Correlation of Computed van der Waals and Molecular Volumes with Apparent Molar Volumes (AMV) for Amino Acid, Carbohydrate and Sulfamate Tastant Molecules. Relationship between Corey–Pauling–Koltun Volumes (V_{CPK}) and Computed Volumes

William J. Spillane,† Gordon G. Birch, Michael G. B. Drew‡ and Ivan Bartolo
Departments of Food Science and Technology and Chemistry, University of Reading, Whiteknights, Reading, RG6 2AP, UK

Good linear correlations have been found between apparent molar volumes (AMV, $^{\phi}V$) and GEPOL-generated van der Waals (V_w), molecular (V_m) and accessible (V_a) volumes for 17 amino acids, 22 carbohydrates and 16 sulfamates. For a larger group of 24 sulfamates Corey–Pauling–Koltun measured volumes (V_{CPK}) gave good linear correlations with V_w (r = 0.925) and V_m (r = 0.935).

For the monosaccharides L-arabinose, D-glucose, L-lyxose, D-mannose and D-xylose the computer volumes and areas are slightly greater for the β- than for the α-anomers. However, for some methylated sugars the reverse occurs.

The use of 'effective' Bondi, rather than Pauling, van der Waals radii has little effect on the computed volumes and areas for the sulfamates. The effect, on the computed volumes and areas, of using a number of different sets of van der Waals radii has been examined for sucrose, glycine and cyclamate. The effect of increasing the probe (sphere) radius from 1.4 Å to 2.0 Å has been examined and it is found that the van der Waals and molecular volumes and the corresponding areas remain virtually unchanged, but the accessible volumes increase by *ca.* 31% and the accessible areas by *ca.* 22%.

The GEPOL program generally gives van der Waals volumes in good agreement with those available in the literature.

Apparent specific and apparent molar volumes (ASV and AMV), which are readily and accurately measurable, give an apparent measure of solute size and reflect displacement or disturbance of water structure. ASV values have proved to be particularly important in interpreting events in the chemoreception of tastant molecules and four appropriate ranges of ASV values have been found to be associated with the four basic tastes: salty, $< \sim 0.33$, sour $\sim 0.33 - \sim 0.52$, sweet $\sim 0.52 - \sim 0.71$ and bitter $\sim 0.71 - \sim 0.93$.[1] With the availability in recent times of a new[2,3] and superior[4] computer program for calculating molecular volumes and areas (surfaces) it was of considerable interest to calculate these molecular measurements for a variety of tastant molecules to see if significant relationships might exist between them and apparent molar volumes. The GEPOL program actually gives three types of volumes and areas: van der Waals, molecular and accessible (*vide infra*). These measurements have been calculated for three groups of molecules: (*i*) 17 amino acids displaying a variety of tastes, (*ii*) 22 carbohydrates, which are mainly sweet and (*iii*) 24 sulfamates, most of which are sweet.

For some years Corey–Pauling–Koltun (CPK) space-filling models have been used to measure dimensions of sulfamate molecules and a volume (V_{CPK}), which has proved very useful in developing structure–taste relationships for different classes of sulfamates, has been derived. These volumes are expected to show good correlations with the GEPOL-generated volumes and this prediction is examined in the present work.

Computer Techniques.—Molecular surfaces (areas) and volumes were generated using the GEPOL/87 program. The structures of amino acids were obtained from the Cambridge Structural Database (CSD)[5] and the molecules were displayed on the molecular graphics system CHEMMOD.[6] For L-proline data from *ab initio* calculations[7] were used as a basis for CHEMMOD modelling and this was also done for the related L-hydroxyproline molecule using the L-proline calculations.

For the monosaccharides studied we used literature data for α-glucopyranose[8] as a basis and other molecules were modelled on CHEMMOD and torsional optimization carried out to attain the lowest possible potential energy. For disaccharides, data from the CSD file was used to construct the CHEMMOD models and in some cases, where the positions of the hydrogens had not been reported, they were included in suitable positions *via* CHEMMOD.

The positions of the hydrogen atoms bonded to oxygen atoms in the carbohydrates are likely to be mobile in solution. We therefore checked the effect of varying from 0° to 300° the C(3)–C(4)–O(4)–H(4) (at 0°, 60°, 150°, 180°, 240° and 300°) and C(4)–C(3)–O(3)–H(3) (at 0°, 75°, 90°, 150°, 180°, 240° and 300°) torsional angles for β-glucopyranose in the $^4C_1$ conformation. For each torsional angle change the change in area is ≤ 2 Å^2 and in volume ~ 3 Å^3.

For the aliphatic sulfamates we used structural data from Kennard[9] on CH_3NHSO_3K. In particular we kept the spatial arrangement of the C–NH–SO_3 moiety and replaced the hydrogen on carbon *trans* to the sulfur atom and the C–N bond by the alkyl group [Fig. 1(*a*)]. For the aromatic sulfamates we used our data from the crystal structures of phenyl- and *m*-bromophenyl-sulfamates from the CSD database. In these two structures the conformation of the $NHSO_3$ group in relation to the phenyl ring is similar and we therefore took the average conformation for all the aryl compounds [Fig. 1(*b*)]. For cyclohexylsulfamate we used data obtained from a crystal structure study of silver cyclohexylsulfamate[10] [Fig. 1(*c*)]. For other cycloalkyls we used minimum energy conformations of

† *Present address:* Department of Chemistry, University College, Galway, Ireland.
‡ Department of Chemistry.

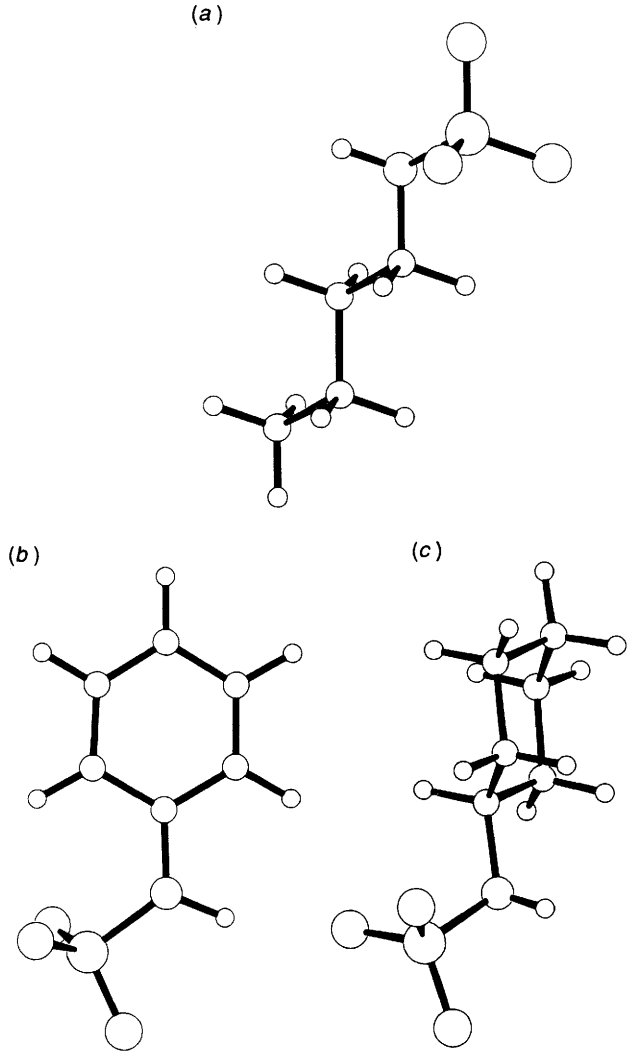

Fig. 1 Structures of (*a*) butylsulfamate, (*b*) phenylsulfamate and (*c*) cyclohexylsulfamate used in the calculations

the rings[11] and constructed models using the attachment of sulfamate to the cyclohexyl ring as a basis. The norbornyl ring system was constructed *via* CHEMMOD and energy-minimised. The sulfamate was then added *endo*- and *exo*- using the above method.

Further input for the GEPOL/87 program was provided by the different sets of van der Waals radii listed in Table 1.

The three numerical variables required by the program were entered as follows: FRAIDO, which controls the new spheres created during the calculation of the molecular surface was 0.55 (cited in the GEPOL/87 manual[2]); NDIV causes the surface of each sphere to be divided up into 960 triangles (with NDIV = 3) or 15 360 triangles (with NDIV = 5); NDIV = 3 gives volumes/surfaces with an accuracy of *ca.* ±0.5%, while with NDIV = 5 the accuracy is *ca.* ±0.05% (volumes) and *ca.* ±0.1 (surfaces).[19] We used NDIV = 5 in order to obtain the highest accuracy. RD is the probe or solvent radius and values of 1.4 Å[20] and 1.5 Å[2,3] have been used. Lee and Richards[20] used 1.4 Å in their study of protein structure and we have used this value in our present studies. For the sulfamates we have also used a value of 2.0 Å to examine the effect of changing the probe radius on the output values for the areas/volumes provided by the program.

Results and Discussion

The GEPOL/87 program calculates the envelope surface for a molecule and computes its corresponding area and volume. The program calculates three types of surface; (*i*) the accessible molecular surface—this is the surface defined by Lee and Richards;[20] (*ii*) the van der Waals surface—this is the external surface resulting from a set of spheres centred on the atoms or groups of atoms forming the molecule; (*iii*) the molecular surface—also defined by Richards.[21] The three types of surface are illustrated in Fig. 2. The program output data have been labelled A_a, V_a (accessible area and volume), A_w, V_w (van der Waals area and volume) and A_m, V_m (molecular area and volume).

Apparent Molar Volume (AMV) and Computed Volume Correlations.—A priori one would not necessarily expect to see a reasonable correlation between AMV and computed volumes for solutes. This is because the latter volumes are related only to the molecular configuration and do not in any way reflect the environment of the particular solutes. The AMV however gives a direct measure of disturbance or displacement of water by solute and reflects compatibility of solute with water structure.

In Table 2 the values for the computed accessible (V_a), van der Waals (V_w) and molecular (V_m) volumes together with the apparent molar volumes (AMV, $^{\phi}V$) for 17 amino acids using Pauling radii (Table 1) are given. The range of taste displayed by these amino acids has been described by Birch and Kemp[22a] and more recently all of them, except hydroxyproline, have been assessed by Haefeli and Glaser.[22b] There is a good agreement between the reports, and the former group have also compared their findings with those of previous workers. The data in Table 2 can be correlated with the following eqns. (1)–(3).

$$\text{AMV} = 0.209(\pm 0.020)V_a - 39.2(\pm 9.1) \quad (1)$$

$$\text{AMV} = 0.871(\pm 0.046)V_w - 12.9(\pm 5.5) \quad (2)$$

$$\text{AMV} = 0.811(\pm 0.045)V_m - 11.4(\pm 5.8) \quad (3)$$

The correlation coefficients (r) were 0.966, 0.980 and 0.978 respectively and were all significant at the 0.001 level.[23] The data correlated in eqn. (3) have been plotted in Fig. 3. Plots of AMVs for the 17 amino acids against accessible area (A_a), van der Waals area (A_w) and molecular area (A_m) also gave quite good correlations with correlation coefficients of 0.952, 0.966 and 0.963 respectively.

An interesting aspect of Fig. 3 is that it tends to divide the L-amino acids into hydrophobic (above the line) and hydrophilic (below the line) groups. Thus, phenylalanine (phe), leucine (leu), isoleucine (ile), methionine (met), valine (val) and proline (pro), which are hydrophobic,[19] lie above the line and arginine (arg), glutamine (gln), glutamic acid (glu), aspartic acid (asp) and serine (ser), which are hydrophilic lie below the line. Trytophan (trp), histamine (his) and glycine (gly) lie on the line and lysine (lys) and alanine (ala) are on the wrong sides of the line.

In Table 3 using Pauling van der Waals radii in the program the computed and apparent molar volumes for a series of carbohydrates that includes mono- and di-saccharides and some methylated sugars, have been calculated. In addition, calculations have been performed for a few sets of anomers. Most of the 22 compounds in Table 3 are sweet. Some display a sweet-bitter taste and β-D-mannose is bitter, though the α-form is sweet.[24–26] As with the amino acids these 22 sugars showed good correlations when plots of AMV *vs.* V_a [eqn. (4)], V_w [eqn. (5)] and V_m [eqn. (6)] were performed.

$$\text{AMV} = 0.289(\pm 0.008)V_a - 30.9(\pm 4.3) \quad (4)$$

$$\text{AMV} = 0.822(\pm 0.026)V_w - 0.79(\pm 4.4) \quad (5)$$

Table 1 Sets of van der Waals radii (in Å)

Set	Radii							
	H	C	O	N	S	F	Cl	Br
Pauling[12]	1.2	—[a]	1.40	1.5	1.85	1.35	1.80	1.95
Effective Bondi[13]	1.2	1.7	1.50	1.70	1.74	1.47	1.76	1.85
Bondi[14]	1.2	1.7	1.5	1.55	1.80	1.47	1.75	1.85
Allinger[15]	1.5	1.75[b]	1.65	1.70	2.0	1.6	1.95	2.10
Motoc and Marshall[16]	1.08	1.53	1.36	1.45	1.70	1.30	1.65	1.80
Bartell[17]	0.92	1.25	1.13	1.14	1.45[c]	1.08	1.44	1.59[d]

[a] Value not listed but we have used a value of 1.6 Å which seems reasonable though Allinger[15] has argued for a slightly higher value. The average of the other five values for C in the Table is 1.59 Å. [b] For sp^2, sp carbon a value of 1.85 Å is given by Allinger. [c] From ref. 18*a*. [d] From ref. 18*b*.

Table 2 Accessible, van der Waals, molecular and apparent molar[a] (AMV) volumes for amino acids

L-Amino acid	V_a/Å^3	V_w/Å^3	V_m/Å^3	AMV/ cm^3 mol^{-1}
asp	405.8	103.1	109.6	66.1
gly	280.5	62.8	65.1	42.7
ser	353.7	86.4	91.3	59.6
glu	457.2	119.5	126.6	86.6
hpro	421.6	110.1	116.0	82.1
his	475.6	127.8	134.3	97.9
gln	466.8	123.7	131.6	92.8
ala	329.7	79.2	83.5	58.0
arg	577.4	157.7	167.0	118.6
trp	600.8	174.2	184.8	139.9
met	482.5	131.8	139.2	104.6
pro	401.3	102.9	108.1	80.6
lys	508.1	139.3	184.8	102.4
phe	526.6	146.0	152.9	118.6
val	414.4	111.4	118.1	89.9
ile	464.6	127.9	136.2	105.0
leu	470.6	127.6	136.2	105.0

[a] Ref. 22*a*. The values listed here were determined at 0.6% w/w.

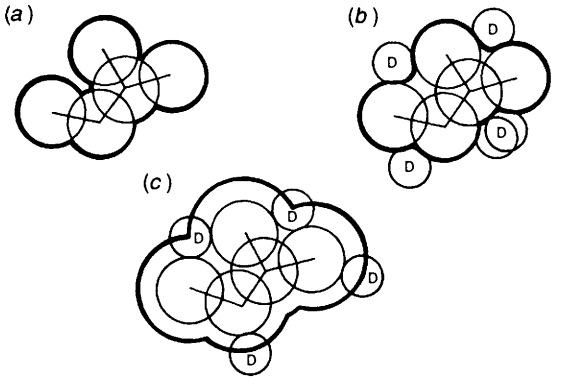

Fig. 2 GEPOL-generated volumes and surfaces: van der Waals (*a*); molecular surface (*b*); accessible surface (*c*).

$$\text{AMV} = 0.750(\pm 0.025)V_m - 1.83(\pm 4.4) \quad (6)$$

The correlation coefficients obtained were 0.993, 0.990 and 0.989 respectively which indicate that the correlations were significant at the 0.001 level. Plots of AMV for these 22 sugars against A_s, A_w and A_m also gave good correlations with r 0.993, 0.990 and 0.992 respectively. The data correlated in eqn. (4) have been plotted in Fig. 4.

There are five sets of anomers in Table 3 and it is interesting to note that for β- and α-L-arabinose, β- and α-D-glucoses and β- and α-D-xylopyranosides the volumes computed for the β-configurations are greater than those for the α-forms. However,

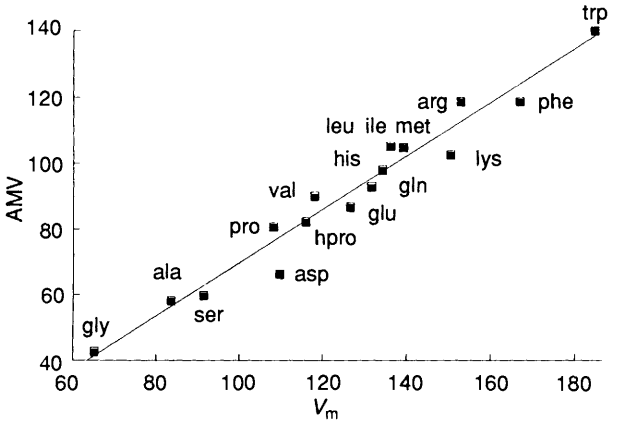

Fig. 3 Plot of AMV *vs.* V_m for 17 amino acids

in the case of methyl-D-galactopyranosides and methyl-D-glucopyranosides the volumes computed for the α-anomers are larger than those for the β-forms. We probed this a little further by calculating volumes for the α-forms of D-xylose, L-lyxose and D-mannose, whose β-form volumes are in Table 3. For the five monosaccharides considered, the β-anomer always has the larger volume (see Table 3 and footnotes).

Table 4 contains V_a, V_w and V_m computed volumes (again using the Pauling radii in Table 1) for 24 sulfamates and AMVs for 16 of these sulfamates. The first 14 sulfamates in the Table are sweet while phenylsulfamate has a sweet aftertaste and 4-bromophenylsulphamate is sour.[27] Plots of AMV *vs.* V_a, V_w and V_m had quite good correlation coefficients (r = 0.977, 0.967 and 0.959 respectively) and were described by eqns. (7)–(9). AMV values were available for 16 compounds only.

$$\text{AMV} = 0.318(\pm 0.187)V_a - 40.9\,(\pm 9.3) \quad (7)$$

$$\text{AMV} = 0.776(\pm 0.055)V_w - 10.4(\pm 7.6) \quad (8)$$

$$\text{AMV} = 0.689(\pm 0.054)V_m + 17.6(\pm 7.9) \quad (9)$$

Plots of AMV for the same 16 sulfamates against the computed surface areas, A_a, A_w and A_m gave correlation coefficients of 0.954, 0.954 and 0.977 respectively. Thus all six plots involving AMV were significant at the 0.001 level. Eqn. (8) is plotted in Fig. 5

These relationships that we have found for the amino acids, carbohydrates and sulfamates (Tables 2–4) show that there is an important connection between the experimentally determined AMVs and the theoretically determined van der Waals volumes. That such a relationship should exist is perhaps to be anticipated since the AMV measures the volume of solvent displaced by the substrate and the van der Waals volumes are clearly related to the size of the substrates. They are not structure–taste relationships.

Table 3 Accessible, van der Waals, molecular and apparent molar[a] (AMV) volumes for carbohydrates

Carbohydrate	V_a/Å^3	V_w/Å^3	V_m/Å^3	AMV/cm^3 mol^{-1}
β-L-arabinose	432.7	118.4	126.1	91.35[e]
α-L-arabinose	429.0	118.3	125.9	91.35[e]
β-D-galactose	497.7	140.7	149.8	109.0[e]
β-D-xylose[b]	436.5	118.4	126.3	93.8[e]
lactose	831.7	252.1	271.8	210.5[e]
sucrose	836.9	264.7	289.2	208.3[e]
xylitol	467.4	124.0	133.4	100.5[e]
maltose	807.0	254.1	271.2	209.4[e]
β-D-glucose	506.0	141.6	151.5	110.8[f]
α-D-glucose	500.7	141.1	150.4	110.8[f]
β-D-fructopyranose (10%)	492.7	141.0	150.1	110.6[f]
β-L-lyxose[c]	434.3	118.6	126.6	93.8[f]
β-D-ribopyranose	437.8	118.7	127.0	94.8[f]
α-D,α-L-trehalose (2.71%)	846.1	254.0	275.5	206.3[g]
3-*O*-Me-β-D-glucopyranoside	571.9	158.3	169.7	133.4[g]
Me-β-D-galactopyranoside	556.3	157.7	170.5	130.2[g]
Me-α-D-galactopyranoside	561.7	158.2	172.4	136.1[g]
Me-β-D-glucopyranoside	547.7	151.5	162.3	136.0[g]
Me-α-D-glucopyranoside	558.7	158.0	170.7	132.1[g]
Me-β-D-xylopyranoside	493.6	134.2	143.5	116.6[g]
Me-α-D-xylopyranoside	467.3	127.3	134.6	115.5[g]
β-D-mannose[d]	506.4	141.6	151.7	110.8[e]

[a] AMVs refer to compounds at 3% (w/w) concentration unless otherwise stated. [b] For α-D-xylose, V_a = 432.7, V_w = 118.3, V_m = 126.1 (all Å^3). [c] For α-L-lyxose, V_a = 428.9, V_w = 118.6, V_m = 126.4 (all Å^3). [d] For α-D-mannose, V_a = 503.4, V_w = 141.5, V_m = 151.5 (all Å^3). [e] Ref. 24. [f] Ref. 25. [g] Ref. 26.

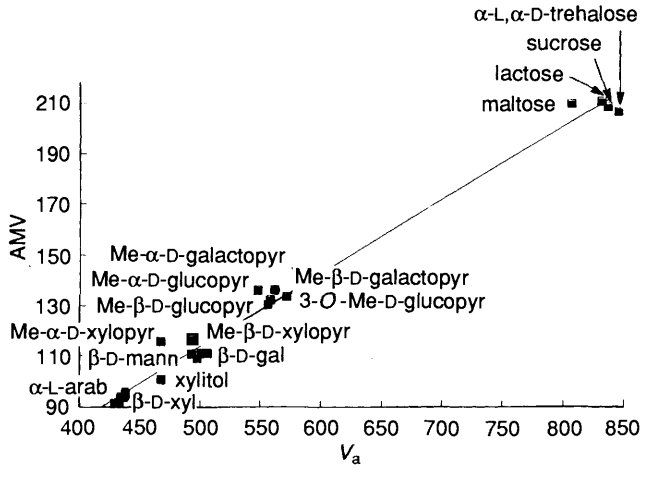

Fig. 4 Plot of AMV *vs.* V_a for 22 carbohydrates. Not all the carbohydrates plotted are labelled on the graph.

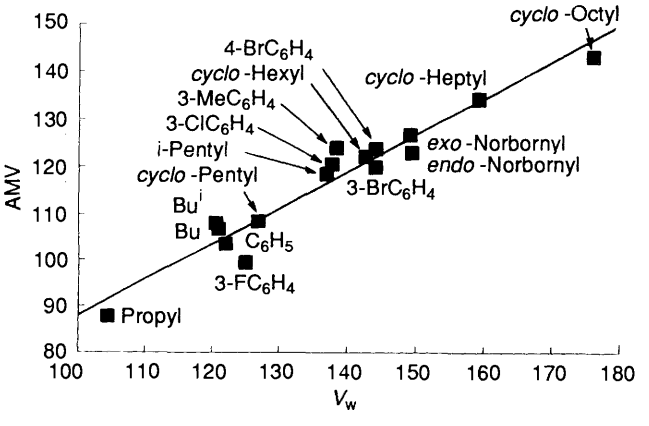

Fig. 5 Plot of AMV *vs.* V_w for 16 sulfamates

It is interesting to note that the slopes of the AMV *vs.* V_w and V_m plots in eqns. (2), (3), (5), (6), (8) and (9) are 0.79 ± 0.05 while the slopes of the AMV *vs.* V_a plots in eqns. (1), (4) and (7) are 0.27 ± 0.04. A further point of interest is the fact that recently[27] we have found a good linear correlation (r = 0.917) between the V_{CPK} volumes for the sulfamates in Table 4 and their apparent molar volumes. This suggests that the present computed volumes might show a relationship with V_{CPK} and we have looked at this for a slightly larger group of 24 sulfamates below.

Correlation of V_{CPK} *Volumes with Computed Volumes and Areas of Sulfamates.*—In Table 4 the accessible (V_a), van der Waals (V_w) and molecular (V_m) volumes are given for all 24 sulfamates together with the corresponding V_{CPK} volumes. The V_{CPK} volumes are measured using Corey–Pauling–Koltun space-filling molecular models where the 'length' (x Å), the 'width' (y Å) and the 'height' (z Å) of R in $RNHSO_3^-$ were measured and $V_{CPK} = x \times y \times z$ Å^3. Thus the V_{CPK} volumes usually describe rectangular spaces into which the R moiety of the sulfamate may be thought of as fitting[29] as a first step in the operation of the chemoreceptor mechanism. The V_{CPK} volumes are therefore generally larger than either the van der Waals or molecular volumes computed for the sulfamates. Measurement of the V_{CPK} volumes are based on the R part of the sulfamate, $RNHSO_3^-$, while the computed volumes and areas in the GEPOL program were calculated for the entire sulfamate ion. This will of course bring the computed volumes closer to the V_{CPK} volumes than might otherwise be the case.

Despite these approximations it was found that quite good correlations exist between V_{CPK} and the volumes V_w and V_m for the 24 sulfamates in Table 4 [eqns. (10) and (11)]. The

$$V_{CPK} = 2.53(\pm 0.22)V_w - 176(\pm 31) \quad (10)$$

$$V_{CPK} = 2.26(\pm 0.18)V_m - 156(\pm 27) \quad (11)$$

correlation coefficients were 0.925 and 0.935 respectively. A V_{CPK} *vs.* V_a plot had a correlation coefficient of 0.894 and plots of V_{CPK} *vs.* A_a, A_w and A_m gave r 0.807, 0.236 and 0.877 respectively. Eqn. (10) is plotted in Fig. 6.

In Table 4 we have included a few sulfamates of special interest *viz. exo*- and *endo*-norbornyl- and *trans*-2- and 4-methyl-

Table 4 Accessible, van der Waals, molecular, CPK[a] (V_{CPK}) and apparent molar[b] volumes for sulfamates[c]

Sulfamate (R in $RNHSO_3^-$)	V_a/Å^3	V_w/Å^3	V_m/Å^3	V_{CPK}/Å^3	AMV/cm^3 mol^{-1}
3-BrC_6H_4	514.7	143.8	148.3	172	120.3
Pr	411.7	104.4	109.7	90	87.6
3-FC_6H_4	462.7	124.7	128.8	133	99.3
Bu^i	453.3	120.4	127.1	158	107.8
3-ClC_6H_4	498.3	137.5	141.9	152	120.5
Bu	465.8	120.9	127.8	113	106.8
Isopentyl	500.7	136.6	144.4	180	119.1
cyclo-Octyl	588.6	175.7	188.8	246	143.1
cyclo-Heptyl	542.8	158.9	169.7	208	134.2
cyclo-Hexyl	508.1	142.5	151.9	181	122.6
cyclo-Pentyl	464.0	126.7	134.1	155	108.1
3-MeC_6H_4	506.2	138.2	143.5	158	123.9
exo-Norbornyl	525.3	149.1	158.7	210	126.9
endo-Norbornyl	523.3	149.1	159.3	210	123.1
C_6H_5	452.9	121.9	125.8	123	102.5
4-BrC_6H_4	515.9	143.9	148.4	177	123.6
2-Methylbutyl	502.8	136.7	145.1	195	
Neopentyl	493.6	136.5	145.1	186	
3,5-$DiFC_6H_3$	472.4	127.7	132.0	143	
Pentyl	519.9	137.3	145.8	142	
trans-4-Me-cyclo-Hexyl	557.3	159.2	170.5	238	
cis-4-Me-cyclo-Hexyl	547.7	158.6	168.6	254	
trans-2-Me-cyclo-Hexyl	552.3	159.0	170.5	236	
cis-2-Me-cyclo-Hexyl	541.0	158.1	167.4	243	

[a] Most of the V_{CPK} values have been given previously.[28] Replicate measurements indicate an accuracy of ±5%. [b] Ref. 27. Apparent molar volumes were determined at 5% concentration (w/w). [c] All the compounds display some degree of sweetness with the exception of 4-bromophenyl- and pentylsulfamates.

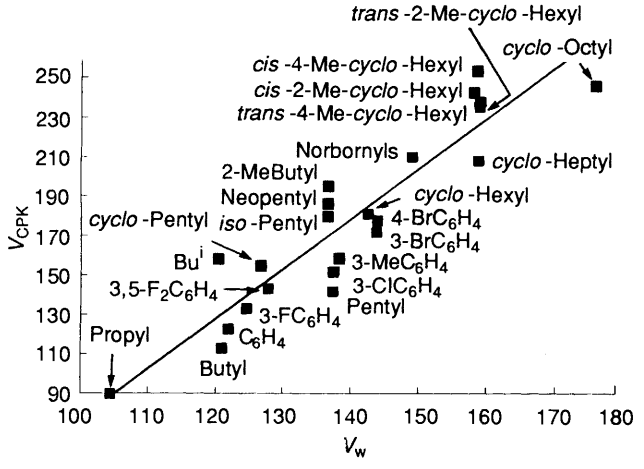

Fig. 6 Plot of V_{CPK} *vs.* V_w for 24 sulfamates

and *cis*-2- and 4-methylcyclohexylsulfamates. There is good evidence that the *exo*- and *endo*-norbornyl compounds differ in their degree of sweetness[28b] though they give the same V_{CPK} value of 210 Å^3. *cis*-2-Methyl- and *trans*-2-methylcyclohexylsulfamates have the same relative sweetness[29] but the *cis*-4-methyl- and *trans*-4-methylcyclohexylsulfamates may differ in sweetness.[30]

The two norbornyl compounds have the same V_{CPK} values and show just a slight variation in the computed volumes; however the methylcyclohexyl isomer sets do show some differences in both the V_{CPK} and computed volumes, a fact that may help to explain differences in relative sweetness, if established, for the 4-methyl compounds. In previous work[28a] we gave V_{CPK} values for the 2- and 4-methylcyclohexylsulfamates of 250 Å^3 and 199 Å^3 respectively. These values were for hybrid structures and the values in Table 4 should now be used.

For some years we have successfully employed V_{CPK} values to develop structure–taste relationships for sulfamates and the volumes measured are very significant in the construction of semi-quantitative SARs for these tastant molecules.[28,31] The approach was pragmatic and its success emphasised clearly the importance of the steric requirements of R in $RNHSO_3^-$ especially for carbosulfamates[28] and for monosubstituted aromatic sulfamates.[31] That spatial considerations are crucial is further underlined by the good correlations now found between, particularly, the van der Waals and molecular computed volumes and the V_{CPK} volumes. Since the computed volumes can be calculated to within an accuracy of ±0.05% it is hoped that calculation of further volumes may help to refine the overall picture and perhaps clear up the few problems that exist with misclassified compounds.

Effect of using Different van der Waals Radii and a Different Probe Radius.—Our sophisticated calculations using the GEPOL program are dependent on the estimates of the van der Waals radii used[19] in the input and further, since the V_{CPK} models are based partially on Pauling's set of van der Waals radii, it was desirable to use an alternative set of radii to see to what extent this affects the output volumes and areas.

There have been recent efforts to revise the Pauling radii based on an analysis of data[13] in the Cambridge Structural Database. In this work data are extracted for the shortest non-bonded distances between specified atoms. It was felt that the shortest distance, d_{xy}, between non-bonded atoms X and Y is a consequence of the intermolecular forces between the molecule containing X and all the other molecules in the crystal containing atom Y. Thus the effective external shapes of the atoms X and Y are dependent on chemical environment and are often non-spherical, being shorter in their 'head-on' contacts than in their 'sideways' contacts. For this reason Nyborg and Faerman[13] have postulated an elliptical rather than a spherical shape for the van der Waals surface in a variety of atoms.

It was not practical to incoporate these ideas into the GEPOL program; however, we tested the dependence of the calculations on the van der Waals radii by using Bondi 'effective' radii[13] instead of Pauling's values (Table 1) and recalculating volumes and areas for the first 22 sulfamates in Table 4. The

Table 5 Effect of using different sets of van der Waals radii (in Å) for selected molecules

Tastant	A_a	V_a	A_w	V_w	A_m	V_m	Set[a]
Glycine	220.7	280.5	97.0	62.8	93.9	65.1	Pauling
	225.0	291.6	99.2	69.4	97.3	70.8	Eff. Bondi
	252.1	346.3	118.9	91.5	117.5	92.4	Allinger
	210.7	260.9	89.8	55.1	87.3	57.3	Motoc and Marshall
	191.9	221.6	76.0	36.9	74.2	40.4	Bartell
Sucrose	490.9	836.9	349.5	264.7	295.4	289.2	Pauling
	494.1	852.3	339.2	280.9	297.3	301.1	Eff. Bondi
	535.9	985.6	368.8	365.6	329.3	382.7	Allinger
	475.8	790.4	336.4	234.7	285.1	260.8	Motoc and Marshall
	450.9	705.5	314.5	166.5	271.7	204.2	Bartell
Cyclohexylsulfamate	340.0	508.1	192.7	142.5	175.6	151.9	Pauling
	346.2	522.5	192.9	151.0	180.9	158.1	Bondi[b]
	345.1	521.2	192.1	150.8	180.6	157.6	Eff. Bondi
	378.2	609.7	218.3	199.0	210.0	202.8	Allinger
	328.4	477.5	179.6	126.5	168.3	135.1	Motoc and Marshall
	305.7	417.6	165.0	88.0	150.7	100.9	Bartell

[a] See Table 1. [b] Not computed for the molecules glycine and sucrose because the difference between Bondi and effective Bondi van der Waals radii are the same for C,H,O and only slightly different for N which has values of 1.55 Å and 1.70 Å and the computed volumes/areas will be within ± 1 Å^3/Å^2.

Table 6 Comparison between GEPOL-generated van der Waals volumes and previously published values

	V_a/Å^3				
Molecule	Bondi[a]	Gavezzotti[b]	Motoc and Marshall[c]	This work	Govers and de Voogt[d]
Methane	28.39	28.01		27.1	
Ethane	45.33	44.63		43.6	
Ethene	39.68	40.25		36.9	
Ethyne	38.35	36.15		31.1	
Benzene	80.36	85.39		79.0	83.4
Naphthalene	122.81		100.1	120.0	127.8
Cyclohexane	100.0	99.1		99.1	
Chloroethane	58.96	59.45		59.6	
Toluene	98.79	101.8		95.4	
Bromobenzene	100.3	106.6		97.3	
2-Methylbut-1-ene			66.5	86.0	
Penta-1,4-diene			63.1	79.1	
Ethylbenzene			89.7	114.0	

[a] Ref. 14. [b] Ref. 32. [c] Ref. 16. [d] Ref. 33.

output volumes and areas were within, at least, $\pm 6\%$ of the previously calculated values. Further plots of $(V_w, V_m, V_a)_{Pauling}$ *vs.* $(V_w, V_m, V_a)_{Bondi}$ and of $(A_m, A_a)_{Pauling}$ *vs.* $(A_m, A_a)_{Bondi}$ had slopes of 1.0 ± 0.05 and $r \geqslant 0.987$ and the plot of $(A_w)_{Pauling}$ *vs.* $(A_w)_{Bondi}$ had a slope of 1.15 ($r = 0.994$). Plots of AMV and V_{CPK} for the reduced sets of 16 and 22 sulfamates respectively using the volumes computed with effective Bondi data as input were therefore similar to those reported above.

The effect of using the different sets of van der Waals radii reported in Table 1 has been examined for a key tastant molecule from each of the groups studied in this work (Table 5). The effect is small when one uses Pauling or effective Bondi radii but, the use of Allinger, Motoc and Marshall or Bartell (Glidewell) sets of radii produce substantial differences.

The effect of changing the probe (solvent) radius from 1.4 Å to 2.0 Å was examined for the first 22 sulfamates in Table 4. The van der Waals and molecular volumes and the corresponding areas were virtually the same but the accessible volumes increased by 31% and the accessible areas by 22%. Normally a probe radius of *ca.* 1.4 Å has been used (*vide supra*) as an estimate of the size of the water molecule, but if one wished to do calculations for a larger solvent molecule then a value of 2.0 Å would be more appropriate and this calculation is useful in that it gives one an idea of the effect of probe radius on the output volumes and areas.

Comparison with Literature van der Waals Volumes.—Finally we have used the present method to calculate van der Waals volumes (V_w) for a number of molecules whose van der Waals volumes have been reported previously in the literature. In Table 6 all the data are brought together and the agreement between the GEPOL-computed V_w values and those of Bondi and Gavezzotti is quite good. The values of Motoc and Marshall are consistently *ca.* 20% lower than our values and, in one case where comparison is possible, that of Bondi for naphthalene. Finally, the values of Govers and de Voogt[33] for benzene and naphthalene are in good agreement with our own and with those of Bondi and Gavezzotti (benzene) and with that of Bondi (naphthalene).

Conclusion

In this work we have shown that apparent molar volumes (AMV) for three types of tastant molecule can be correlated with the different types of volumes and areas generated by the GEPOL/87 program. Corey–Pauling–Koltun volumes (V_{CPK}) for sulfamates can also be correlated with the GEPOL volumes/areas. The volumes and areas produced by the program have been found in some cases to be very dependent on the particular set of van der Waals radii used in the input. For the Pauling and effective Bondi sets we have shown that the output volumes and

areas are similar. From Table 5 it is clear that ratios of volumes or areas calculated for any two molecules using any of the five sets of van der Waals radii fall within 10% even though the individual values are very different. This suggests that any of these sets of van der Waals radii could be used to obtain suitable fits to experimental data. However, it is crucial that only one set be used in a specific study.

Acknowledgements

W. J. S. thanks the Royal Society (London) and the Royal Irish Academy for the award of a Royal Society European Science Exchange Fellowship. Mrs. C. Usher, Department of Chemistry, University of Reading is thanked for help with computing and graphics. W. J. S. also thanks the Departments of Food Science and Technology and Chemistry for facilities and assistance during a period of sabbatical leave. G. G. B. thanks the Royal Society and the Reading University Research Board for travel grants in support of this work.

References

1 S. Shamil, G. G. Birch, M. Mathlouthi and M. N. Clifford, *Chem. Senses*, 1987, **12**, 397; S. Shamil and G. G. Birch, *Endeavour*, 1990, **14**, 191.
2 J. L. Pascual-Ahuir, E. Silla, J. Tomasi and R. Bonaccorsi (1987), Quantum Chemistry Program Exchange, QCPE Program No. 554. GEPOL/87: Areas and volumes of molecules. Department of Chemistry, Indiana University, Bloomington, Indiana, USA.
3 J. L. Pascual-Ahuir and E. Silla, *J. Comput. Chem.*, 1990, **11**, 1047.
4 E. Silla, F. Villar, O. Nilssson, J. L. Pascual-Ahuir and O. Tapia, *J. Molec. Graphics*, 1990, **8**, 168.
5 Cambridge Crystallographic Data Centre, January 1991 update.
6 CHEMMOD, Chemical molecular modelling system, IBM PC version, Software version 2.0, U-Microcomputers Ltd., 1987. Available from Fraser-Williams Ltd.
7 A.-M. Sapse, L. Mallah-Levy, S. B. Daniels and B. W. Erickson, *J. Am. Chem. Soc.*, 1987, **109**, 3526.
8 S. N. Ha, A. Giammona, M. Field and J. W. Brady, *Carbohydr. Res.*, 1988, **180**, 207.
9 A. J. Morris, C. H. L. Kennard, J. R. Hall, G. Smith and A. J. White, *Acta Crystallogr., Sect. C*, 1983, **39**, 1236; A. J. Morris, C. H. L. Kennard, J. R. Hall and G. Smith, *Inorg. Chem. Acta*, 1982, **62**, 247.
10 W. J. Spillane, D. Cunningham, P. McArdle and C. Ryder, unpublished work.
11 U. Burkert and N. L. Allinger, *Molecular Mechanics*, ACS Monograph 177, American Chemical Society, Washington DC, 1982.
12 L. Pauling, *The Nature of the Chemical Bond*, Cornell University Press, Ithaca, NY, 1960, pp. 257–264.
13 S. C. Nyborg and C. H. Faerman, *Acta Crystallogr., Sect. B*, 1985, **41**, 274; S. C. Nyborg, C. H. Faerman and L. Prasad, *Acta Crystallogr., Sect. C*, 1987, **43**, 106.
14 A. Bondi, *J. Phys. Chem.*, 1964, **68**, 441; A. Bondi, *Physical Properties of Molecular Crystals, Liquids and Glasses*, Wiley, N.Y., 1968, pp. 452 *et seq.*
15 N. L. Allinger, *Adv. Phys. Org. Chem.*, vol. 13, eds. V. Gold and D. Bethell, Academic Press, 1976, pp. 1–82.
16 I. Motoc and G. R. Marshall, *Chem. Phys. Lett.*, 1985, **116**, 415.
17 L. S. Bartell, *J. Chem. Phys.*, 1960, **32**, 827.
18 (*a*) C. Glidewell, *Inorg. Chem. Acta*, 1975, **12**, 219; (*b*) 1976, **20**, 113.
19 I. Bartolo, M.Sc. Thesis, University of Reading, 1989.
20 B. Lee and F. M. Richards, *J. Mol. Biol.*, 1971, **55**, 379.
21 F. M. Richards, *Ann. Rev. Biophys. Bioeng.*, 1977, **6**, 151.
22 (*a*) G. G. Birch and S. E. Kemp, *Chem. Senses*, 1989, **14**, 249; (*b*) R. J. Haefeli and D. Glaser, *Lebensmittel-Wissenschaft u. Technol.*, 1990, **23**, 523.
23 R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research*, 6th edn., Longman, 1974, p. 63.
24 S. Shamil, G. G. Birch and S. Njoroge, *Chem. Senses*, 1988, **13**, 457.
25 G. G. Birch and S. Shamil, *J. Chem. Soc., Faraday Trans. 1*, 1988, **84**, 2635.
26 S. Shamil, G. G. Birch, M. Mathlouthi and M. N. Clifford, *Chem. Senses*, 1987, **12**, 397.
27 W. J. Spillane, G. Morini and G. G. Birch, *Food Chem.*, in the press.
28 (*a*) W. J. Spillane and G. McGlinchey, *J. Pharm. Sci.*, 1981, **70**, 933; (*b*) W. J. Spillane and M. B. Sheahan, *J. Chem. Soc., Perkin Trans. 2*, 1989, 741.
29 B. Unterhalt and L. Boschemeyer, *Z. Lebensm. Unter. Forsch.*, 1975, **158**, 35.
30 W. J. Spillane, G. A. Benson and G. McGlinchey, unpublished work.
31 W. J. Spillane, M. B. Sheahan, J. M. Simmie, D. Cunningham, P. McArdle and T. Higgins, *J. Chem. Soc., Chem. Commun.*, 1989, 545.
32 A. Gavezzotti, *J. Am. Chem. Soc.*, 1983, **105**, 5220.
33 H. Govers and P. de Voogt, *Quant. Str. Act. Relat.*, 1989, **8**, 11.

Paper 1/05371J
Received 22nd *October* 1991
Accepted 3rd *January* 1992