

Use of solvent cavity area and number of packed solvent molecules around a solute in regard to hydrocarbon solubilities and hydrophobic interactions

(molecular surface area)

ROBERT B. HERMANN

The Lilly Research Laboratories, Division of Eli Lilly and Company, Indianapolis, Indiana 46206

Communicated by Michael Kasha, August 9, 1977

ABSTRACT Two currently used methods of calculating the area of contact between a solute molecule and the solvent are examined and some important differences are pointed out. Some pitfalls and uncertainties in the application of the concept of molecular surface area to the quantitative estimation of unitary free energies of solution are discussed.

Hydrocarbon solubility in water, a subject that is important for hydrophobic interactions, can be empirically correlated with the degree of contact between the solute molecule and the solvent. This quantity can be represented as a cavity area and, as such, has been related empirically to the logarithm of the solubility (1) and to hydrophobic interaction energies (2-10). Another quantity, the number of solvent molecules that can be packed about the solute molecule, has been related to solubility by Nemethy and Scheraga (11), Harris *et al.* (12),* and Reynolds *et al.* (13).

In analyzing the two approaches, Reynolds *et al.* (13) made the erroneous assumption that the cavity surface area as defined[†] in ref. 1 is strictly *proportional* to ΔG , the unitary free energy of solution, rather than simply *linearly* related, and as a result arrived incorrectly at a proportionality constant of 20-25 cal mol⁻¹ Å⁻².

Reynolds *et al.* (13) made relative area measurements by manually packing spheres around molecular models and proposed to convert from number of packed spheres to absolute cavity area by devising a proportionality constant based on the evaluated area for methane (1). However, the conversion cannot be based on such a single area calculation because a free energy-versus-area plot does not go through the origin as they assumed. In order to convert, one must calibrate by using incremental areas along the line and not by using a single total area as they did. For an *incremental* area, the contribution to the unitary free energy is 33 cal mol⁻¹ Å⁻² (of cavity area) for the alkanes at 298 K, as given in ref. 1.

In regard to the assumption by Reynolds *et al.* (13) that an extrapolated line through the cavity area data of ref. 1, when plotted against unitary free energy of solution, should pass through the origin, it should be noted that (i) there is as yet no theoretical justification for assuming that either linear plot when extrapolated passes through the origin and (ii) it is erroneous to ignore the geometrical distinction between the two methods of measuring cavities and assume that *both* extrapolated lines can have the origin in common. It is obvious that both methods cannot be strictly proportional to the free energy of solution because the packing of spheres itself is not strictly proportional

to area: a sphere covers a different amount of area on a cavity surface, depending on the curvature of the surface. A solvent sphere covers or is associated with a larger portion of the cavity surface in the case of smaller, more highly curved molecules so that the relationship would vary in going from smaller to larger systems. In addition, the intercept in the case of calculated cavity area varies depending on the nature of the molecular model. Reynolds *et al.* (13) used CPK models whereas Bondi radii (14) were used in ref. 1. In particular, R_{CPK} (carbon) = 1.25 Å while R_{Bondi} (carbon) = 1.7 Å.

Without further considerations of the interaction energy between solute and solvent, cavity size-versus-solubility correlations are limited to series of compounds containing similar functional groups. The cavity area can be correlated with the cavity formation energy, but the interaction energy between the solute and solvent depends on other quantities that cannot easily be related to molecular area, such as the polarizability of the solute (2). For example, the aromatic systems of ref. 1 do not fit on the same line as the saturated hydrocarbons, because systems containing the benzene nucleus are more soluble than are aliphatic systems of equal area, presumably due to a higher solute-solvent interaction energy of the benzene nucleus. Reynolds *et al.* (13) neglected a similar consideration and assumed that He and Ne should fall on the same line as the alkanes.

Finally, it is the canonical average of the surface area or average number of packed spheres taken over the distribution of available conformations that is related to solubility and not just a consideration of one conformation, when several are present in solution. The *n*-alkanes illustrate this point. If 0.8 kcal is used for a *gauche* interaction (15), calculations indicate that the extended conformations of the *n*-alkanes, which were the only ones considered by Reynolds *et al.* (13), are present to an extent of at most only 32%, 22%, and 15% for hexane, heptane, and octane, respectively, and thus are not representative.[‡]

For *n*-hexane, *n*-heptane, and *n*-octane, there are 27, 81, and 243 mathematically possible fully staggered conformations, respectively, but only 17, 41, and 99 are sterically allowable, respectively, ranging in energy from 0 for the extended con-

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

* In order to measure the degree of contact between the aqueous solvent and the solute, Harris *et al.* (12) packed spheres representing the hydrogen atoms of the solvent water molecules rather than spheres representing the water molecules themselves.

† The cavity surface as defined in ref. 1 is the surface traced out by the center of a spherical solvent molecule that is sliding over a solute molecule composed of spherical atoms.

‡ It is assumed here that the degree of association between hydrocarbon molecules in water is small and does not significantly affect the conformation distribution.

Table 1. Cavity areas and distribution of conformations for *n*-heptane at 298 K

Conformation ^a	Conformational degeneracy ^b , <i>n</i>	Area ^c	$n \cdot e^{-E/RT}$ ^d	$n \cdot e^{-(E-E_s)/RT}$ ^e
1 Extended	1	350.6	1.00	1.00
2 (5,6R)	4	344.2	1.04	1.48
3 (4,5R)	4	344.4	1.04	1.47
4 (4,5R)(5,6R)	4	335.9	0.27	0.60
5 (3,4R)(5,6R)	4	337.1	0.27	0.57
6 (3,4R)(5,6L)	4	338.0	0.27	0.54
7 (3,4R)(4,5R)	2	334.9	0.13	0.32
8 (3,4R)(4,5R)(5,6R)	4	328.5	0.07	0.24
9 (2,3R)(5,6R)	2	338.0	0.13	0.27
10 (2,3R)(5,6L)	2	336.7	0.13	0.29
11 (2,3R)(4,5R)(5,6R)	4	325.9	0.07	0.28
12 (2,3R)(4,5L)(5,6L)	4	329.6	0.07	0.22
13 (2,3R)(3,4R)(4,5R)(5,6R)	2	322.1	0.01	0.04

^a (x,yR) means a gauche interaction about bond x,y with a dihedral angle of 60°; and (x,yL) similarly implies a dihedral angle of 300°. Each gauche interaction contributes an additional energy of 0.8 kcal above the extended conformation.

^b Enantiomers are considered to be conformationally degenerate.

^c In Å² per molecule.

^d *E* represents the sum of the gauche interaction energies.

^e *E_s* represents the area-dependent solvation energy $b_0(A_1 - A_i)$ in which b_0 is 33 cal mol⁻¹ Å⁻², and A_i is the conformational area.

formations to 2.4, 3.2, and 4.0 kcal/mol for the corresponding, energetically highest conformations. The areas range from 319.0, 350.6, and 382.3 for the extended conformations of *n*-hexane, *n*-heptane, and *n*-octane, respectively, to 297.6, 322.1, and 344.3 for the most folded conformations.

Upon evaluation of the canonical average of the area taken over the distribution of available conformations as previously described (1), the average areas are found to be 311.5 Å² for *n*-hexane, 340.7 Å² for *n*-heptane, and 369.8 Å² for *n*-octane. Table 1 gives detailed results for *n*-heptane. If the average areas for these *n*-alkanes together with the average areas found for *n*-butane and *n*-pentane (1) are plotted against their (liquid) solubilities at 298 K (16), a line with a slope of 31 cal mol⁻¹ Å⁻² is obtained. A consideration of only the extended conformations would lead to a slope of 28 kcal mol⁻¹ Å⁻².

The foregoing comments apply to the proper use of the two methods of measuring area for quantitatively estimating solubility and hydrophobicity. They do not contribute to a critique of the relative usefulness of the two methods. Both methods involve difficulties in interpretation. The sphere-packing method is not representative of the way simple solvent spheres pack in a liquid but rather represents some high-density distribution of low statistical significance. The cavity area-measuring method is based on the concept of the collision diameter but has been extended nonrigorously to nonspherical systems. The relative merits of the two methods will have to be judged

by their agreement with experimental data and their ease of applicability.

- Hermann, R. B. (1972) *J. Phys. Chem.* **76**, 2754-2759.
- Hermann, R. B. (1975) *J. Phys. Chem.* **79**, 163-169.
- Hermann, R. B. (1974) *Seventh Jerusalem Symposium on Molecular and Quantum Pharmacology*, eds Bergman, E. & Pullman, B. (D. Reidel Publishing Co., Dordrecht-Holland), pp. 441-455.
- Amidon, G. L., Yalkowski, S. H., Anik, S. T. & Valvani, S. C. (1975) *J. Phys. Chem.* **79**, 2239-2245.
- Amidon, G. L., Yalkowski, S. H. & Leung, S. (1974) *J. Pharm. Sci.* **63**, 1858-1866.
- Lee, B. & Richards, F. M. (1971) *J. Mol. Biol.* **55**, 379-400.
- Chothia, C. & Janin, J. (1975) *Nature* **256**, 705-708.
- Chothia, C. (1975) *Nature* **254**, 304-308.
- Chothia, C. (1974) *Nature* **248**, 338-339.
- Shrake, A. & Rupley, J. A. (1973) *J. Mol. Biol.* **79**, 351-372.
- Nemethy, G. & Scheraga, H. A. (1962) *J. Chem. Phys.* **36**, 3401-3417.
- Harris, M. J., Higuchi, T. & Rytting, J. H. (1973) *J. Phys. Chem.* **77**, 2694-2703.
- Reynolds, J. A., Gilbert, D. B. & Tanford, C. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 2925-2927.
- Bondi, A. (1964) *J. Phys. Chem.* **68**, 441-445.
- Eliel, E. L., Allinger, N. L., Angyal, S. J. & Morrison, G. A. (1965) *Conformational Analysis* (Wiley, New York).
- McAuliffe, C. (1966) *J. Phys. Chem.* **70**, 1267-1275.