The Molecular Free Energy Surface Density (MOLFESD) approach for a quantitative treatment of hydrophobicity

J. Brickmann and R. Jäger

Institut für Physikalische Chemie and Darmstädter Zentrum für Wissenschaftliches Rechnen (DZWR), Technische Universität Darmstadt Petersenstr. 20, D-64287 Darmstadt, Germany

A method for the localization, the quantification, and the analysis of hydrophobicity of a molecule or a molecular fragment is presented which is based on the concepts that the overall hydrophobicity of a molecule can be obtained as a superposition of fragment contributions, and that the corresponding free energy can be calculated as an integral of the *molecular free energy surface density (MOLFESD)*, ρ, over the molecular solvent accessible surface [1-5]. The MOLFESD is composed as a sum of atomic increment functions. The concept can be well used for the quantification of the hydrophobic contribution to the (relative) binding constants for molecule-receptor complexes. This is demonstrated with the sweetener sucrose and its various chlorine derivatives.

The interaction of a non-polar molecule with an aqueous solvent or the quasi-attractive interaction of organic compounds or non-polar groups in the water phase is known as the *hydrophobic effect* or *hydrophobic interaction*, respectively [6,7]. Hydrophobic bonding is one of the most important forces that arise from the interaction of nonpolar regions of drugs with their receptor. From a thermodynamic point of view, the hydrophobic effect can be quantified using the solvation free energy of a molecular compound, ΔG, or the Gibbs free energy change $\Delta\Delta G = \Delta G_{\text{transfer}}$ of transfer of the compound between an aqueous and an apolar liquid phase. A similar approach can be used to study the hydrophobic interaction of two molecules in aqueous solution. Up to now, there is still no *simple* physical model available for hydrophobicity and hydrophobic interaction (for an extensive discussion see [6]).

A simple thermodynamic measure for a quantitative comparison of the overall hydrophobicity of different molecules can be obtained from the partition coefficients P or the hydrophobic index log P [8,9] of the sample in polar-apolar heterogeneous reference systems. The two phases are in equilibrium if the relation

$$\log P = \frac{\Delta\Delta\bar{G}_{\text{I,II}}}{RT} = \frac{\Delta\bar{G}_{\text{transfer}}}{RT} \qquad (1)$$

holds. The hydrophobicity of a molecule - measured, for example, by its hydrophobicity index (log P-value) - can be regarded as the sum of increments related to the hydrophobicities of its fragments, f_i,

$$\log P = \sum_i f_i \; . \qquad (2)$$

Many authors (see [9] for references) published sets of fragmental values for partition coefficients (dominantly for the n-octanol/water system). In place of these we refer to an empirical quantification of lipophilicity contributions of individual atoms which has been established by Ghose and Crippen [10]. In their treatment the atomic contributions were defined by using a classification of each atomic fragment according to the number and nature of their next (and second next) connected neighbors. The latest update of this classification lists roughly 120 constitutive atom types, describing the atoms in their individual structural environment [11]. A modification and extension of this set has been elaborated in the group of the authors [3-5].

The subdivision of molecular lipophilicity into fragmental contributions seems to be a prerequisite for the manifestation of local hydrophobicity, and a surface map of this quantity. It seems to be reasonable to postulate a distance-dependent rule for the influence of different fragments on the lipophilicity at a certain surface point. One possible function for this rule has been proposed by Audry et al. [12]. They introduced the name "molecular lipophilicity potential" (MLP), and postulated (in analogy to the electrostatic potential function) a functional form

$$V_{\text{hydrophobic}} = \sum_i f_i \, / \, (1 + d_i). \qquad (3)$$

where f_i is the partial lipophilicity of the i-th fragment of a molecule and d_i is the distance of the measured point in 3D-space from the center of fragment i.

A variety of different distance functions have been proposed by other authors (see [2] for references). Despite of

the success of the MLP approach in qualitative discussions there is no physical reason for the use of one or the other distance dependencies in this "potential" and at the same time using the partial log P-values of Ghose and coworkers as "partial charges" located at the positions of the atoms in a molecule.

In an earlier paper from the group of the authors [13] a molecular hydrophobicity mapping (MHM) approach was suggested which was particularly designed for projecting atomic increment values on the molecular surface (a solvent accessible surface or contact surface of a molecule) [13-17]. The MHM approach is well suited for a qualitative discussion of local hydrophobicity on the basis of incremental contributions because the formalism induces a weighted projection of the incremental f_i-values of Ghose and coworkers [10,11] onto the molecular surface and so allows an easy visualization of this quantity [14-17]. However, there is no way to compute quantitative data like $\Delta G_{transfer}$ or the partition coefficient P from molecular hydrophobicity "potentials" or maps. This can be done using the molecular free energy density concept (MOLFESD) which was recently developed in the group of the authors [1-5].

In this development we followed the argumentation of Richards [18-19] and coworkers, i.e. we started from the assumption that the transfer free energy $\Delta G_{transfer}$ for one mole substance from one solvent to the other can be composed of additive contributions which are related to the solvent accessible surface of the molecule. In this approach the $\Delta G_{transfer}$-value does not occur as a sum but as a surface integral over a free energy surface density ρ. The concept is based on the idea that the overall hydrophobicity index is related to the solvent accessible surface [18,19] and a quantity which measures to what extent a certain surface element contributes to the $\Delta G_{transfer}$-value per unit surface.

We restricted our considerations to rigid molecules. The central model assumption is that the total free energy $\Delta G_{transfer}$ for the transfer of one mole solute molecules from solvent I (water) to solvent II (octanol) can be represented as a surface extensive property (a surface integral)

$$\Delta G_{transfer} = \oint_0 \rho(r)\, dO \qquad (4)$$

with the local function $\rho(r)$ which may be defined as a function in the whole space but has a physical meaning only on the molecular surface. Here it represents the transfer free energy per unit surface which was termed as molecular free energy surface density (MOLFESD). Within this model approach the MOLFESD can be represented as a superposition of local contributions

$$\rho(r_s) = \sum_i \rho_i(r_s) \qquad (5)$$

where the summation is taken over atomic or molecular increments i.

It is obvious, that the global hydrophobicity index log P can be decomposed on the basis of the representation of the FESD (Eq. 5) into incremental contributions

$$\Delta G_{transfer} = RT \log P = \sum_i \Delta G_i \qquad (6)$$

with

$$\Delta G = \oint_o \rho_i(r)\, dO. \qquad (7)$$

The MOLFESD is represented as a superposition of incremental contributions $\rho_i(r_s)$ which are chosen in a similar way as the mapping functions considered earlier [13], i.e. only atomic increments were considered and it was assumed that ρ_i can be adequately modeled by [3,4]

$$\rho_i(r_s) = F_i\, \mu_i(r_s) \qquad (8)$$

with an increment specific constant F_i and a membership function [20]

$$\mu_i(r_s) = \mu\left(\left|r_i - r_s\right|; c_i, \delta_i\right) \qquad (9)$$

indicating to what extent a surface point r_s belongs to the increment i [2-4]. The normalization of the membership functions is chosen in such a way that

$$\sum_i \mu_i(r_s) = 1 \qquad (10)$$

is fulfilled for all surface points. The function $r < c$ where c is a cutoff value which is termed the *proximity distance* of an increment. For distances $r > c$ the function values rapidly tends towards zero within a distance range δ, i.e. the corresponding increment does not significantly contribute to the overall value of the MOLFESD. The function represents a weighted average of all the F_i values of atomic increments i for which $d_i < c_i$ is fulfilled, where d_i is the distance from a surface point to the center of the i-th increment. All atoms which are further away from the surface point do not contribute significantly.

The discussion of these simple expressions is based on the assumption that the proximity parameters c_i and δ_i for all atomic increment types are known and that the increments' FESD-values are well established. This establishment has been done in the group of the authors [3-5] (according to Ghose's classification scheme [10,11]) within an empirical statistical scheme where a large number of experimental data are incorporated. From equation (6) and equation (7) one obtains with the explicit form of $\rho_i(r)$ equation (8)

$$\Delta G_i = F_i O_i \qquad (11)$$

with

$$O_i = \oint_o \mu_i(r_s)\, dO. \qquad (12)$$

The expression (12) represents the contribution of the increment i to the total molecular surface O, and one simply has

$$O = \sum_i O_i \cdot \qquad (13)$$

One should note here that although the total surface is given as a sum of increment surfaces O_i the latter are not strictly local properties, i.e. a given surface point can only be associated to an increment with a certain probability. There is no sharp borderline between the individual O_i areas in contrast to the treatment of Richards et al. [18,19].

The surface increments O_i are in this representation not only dependent on properties of the increment but on the relative position of all the other increments of the molecule relative to the increment i, i.e. O_i is explicitly dependent on the molecular conformation, i.e. the same atomic increment i will, in general, contribute to the total $\Delta G_{transfer}$-value in a different way, even if the F_i-value of the increment is the same. This fact opens the possibility to include conformation of a given compound in an empirical parametrization strategy.

The FESD approach, however, is not only very useful for the prediction of unknown log P values, it can form the basis of an understanding and quantitative treatment of local hydrophobic interaction. In the integral representation of the log P value there are, in general, surface areas which contribute positively and there are such which contribute negatively. Areas of the first type are those which would prefer to be surrounded by water while those of the second type preferably would like to be surrounded by the less polar solvent. When we call the first hydrophilic and the second hydrophobic, the molecular surface can be subdivided according to these two classes. The FESD approach allows to take the surface integral over each of these areas in order to establish local contributions to the total $\Delta G_{transfer}$.

This has been demonstrated with the sweetener sucrose and its various chlorine derivatives. The present structure sweetener relationship rests on the idea that all sweet tasting molecules have a common tripartite structural element: a hydrogen bond donor, a hydrogen bond acceptor functionality, and a more or less extended hydrophobic binding area, all arranged in a triangle (the sweetness triangle) (see [21] for references). If one assumes that these binding sites contribute independently to the ΔG_i value for the binding of the sweetener to the (still unknown) receptor, one should expect a linear relationship between the sweetness and a local ΔG_i value if a sweetener is only modified at the binding site i. This is indeed found [3] in studies of the sweetener sucrose and sucralose and their modifications at the hydrophobic site. The integral over the hydrophobic part over the MOLFESD of these molecules (see Fig. 2) nicely correlates with the experimentally found sweetener (see Fig. 1).

It has been demonstrated [1-5] that the transfer free energy for bringing a molecule from one solvent (n-octanol) to another (water) can be well represented as a surface inte-
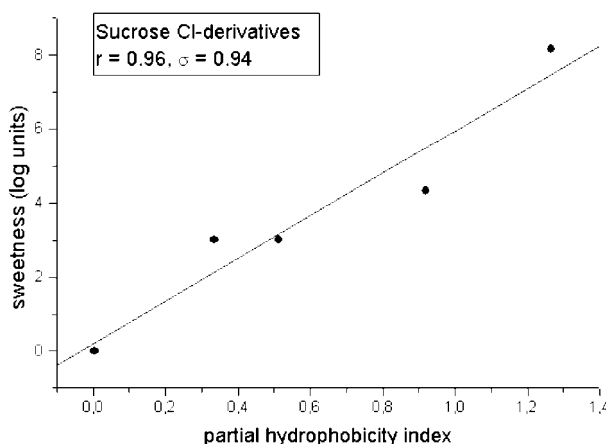

Figure 1. Sweetness of sucrose and various chlorine derivatives (as in Fig. 2) as a function of the partial hydrophobicity index (log P-value) which was calculated as a surface integral of the MOLFESD over the hydrophobic surface areas shown in figure 2.

gral over the molecular contact surface with the solvent. A molecular free energy surface density (MOLFESD) concept has been formulated on the basis of atomic increment contribution. It was shown that the concept of the MOLFESD function can be very well used for an empirical fit of parameters in such a way that the surface integral can be linearly related to the experimentally determined log P values in the case of n-octanol/water solvents. In addition, the formalism allows to study the influence of conformational changes on hypothetical log P values [1]. It has been demonstrated that the MOLFESD approach opens new possibilities in order to generate correlation parameters in QSAR studies.

Acknowledgements

References

1. Pixner, P.; Heiden, W.; Merx, H.; Möller, A.; Moeckel, G.; Brickmann, J. Empirical Method for the Quantification and Localization of Molecular Hydrophobicity, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1309-1319.
2. Brickmann, J. Localisation of Hydrophobicity, in: Software Development in Chemistry, Jochum C., Ed., GDCh-Publication, Frankfurt, 1994; pp. 139-156.
3. Schmidt, F.; Schilling, B.; Jäger, R.; Brickmann, J. Localisation and Quantification of Hydrophobicity:

The Molecular Free Energy Density (MOLFESD) Concept and its Application to the Sweetness Recognition, in preparation.

4. Jäger, R.; Brickmann, J. Empirical quantification of hydrophobicity with the MOLFESD strategy, 1998, in preparation.

5. Program package MOLFESD (MOLecular Free Energy Surface Density), Brickmann J., Darmstadt University of Technology, Germany.

6. Blokzijl, W.; Engberts, J. B. F. N. Hydrophobe Effekte – Ansichten und Tatsachen, *Angew. Chemie Intern. Edit.* **1993**, *32*, 1610-1648.

7. Tanford, C. The Hydrophobic Effect: Formation of Micelles and Biological Membranes, Wiley, New York, 1973.

8. Rekker, R. F.; Mannhold, R. Calculation of Drug Lipophilicity, VCH, Weinheim, 1992.

9. Mannhold, R.; Kubinyi, H.; Timmermann, H., Lipophilicity in Drug Action and Toxicology, VCH, Weinheim, 1995.

10. Ghose, A. K.; Crippen, G. M. Atomic Physicochemical Parameters for Three Dimensional Stucture Directed Quantitative Structure-Activity. Relationships II: Modelling Dispersive and Hydrophobic Interactions, J. Chem. Inf. Comput. Sci. **1987**, *27*, 21-35.

11. Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. Prediction of Hydrophobic (Lipophilic) Properties of Small Organic Molecules Using Fragmental Methods: An Analysis of ALOGP and CLOGP Methods, *J. Phys. Chem A* 1998, *102*, 3762-3772.

12. Audry, E.; Dubost, J.-P.; Colleter, J.-C.; Dallet, P. Une nouvelle approche des relations structure activité: Le potentiel de lipophilie moléculaire, *Eur. J. Med. Chem. - Chim. Ther.* **1986**, *21*(1), 71.

13. Heiden, W.; Moeckel, G.; Brickmann, J. A New Approach to Analysis and Display of Local Lipophilicity/Hydrophilicity Mapped on Molecular Surfaces (MLP), J. Comp. Aided Design **1993**, *7*, 503-514.

14. Waldherr-Teschner, M.; Goetze, T.; Heiden, W.; Knoblauch, M.; Vollhardt, V.; Brickmann, J. MOLCAD - Computer Aided Visualization and Manipulation of Models in Molecular Science, in: Advances in Scientific Visualization, Post F. H., Hin A. J. S. Eds., Springer, Heidelberg 1992; pp 58-67.

15. Brickmann, J.; Goetze, T.; Heiden, W.; Moeckel, G.; Reiling, S.; Vollhardt, H.; Zachmann, C.-D. Interactive Visualization of Molecular Scenarios with MOLCAD/SYBYL in: Data Visualization in Molecular Science, Tools for Insight and Innovation; Bowie, J. E., Ed., Addison-Wesley Publishing Company Inc., Reading, Mass.1995, pp. 83-97.

16. Brickmann, J. Guest Host Interaction: Model Scenarios for the Understanding of Molecular Recognition, *Pure Appl. Chem.* **1997**, *69*, 55-69.

17. Brickmann, J.; Exner, T.; Keil, M.; Marhöfer, R.; Moeckel, G. Visualization of Molecular Models. Encyclopedia of Computational Chemistry, Ragué Schleyer P. von, Schreiner P. R. Eds., Wiley J., 1998, in press.

18. Richards, N. G. J.; Williams, Ph. B.; Tute, M. Empirical Methods for Computing Molecular Partition Coefficients. I. Upon the Need to Model the Specific Hydration of Polar Groups in Fragment-Based Approaches. *Int. J. Quantum Chem. Quantum Biol. Symp.* **1991**, *18*, 299-316.

19. Richards, N. G. J.; Williams, Ph. B.; Tute, M. Empirical Methods for Computing Molecular Partions Coefficients: II. Inclusion of Conformational Flexibility within Fragment-Based Approaches, *Int. J. Quantum Chem.* **1992**, *44*, 219-33.

20. Zimmermann, H. -J. Fuzzy set theory and its applications, Kluwer, Boston, 1991.

21. Lichtenthaler, F. W.; Immel, S. Computer simulation of chemical and biological properties of sucrose, the cyclodextrins and amylose, *Int. Sugar J.* **1995**, *97*, 12-22.