International Journal of Pharmaceutics, 21 (1984) 123–128
Elsevier

IJP 00716

Calculation of molecular surface areas and aqueous solubilities at ambient temperatures

K.C. James

Welsh School of Pharmacy, University of Wales Institute of Science and Technology, Cardiff CF1 3XF (U.K.)

(Received April 5th, 1984)
(Accepted May 2nd, 1984)

It is not possible to predict aqueous solubilities with any precision, even with the simplest molecules and the most sophisticated calculations. However, frequently all that is required is an indication of where a solubility lies at ambient temperature. A simple method for estimating approximate aqueous solubilities is therefore desirable.

Solubility can be considered in terms of the energy involved in creating an interface between solute and solvent. One such approach (Amidon et al., 1974) used Eqn. 1, in which HYSA represents the surface area of the hydrocarbon groups in the solute molecule, and FGSA is the surface area of the hydrophilic functional group. FHI has the value of zero for hydrocarbons, and of unity when the solute molecule contains a hydrophilic functional group. The coefficients of the equation have been changed for use with S.I. units (nm^2), rather than Å^2, employed in the original publication.

$$\ln \text{molal solubility} = -4.3\ \text{HYSA} - 5.86\ \text{FGSA} + 8.003\ \text{FHI} + 4.42 \qquad (1)$$

Total and functional group surface areas have been calculated by computer for a range of compounds (Herman, 1972; supplementary data to Amidon, et al., 1975). This information is used below to calculate substituent constants for the surface areas of some chemical groups. These constants can be summed to give hydrocarbon and functional group surface areas for new compounds, and substitution in Eqn. 1 provides a rapid means of estimating aqueous solubilities. Some of the group surface areas were taken directly from the supplementary data to Amidon et al. (1975). The remaining substituent constants were calculated using total surface areas (TSA) from the same source, by subtracting the total surface area of one compound from that of an identical compound, except that a hydrogen atom had been substituted by the chemical group whose surface area was required. Thus the methylene group partial

---

Correspondence: K.C. James, Welsh School of Pharmacy, University of Wales Institute of Science and Technology, P.O. Box 13, Cardiff CF1 3XF, U.K.

surface area (PSA) can be calculated from,

$$\mathrm{PSA}_{\mathrm{CH_2}} = \mathrm{TSA}_{n\text{-hexane}} - \mathrm{TSA}_{n\text{-pentane}} = 3.190 - 2.870 = 0.320$$

The value given in Table 1 is the mean from 55 pairs of compounds, comprising hydrocarbons, alcohols, aldehydes, ethers, esters, ketones and fatty acids. Methylene groups adjacent to functional groups (α-CH_2) gave substituent constants which were significantly greater, and were assigned a different value. Methyl group surface areas were obtained from calculations of the type,

$$\mathrm{PSA}_{\mathrm{CH_3}} = \left(\mathrm{TSA}_{n\text{-butane}} - 2\ \mathrm{PSA}_{\mathrm{CH_2}}\right)/2 = (2.552 - 0.636)/2 = 0.958$$

Only compounds containing normal alkyl groups were considered for the above calculations. The effect of methyl branching was assessed by comparing pairs of molecules such as 1-pentanol and 3-methyl-1-butanol, and gave a mean value of 0.126 nm^2 for 6 pairs of compounds. Molecular surface areas for compounds containing a branching methyl group were then obtained by calculating the molecu-

TABLE 1

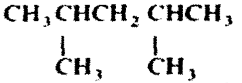
PARTIAL SURFACE AREAS FOR SUBSTITUENT GROUPS AND CONFIGURATIONAL CHANGES

Group or change	Number of results	Mean partial surface area (nm^2)	Confidence limits ($P^1 = 0.01$)
(1) Taken directly from Amidon et al. (1975)			
-OH	9	0.592	0.000
-O- in OCH_3	2	0.161	0.000
-O-	4	0.091	0.000
CHO	4	0.498	0.000
C = O in -$COCH_3$	4	0.376	0.000
C = O	4	0.265	0.000
-COOH, except formic and acetic	8	0.684	0.000
Ester -COO- in formates — except methyl	4	0.702	0.000
Ester -COO- in acetates — except methyl	4	0.490	0.000
Ester -COO- in methyl esters —	2	0.450	0.000
except formate and acetate -COO-	8	0.380	0.000
(2) Calculated from data of Amidon et al. (1975)			
CH_3-	6	0.957	0.001
-CH_2-	55	0.318	0.0003
α-CH_2-	9	0.441	0.002
Change from terminal OH			
to 2-OH	4	-0.081	0.006
to 3- or higher -OH	7	-0.104	0.000
Change from *n*-alkyl to CH_3- branching			
For each branching CH_3-	6	-0.126	0.002

TABLE 2

CALCULATION OF TOTAL SURFACE AREAS AND AQUEOUS SOLUBILITIES

(1) 2,4-Dimethylpentane



$\mathrm{CH_3\underset{CH_3}{CH}CH_2\underset{CH_3}{CH}CH_3}$

The compound is first considered to be normal, i.e. *n*-heptanol, and then corrected to allow for branching.

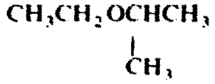
2 CH_3	$=2\times0.96$ =	1.92
5 CH_2	$=5\times0.32$ =	1.60
2 Branching methyl groups	$=2\times(-0.13)=$	-0.26
	Total	3.26 nm²

Observed molecular surface area = 3.25 nm² (Amidon et al., 1975). Substitution in eqn. 1 gives:

$\ln$ molal solubility $= -4.3\times3.26+4.42 = -9.598$

$\therefore$ molal solubility $= 6.8\times10^{-5}$ Observed molal solubility $= 4.1\times10^{-5}$

(2) Ethylisopropyl ether



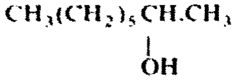
$\mathrm{CH_3CH_2O\underset{CH_3}{CH}CH_3}$

2 CH_3		$=2\times0.96=$	1.92
CH_2		$=0.32=$	0.32
2 α-CH_2		$=2\times0.44=$	0.88
Branching methyl group		$=-0.13=$	-0.13
	HYSA =		2.99
0	= FGSA =	0.09	0.09
	TSA –		3.08 nm²

Observed molecular surface area = 3.06 nm²
Substitution in eqn. 1 gives:
$\ln$ molal solubility $= -4.3\times2.99-5.86\times0.09+8.003\times1+4.42$
$= -0.916$

$\therefore$ molal solubility = 0.38 Observed molal solubility = 0.28.

(3) 2-Octanol



$\mathrm{CH_3(CH_2)_5\underset{OH}{CH}.CH_3}$

CH_3	=	0.96
6 CH_2	$=6\times0.32$	
α-CH_2	=	0.44
2-OH branching	=	-0.08
	HYSA =	3.24
OH	= FGSA =	0.59
	TSA =	3.83

Observed molecular surface area = 3.91 nm²

TABLE 2

CALCULATION OF TOTAL SURFACE AREAS AND AQUEOUS SOLUBILITIES

Substitution in Eqn. 1 gives,

$$\ln \text{molal solubility} = -4.3\times 3.24 - 5.86\times 0.59 + 8.003\times 1 + 4.42$$
$$= -4.966$$

$$\text{molal solubility} = 7.0\times 10^{-3}$$

$$\text{Observed molal solubility} = 7.9\times 10^{-3}$$

---

lar surface area of the corresponding straight-chain compound, and subtracting 0.13 (the rounded off value of 0.126).

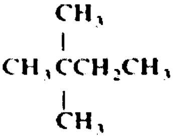
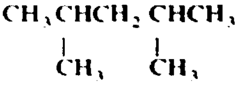
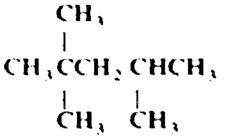
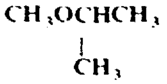
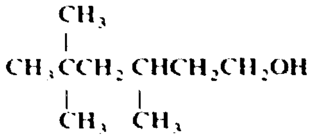
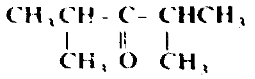
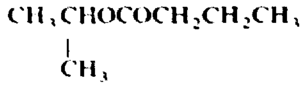
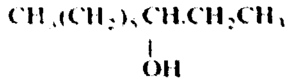
Substituent constants, together with confidence limits, and the number of pairs of compounds used to obtain each constant, are given in Table 1. Some examples of calculation of cavity surface areas are given in Table 2, and are substituted into Eqn. 1 to obtain an estimate of the aqueous solubility. Results for other compounds are given in Table 3, to show the versatility of the method. None of these compounds was used in the calculation of the partial surface areas of substituent groups given in Table 1. Rounded off values were used in the calculations, because no improvement in prediction resulted from considering the digits beyond the second decimal place.

Partial surface areas for substituent groups were calculated from the molecular surface areas of straight-chain compounds only, so that the scheme could be tested with the more complicated molecules for which data were available. The estimates obtained, considering the small numbers of compounds used to derive the partial surface areas and the rounding off of numbers, is remarkably good, and adequate for situations where an approximate idea of a solubility is all that is required. The procedure could be extended, whereby groups could be assigned different partial surface areas according to the environment in which they reside, but there are at present insufficient raw data available. It could also be refined by introducing amended forms of Eqn. 1, applicable to specific types of compounds. This approach has been attempted (Amidon et al., 1975).

An interesting outcome of this work is that estimates for solids have been found to be reasonably good. Tetradecanol, pentadeconal and hexadecanol, for example, yield estimates of 1.3×10^{-6}, 3.2×10^{-7} and 8.2×10^{-8} respectively, in comparison with observed molal solubilities of 1.5×10^{-6}, 4.5×10^{-7} and 1.7×10^{-7}. It is also interesting to note that the estimated values are lower than the observed values, so that an ideal solubility correction involving enthalpy of fusion and/or melting point would throw the estimate even further out. It has been suggested that the influence of solid liquid transition on aqueous solubilities of non-electrolytes is small and is overestimated by the ideal solubility equation (James, 1976). The examples cited here support this observation.

TABLE 3

OBSERVED AND PREDICTED SURFACE AREAS AND MOLAL SOLUBILITIES

Compound	Total surface area (nm^2)		Molal solubility	
	Observed	Predicted	Observed	Predicted
2,2-Dimethylbutane  $CH_3\overset{CH_3}{\underset{CH_3}{C}}CH_2CH_3$	2.91	2.94	6.6×10^{-4}	2.7×10^{-4}
2,4-Dimethylpentane  $CH_3\underset{CH_3}{CH}CH_2\underset{CH_3}{CH}CH_3$	3.25	3.26	4.1×10^{-5}	6.8×10^{-5}
2,2,4-Trimethylpentane  $CH_3\overset{CH_3}{\underset{CH_3}{C}}CH_2\underset{CH_3}{CH}CH_3$	3.39	3.45	7.5×10^{-5}	3.0×10^{-5}
Methylisopropyl ether  $CH_3OCH\underset{CH_3}{H_3}$	2.68	2.71	0.9	1.7
3,5,5-Trimethyl-1-hexanol  $CH_3\overset{CH_3}{\underset{CH_3}{C}}CH_2\underset{CH_3}{CH}CH_2CH_2OH$	3.77	3.84	3.1×10^{-3}	6.7×10^{-3}
2,4-Dimethyl-3-pentanone  $CH_3\underset{CH_3}{CH}-\underset{O}{\overset{\parallel}{C}}-\underset{CH_3}{CH}CH_3$	3.30	3.45	5.0×10^{-2}	5.9×10^{-2}
Isopropyl *n*-butyrate  $CH_3\underset{CH_3}{CH}OCOCH_2CH_2CH_3$	3.71	3.69	1.2×10^{-2}	1.8×10^{-2}
3-Nonanol  $CH_3(CH_2)_5\underset{OH}{CH}CH_2CH_3$	4.21	4.13	2.0×10^{-3}	1.9×10^{-3}

References

Amidon, G.L., Yalkowsky, S.H., Anik, S.T. and Valvani, S.C., Solubility of non-electrolytes in polar solvents. V. Estimation of the solubility of aliphatic monofunctional compounds in water using a molecular surface area approach. J. Phys. Chem., 79 (1975) 2239–2246 (plus supplementary data).

Amidon, G.L., Yalkowsky, S.H. and Leung, S., Solubility of non-electrolytes in polar solvents. II. Solubility of aliphatic alcohols in water. J. Pharm. Sci., 63 (1974) 1858–1866.

Herman, R.B., Theory of hydrophobic bonding. II. The correlation of hydrocarbon solubility in water with solvent surface area. J. Phys. Chem., 76 (1972) 2754–2759.

James, K.C., Free energies of solution in water of some androstanolone, nandrolone and testosterone esters. J. Pharm. Pharmacol., 28 (1976) 729–731.