sub> H <sub>8</sub> N <sub>2</sub>	162	149
terpyridine	C <sub>15</sub> H <sub>11</sub> N <sub>3</sub>	214	207

<sup>a</sup> The van der Waals volume calculated from eq 4 for each salt is to be considered as two single molecular species.



**FIGURE 1.** Plot of  $V_{\text{vdW}}$  calculated from TSAR against those from eq 4 for 677 compounds.

correlation coefficient is 0.992, and the standard error is 8.6 (eq 6).

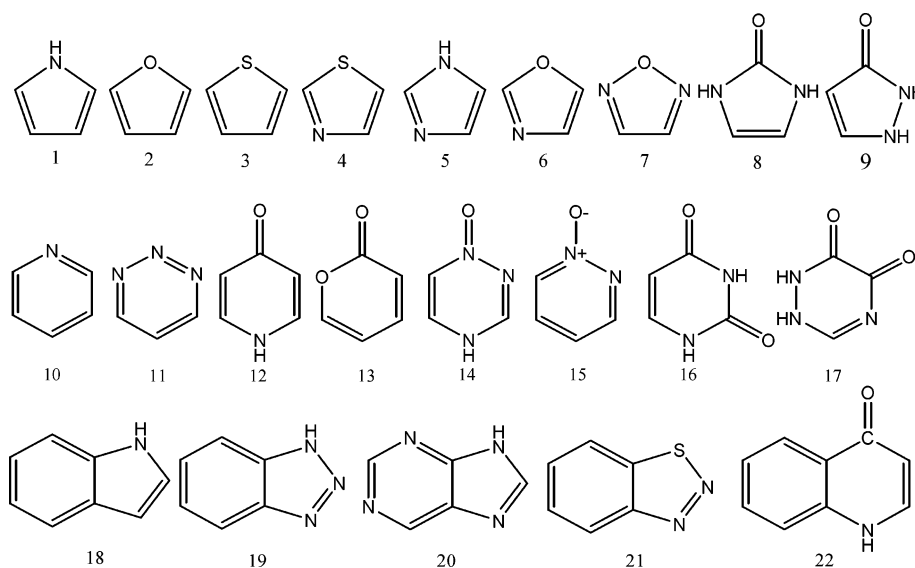
$$V_{\text{vdW}}(\text{TSAR}) = 0.801 V_{\text{vdW}}(\text{eq 4}) + 0.18 \quad (6)$$

$$n = 677; S = 8.65; r = 0.992$$

Table 5 lists some van der Waals volumes for aromatic rings that commonly appear in drug compounds.<sup>23</sup> The corresponding structures are shown above the table. To compare the van der Waals volumes from the TSAR program and the method developed in this paper, van der Waals volumes calculated from eq 4 were converted into the equivalent TSAR values through eq 6. The results show that van der Waals volumes calculated from the computer program TSAR are close to the values calculated by eq 4 after conversion by eq 6. This result

(23) Lewell, X. Q.; Jones, A. C.; Bruce, C. L.; Harper, G.; Jones, M. M.; Mclay, I.; Bradshaw, J. *J. Med. Chem.* **2003**, *46*, 3257–3274.



**TABLE 5.** Comparison of van der Waals Volumes in Å<sup>3</sup>/Molecule Calculated by TSAR and through Eq 4 for Aromatic Rings in Drugs

compd no.	formula	$V_{\text{vdW}}(\text{TSAR})$	$V_{\text{vdW}}(\text{eq 6})$	$V_{\text{vdW}}(\text{eq 4})$
1	C <sub>4</sub> H <sub>5</sub> N	48	48	60
2	C <sub>4</sub> H <sub>4</sub> O	47	47	58
3	C <sub>4</sub> H <sub>4</sub> S	58	54	68
4	C <sub>3</sub> H <sub>3</sub> NS	56	49	61
5	C <sub>3</sub> H <sub>4</sub> N <sub>2</sub>	46	43	54
6	C <sub>3</sub> H <sub>3</sub> NO	45	41	52
7	C <sub>2</sub> H <sub>2</sub> N <sub>2</sub> O	42	36	45
8	C <sub>3</sub> H <sub>4</sub> N <sub>2</sub> O	52	50	63
9	C <sub>3</sub> H <sub>4</sub> N <sub>2</sub> O	52	50	63
10	C <sub>5</sub> H <sub>5</sub> N	56	60	75
11	C <sub>3</sub> H <sub>3</sub> N <sub>3</sub>	53	50	62
12	C <sub>5</sub> H <sub>5</sub> NO	62	67	84
13	C <sub>5</sub> H <sub>4</sub> O <sub>2</sub>	61	65	81
14	C <sub>4</sub> H <sub>4</sub> N <sub>2</sub> O	60	62	77
15	C <sub>4</sub> H <sub>4</sub> N <sub>2</sub> O	61	62	77
16	C <sub>4</sub> H <sub>4</sub> N <sub>2</sub> O <sub>2</sub>	66	69	86
17	C <sub>3</sub> H <sub>3</sub> N <sub>3</sub> O <sub>2</sub>	66	64	80
18	C <sub>8</sub> H <sub>7</sub> N	79	81	101
19	C <sub>6</sub> H <sub>5</sub> N <sub>3</sub>	75	71	88
20	C <sub>5</sub> H <sub>4</sub> N <sub>4</sub>	73	66	82
21	C <sub>7</sub> H <sub>5</sub> NS	87	82	102
22	C <sub>9</sub> H <sub>7</sub> NO	93	100	124

suggests that the method developed in this paper can successfully be used to calculate the van der Waals volumes for organic compounds based on atomic volumes and number of bonds.

## Conclusions

Our results show that the calculation of van der Waals volume developed in the paper is equivalent to the TSAR and MacroModel computer-calculated van der Waals volumes for organic compounds. van der Waals volumes can simply be calculated by addition of the atomic volumes for all the atoms in the molecule and subtraction of 5.92 Å<sup>3</sup>/molecule for each bond ( $N_B$ ) regardless of whether it is a single, double, or triple bond. The general formula for calculating the van der Waals volume of a molecule is  $V_{\text{vdW}} (\text{Å}^3/\text{molecule}) = \sum \text{all atom contributions} - \sum 5.92B - 14.7R_A - 3.8R_{NA}$ . Atomic volumes used in the calculation can be found in Table 2. An Excel spreadsheet for the calculation of van der Waals volume is

available from the authors (yuan.zhao@ucl.ac.uk and m.h.abraham@ucl.ac.uk)

## Experimental Section

**Bondi Atomic Radii and Volume.** Bondi radii ( $R_{\text{Bondi}}$ ) were obtained from ref 12 directly. van der Waals volumes ( $V_{\text{vdW}}$ ) are calculated from Bondi radii by the following:  $V_{\text{vdW}} = 4\pi R_{\text{Bondi}}^3/3$ . The atomic radii and volume values are listed in Table 2.

**Intrinsic Molecular Volume.** Intrinsic molecular volumes shown in this paper have been obtained from different sources. van der Waals volume calculated from the Bondi method is based on the group contributions and functional group contributions listed in Tables 15 and 16 of ref 12. The contributions are added to those of carbon to calculate van der Waals volume for different kinds of molecules. The volumes values listed in Tables 1 and 3 are given in Å<sup>3</sup>/molecule converted from cm<sup>3</sup>/mol by dividing by 0.602. van der Waals volumes for 27 organic compounds (Table 4) calculated from a computer program were obtained from Tran et al.<sup>14</sup> These workers calculated volumes using the volume function of MacroModel

(version 3.0). The atomic radii (Å) used in the calculation are as follows: H, 1.2; C, 1.7; N, 1.6; O and F, 1.4; P, S, and Cl, 1.8; Br, 2.0; I and Si, 2.1; all other atoms, 1.2. van der Waals volumes for 440 general organic compounds and 237 drug compounds were calculated from the computer program TSAR (SMILES and van der Waals volumes are available from the authors). For the TSAR program, Corina is used to convert from SMILES to structures. No energy minimization was performed before the van der Waals volume calculation.

**Statistical Analysis.** The data set was analyzed by using Excel 97. The regression coefficients were obtained by least-squares regression analysis. For each regression, the following information is provided: number of observations used in the analysis ( $n$ ), square of the coefficient of determination ( $r^2$ ), and standard error ( $S$ ).

**Acknowledgment.** The authors thank Anne Hersey and Sami X. Paza (GSK) for calculating the van der Waals volume (TSAR program) and supporting structures of aromatic rings used in drugs. We thank the EPSRC for postdoctoral fellowships to Y.H.Z. and A.M.Z.

**Supporting Information Available:** van der Waals volumes calculated from eq 4 and TSAR and used in Figure 1 of this paper, together with their names, SMILES and  $V_{\text{vdw}} - \text{VABC}$ . This material is available free of charge via the Internet at <http://pubs.acs.org>.

JO034808O