3D molecular lipophilicity potential profiles: a new tool in molecular modeling

P. Furet, A. Sele and N.C. Cohen*

Ciba-Geigy Ltd., Pharmaceuticals Division, R-1058.P.44 CH-4002, Basel, Switzerland

This paper proposes a three-dimensional molecular representation visualizing in the same frame the hydrophobic and geometric properties of a molecule. Based on the concept of molecular lipophilicity potential, the representation consists of the variations of a probability function PF displayed on the molecular van der Waals surface. The values taken by this function reflect the degree of lipophilicity or hydrophilicity of the different parts of the molecule. The utility of the new representation is illustrated by a study of the common amino-acid residue hydrophobicity. The results obtained suggest that this approach has the potential to become a useful tool in areas such as molecular modeling and drug design, where molecular properties are considered in a 3D perspective.

Keywords: *hydrophobicity, molecular modeling, drug design, amino acids, lipophilicity potential*

Received 16 July 1988
Accepted 2 August 1988

Hydrophobic interactions, together with hydrogen bonding and electrostatic interactions, constitute the main forces involved in the folding of proteins and in ligand-biological receptor (-enzyme) binding associations. But, unlike the two latter types of forces, hydrophobic forces are neither directional nor do they correspond to strong interactions between individual atomic or chemical groups. However, their number is usually large, and although their precise contribution to the total stabilization energy of a ligand-macro molecule complex or of a protein in water is difficult to assess, there is ample evidence that they add up to significant amounts of energy and produce strong overall stabilizing interactions.[1] Hydrophobicity is therefore regarded as an important molecular property in the study of protein folding and in the search of relationships between chemical structure and biological activity, as illustrated by the considerable effort that has been devoted to the subject in these two areas of research during the last 20 years (for reviews, see references 2 and 3).

Still, despite this wealth of studies, very little progress has been made in devising methods for characterizing the hydrophobicity of a molecule or a chemical fragment that does more than represent this molecular property by a single number. The parameter usually used is logP, the partition coefficient of the molecule in the water-octanol system, or, in the case of chemical substituents, Hansch's π constant derived from logP data. This simple "one-dimensional" representation of hydrophobicity, although appropriate for classical quantitative-structure-activity-relationships (QSAR) methods, reflects only an overall property. It becomes insufficient when more detailed insight is needed, particularly in current molecular modeling studies where precise three-dimensional (3D) structural and stereochemical features of molecules are analyzed in the context of intermolecular interactions with macromolecular biological systems (receptors, enzymes, etc.).

Recently, the concept of "hydrophobic dipole moment," which goes beyond the simple logP characterization of hydrophobicity, has been introduced by Eisenberg and McLachlan.[4] This quantity was proposed as a measure of the asymmetry of molecular hydrophobicity, the so-called amphiphilicity, and was applied to a study of the amino-acid residue hydrophobicity. According to this idea, the hydrophobicity of a molecule can be characterized by two indexes. One (logP or a related constant) gives the overall magnitude of the property. The other is a measure of the extent of hydrophobic polarization existing in the molecule and reflects the fact that molecules are made of more or less polar and apolar moieties contributing differently to the overall hydrophobicity of the molecule.

In the present article, the above approach is extended by proposing a tridimensional molecular representation that makes it possible to visualize in the same frame the 3D geometrical features of a compound and a detailed description of its hydrophobic characteristics. Taking into consideration that hydrophobicity is in fact a property that is distributed all over the different parts of a molecule, the treatment allows one to reveal the

*To whom correspondence should be addressed

lipophilic and hydrophilic regions on the van der Waals molecular surface and also provides semiquantitative information concerning the degree of hydrophobicity of each individual area. The new representation is based on the concept of molecular lipophilicity potential introduced by Audry *et al.*[5] and can be viewed as the pendant for hydrophobic properties to the now classical electrostatic potential molecular surface representations, which have proven to be valuable tools in the analysis of the electronic properties of molecules.[6] As such, it has potential utility in molecular modeling and drug design for further revealing the 3D mimetism of molecules exhibiting a common activity, in addition to the more usual geometrical and electrostatic aspects, or for assessing the extent of hydrophobic complementarity between a compound and a receptor binding site.

Being based on a not well-defined chemical fragment formalism, in its original form the concept of molecular lipophilicity potential is of limited practical use. In general, several decompositions into chemical fragments are possible for a molecule, and there is some arbitrariness in the choice of one particular decomposition. To circumvent this problem and render the concept widely applicable, the present authors envisaged a modification in an atomic formalism suppressing the ambiguity associated with molecular fragments. But this required hydrophobicity atomic parameters that were not available at that time. However, very recently Ghose and Crippen[7] have derived such atomic parameters from experimental logP data for QSAR purposes not directly related to the approach described here. This progress allowed us to proceed along the lines mentioned above.

In the following sections, the definition of the lipophilicity potential is first recalled, then the new 3D molecular representation is presented and illustrated by its application to characterize the hydrophobicity of the side chains of the 20 naturally occurring amino acids.

THE MOLECULAR LIPOPHILICITY POTENTIAL

The molecular lipophilicity potential is defined by considering a molecule A surrounded by organic solvent molecules of low polarity and assuming that its overall lipophilicity — measured, for example, by logP — can be decomposed into fragmental or atomic contributions represented by parameters f_i whose values measure the individual hydrophobic characters of the fragments or atoms:

$$\text{logP} = \sum_i f_i \tag{1}$$

Intuitively, the arrangement of the solvent molecules around A is expected to vary from a random distribution at far distances to more ordered states as one gets closer to A, due to the respective tendencies of the lipophilic and hydrophilic parts of the molecule (measured by the f_i's) to attract or repel the solvent molecules. Thus, according to this picture, the distribution of the solvent molecules around A can be considered as depending on the hydrophobic fragmental or atomic contributions f_i and the distances at which one is from the corresponding fragments or atoms forming A. To express this idea in a quantitative manner, Audry *et al.* have introduced the notion of "molecular lipophilicity potential" (MLP), which is defined in a point M of the space surrounding A as follows:

$$\text{MLP} = \sum_i f_i/(1 + d_i) \tag{2}$$

where d_i is the distance between M and fragment or atom i (as proposed in the present approach).

The above formula accounts for the fact that the solvent molecules around A experience a hydrophobic force whose attractive and repulsive components originate from interactions with the lipophilic ($f_i > 0$) and hydrophilic ($f_i < 0$) groups of A, respectively, and whose magnitude is maximal when the d_i's are equal to zero and progressively vanishes as these distances get larger. Given a set of f_i values, it becomes therefore possible with Equation (2) to quantitatively characterize the hydrophobic properties of a molecule by the "hydrophobic force field" that it creates in its surroundings, much as the electronic properties of this molecule can be characterized by the electrostatic force field created by its charge distribution. In this context, the MLP appears to be a natural extension of the concept of hydrophobic dipole moment previously mentioned and defined by analogy to the electrostatic dipole moment as:

$$\vec{\mu} = \sum_i f_i \vec{r}_i \tag{3}$$

where the f_i's have the same meaning as in Equation (2) and $\vec{r}_i$ is the vector joining a given origin to atom of fragment i.

One should notice at this point that the MLP constitutes an heuristic notion, conceived to give a 3D extension to the simple logP representation of hydrophobicity. As such, other mathematical forms than Equation (2) may well be envisaged for its expression. The only constraints a proper formula should satisfy concern the boundary conditions and the physical content that the MLP must continuously decrease as the distances d_i increase. Various expressions were considered in the present work, but it was found that they offered no particular advantage over Equation (2), which appears to be of good information content and correctly conveys chemical intuition and current ideas concerning hydrophobicity as shown in the next sections.

It is worth stressing two advantages of an atomic-level definition of the MLP over a chemical fragment definition that are apparent looking at the form of Equation (2):

- By increasing the number of terms in the summation forming the right hand side of the formula, an atomic level definition provides a more precise description of molecular hydrophobicity.
- The distance d_i between an atom and a point M of space is a more natural notion than the distance between a chemical fragment and M, which is a somehow ambiguous notion.

3D MOLECULAR LIPOPHILICITY POTENTIAL PROFILES

Having a mathematical formula to express MLP — a quantity proposed as a measure of the hydrophobic force exerted by a molecule in its surroundings — the variations of MLP in the space around the molecule can be computed and visualized in order to obtain a 3D representation of the hydrophobic features of the molecule. But before presenting this aspect, a discussion concerning the atomic hydrophobicity parameters f_i appearing in Equation (2) is desirable.

Atomic hydrophobicity constants f_i

Hydrophobic bonding is a complex force resulting primarily from entropic effects related to the change in the organization of the solvent molecules around the solute ones. It does not lend itself to a simple theoretical treatment, unlike the electrostatic or polarization forces. Although a very promising theoretical method, namely the thermodynamic cycle-perturbation method,[8] has recently emerged for computing free energy differences (and thus related quantities like partition coefficients, equilibrium constants, etc.), to our knowledge no method exists yet to access atomic hydrophobicity constants such as the f_i's by theoretical calculations as one would determine electronic atomic charges from population analyses on *ab initio* or semiempirical wave functions. Therefore, one has to resort to experimental data to derive values for the f_i's.

Assuming, as usual, that logP is a measure of the overall hydrophobicity of a compound, values for the parameters f_i can be obtained by postulating Equation (1), which implies that logP is an additive property. Starting from Equation (1) in conjunction with the logP data for about 500 representative organic molecules and using least-square techniques, Ghose and Crippen[7] have succeeded in deriving atomic f_i values for the most common atoms encountered in organic chemistry. They have classified the atoms in 110 different types according to two principal topological criteria: the number and nature of the atoms directly connected to the atom under consideration. These criteria were chosen for their relevance to control the electron distribution around the atom and the approachability of the solvent to the atom, the main factors considered to influence atomic hydrophobicity.

Eisenberg and McLachlan[4] took a similar approach to the evaluation of atomic hydrophobicity constants in connection with the development of a method for calculating the solvation energy of proteins. These authors assumed that the free energy of transfer of a given amino-acid residue from the interior of a protein to aqueous solution (taken as a measure of its hydrophobicity) could be written as a sum of atomic contributions proportional to the solvent-accessible surface areas A_i of the atoms forming the residue:

$$\Delta G_R = \sum_i \Delta \sigma_i A_i \qquad (4)$$

and by a fit to experimental logP data of small amino-acid residue analogues, the authors were able to obtain values for the proportionality constants $\Delta\sigma_i$ within a classification comprising five type of atoms: carbon, neutral oxygen and nitrogen, charged oxygen (O^-), charged nitrogen (N^+) and sulphur.

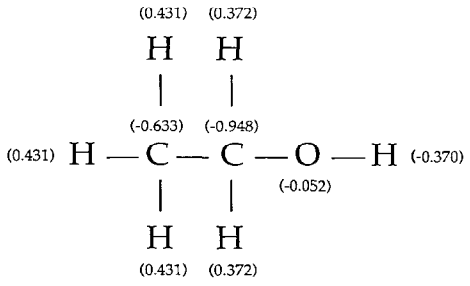
For the 3D lipophilicity potential representation described in the present work, the atomic hydrophobicity constants of Ghose and Crippen have been adopted and a computer program has been written for the automatic recognition of the class to which each atom of a molecule belongs and the subsequent assignment of the corresponding f_i value. The parameters of Ghose and Crippen were chosen because they offer two major advantages in the present application over those of Eisenberg and McLachlan: They are applicable to all organic molecules, not only to peptides, and their number is 110, versus 5 for the latter parameters. They constitute, therefore, a much more precise description of atomic hydrophobicity in which subtle environmental features of the atoms are differentiated. In particular, they include values for hydrogen atoms, whereas Eisenberg and McLachlan used a united atom classification.

However, it should be mentioned that in the approach of Ghose and Crippen, the accessibility of the solvent to the atoms is only implicitly taken into account on a topological basis, while in that of Eisenberg and McLachlan, this factor is explicitly considered through the solvent-accessible surface areas A_i present in Equation (4), which are quantities depending on the 3D coordinates of the atoms. Thus, by using the former approach, the dependence of the hydrophobicity parameters f_i on the conformation of the molecule is lost.

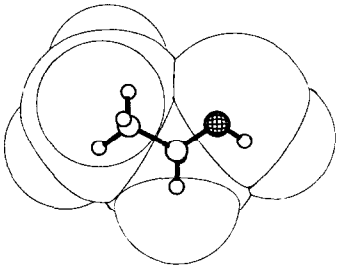
The representation

The MLP can be visualized in different manners. These include 2D isopotential contour maps,[5] which are only really appropriate for planar molecules or color-coding the values of the potential calculated on a Connolly molecular surface. In this work, the MLP is visualized using the "four-dimensional" representation of molecules proposed by Cohen,[9] which consists of a molecular envelope on whose surface isopotential contour lines are drawn. This representation provides an efficient way to visualize in the same frame the 3D geometrical features of a compound (steric volume, van der Waals surface, solvent-accessible surfaces, etc.) and any molecular property that can be expressed mathematically in the form of a function of the 3D coordinates of a point of space. The corresponding computer program VISPOT, originally conceived for visualizing the molecular electrostatic potential, has been adapted to that of the MLP.

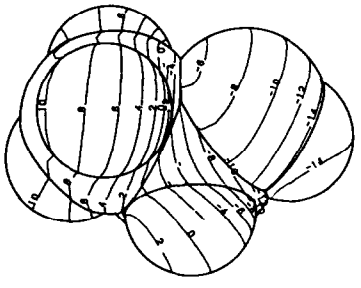
An example of representation obtained with the resulting program LIPOT is given in Figure 1 for the ethanol molecule, along with an illustration of the process of constructing the representation. The molecular envelope represented here is the van der Waals surface of the molecule, although the isopotential contour lines can also be calculated and displayed on more expanded surfaces, such as the solvent-accessible surfaces. One observes in this figure a decrease from 10 to −14 in



1



2



3

Figure 1. Construction of a 3D molecular lipophilicity potential representation, illustrated here with the ethanol molecule.

Step 1: Assignment of the f_i values on the basis of the toplogical structure of the molecule logP = $\sum_i f_i$

Step 2: Construction of the 3D geometrical structure of the molecule (van der Waals surface)

Step 3: Calculation and visualization of isovalue contour lines of the MLP on the van der Waals surface

the values of the MLP, reflecting the expected gradual decrease in hydrophobicity, as one goes from the methyl part of the molecule to the hydroxyl one: the more hydrophobic a region, the greater the MLP. This demonstrates the usefulness of the representation for providing a detailed quantitative characterization of the hydrophobicity of a molecule.

Experience has shown that a more useful representation can be obtained by visualizing, instead of the actual MLP, contour lines corresponding to the function expressed by:

$$PF = 100 \times (PL - \tfrac{1}{2}) \tag{5}$$

where PL is a probability function connected to the MLP by the following equation:

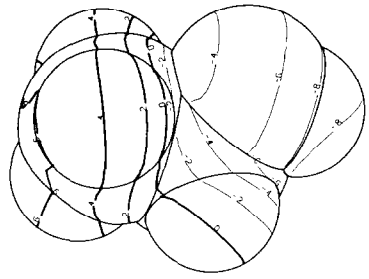


Figure 2. Isovalue contour lines of the PF function defined by Equation (5) for the ethanol molecule. Thick and thin lines correspond to positive and negative values of PF, respectively

$$PL = 10^{MLP}/(1 + 10^{MLP}) \tag{6}$$

According to Audry *et al.*,[5] the value of PL in any point of the space surrounding a molecule placed in a solvent of low polarity can be given the direct physical meaning of representing the probability that a lipophilic solvent molecule exists at this very point. The PF function was conceived to obtain a quantity that takes positive values for lipophilic regions of the molecular surface (probability PL greater than 0.5), while negative values are taken for hydrophilic regions (probability less than 0.5). This is illustrated in Figure 2 for ethanol, where one can distinguish two well-defined lipophilic (positive values = thick lines) and hydrophilic (negative values = thin lines) regions of the molecular envelope corresponding, respectively, to the methyl and hydroxyl groups of the molecule. This figure clearly shows that the PF representation not only provides qualitative information about the repartition of lipophilic and hydrophilic zones in the molecule, but also quantitatively indicates the degree of lipophilicity or hydrophilicity of each region as measured by the intensity of the isoprobability contour lines.

Other views concerning some simple common organic compounds are given in Figure 3 to further illustrate the PF representation, which will be referred to as an MLP profile. One can notice that when a molecule has a marked overall lipophilic (hydrophilic) character corresponding to a significantly positive (negative) value of logP, only positive (negative) PF contour lines appear. Taking up again the analogy with the molecular electrostatic potential, the hydrophobicity constants f_i can be compared to net atomic charges and logP to the total charge of the molecule (Equation (1)). In this context, the MLP profiles of molecules whose logP value differs significantly from zero can be compared to the electrostatic potential profiles of anions or cations. In such cases, which are the most frequently encountered, the profiles are useful in that they visualize where the less and most lipophilic (hydrophilic) parts of the molecule are located.

MLP PROFILES OF THE 20 NATURALLY OCCURRING AMINO-ACID RESIDUES

Because hydrophobic interactions play an important role in the energetics of protein folding in water, the hydro-

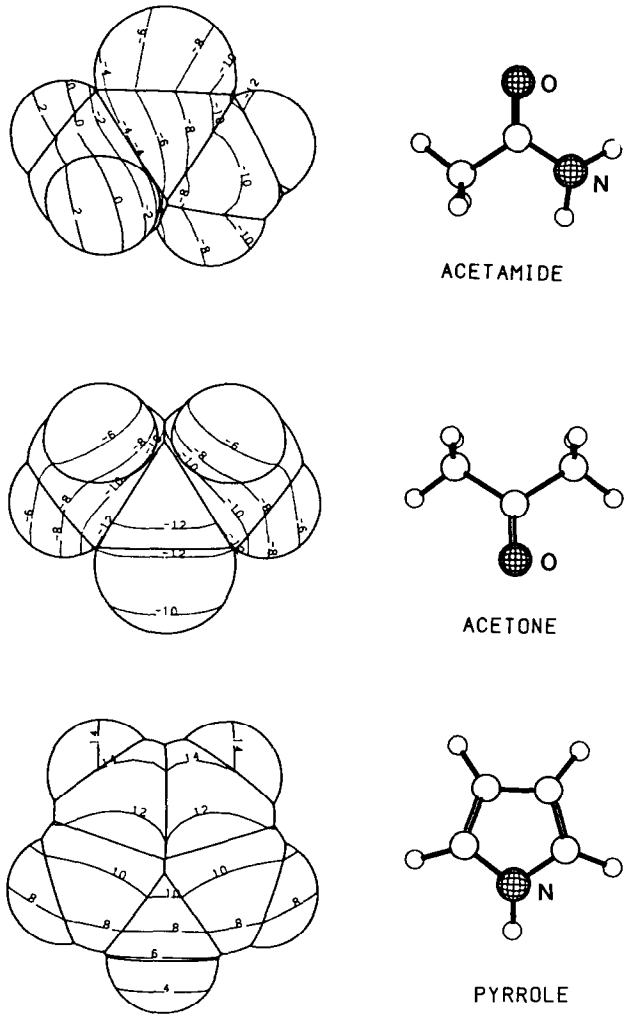


Figure 3. MLP profiles of acetamide, acetone and pyrrole. The corresponding dreiding models are represented at the right of the profiles. Acetone, which has a negative logP value (calculated value of −0.42), presents a profile with only negative PF contour lines. The profile of pyrrole reflects the overall lipophilic character of this molecule (calculated logP value of 0.67) and indicates that the less lipophilic part of the molecule is the N–H region (PF values of 4–6)

phobicity of the common amino acids has been the subject of numerous studies. Several residue hydrophobicity scales have been proposed on the basis of various approaches (for a review, see Reference 10). As mentioned in the introduction of this article, common to most of these studies is the characterization of the residue hydrophobicities by a single constant. As an illustration of the new molecular representation and as an attempt to give more complete descriptions of the hydrophobic characters of the amino-acid residues, their 3D MLP profiles are now presented. These are shown in Color Plates 1–20.

For each residue, two profiles are given. The first one concerns the side-chain fragment in which the α-carbon attachment has been simply replaced by a hydrogen atom. For instance, the corresponding molecule in the case of ALA is CH_4. This profile is assumed to reflect the intrinsic hydrophobic properties of the residue. The second one is the profile of the N-acetyl-amino-acid amide analogue of the residue: $CH_3CO-NH-CH(R)-CONH_2$ (R = given residue). It is supposed to represent the hydrophobicity of the residue in a peptidic or protein environment. The molecule is represented in its lowest energy conformation as determined by Vasquez *et al.*[11] using the ECEPP/2 molecular mechanics method. A comparison of the two types of profiles permits an appreciation of the effect of the peptidic chain on the intrinsic hydrophobicity of the residue. It should be stressed in this regard that a valuable feature of the MLP representation is its ability to account for the influence of the neighboring atoms or groups of atoms on the hydrophobicity of any particular region of a molecule. This is in contrast to a representation in which the atoms would be color-coded solely on the basis of the value of their hydrophobicity constant f_i, without making use of Equation (2).

To give an idea of the reliability of the atomic hydrophobicity parameters Ghose and Crippen used for constructing the profiles, the logP values of the 20 N-acetyl-amino-acid amides calculated with Equation (1) are compared to their corresponding experimental values, as measured by Fauchère and Pliska[12], in Table 1. Also included in the table are the logP values computed with the fragmental methodology of Hansch and Leo[13] (CLOGP program), which have been recently published by Abraham and Leo.[14] Excluding the four ionized residues GLU, ASP, LYS and ARG (see footnote of Table 1), the atomic constants of Ghose and Crippen reproduce the experimental logP values with an average deviation of 0.44. This compares well with the results obtained with the more conventional chemical fragment approach of Hansch and Leo, the corresponding average deviation being 0.46. With both methods, the same trends are observed. While the calculated values are generally in good agreement with experimental data for the apolar residues, the agreement is poorer for the polar ones, the calculated logP values of these residues being consistently too negative. Both methods also overestimate by more than one logP unit the hydrophobicity of the N-acetyl-cysteine amide.

Some possible explanations for the discrepancies observed for the polar residues are discussed in the article of Abraham and Leo in the context of the Hansch approach. These include the possibility that polar residues are transferred to octanol in the form of hydrates, which makes the molecules more hydrophobic than one would predict on the basis of calculations concerning the nonhydrated form of the molecules. This remains to be established, but the fact that both methods present the same shortcomings for the series of compounds examined here might be due to the use of the logP data of the same, not enough representative, set of compounds to derive the atomic and fragmental constants. It might also mean an inherent limitation in the calculation of logP as an additive property or a lack of sophistication in the methods. In any case, the method of Ghose and Crippen has been shown to work well for

Table 1. Calculated and observed logP values of the N-acetyl-amino-acid amides

Amino acid	logP obs.[a]	logP calc.[b]	Δ[d]	logP calc. (CLOGP)[c]	Δ[d]	Degree of hydrophobic polarization[e]
TRP	0.42	0.17	−0.25	−0.12	−0.54	4
PHE	−0.04	0.25	0.29	−0.12	−0.08	4
ILE	−0.03	−0.19	−0.16	−0.23	−0.20	2
LEU	−0.13	−0.19	−0.06	−0.23	−0.10	2
CYS	−0.29	−1.68	−1.40	−1.57	−1.28	6
MET	−0.60	−0.86	−0.26	−1.21	−0.61	4
VAL	−0.61	−0.65	−0.04	−0.76	−0.15	0
TYR	−0.87	−0.02	0.85	−0.78	0.09	10
PRO	−1.34	−1.04	0.30	−1.03	0.31	12
ALA	−1.52	−1.47	0.05	−1.68	−0.16	0
THR	−1.57	−2.46	−0.89	−2.31	−0.74	10
GLY	−1.83	−1.89	−0.06	−1.99	−0.17	–
SER	−1.87	−2.75	−0.88	−2.62	−0.75	6
HIS	−1.70	−2.18	−0.48	−2.31	−0.61	20
GLN	−2.05	−2.52	−0.47	−3.10	−1.11	18
ASN	−2.41	−2.98	−0.57	−2.97	−0.56	8
GLU	−2.47	−2.66[f]	−0.19	–	–	20
ASP	−2.60	−3.12[f]	−0.52	–	–	14
LYS	−2.82	−2.46[f]	0.36	−3.77	−0.95	32
ARG	−2.84	−3.44[f]	−0.60	−5.01	−2.17	30
Average deviation[g]:			0.44		0.46	

(a) Experimental values from Ref. 12
(b) Values calculated with the method of Ghose and Crippen
(c) Values calculated with the method of Hansch and Leo. From Ref. 14
(d) Δ = logP calc. − logP obs.
(e) Defined as the absolute value of the difference between the maximum and the minimum value of the PF contour lines in the corresponding R–H profile. The greater this number, the more polarized the residue
(f) These residues were considered in their ionized form. Since at the present time, the classification of Ghose and Crippen does not provide f_i values for the charged species, the O^- and N^+ $\Delta\ \sigma_i$ parameters of Eisenberg and McLachlan[4] were used to obtain f_i values for the charged oxygens and nitrogens present in these residues. The logP values appearing in the table were calculated using these f_i values, which are equal to −1.51 and −0.74 for N^+ and O^-, respectively. While there is some inconsistency in mixing two different classifications, the reasonable agreement obtained with the experimental logP values can justify such a procedure
(g) Mean value of absolute value of Δ, excluding the ionized residues, which were calculated with a nonhomogeneous method

many molecules,[7] and this is confirmed here for a number of N-acetyl-amino-acid amide compounds.

Having discussed the reliability of the constants f_i used in this study, the MLP profiles of the amino acids presented in Color Plates 1–20 can now be analyzed. The following observations can be made:

R–H profiles: intrinsic amino-acid side-chain hydrophobicities

- ALA, VAL, LEU, ILE, TRP, PHE, TYR, MET and CYS are predicted to be lipophilic residues, only red (positive) isoprobability contour lines appearing on their van der Waals surface. For the aromatic residues PHE, TYR, TRP, the region below the center of the ring on the $-CH_3$ substituent side appears to be the most lipophilic, while for the MET and CYS residues the most lipophilic region is that corresponding to the van der Waals surface of the sulphur atom. The decrease in lipophilicity brought up by the hydroxyl substituent is apparent when comparing the values of the contour lines of the PHE and TYR residue profiles.
- PRO, THR, ASP, GLN, GLU, HIS, LYS and ARG are predicted to be amphiphilic residues. One portion of the molecule corresponding approximately to the hydrocarbon moiety is lipophilic, while the other portion containing the polar moiety is hydrophilic.
- SER and ASN are predicted to be hydrophilic residues (only blue (negative) contour lines). In the case of these two residues, the lipophilicity of the methyl moiety is not sufficient to offset the hydrophilicity of the polar alcohol and amide functions. This is in contrast to the ethyl analogues THR and GLN, which appear as amphiphilic.
- Taking as a rough measure of the degree of hydrophobic polarization of a residue the absolute value of the difference between the maximum and the minimum value of the PF contour lines (see values reported in Table 1), the profiles indicate that ARG, LYS, GLU, GLN and HIS are the most polarized residues, which reflects their pronounced amphiphilic

character, while the lipophilic ALA, VAL, LEU and ILE residues are the less polarized. This is in agreement with the results obtained by Eisenberg and McLachlan,[4] using the concept of hydrophobic dipole moment as defined by Equation (3).

CH_3CO–NH–CH(R)–$CONH_2$ profiles: effect of the peptidic chain

- The intrinsic hydrophilicity of the peptidic chain is revealed by the profile of the N-acetyl-glycine-amide, which consists of only blue contour lines.
- For all residues, the peptidic linkage causes a decrease in their hydrophobicity. This effect can be appreciated by comparing the intensity of the contour lines in the R–H profile with that in the corresponding portion of the N-acetyl-amide profile. The residues that were amphiphilic in the R–H representation become purely hydrophilic. Only the most intrinsically hydrophobic residues like PHE, TRP or ILE conserve their identity as lipophilic residues. The peptidic chain renders the VAL and MET residues amphiphilic, while the small ALA residue turns hydrophilic.
- This points to a second effect: the peptidic environment polarized the hydrophobicity of the nonpolar residues. This is clearly demonstrated by the increase in the number of contour lines going from the R–H profile to the N-acetyl-amide profile for a residue like VAL, for example.
- Conversely, intrinsically lipophilic residues cause a decrease in the hydrophilicity of the peptidic chain. This effect is most apparent for TRP in comparison with the profile of GLY.

One question that naturally arises concerns the validity of the picture of the amino-acid residue hydrophobicity given by the above MLP profiles. In this respect, one should first recall that a profile is correct within the limits of the ability of the constants f_i used for its construction to reproduce the measured logP value of the molecule under consideration. Thus, in their N-acetyl-amide profile representations, the polar residues are certainly portrayed as too hydrophilic, which is a consequence, as discussed earlier, of the inability of the additive schemes of estimation of logP to correctly handle such molecules in which several polar functional groups are in close proximity.[15]

A second point is that the MLP is a purely heuristic notion with no theoretical basis. There is, therefore, no rigorous way to check the validity of the representation it gives of the hydrophobicity of a compound. However, one can notice that overall the MLP profiles offer a plausible description of the hydrophobicity of the amino acids that is in agreement with chemical intuition and common ideas regarding how hydrophobicity is distributed over the different polar and nonpolar parts of a molecule. For instance, it is satisfying to obtain in the R–H profile of glutamine (GLN) two distinct lipophilic and hydrophilic zones corresponding, respectively, to the ethyl and amide portion of the residue. The fact that the N-acetyl-amide profiles of the nonpolar residues present a hydrophilic region corresponding approximately to the polar peptidic chain while the side chain region is lipophilic is another example of full consistency with what is expected from chemical sense. In addition, the MLP profiles account for the expected dependence of the hydrophobicity of the amino-acid side chains on their environment. For a given residue R, significant differences are observed in the regions corresponding to R when comparing the R–H and CH_3CO–NH–CH(R)–$CONH_2$ profiles, which reflects the polarizing effect of the peptidic chain particularly sensitive for the most lipophilic residues.

CONCLUSION

The examples presented in this article show that the MLP profiles constitute an interesting approach for giving a 3D content to the simple logP characterization of molecular hydrophobicity. In fact, to the authors' knowledge, this approach can be considered as the first attempt made so far at representing hydrophobicity as a function varying continuously over the different parts of a molecule. In more conventional approaches, hydrophobicity is simply partitioned in a discrete manner into values characterizing chemical fragments or atoms. One of the most distinctive features of the MLP representation lies in its ability to provide a way to account for the influence of the atomic environment on the hydrophobicity of any particular fragment of a compound. This may make the representation useful to study substituent effects.

Another important feature of the MLP profiles is that they provide a visualization of the 3D geometrical characteristics of a molecule. By combining a representation of the geometric and hydrophobic aspects of molecular structure, they have the potential to become a useful tool in molecular modeling and drug design for the analysis of the 3D mimetism of molecules from the point of view of hydrophobicity. They can also find applications in ligand-macromolecule docking studies to analyze the complementarity of the ligand with the active site in terms of shape and hydrophobicity. For instance, one can envisage comparing the MLP created by the ligand on its solvent-accessible surface to the MLP created by the macromolecule on this same surface in a similar fashion to what has been done by Nakamura *et al.*[16] for the electrostatic potential. Analysis of these profiles would suggest ways to improve the complementarity. Studies along these lines have been undertaken by the authors to explore the utility of such an approach and are currently in progress.

Finally, it is worth stressing that the method to construct a 3D MLP profile is simple and does not involve large computation times. To generate a profile, one needs only the 3D coordinates of the atoms forming the molecule and the atomic hydrophobicity parameters f_i as input.

REFERENCES

1 Tanford, C. *The Hydrophobic Effect*. Wiley, New York, 1980
2 Rose, G. D., Gierasch, L. M., and Smith, J. A. *Advan. Protein Chem.* 1985, **37**, 1–109
3 Van de Waterbeemd, H., and Testa, B., in *Advances in Drug Research*, Ed. B. Testa. Academic Press, London, 1987, **16**, 85–225
4 Eisenberg, D., and McLachlan, A. D. *Nature* 1986, **319**, 199–203
5 Audry, E., Dubost, J. P., Colleter, J. C., and Dallet, P. *Eur. J. Med. Chem.–Chim. Ther.* 1986, **21**, 71–72
6 See, for example: Weiner, P.K., Langridge, R., Blaney, J. M., Schaeffer, R., and Kollman, P. A. *Proc. Natl. Acad. Sci. USA* 1982, **79**, 3754–3758 and references therein
7 Ghose, A. K., and Crippen, G. M., *J. Comp. Chem.* 1986, **7**, 565–577
8 Review: McCammon, J. A. *Science* 1987, **238**, 486–491
9 Cohen, N. C. *ACS Symp. Ser.* 1979, **112**, 371–381
10 Eisenberg, D. *Ann. Rev. Biochem.* 1984, **53**, 595–623
11 Vásquez, M., Némethy, G., and Scheraga, H. A. *Macromol.* 1983, **16**, 1043–1049
12 Fauchère, J. L., and Pliska, V. *Eur. J. Med. Chem. — Chim. Ther.* 1983, **18**, 369–375
13 Hansch, C., and Leo, A. J. *Substituent Constants for Correlation Analysis in Chemistry and Biology*. Wiley, New York, 1979
14 Abraham, D. J., and Leo, A. J. *Proteins* 1987, **1**, 130–152
15 Roseman, M. A. *J. Mol. Biol.* 1988, **200**, 513–522
16 Nakamura, H., Komatsu, K., Nakagawa, S., and Umeyama, H. *J. Mol. Graph.* 1985, **3**, 2–11