Colloid Polym Sci (2006) 284: 575–585
DOI 10.1007/s00396-005-1447-6

ORIGINAL CONTRIBUTION

Ivan M. Okhapkin
Andrei A. Askadskii
Vladimir A. Markov
Elena E. Makhaeva
Alexei R. Khokhlov

Two-dimensional classification of amphiphilic monomers based on interfacial and partitioning properties. 2. Amino acids and amino acid residues

Received: 22 September 2005
Accepted: 9 December 2005

I. M. Okhapkin · A. A. Askadskii
Nesmeyanov Institute
of Organoelement Compounds,
Vavilov Street 28,
119991 Moscow, Russia

I. M. Okhapkin · A. R. Khokhlov (✉)
Department of Polymer Science,
University of Ulm,
Albert-Einstein-Allee 11,
89069 Ulm, Germany
e-mail: khokhlov@polly.phys.msu.ru
Tel.: +7-095-9392959
Fax: +7-095-9392988

V. A. Markov · E. E. Makhaeva ·
A. R. Khokhlov
Physics Department,
Moscow State University,
119992 Moscow, Russia

Abstract A new classification for amino acid residues is proposed, based on interfacial and partitioning properties. Amino acid residues are characterized by two parameters: standard free energy of adsorption at the hexane/water interface and standard free energy of the partition between hexane and water. Theoretical approaches are used for the estimation of these parameters. As a result, several groups of amino acid units having close values of the parameters are distinguished. Classification of the amino acid units by such a method is believed to be able to provide promising results in the search of correlations in protein sequences. The comparison of experimental and theoretical studies of free amino acids at the hexane/water interface confirm the predictive power of the theory used, because a very good coincidence between theoretical and experimental data points is observed. The obtained values of standard free energies of adsorption are consistent with literature data.

Keywords Partition · Liquid–liquid interfaces · Amino acids · Residues · Classification

Introduction

In the preceding papers [1, 2], we introduced a two-dimensional diagram for the classification of amphiphilic monomers, which involved accounting for both interfacial and partitioning properties. It has been shown that a compound can be related to a point in a two-dimensional space; the abscissa and ordinate positions of this point represented quantitative characteristics for the above-mentioned properties. Namely, the abscissa coordinate corresponded to the affinity of a given compound to polar (water) or nonpolar (hexane) phase. The ordinate dimension corresponded to the interfacial activity at the interface between the two phases. The standard free energy of the partition between two phases was used as a quantitative parameter for the abscissa axis (Δf_{part}), whereas the standard energy of adsorption at the interface defined the position of a point along the ordinate axis (Δf_{ads}). Both parameters were normalized by kT factor. For the illustration of the proposed diagram, four monomers of synthetic water-soluble polymers were classified according to the above principle. The experiments on interfacial and partitioning properties were carried out for *N*-isopropylacrylamide (NIPA), 1-vinylimidazole (Vim), *N*-vinylpyrrolidone (VP) and *N*-vinylcaprolactam (VC). Water was chosen as a hydrophilic phase, while hexane represented a hydrophobic one. Δf_{ads} and Δf_{part} were calculated using the following formulae:

$$\Delta f_{part} = \ln\left(\frac{c_h}{c_w}\right) \tag{1}$$

$$\Delta f_{ads}^{w} = \ln\left[1 + \frac{1}{RT\tau}\left(\left.\frac{\gamma^0 - \gamma}{c_w}\right|_{c_w \to 0}\right)\right] \tag{2}$$

$$\Delta f_{ads}^{h} = \ln\left[1 + \frac{1}{RT\tau}\left(\left.\frac{\gamma^0 - \gamma}{c_h}\right|_{c_h \to 0}\right)\right] \tag{3}$$

where c_{h} is the concentration of dissolved molecules in hexane phase, c_{w} is the concentration of dissolved molecules in water phase, τ is the film thickness, γ^0 is the surface tension for the pure hexane/water boundary, γ is the surface tension in the presence of solute, Δf_{ads}^{h} is the standard free energy of adsorption from hexane, and Δf_{ads}^{w} is the standard free energy of adsorption from water. From the two values, Δf_{ads}^{h} and Δf_{ads}^{w}, the free energy of adsorption from the phase in which the compound is more soluble was chosen to characterize the position of this compound along the ordinate axis. Δf_{part} was postulated to be positive for hydrophilic compounds, Δf_{ads} was postulated to be positive for interfacially active compounds.

The proposed diagram was assumed to be helpful in predicting energetic preferences of chemicals for interfaces or bulk phases in systems with developed interfaces. Such classification should be relevant, for example, for biological systems, in which there are many interfaces between hydrophobic and hydrophilic areas (e.g., in organelles). Normally, the affinity of a compound to those areas is estimated via the octanol/water partition coefficient [3]; however, this quantity does not take into account the possibility for molecules to be located at the interface.

Besides much experimental work done for determination of n-octanol/water partition coefficients (over 10,000 compounds were studied), a series of theoretical approaches to the problem have been developed. The most well-known are the "group" contribution and the "atomistic" additive schemes. In these approaches, the estimation of the partition coefficient is performed by summation of contributions over groups of atoms or single atoms, respectively. Adequate summation over groups of atoms normally requires correction factors, whereas the atomistic approach needs none. However, very large data sets for different atom types should be used for correct application of the atomistic method, because for each atom, that atom's chemical environment should be taken into account. A special method for estimating partition coefficients was worked out by Ruelle [4] on the basis of mobile order and disorder (MOD) thermodynamics. In this method, contributions of various energetic effects to the partition process were taken into account, in particular hydrophobic effect, H-bonding, and nonideality of mixing entropy. Unlike the above-mentioned additive schemes, this method possesses an advantage to evaluate partition coefficients not only in the n-octanol/water system, but also in any biphasic system consisting of solvents with limited mutual solubility. The results on hexane/water and n-octanol/water partition coefficients obtained by that method corroborated very well a comprehensive set of experimental data.

Whereas the procedure for nonempirical estimation of partition coefficients (also for free energy of partition) is well established, the corresponding calculations for the free energy of adsorption are yet to be elaborated. This would allow the assessment of both parameters of the two-dimensional diagram by theoretical approaches.

A new classification of amino acid residues according to their values of Δf_{ads} and Δf_{part} is one of the attractive applications for the theoretical prediction of Δf_{ads}. Relative hydrophobicity of the residues has long been a subject of considerable interest. Nozaki and Tanford [5] were the first to evaluate quantitatively the hydrophobicity of amino acid residues. They measured partition coefficients of amino acids between water and organic solvent, from which the free energies of transfer from water to organic phase were further calculated. Glycine was used as a reference, so that its free energy was subtracted from the corresponding values for other amino acids. Later, Bull and Breese [6] constructed another hydrophobicity scale of amino acids based on the adsorption properties of the natural amino acids at the air/water interface. Free energy of adsorption was taken as a criterion for hydrophobicity. Wolfenden et al. [7] used free energy of transfer from water to the vapor phase of compounds with the general formula R–H, where R is amino acid residue. Leodidis and Hutton [8] found very close correlations (with correlation coefficient up to 0.99) between the free energies of transfer of amino acids from water to the interface formed by isooctane, water, and AOT surfactant [bis-(2-ethylhexyl) sodium sulfosuccinate]. The classifications mentioned above are mainly one-dimensional. Leodidis and Hutton proposed two-dimensional correlations, yet they were related not to amino acid residues but rather to entire amino acid molecules. We believe that the two-dimensional classification proposed here can significantly improve the previous methods of description of amino acid residues.

The chemical structures of the residues will be approximated similar to the approach of Wolfenden [7], in which compounds with the general formula of R−H were considered, where R is amino acid residue (Scheme 1). The corresponding values of Δf_{part} and Δf_{ads} for such compounds will be found.

In the present paper, the Askadskii theory [9, 10] is implemented for the calculation of interfacial tension isotherms (at the water/hexane interface) and free energies of adsorption. The theory is based on the representation of a compound as a set of anharmonic oscillators, which describes the thermal motion of atoms in the field of intra- and intermolecular forces (including weak dispersion forces, dipole–dipole interaction, hydrogen, and valence bonds). The incorporation of a few simplifying assumptions in the theory results in an additive calculation technique, which utilizes a set of atomic constants and a small number of empirical parameters independent of the chemical structure of the compounds studied. This approach is used for predicting the interfacial behavior of several monomers of synthetic polymers, as well as the interfacial behavior of amino acids. Theoretical consideration is supported by experimental results obtained earlier in Okhapkin et al. [1], as well as by some new results reported in this paper. Taking into account the satisfactory coincidence of experimental and theoretical data for the monomers and amino acids, the theoretical approach is applied to obtain the two-dimensional diagram characterizing affinity of the residues of 20 natural amino acids both to bulk and to interface. Affinity to bulk phases is quantified using the method described by Ruelle [4], which allows the calculation of partition coefficients and free energy of partition.

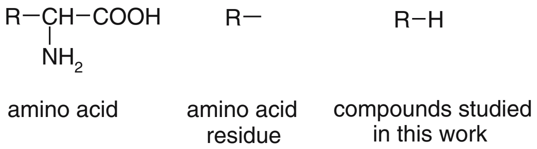
Theory

Interfacial tension at a liquid/liquid interface, $\gamma_{L_1L_2}$, can be calculated using the relation proposed by Girifalco and Good [11]:

$$\gamma_{L_1L_2} = \gamma_{L_1} + \gamma_{L_2} - F \tag{4}$$

where γ_{L_1} and γ_{L_2} are the surface tension values of the liquids. F is calculated as:

$$F = 2\left(\gamma^d_{L_1}\gamma^d_{L_2}\right)^{1/2} \tag{5}$$



Scheme 1 General chemical formulae for amino acids, amino acid residues, and related compounds

where $\gamma^d_{L_1}$ and $\gamma^d_{L_2}$ are dispersion components of surface tension of the liquids. F has a physical meaning of the free energy of adhesion of the phases. Equation 5 applies for interactions typical of regular solutions with ideal entropy of mixing. Possible deviations from the "regularity" are taken into account via factor Φ as a correction term for differences in molecule sizes:

$$F = 2\Phi\left(\gamma^d_{L_1}\gamma^d_{L_2}\right)^{1/2} \tag{6}$$

$$\Phi = \frac{4\left(\mathrm{V}_{\mathrm{L}_1}\mathrm{V}_{\mathrm{L}_2}\right)^{1/3}}{\left(\mathrm{V}_{\mathrm{L}_1}^{1/3} + \mathrm{V}_{\mathrm{L}_2}^{1/3}\right)^2} \tag{7}$$

where V_{L_1} and V_{L_2} are the molar volumes of the liquids.

Surface tension of a low molecular weight liquid is calculated as [9, 10]:

$$\gamma_L = A_j \frac{\sum_i \Delta E_i^*}{\left(\sum_i \Delta V_i\right)^{2/3}} \tag{8}$$

$$\sum_i \Delta E_i^* = \left(\sum_i \Delta E_{i,d}^* + \sum \Delta E_{i,p}^* + \sum_i \Delta E_{i,h}^*\right) \tag{9}$$

where A_j is the coefficient which assumes different values for different types of liquids. For hydrocarbons, perfluorocompounds, halide-containing compounds, esters, aldehydes, ketones, and nitro-compounds, A_1=0.0287; for alcohols, amides, and organic acids (all those compounds that can form hydrogen bonds), A_2=0.0181; for compounds containing nitrile groups, A_3=0.0229; $\sum_i \Delta E_i^*$ is the effective cohesion energy of a liquid; $\sum_i \Delta E_{i,d}^*$, $\sum_i \Delta E_{i,p}^*$, $\sum_i \Delta E_{i,h}^*$ are the dispersion interaction, dipole–dipole interaction, and hydrogen bond components of the effective cohesion energy, respectively; $\sum_i \Delta V_i$ is the Van-der-Vaals volume of the liquid. The estimation of molar volumes of pure compounds, Van-der-Vaals volumes, and cohesion energies for Eqs. 7, 8, 9 is performed using additive procedures developed by Askadskii [9, 10]. The corresponding numerical values for the compounds studied in this paper are available in Appendix A. Note that the chemical structure is the mere basis for calculation of those values. Thus, A_j is the only empirical parameter used in the model.

The surface tension of a mixture of organic liquids is:

$$\gamma_L = \left(\sum_j \sigma_j A_j\right) \frac{\alpha_1\left(\sum_i \Delta E_i^*\right)_1 + \alpha_2\left(\sum_i \Delta E_i^*\right)_2 + \ldots + \alpha_n\left(\sum_i \Delta E_i^*\right)_n}{\left[\alpha_1\left(\sum_i \Delta V_i\right)_1 + \alpha_2\left(\sum_i \Delta V_i\right)_2 + \ldots + \alpha_n\left(\sum_i \Delta V_i\right)_n\right]^{2/3}} \tag{10}$$

where α_1, α_2,…, α_n are the molar fractions of the components 1, 2,…, n., while σ_j is the molar fraction of liquids with identical A_j values.

Using Eqs. 4, 5, 6, 7, 8, 9, 10, it is possible to calculate interfacial tension between two immiscible liquids, both in the case of pure liquids and in the presence of admixtures, including polymeric ones. For the application of Eqs. 4, 5, 6, 7, 8, 9, 10 to the biphasic system of water and hexane studied in this work, see Appendix B. The main assumption for the calculation of surface tension using Eq. 10 is that none of the components of the mixture are considered to be adsorptionally active, that is, the concentration in the surface layer is equal to the concentration in bulk. The way to take the adsorption of a mixture component into account is described in Appendix C.

Experimental section

The experimental procedure for interfacial tension measurements was described elsewhere [1]. Interfacial tension measurements were performed using the drop-weight method. The solutions of amino acids for the interfacial measurements at different pH values were prepared by adding an adequate amount of either sodium hydroxide or hydrochloric acid stock solutions (both 10% wt) to the aqueous solutions of amino acids. The pH values of the solutions were controlled using the Hanna pH-300 pH-meter.

Results and discussion

To analyze the predictive power of the theory, the theoretical calculation of interfacial-tension isotherms for NIPA, Vim, VP, and VC was performed. The experimental isotherms were obtained by Okhapkin et al. [1], and the experimental isotherms are to be compared with the theoretical isotherms. Also, the behavior of several amino acids at the water/hexane interface was studied and the experimental results were compared with the theoretical predictions.

In Fig. 1, a comparison of experimental and theoretical interfacial tension vs concentration dependences is presented for NIPA, Vim, VP, and VC. A very good consistency of results was observed for Vim, VP, and VC. For NIPA, the coincidence of experimental and theoretical points was not as good as those of Vim, VP, and VC, yet, it was acceptable.

Table 1 shows the values of Δf_{ads} which were calculated from the theoretical curves by extrapolating them to zero concentration. Accuracies within 6% were observed for Vim, VP, and VC. For NIPA, the difference between experimental and theoretical value was equal to 15%.

Figure 2 shows the experimental data for the location of eight amino acids in the two-dimensional diagram. Here, Δf_{ads} is the free energy of adsorption to the hexane/water interface and Δf_{part} is the free energy of partition calculated from the octanol–water partition coefficients of the amino acids [12]. The experimental determination of the hexane/water partition coefficients of amino acids encountered great difficulties due to extremely low solubility of amino acids in hexane (e.g., 0.00036% for isoleucine [13]).

All the amino acids studied possess hydrophobic residues (for glycine, R=H, and glycine may be considered as a neutral residue; i.e., neither hydrophobic, nor hydrophilic, as was considered by Wolfenden et al. [7]). However, the contribution of the hydrophilic amino and carboxyl groups for the case of glycine and alanine are so large (the zwitterionic character of the amino acids is also important here) that their behavior is close to that of simple inorganic salts, which is similar to previous studies of the adsorption of

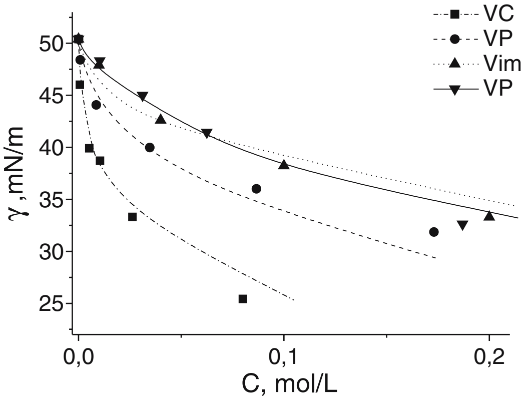


Fig. 1 Comparison of theoretical and experimental isotherms of interfacial tension for NIPA, Vim, VP, and VC at 25°C

Table 1 Comparison of theoretical and experimental data on free energy of partition for monomers studied in Okhapkin et al. [1]

	Δf_{ads}, experiment	Δf_{ads}, theory
N-isopropylacrylamide	5.1	5.92
N-vinylimidazole	5.36	5.68
N-vinylpyrrolidone	7.02	6.71
N-vinylcaprolactam	8.3	8.11

these amino acids at the air/water interface [6]. The other six amino acids are interfacially active at the water/hexane interface. Figure 2 shows also the values of Δf_{ads} and Δf_{part} which were calculated theoretically for Trp, Phe, Leu, Ile, Met, Val, and for two other amino acids with hydrophobic residues (Pro and Cys), which were not studied in this experiment.

As can be seen from Figs. 2 and 3, a very small deviation of theoretical points from the experimental ones was observed for Trp, Phe, Ile, Leu, and Met, while for Val the discrepancy is slightly larger. The main feature of the experimental data points for the eight amino acids is that Δf_{ads} shows satisfactory linear correlation with Δf_{part}. The regression parameters are presented in Table 2. Adsorption at the interface becomes stronger as hydrophobicity of the amino acids increases. Because the amino acids have very hydrophilic amino and carboxyl groups in composition, this increase of interfacial activity is caused by the enhancement of the amphiphilic character of the amino acids.

The experimental diagram obtained is consistent with the literature data. Leodidis and Hutton [8] observed a linear correlation between the octanol–water partition coefficients and the free energies of transfer of Pro, Val, Met, Leu, Ile, Tyr, Phe, and Trp to AOT–water–isooctane interface. Both in the study by Leodidis and Hutton [8] and in the present study the slopes of correlation lines exceed 1.00 and the correlation coefficients are close to unity. The absolute values are also close to each other as seen in Table 3.

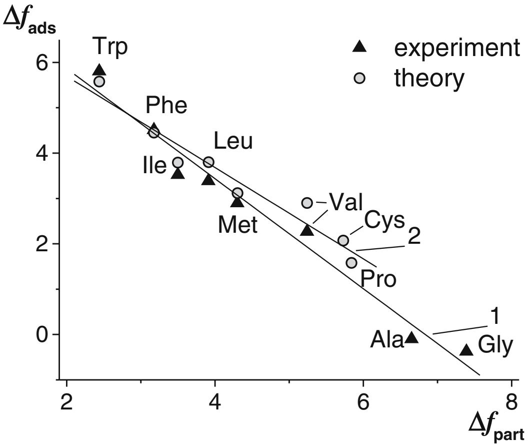


Fig. 2 Comparison of experimental and theoretical data points for amino acids plotted in the two-dimensional classification diagram at 25°C. For amino acids with hydrophobic residues (*triangles*, *correlation line 1*: experimental data points; *circles*, *correlation line 2*: theoretical data points), good linear correlation of standard free energies of partition and adsorption is typical

As far as the theoretical data points are concerned, the corresponding Δf_{ads} values show the linear correlation with Δf_{part}, similar to that of the experimental Δf_{ads} values (Table 2). This correlation is logical because the Δf_{ads} values are close for both theoretical and experimental data points (Table 3). Proline and cysteine, for which only theoretical data are available, also fit to this correlation. The common property of the residues of the amino acids studied is that they consist of hydrophobic aliphatic, aromatic, or sulfur-containing groups, for which dispersion interactions dominate. It is also worth mentioning that the aromatic Trp residue contains a hydrophilic NH group, and yet Trp fits the correlation very well, possibly due to the fact that the contribution of the NH group to the interaction with the water molecules is small compared to the contribution of the hydrophobic aromatic rings.

The correlations observed appear to be logical, as it is well known that the compounds with similar molecular structures may demonstrate functional properties (e.g., biological activity in quantitative structure–activity relationships) which are linearly correlated with certain molecular properties (e.g., hydrophobicity, molecular refractivity, electronic structure, etc). In our case, we observed the correlation of interfacial activities of the structurally similar amino acids (i.e., those containing hydrophobic residues) with their partitioning properties.

For several amino acids, pH-dependencies of Δf_{ads} were studied. For the two-dimensional diagram, the measurements were made at the pH of the isoelectric point of amino acids. For Gly, Val, Leu, and Phe, interfacial tension was measured at both high and low pH values, when the amino acids were either positively or negatively charged (Fig. 4). Despite partial differences in the shape of dependencies, the general feature of the amino acids is that Δf_{ads} and the interfacial activity counter intuitively increase as the degree of ionization is increased. Indeed, the ionization should increase the hydrophilicity of the amino acids and decrease the affinity of the amino acids to the interface, which is in fact a hydrophobic environment. However, such behavior corroborates the literature data. In the report by Rosen [14], a summary table for a series of surfactants is given, where the data on the free energies of adsorption are presented. It turns out for surfactants with similar hydrophobic parts (Table 4) that the free energy of adsorption is larger for ionic surfactants than for nonionic or zwitterionic surfactants, both at air/water and hydrocarbon/water interfaces. A possible explanation for this effect is the following: for a compound of X–Y type, where X is hydrophobic and Y is hydrophilic, both the enhancement of X hydrophobicity and the enhancement of Y hydrophilicity increase the

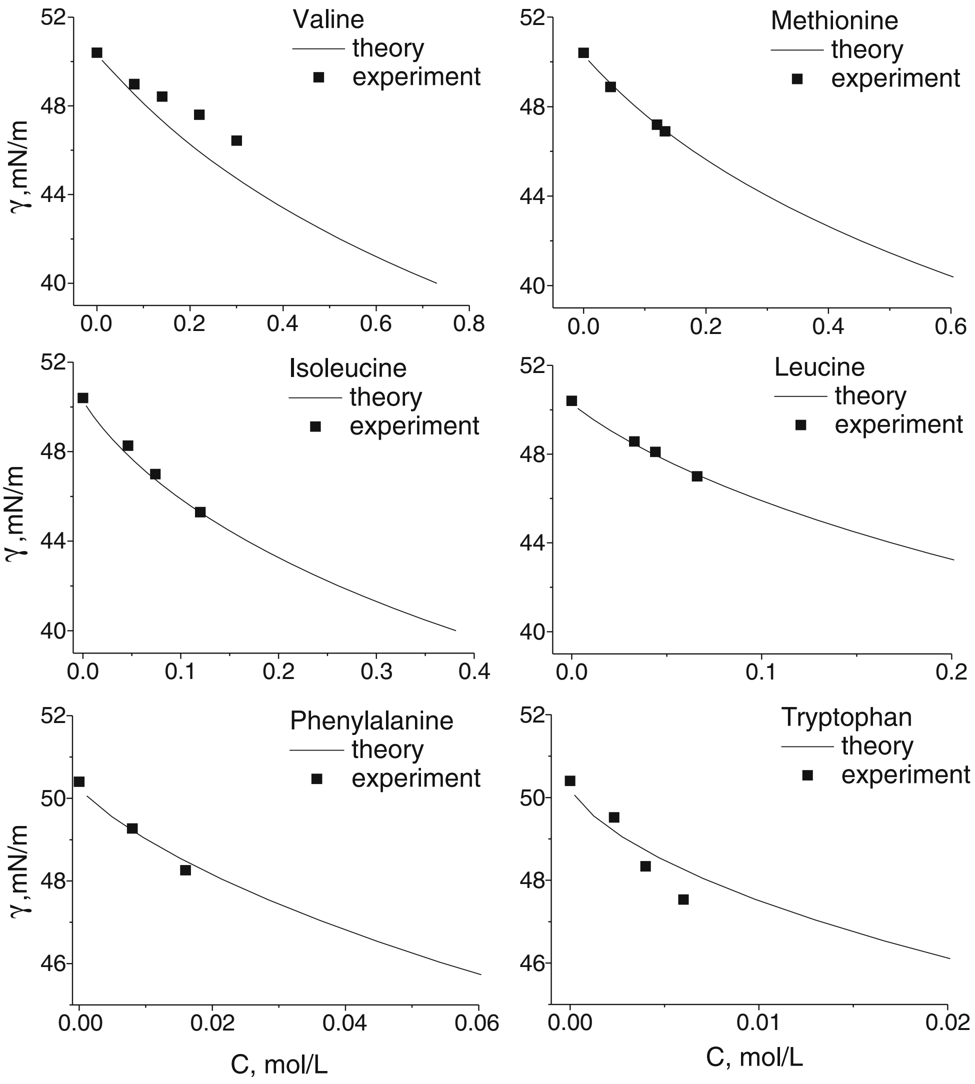


Fig. 3 Comparison of theoretical and experimental isotherms of interfacial tension for Val, Leu, Ile, Phe, Met, and Trp at 25°C

amphiphilicity (and, correspondingly, the interfacial activity) of the compound.

In summary, it should be pointed out that the Askadskii theory is successful in predicting interfacial properties of several amphiphilic monomers and amino acids. Therefore, it is challenging to extend the theoretical approach used to a series of compounds modeling the residues of all natural amino acids, which will serve to develop a new type of classification of the residues. In a further consideration, those compounds will be referred to as RH (or RH-compounds) and by the symbols of each compound's corresponding amino acid residue marked with an accent (e.g., Ala′). In Fig. 5, the results of the calculation of Δf_{part} and Δf_{ads} for RH are presented in the form of a two-dimensional diagram. Partition coefficients (and therefore the values for Δf_{part}) for RH were calculated using the method of Ruelle [4] (MOD thermodynamics).

Table 2 Regression parameters for correlation of Δf_{part} and Δf_{ads} for hydrophobic amino acids (this work and literature data)

	Correlation coefficient	Slope
Experiment	0.987	1.22
Theory	0.976	1.01
Data of Leodidis and Hutton [8]	0.998	1.52

Most of the RH-compounds reveal a certain affinity to the water/hexane interface as indicated by the positive values of Δf_{ads}. Two of them (Phe′ and Trp′) should be adsorbed at the hexane/water interface from the organic phase, in which they are preferentially soluble. For Phe′ and Trp′, dipole–dipole interactions between water molecules and the solutes are the driving force for the adsorption. However, such adsorption could not be characterized quantitatively, as the dipole–dipole interaction component of the cohesion energy of water was neglected in the calculations (Appendix B). Such approximation was justified for compounds having significant

Table 3 Theoretical and experimental values of amino acid free energies of adsorption (Δf_{ads}) compared to the literature data. The literature values are normalized by kT and their signs are inversed to bring them to our scale

Amino acid	Δf_{ads}, theory	Δf_{ads}, experiment	Δf_{ads}, data of Leodidis and Hutton [8]
Val	2.90	2.26	1.51
Leu	3.80	3.38	3.25
Ile	3.79	3.50	2.84
Pro	1.58	–	0.78
Phe	4.45	4.51	4.49
Tyr	–	–	2.20
Trp	5.58	5.80	5.64
Met	3.12	2.89	2.95
Ala	–	−0.1	–
Gly	–	−0.38	–
Cys	2.07	–	–

aqueous solubility, for which hydrophobic interactions are the main driving force of adsorption. However, approximation led to incorrect results for the compounds that concentrate in hexane phase, such as Phe′ and Trp′. Nevertheless, as it is appropriate to include Phe′ and Trp′ in the two-dimensional classification which comprises 20 basic natural amino acids, one may ascribe them with certain positive values of Δf_{ads} (e.g., 1.00) indicative of the ability to adsorb at the interface. It will be demonstrated in this paper that the absolute values of Δf_{ads} of Phe′ and Trp′ do not play a great role in the classification. The calculations showed that the other six RH-compounds with negative Δf_{part} (Leu′, Ile′, Val′, Pro′, Met′, and Ala′) would not adsorb at the hexane/water interface. Indeed, these six compounds do not contain chemical groups that should be capable of effectively interacting with water molecules. Thus, the Δf_{ads} values of the six RH-compounds were taken to be zero, which points to the fact that they do not adsorb at the interface. The values of Δf_{part} and Δf_{ads} for Gly′ (hydrogen) could not be estimated using both MOD thermodynamics and Askadskii theory. However, it was logical to use Gly′ as a reference item by ascribing it with zero energies, similar to what was done in literature [5–7].

The RH-compounds, each assigned with Δf_{part} and Δf_{ads} values, form a special pattern of the data points in the diagram (Fig. 5). From this pattern it is possible to select groups of residues with close values of Δf_{part} and Δf_{ads}. It is presumed that the amino acid residues inside each group should respect similar properties in protein sequences, which are connected with the ability of amino acid residues to be located either inside protein structures or at their interfaces. This ability is thus represented with the help of the two-dimensional diagram.

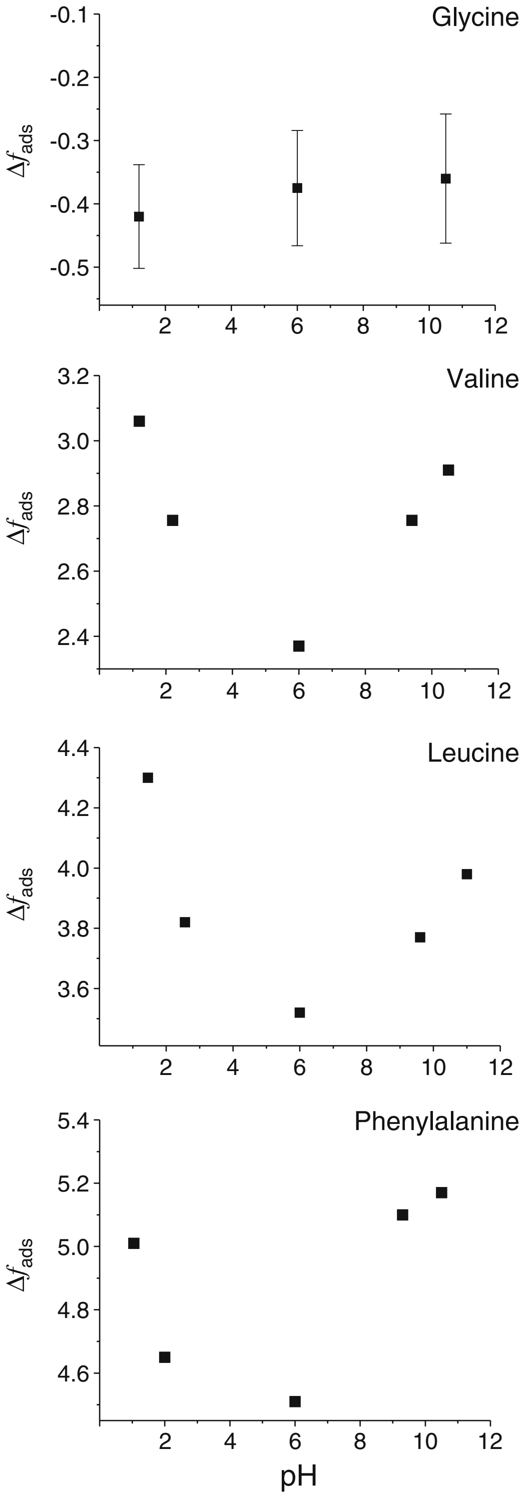


Fig. 4 pH-dependencies of standard free energy of adsorption for Gly, Val, Leu, and Phe at 25°C

The RH-compounds can be grouped according to the values of Δf_{part} and Δf_{ads}, as shown in Fig. 5 by ovals enclosing several sets of the data points. The compounds in each group have typical values of Δf_{part} and Δf_{ads} (Table 5). Indeed, Lys′, Cys′, and Tyr′ have the highest Δf_{ads} values

Table 4 Structure of hydrophilic moiety as a factor affecting standard free energy of adsorption (Δf_{ads}). Data taken from Rosen [14]

Compound	Interface	$\lvert\Delta f_{ads}\rvert$, kJ/mol
$C_{12}H_{25}SO_4Na$	Water/cyclohexane	58
$C_{12}H_{25}N^{+}(CH_3)\ (CH_2C_6H_5)CH_2COO^{-}$	Water/cyclohexane	48
$C_{10}H_{21}OC_2H_4SO_3Na$	Air/water	50
$C_{10}H_{21}CH(OH)C_2H_4OH$	Air/water	40

and Δf_{part} values close to zero (group 1). The value of Δf_{ads} of Arg′ is also among the highest, but its Δf_{part} value is large, indicating high hydrophilicity (group 2). Glu′, Gln′, Asp′, Asn′, His′, Thr′, and Ser′ are characterized by moderate values of Δf_{part} and Δf_{ads} (group 3). Leu′, Ile′, Val′, Pro′, Met′, and Ala′ are interfacially inactive ($\Delta f_{ads}=0$) and hydrophobic ($\Delta f_{ads}<0$) (group 4). Phe′ and Trp′ are hydrophobic and moderately interfacially active (group 5). For Gly′, Δf_{part} and Δf_{ads} values are both equal to zero indicating no specific affinity to bulk or interface (group 6). The same grouping method should be valid for the corresponding amino acid residues.

It may be noticed that the residues of almost every group have similar chemical structures. Indeed, most of the residues from group 3 contain a hydrophilic group and a small amount of hydrophobic ones, those from group 4 have purely aliphatic structures, and those from group 5 are aromatic. Groups 2 and 6 are composed of single residues of unique structure. The residues of group 1 are not chemically similar, but they have a special balance of hydrophilic and hydrophobic groups, which allows them to be mostly located at the interface rather than in bulk phases because their Δf_{part} values are much lower than their Δf_{ads} values. Indeed, as was shown in our previous works [1, 2], such a ratio between Δf_{part} and Δf_{ads} for a compound may lead to the fact that that compound's molecules tend to position themselves at interface rather than in bulk, provided interfaces are abundant in the system.

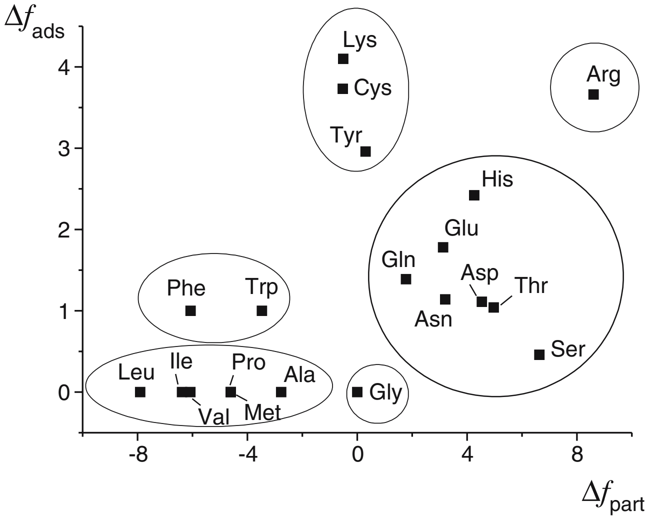


Fig. 5 Two-dimensional diagram for the compounds modeling amino acid residues

Table 5 Grouping of compounds which model amino acid residues according to the values of free energies of adsorption and partition

RH	Typical values of Δf_{part} and Δf_{ads}
1. Lys′, Cys′, Tyr′	$\Delta f_{part}\approx 0$, $\Delta f_{ads}>>0$
2. Arg′	$\Delta f_{part}>>0$, $\Delta f_{ads}>>0$
3. Glu′, Gln′, Asp′, Asn′, His′, Thr′, Ser′	$\Delta f_{part}>0$, $\Delta f_{ads}>0$
4. Leu′, Ile′, Val′, Pro′, Met′, Ala′	$\Delta f_{part}<0$, $\Delta f_{ads}=0$
5. Phe′, Trp′	$\Delta f_{part}<0$, $\Delta f_{ads}>0$
6. Gly′	$\Delta f_{part}=0$, $\Delta f_{ads}=0$

In summary, the RH-compounds (and the corresponding amino acid residues) are classified according to the values of free energy of adsorption and partition being divided into six groups. For the residues of each group, the specific behavior at the interface is typical; therefore, it is probable that their location in protein sequences will correlate to their three-dimensional location with respect to the interface of a protein globule.

Conclusion

The Askadskii theory is successful in predicting interfacial properties of amphiphilic monomers and amino acids because the predictions of the Askadskii theory match the experimental results. For hydrophobic amino acids, both theoretical and experimental values of the free energy of adsorption linearly correlate with the values of the free energy of adsorption calculated from n-octanol/water partition coefficients. Thus, it may be stated that for the amino acid units with similar structures of residues, the interfacial activity is reciprocally determined by the partitioning properties.

Using the predictions of the theory, the compounds modeling residues of 20 natural amino acids were characterized with the help of a two-dimensional diagram. Several groups of such compounds (and amino acid residues, correspondingly) were distinguished, for which the behaviors of Δf_{part} and Δf_{ads} are similar. It is supposed that such grouping of amino acid residues will be useful for the search of correlations in amino acid distributions in protein sequences.

Acknowledgements The financial support from the German Research Foundation, project Sonderforschungsbereich 569, and the Russian Foundation for Basic Research, project 05-03-33077, is highly appreciated. We also thank Prof. Valerii Grinberg and Prof. Yuri Davidovich for useful discussions.

Appendix A

Table 6 Molar volume, Van-der-Vaals volume, and cohesion energy values of the compounds studied

Compound (according to the abbreviations in the text)	Van-der-Vaals volume, Å^3	Molar volume, cm^3/mol	Cohesion energy, J/mol
NIPA	124	136	41,800
Vim	97	107	20,000
VC	146	160	26,300
VP	112	123	32,000
Cys	101	104	43,200
Ile	135	138	40,700
Leu	135	138	42,500
Met	136	140	55,400
Phe	159	161	53,500
Pro	106	109	35,900
Trp	189	191	62,000
Val	118	121	39,700
Ala′	29	32	3,100
Arg′	111	122	33,200
Asn′	59	65	34,500
Asp′	55	57	23,000
Cys′	47	51	10,500
Gln′	76	84	37,300
Glu′	72	74	25,700
Gly′	–	–	–
His′	84	92	35,500
Ile′	80	88	11,200
Leu′	80	88	9,500
Lys′	92	100	23,200
Met′	82	89	15,900
Phe′	105	106	34,800
Pro′	51	56	8,100
Ser′	36	37	20,200
Thr′	54	55	22,900
Trp′	135	136	29,400
Tyr′	111	113	37,600
Val′	63	69	8,500

If used in Eq. 13, the dimensionalities of all the parameters should be such as shown in the table. In that case, dimensionless proportionality coefficient A_j assumes the values stated in the theory section, and the resulting γ_L is expressed in mN/m

Appendix B

The general scheme for calculation of interfacial tension values is here adapted for the biphasic water/hexane system with a solute partitioned between the phases.

To calculate interfacial tension γ for such a system, surface tension at (water–solute)/air ($\gamma_{s,w}$) and (hexane–solute)/air ($\gamma_{s,h}$) boundaries should be calculated first, the concentrations of solute in water and hexane being connected by partition coefficient.

The use of Eq. 10 gives:

$$\gamma_{s,w} = [\alpha_w A_w + \alpha_{s,w} A_s] \times \frac{\alpha_w \left(\sum_i \Delta E_i^*\right)_w + \alpha_{s,w}\left(\sum_i \Delta E_i^*\right)_s}{\left\{\alpha_w\left(\sum_i \Delta V_i\right)_w + \alpha_{s,w}\left(\sum_i \Delta V_i\right)_s\right\}^{2/3}} \tag{11}$$

$$\gamma_{s,h} = [\alpha_h A_h + \alpha_{s,h} A_s] \times \frac{\alpha_h \left(\sum_i \Delta E_i^*\right)_h + \alpha_{s,h}\left(\sum_i \Delta E_i^*\right)_s}{\left\{\alpha_h\left(\sum_i \Delta V_i\right)_h + \alpha_{s,h}\left(\sum_i \Delta V_i\right)_s\right\}^{2/3}} \tag{12}$$

where, α_w and $\alpha_{s,w}$ are the molar fraction of water and solute in an aqueous phase, α_h and $\alpha_{s,h}$ are the molar fractions of hexane and solute in an organic phase. Here and throughout the text, "w", "h", and "s" indices denote "water", "hexane", and "solute", respectively.

Estimation of $\alpha_{s,w}$ and $\alpha_{s,h}$ involves taking into account the value of the partition coefficient:

$$P = \frac{C_{s,w}}{C_{s,h}};\ P = \frac{C_{s,w}}{C - C_{s,w}};\ C_{s,w} + C_{s,h} = C \tag{13}$$

where C is given as ($C_{s,w}$+$C_{s,h}$), $C_{s,w}$ is the equilibrium concentration of solute in water phase, and $C_{s,h}$ is the equilibrium concentration of solute in hexane phase, provided the volumes of both phases are equal.

$C_{s,w}$ and $C_{s,h}$ are expressed via partition coefficient as:

$$C_{s,w} = \frac{PC}{P+1} \tag{14}$$

$$C_{s,h} = \frac{C}{P+1} \tag{15}$$

Bearing in mind that

$$\alpha_{s,w} = \frac{C_{s,w}}{C_{s,w} + C_w} \tag{16}$$

$$\alpha_{s,h} = \frac{C_{s,h}}{C_{s,h} + C_h} \tag{17}$$

for dilute solutions, one may obtain

$$\alpha_{s,w} = \frac{1}{1 + \frac{1000 \cdot V_w^{-1}}{C_{s,w}}} \tag{18}$$

$$\alpha_{s,h} = \frac{1}{1 + \frac{1000 \cdot V_h^{-1}}{C_{s,h}}} \tag{19}$$

where V_w and V_h are the molar volumes of water and hexane, respectively. Finally, substitution of $C_\mathrm{s,w}$ and $C_\mathrm{s,h}$ into Eqs. 18 and 19 gives:

$$\alpha_{s,w} = \frac{PC}{PC + 1000 \cdot V_w^{-1} P + 1000 \cdot V_w^{-1}} \tag{20}$$

$$\alpha_{s,h} = \frac{C}{C + 1000 \cdot V_h^{-1}(P+1)} \tag{21}$$

After the calculation of $\gamma_\mathrm{s,h}$ and $\gamma_\mathrm{s,w}$, interfacial tension γ between hexane and water phases should be assessed using Eq. 4:

$$\gamma = \gamma_{s,h} + \gamma_{s,w} - F \tag{22}$$

The F value for Eq. 22 is calculated from Eq. 6. Correction factor Φ is now given as:

$$\Phi = \frac{4(\mathrm{V}_{\mathrm{s,w}} \cdot \mathrm{V}_{\mathrm{s,h}})^{1/3}}{\left(\mathrm{V}_{\mathrm{s,w}}^{1/3} + \mathrm{V}_{\mathrm{s,h}}^{1/3}\right)^2} \tag{23}$$

where $V_\mathrm{s,w}$ is the molar volume of a water/solute mixture; $V_\mathrm{s,h}$ is the molar volume of a hexane/solute mixture. $V_\mathrm{s,w}$ and $V_\mathrm{s,h}$ are given as linear combinations of molar volumes of solvent (V_w, V_h) and solute (V_s):

$$V_{s,w} = \alpha_w V_w + \alpha_{s,w} V_s \tag{24}$$

$$V_{s,h} = \alpha_w V_w + \alpha_{s,h} V_s \tag{25}$$

Similar linear combinations are suitable for estimation of the dispersion components of surface tension (DCST), which occur in the expression for F:

$$\gamma_{s,w}^d = \alpha_w \gamma_w^d + \alpha_{s,w} \gamma_s^d \tag{26}$$

$$\gamma_{s,h}^d = \alpha_h \gamma_h^d + \alpha_{s,h} \gamma_s^d \tag{27}$$

where γ_h^d, γ_w^d, and γ_s^d are the DCSTs for pure hexane, water, and solute, respectively. Estimation of DCST is performed using Eq. 28:

$$\gamma^d = A_j \frac{\sum\limits_i \Delta E_{i,d}^*}{\left(\sum\limits_i \Delta V_i\right)^{2/3}} \tag{28}$$

where $\sum\limits_i \Delta E_{i,d}^*$ is the energy of weak dispersion interactions. The DCST of hexane, γ_h^d, is assumed to be equal to the value of surface tension of pure hexane at 25°C (18.41 mN/m). The DCST of water, γ_w^d, is calculated using the following approach.

Cohesion energy of water is taken to consist of two additives: the dispersion interaction component and the hydrogen bond component:

$$\sum_i \Delta E_i = \left(\sum_i \Delta E_{i,d}^* + \sum_i \Delta E_{i,h}^*\right) \tag{29}$$

where $\sum\limits_i \Delta E_{i,h}^*$ is the cohesion energy of hydrogen bonds. The value of hydrogen bond component of cohesion energy is known (16.4 kJ/mol) [8, 9], as well as the value of surface tension of water (72.2 mN/m); therefore, it is possible to calculate DCST of water, combining Eqs. 8, 28, and 29:

$$\gamma_w^d = \gamma_w - A_j \frac{\sum\limits_i \Delta E_{i,h}^*}{\left(\sum\limits_i \Delta V_i\right)^{2/3}} \tag{30}$$

at 25°C $\gamma_h^d = 28.2 mN/m$

Appendix C

The values of interfacial tension obtained using Eq. 4 are calculated with the help of a special concentration scale (capital C in the text). A hypothetical situation is considered when no adsorption of solute takes place from the bulk phases to the interface, and the distribution of solute molecules in bulk and at interface are the same. In that case, the contribution from water phase to the solute's interfacial concentration $C_{\text{interface,w}}$ can be calculated directly from the bulk concentration $C_{\text{s,w}}$. Indeed, there are $10^3 C_{\text{s,w}}$ moles of monomer in 1 m^3 of water solution, and the number of molecules in 1 m^3 is equal to $N{=}10^3 C_{\text{s,w}} N_a$. The number of molecules in 1 m^2 of a monolayer is obviously equal to $N_{\text{mono}}{=}N^{2/3}$ and interfacial concentration is $C_{interface,w} = \frac{N^{2/3}}{N_a}$. The final expression for the interfacial concentration is:

$$C_{interface,w} = 10^2 f N_a^{-1/3} C_{s,w}^{2/3} \tag{31}$$

where f is the ratio of molecular packing coefficient in bulk to molecular packing coefficient at interface, $f{=}1.2$ [9, 10]. Thus, at each value of C and $C_{\text{s,w}}$, interfacial tension and interfacial concentration can be calculated from Eqs. 22 and 31, respectively.

After interfacial tension and interfacial concentration are calculated, a system with adsorption should be considered. For further calculations, an important assumption is made that if other factors are the same, a system with adsorption and a system with no adsorption should have identical interfacial tension values at each value of interfacial concentration.

In a real system with adsorption, the Gibbs interfacial excess Γ can be calculated from the Gibbs equation using bulk concentration and corresponding values of interfacial tension. In our system, for the given interfacial concentration, the corresponding interfacial tension is known. Thus, for the full description of adsorption phenomena in that system, a reverse problem should be solved, that is for each value of interfacial tension a corresponding value of bulk concentration should be found. This can be done by solving the Gibbs equation:

$$\Gamma RT = -\frac{d\gamma}{dc} c \tag{32}$$

where Γ is assumed to be equal to $C_{\text{interface,w}}$.

To use Eq. 32, the dependence of ΓRT on γ should be fitted with a certain correlation line. For this purpose, the values of ΓRT and γ were calculated at corresponding $C_{\text{s,w}}$ values. The appropriate dependencies of ΓRT on γ were fitted by several functions. The most suitable one is as follows:

$$\Gamma RT = \frac{\gamma_0 - \gamma}{P_1\gamma + P_2} \tag{33}$$

When the Eq. 33 is used, the correlation coefficient is close to 1.00 for all the systems studied.

Combination of Eqs. 32 and 33 leads to the following differential equation:

$$-\frac{d\gamma}{dc} c = \frac{\gamma_0 - \gamma}{P_1\gamma + P_2} \tag{34}$$

Separation of the variables and subsequent integration results in:

$$\ln c = P_1\gamma + (P_1\gamma_0 + P_2)\ln(\gamma_0 - \gamma) + c' \tag{35}$$

ΓRT dependence on γ is well approximated by Eq. 33. As a result, it is easy to find P_1 and P_2 as regression parameters. To compute integration constant c', the interfacial tension at the pure solute/hexane interface is taken as a boundary condition: in that case, $c = \frac{1}{V_s}$, whereas γ at the pure solute/hexane interface is estimated using Eq. 22, assuming that solute does not penetrate hexane.

References

1. Okhapkin IM, Makhaeva EE, Khokhlov AR (2005) Colloid Polym Sci 248:117
2. Okhapkin IM, Makhaeva EE, Khokhlov AR (2006) Adv Polym Sci (in press)
3. Sangster J (1997) Octanol–water partition coefficients: fundamentals and physical chemistry. Wiley, Chichester
4. Ruelle P (2000) Chemosphere 40:457
5. Nozaki Y, Tanford C (1971) J Biol Chem 246:2211
6. Bull HB, Breese K (1974) Arch Biochem Biophys 161:665
7. Wolfenden R, Anderson I, Cullis PM, Southgate CCB (1981) Biochemistry 20:849
8. Leodidis EB, Hutton TA (1990) J Phys Chem 94:6411
9. Askadskii AA (1996) Physical properties of polymers. Prediction and control. Gordon and Breach Publishers, Amsterdam
10. Askadskii AA (2003) Computational materials science of polymers. Cambridge International Science of Polymers, Cambridge
11. Girifalco LA, Good RJ (1957) J Phys Chem 61:904
12. Fendler JH, Nome F, Nagyvary J (1975) J Mol Evol 6:215
13. Howard PH (ed) (1997) Handbook of physical properties of organic chemicals. Lewis, Boca Raton
14. Rosen MJ (1989) Surfactants and interfacial phenomena. Wiley, New York