CHROM. 14,027

PARTITION COEFFICIENTS OF AMINO ACIDS AND HYDROPHOBIC PARAMETERS π OF THEIR SIDE-CHAINS AS MEASURED BY THIN-LAYER CHROMATOGRAPHY[*,**]

VLADIMIR PLIŠKA*, MANFRED SCHMIDT and JEAN-LUC FAUCHÈRE

Institut für Tierproduktion, ETH-Zentrum, CH-8092 Zurich (Switzerland) and *Institut für Molekular-biologie und Biophysik, ETH-Hönggerberg, CH-8093 Zurich (Switzerland)*

(First received November 18th, 1980; revised manuscript received May 22nd, 1981)

SUMMARY

Partition coefficients in *n*-octanol–water have been determined for the naturally occurring and some synthetic α-amino acids using thin-layer chromatography in different solvent systems. Both literature and newly measured values of partition coefficients in *n*-octanol–water have been used as reference values. The relationship between R_F values and partition coefficients (P) has been established. Hansch hydrophobic parameters π have been evaluated for the side-chains. Fragmental group contributions to the overall π-value have also been evaluated. In this way P and π values of any non-listed amino acid become easily accessible either by thin-layer chromatography or by computation.

INTRODUCTION

Hydrophobicity is an important physico-chemical property of protein and peptide molecules. It is manifested in their conformational features (folding of polypeptide chains), in their kinetic behaviour (ligand–macromolecule or macromolecule–macromolecule interactions) and it determines their organization into complex entities such as cell membranes or organelles. The importance of hydrophobicity for various biological systems and processes has been emphasized with increasing intensity throughout the past two decades.

In spite of this importance, attempts to express the hydrophobicity of a polypeptide chain in quantitative terms have so far been unsatisfactory. The problem is not simple and straightforward. The "overall" hydrophobicity of such moieties can

___

* Dedicated to the memory of Professor F. Šorm.
** A preliminary account was presented at the *Sixth American Peptide Symposium, Washington, DC, June 1979*[1]. Abbreviations according to the IUPAC–IUB Commission on Biochemical Nomenclature. In addition: Abu = L-β-methylalanine; Ada = adamantyl-L-alanine; Bug = L-*tert.*-butylglycine; Car = *o*-carboranyl-L-alanine; Cit = L-citrulline; Har = L-homoarginine; Lys(CHO) = L-formyllysine; Lys(Me) = Nᵉ-L-methyllysine; Nap = L-naphthylalanine; Neo = L-neopentylglycine; Nle = L-norleucine; Nva = L-norvaline; Pen = DL-penicillamine (β-mercapto-L-valine); Pro(OH) = 4-L-hydroxyproline.

be assessed, for example, by the partition coefficient for water and a lipophilic solvent[2–5] (or related chromatographic values), or by thermodynamic parameters for transfer from aqueous to organic phase[6–8], etc. The "regional" or "side-chain oriented" hydrophobicities[1], which are usually of the greatest interest to biologists, cannot be established by such simple procedures and have to be estimated from the primary structure. Although such hydrophobicities of smaller segments are no doubt reflected in that of larger regions, a simple additivity can be scarcely expected in this complex situation.

However, the hydrophobicity of the side-chain at a certain position in the polypeptide molecule may itself be important. We are referring here to the case of pharmacologically active peptides for which the biological potency is frequently associated with such a feature[4,9,10]. The use of hydrophobicity characteristics derived from partition coefficients has been suggested in a number of instances by Hansch and co-workers[2,11]. As this approach may also be productive for peptides, we have attempted to calculate the appropriate hydrophobicity parameters for the side-chains of their constituent components, the amino acids.

To conform with the Hansch definition[2] of the "hydrophobic parameter π", the partition coefficients P (cf., eqn. 1) have to be obtained in n-octanol–water. However, this system is not suitable for direct partitioning of amino acids, owing to their low solubility in the organic phase. Our preliminary communication[1] on this subject dealt with the possibilities of employing chromatographic methods for this purpose. The application of thin-layer chromatography (TLC) of free amino acids in various solvent systems is reported in this paper. The reference P values were in part taken from the literature and in part measured in our laboratory. Similar experiments have recently been carried out with reversed-phase high-performance liquid chromatography[12].

In this paper we describe first the theoretical relationship between R_F and P values and the details of our computational procedure. Then we present the statistical treatment of the hydrophobic parameters π, derived from these values, and the computation of the fragmental group contributions to π. As the relationship between partition coefficients and R_F values does not follow its theoretical course even in a narrow hydrophobicity range, we suggest another, empirical relationship and use this for the R_F to P conversion. It is of crucial importance to avoid extrapolation beyond the hydrophobicity range of the available standards. The recently established P values for the extremely hydrophobic carboranylalanine and adamantylalanine[13], together with new P estimates for arginine (see above), help to validate the computations also for highly hydrophobic or highly hydrophilic side-chains.

In addition to presenting a reliable list of π-values for the amino acid side-chains, the experiments demonstrate the general usefulness of TLC for their estimation. With few exceptions, our attention was focused on the naturally occurring amino acids. The special cases of proline, hydroxyproline and cystine require a separate discussion. Otherwise, any α-amino acid can be characterized by its π-value in the suggested way.

EXPERIMENTAL

Amino acids

L-Citrulline, DL-penicillamine and L-norleucine were products of Fluka (Buchs,

Switzerland). L-Norvaline and ε-formyl-L-lysine were purchased from Bachem (Bubendorf/BL, Switzerland) and Chemalog (South Plainfield, NY, U.S.A.), respectively. L-Carboranylalanine, L-adamantylalanine, L-*tert*.-butylglycine and L-neopentylglycine were prepared in our Institute[13]. α-Naphthyl-L-alanine and β-naphthyl-L-alanine were gifts from Dr. A. Loffet, UCB (Brussels, Belgium). All other amino acids (L-forms) were taken from Kit No. 721 of Mann Research Laboratories (New York, NY, U.S.A.). All samples were chromatographically pure except for DL-penicillamine and occasionally cysteine, which displayed minor secondary spots.
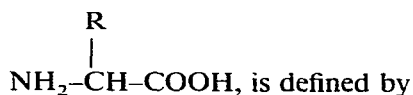
Chromatography
    Thin-layer chromatography was carried out on pre-coated Merck silica gel and cellulose plates; the spots were detected by the standard ninhydrin technique. Three types of solvent systems were employed (all ratios expressed as v/v): (1) alcohol–50 % acetic acid (2:1); (2) alcohol–25 % ammonia (2:1); (3) alcohol–0.05 M ammonium acetate (pH 7.2) (2:1). One of the following alcohols was used: *n*-propanol, isopropanol, *n*-butanol, isobutanol, *sec*.-butanol and *tert*.-butanol. The components were shaken for 5 h in a separating funnel at 22°C. The saturated alcoholic layer was separated by centrifugation and used as the mobile phase. Samples (5 µl of a 0.05–0.1 % solution in water) were applied as narrow lines 1.5 cm from the lower edge. A horizontal groove was cut in a distance of 15 cm from the start and the chromatography in a closed glass trough saturated with water vapour was terminated when the solute front reached the groove. The standards Gly, Ala, Val, Leu, Ile, Met, Phe, Trp, Bug, Ada, Car, Nva and Nle were present in each run.

MATHEMATICAL TREATMENT OF THE DATA

Conversion of chromatographic data into partition values: theory
    The Hansch hydrophobic constant[11], π_R, of an amino acid-side chain R,

$$\begin{matrix} \text{R} \\ | \\ \text{NH}_2\text{–CH–COOH,} \end{matrix}$$ is defined by

$$\pi_R = \log P_{OW} \text{ (amino acid)} - \log P_{OW} \text{ (glycine)} \qquad (1)$$

where P_{OW} is the partition coefficient in *n*-octanol–water. Primarily, therefore, the two P values have to be estimated. For paper chromatography, a relationship between R_F values in a given solute system and the partition coefficients, P_c, in the same system was derived by Martin and Synge[14] and Consden *et al.*[15]; in a slightly modified form it is given by

$$1/P_c = \varrho \,(1/R_F - 1) \qquad (2)$$

Despite its physical significance, *i.e.*, the ratio of the cross-sectional areas of the mobile and stationary phases[15], ϱ should be considered as an empirical parameter for a given chromatographic system, a set of conditions and a group of substances with certain structural and/or physico-chemical similarities. We assume that the same relationship applies to thin-layer chromatography.

The relationship between P (e.g., in n-octanol–water) and P_c (in another system) follows from the study of Collander[16], who demonstrated a linear relationship between the logarithms of the partition coefficients, P_a and P_b, of a substance in two partition systems a and b:

$$\log P_a = \alpha + \beta \log P_b \tag{3}$$

The constants α and β are known for a number of partition system combinations[16,17].

Taking P for P_a and P_c for P_b and substituting eqn. 2 into eqn. 3, one obtains

$$\log P = \alpha' - \beta R_M \tag{4a}$$

where

$$\alpha' = \alpha - \beta \log \varrho \tag{4b}$$

and

$$R_M = \log (1/R_F - 1) \tag{4c}$$

Eqn. 4c was introduced by Bate-Smith and Westall[18]. These equations describe the relationship between R_F values and partition coefficients of essentially any pair of unrelated solute systems. Moreover, when the constants α and β for a certain chromatographic system and two different partition systems (p and q) are available,

$$\log P_p = \alpha'_p - \beta_p R_M \tag{5a}$$

$$\log P_q = \alpha'_q - \beta_q R_M \tag{5b}$$

the Collander constants (eqn. 3) for the systems p and q can be obtained by

$$\beta = \beta_p/\beta_q \tag{6a}$$

$$\alpha = \alpha'_p - \alpha'_q \beta \tag{6b}$$

Such an approach may find an application when, for instance, Collander parameters should be derived from an incomplete set of partition coefficients P_p, P_q (i.e., P_p and P_q are available for non-identical substances and therefore cannot be subjected to the regression analysis based on eqn. 3).

Statistical evaluation of the π_R data

Routinely, we have pooled all π_R and log P data for each side-chain R obtained in alkaline, acidic and neutral solute systems into separate groups. For data samples containing more than twelve values the normality of the data distributions was tested by both Kolmogorov–Smirnov and χ^2 tests (significance level 95%). Extreme values in each group were tested for possible outliers using various criteria for outlying observations[19,20]; the outliers were then excluded.

Using the arithmetic means and standard deviations, we have investigated differences between values for alkaline, acidic and neutral media by means of Student's t-test combined with the F-test for difference between variances of the compared groups. If a significant F value was obtained, an appropriate modification of the t-test was undertaken[21]. The critical probability level for testing all these differences was taken as 95%. Groups with mean values not significantly different on this level were further pooled, and an overall arithmetic mean was taken as a characteristic parameter. In the opposite case, the π and/or log P values were considered as pH dependent and treated separately.

Fragmentation of π-values: group contributions

It is assumed that individual parts of a molecule ("functional groups") and some structural features contribute to the overall hydrophobicity in a linear manner. Additivity of corresponding group contributions to R_M values (eqn. 4c) was assumed many years ago by Bate-Smith and Westall[18] and by Martin[22]. According to the nomenclature of Rekker[23] and Hansch and Leo[11], such contributions will be called "fragmental" in the following text. Our π_R value is then defined as a sum of fragmental contrifutions, π_f, of all participating groups and features, each one appearing p_f times in the given side-chain R:

$$\pi_R = \sum_f p_f \, \pi_f \tag{7}$$

Conversely, eqn. 7 is applicable for computation of the unknown group contributions in a given set of R values. This set is first decomposed into functional groups or molecular fragments and additive fragmental contributions, π_f, are assigned to them. Optimized π_f values are obtained by multiple regression analysis according to eqn. 7, applied to the known set of π_R and p_f values (when a certain functional group is missing in the side-chain, its $p_f = 0$). Understandably, the number of side-chains used for the analysis must be larger than the number of fractional contributions considered. The coefficient of multiple correlation can be taken as a measure of the validity of the linear hypothesis expressed by eqn. 7.

Computational procedure

The final protocol for the evaluation of the chromatography results consists of following steps:

(i) R_F values of the substances in a given solute system were evaluated.

(ii) Available partition coefficients in n-octanol–water, P_{OW}, of the reference substances, together with corresponding R_F values of these substances, were used to carry out the parameter fit in eqn. 4: constants α' and β were estimated by the least-squares procedure. The same procedure was applied to the reference P values in n-butanol–water, P_{BW}.

(iii) Using eqn. 4 and the estimated parameters α' and β, the R_F values of the amino acids investigated were converted into the partition coefficients P_{OW} and P_{BW} (n-butanol–water), respectively.

(iv) The π_R values were then obtained from eqn. 1, using a local P_{OW} value of glycine (or an average value when several such values were available on the same chromatogram) as reference.

(v) The Collander parameters for the partition system n-octanol–water (P_a in eqn. 3) and n-butanol–water (P_b in eqn. 3) were evaluated for each chromatographic run by means of eqns. 6a and 6b. This procedure was taken mainly as a test of the correctness of our theoretical considerations (see Discussion).

RESULTS

Thin-layer chromatography

Chromatography in an acidic solute system and at neutral pH did not pose any problems and yielded for all systems R_F values with a coefficient of variation of at most 3 %. The spots were narrow and easily measurable without density scanning. The substances employed displayed a single, non-diffuse spot, except for cysteine and penicillamine, where by-products were clearly seen. However, the major spot could always be easily recognized.

On the other hand, in systems containing ammonia the spots frequently showed a "rocket" form, with a sharp concave front. Plates with an extreme appearance of this type were discarded. In general, the reproducibility of R_F values was poorer than with acidic systems (variation 7.5 %). Replacement of ammonia with a less volatile organic base (diethanolamine, aminoethanol) did not bring any visible improvement, but rather some additional problems.

Although we could not observe any significant numeric differences in the P and π values between experiments on silica gel and cellulose plates, the latter material is not only more suitable when working in basic solute systems but also offers sharper spots and is generally more recommendable than silica gel.

The pH of the organic phase measured with a glass electrode before and after the chromatography was 2.6–3.1 for acidic, 12.1–12.9 for alkaline and 7.2–7.6 for neutral systems. Running times for a given distance (15 cm) were between 4 and 7 h, depending on solvent, carrier and pH. In the final selection of solvents we chose those for which the R_F values were not less than 0.05 and not larger than 0.8 (*cf.*, Experimental).

Relation between partition and chromatographic data

Partition coefficients of the standards used in our investigations are given in Table I. These values are partly literature data[11,13,24,25], and partly unpublished data obtained by partitioning carried out in our laboratory. The value for arginine (protonated side-chain) was calculated as follows.

The partition coefficients for arginine, lysine and ornithine are known for the system *sec.*-butanol–0.1 M hydrochloric acid[26]. Starting with the known P values for glycine, alanine, 2-aminobutyric acid and norvaline in the same system[26] and in n-butanol–water[24], we first computed the constants α and β in eqn. 3 for interconversion of P values for *sec.*-butanol–0.1 M hydrochloric acid into those for n-butanol–water. Eqn. 3 was then used a second time with known α and β values[17] for the conversion n-butanol–water into n-octanol–water partition coefficients. The computed "overall" Collander constants for n-octanol–water (system a) and *sec.*-butanol–0.1 M hydrochloric acid (system b) are $\alpha = 0.55$ and $\beta = 2.49$. Inserting these values into eqn. 3, the following $\log P_{OW}$ values for the charged species were obtained: arginine -4.23, lysine -4.68 and ornithine -4.76. The literature values[3] appear to be too high.

TABLE I

PARTITION COEFFICIENTS OF L-AMINO ACIDS IN n-OCTANOL–WATER (P_{OW}) AND IN n-BUTANOL–WATER (P_{BW}) AT pH 7 OBTAINED BY PARTITIONING

Amino acid*	Log P_{OW}	Log P_{BW}**
1 Car	0.99***	
2 Ada	0.43***	
3 Trp	−1.06 §	
4 Phe	−1.35 §§	
5 Neo	−1.42***	
6 Ile	−1.69 §§	
7 Leu	−1.76 §§§	−0.74
8 Bug	−1.77***	
9 Met	−1.87 §§	
10 Val	−2.23 §§	−1.14
11 Tyr	−2.26 §	
12 Ala	−2.72 §§	−1.60
Ala	−2.96 §	
13 Gly	−3.21***	−1.81
14 Arg	−4.23 †	
15 Nle		−0.51
16 Nva		−0.98
17 Abu		−1.34
18 α-Nap	−0.13 ††	
19 β-Nap	−0.06 ††	

    * For abbreviations of less common amino acids see footnote on first page. All amino acids are in L-form.
    ** England and Cohn[24].
    *** Fauchère et al.[13].
    § C. Church, unpublished results (see ref. 11).
    §§ Klein et al.[25].
    §§§ Arithmetic mean of several literature data.
    † Calculated value, see text.
    †† Own unpublished values.

For the n-octanol–water partitioning system, the relationship between log P and R_M values shows a conspicuously regular deviation from the linearity predicted by eqn. 4 (Fig. 1). Therefore, in all experiments we tested the difference between the fits to the linear eqn. 4 and to an empirical polynomial relationship:

$$\log P = \gamma + \delta R_M + \varepsilon R_M^2 \tag{8}$$

The significance of this difference was investigated by two F-test procedures, one proving the significance of an increase in the multiple correlation coefficient due to the addition of the quadratic term[27], the other testing the variances of experimental and predicted values. Significant differences exist in all instances on at least a 95% probability level.

For conversion of R_F to log P data, we applied eqn. 8 separately for a group of more hydrophobic (standards 1–9 in Table I) and less hydrophobic amino acids (7–14) (see Fig. 1C and D) and selected the proper region for interpolation according to
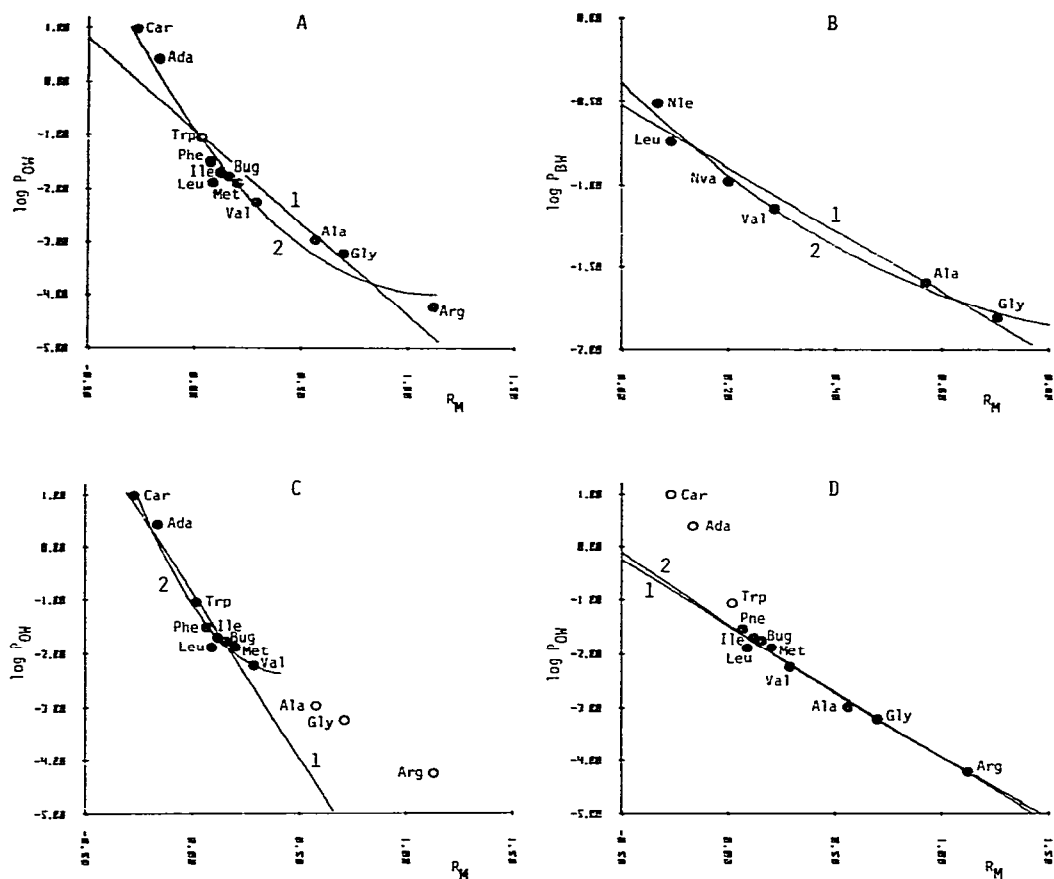
Fig. 1. Relationship between R_M values and partition coefficients, P, for amino acids using thin-layer chromatography in n-butanol–50% acetic acid (2:1). Curves 1 and 2 are linear and second-order polynomial fits, respectively. A, C, and D, R_M versus P relationships for n-octanol–water (pH 7); open symbols in C and D were not considered in the regression analysis, thus comprising more hydrophobic (C) and more hydrophobic (D) regions of side-chains. B, R_M versus P relationship for n-butanol–water (pH 7).

the actual R_F value. An average value was taken for R_Fs from the overlapping regions.

Partition coefficients and hydrophobic parameters π

The results of our calculations are summarized in Table II, which includes the values of log P_{OW} and π for the twenty naturally occurring amino acids and for several other α-L- and DL-amino acids. Whenever statistically different, values under basic, neutral and acidic conditions are reported separately. In all instances, the log P and π values followed a Gaussian frequency distribution. The standard deviation of log P_{OW} lies between 0.08 and 0.3. Exceptions are the unstable cysteine and some residues with extremely polar or extremely hydrophobic side-chains. Standard deviations of π values computed with the aid of the local glycine standard are in the same range of magnitude. Five outlying values were found among our π data by means of the tests

TABLE II

PARTITION COEFFICIENTS IN n-OCTANOL–WATER (P_{OW}) OF L-AMINO ACIDS AND HYDROPHOBIC CONSTANTS π FOR AMINO ACID SIDE-CHAINS

Amino acid[*]	No. of values	Log P_{OW}		π		Predicted π value[***]
		Mean[**]	S.D.[**]	Mean[**]	S.D.[**]	
Ada	15	−0.08	0.22	3.24	0.22	3.24
Ala	24	−2.89	0.09	0.40	0.15	0.42
Arg/charged [§]	10	−4.20	0.02	−0.90	0.07	−0.90
Asn	18	−3.41	0.14	−0.16	0.14	−0.26
Asp/uncharged	5	−3.38	0.03	−0.07	0.04	−0.06
Asp/charged	9	−.425	0.22	−1.05	0.19	−1.24
Bug	15	−1.76	0.10	1.51	0.14	1.64
Car	14	0.86	0.26	4.14	0.26	4.14
Cit	12	−3.19	0.11	−0.02	0.14	−0.02
Cys/uncharged	6	−2.49	0.16	0.82	0.13	0.73
Cys/charged	11	−3.63	0.38	−0.39	0.39	−0.17
Gln	18	−3.15	0.21	0.10	0.21	0.15
Glu/uncharged	10	−2.94	0.04	0.36	0.05	0.36
Glu/charged	12	−4.19	0.08	−0.98	0.21	−0.83
Gly	26	−3.25	0.12	0		0
Har/charged	5	−3.90	0.10	−0.44	0.11	−0.43
His/uncharged	13	−2.84	0.13	0.39	0.12	0.39
His/pH 7 [§§]	3	−3.56	0.14	−0.40	0.16	
His/charged	10	−4.15	0.12	−0.84	0.11	−0.84
Ile	26	−1.72	0.06	1.55	0.08	1.60
Leu	26	−1.61	0.08	1.64	0.14	1.60
Lys/uncharged [§]	4	−3.31	0.11	−0.12	0.16	−0.12
Lys/charged [§]	8	−4.44	0.10	−1.14	0.08	−0.99
Lys(CHO)	6	−2.84	0.13	0.41	0.17	0.41
Lys(Me)/uncharged	4	−2.77	0.08	0.41	0.13	0.41
Lys(Me)/charged	3	−4.21	0.03	−0.78	0.10	−0.78
Met	26	−1.84	0.15	1.42	0.14	1.42
Nle	18	−1.53	0.06	1.70	0.11	1.66
Nva	18	−1.86	0.11	1.37	0.11	1.24
Orn/charged [§]	5	−4.22	0.09	−1.11	0.08	−1.41
Pen/uncharged	8	−1.78	0.09	1.52	0.09	1.54
Pen/charged	11	−2.41	0.19	0.82	0.23	0.64
Phe	26	−1.63	0.09	1.63	0.17	1.63
Pro	7	−2.50	0.12	0.77 [§§§]	0.10	0.77 [§§§]
Pro(OH)	7	−3.17	0.19	0.09 [§§§]	0.19	0.09 [§§§]
Ser	18	−3.30	0.12	−0.08	0.16	−0.05
Thr	17	−2.91	0.10	0.33	0.11	0.31
Trp/uncharged	4	−1.75	0.31	1.41	0.24	1.41
Trp/charged	3	−2.41	0.22	0.83	0.01	0.83
Tyr	6	−2.42	0.23	0.88	0.27	0.81
Tyr(3-OMe)	6	−2.54	0.20	0.83	0.18	0.83
Tyr(Me)	3	−1.89	0.06	1.61	0.21	1.53
Val	26	−2.08	0.13	1.18	0.19	1.18

[*] All amino acids are in the L-form except for penicillamine, which is DL. For abbreviations of less common amino acids see footnote on first page.

[**] Arithmetic mean and standard deviation (S.D.).

[***] Sum of fragmental contributions, π_f, obtained for a particular data set (Table III contains average values of π_f which may be universally employed).

[§] P_{OW} and π values for basic amino acids and corresponding π_f values indicated in our tentative list[1] based on literature data are apparently incorrect.

[§§] Tentative values.

[§§§] The π values of imino acids containing a pyrrolidine ring are *stricto sensu* not defined; the given values have solely a comparative character.

mentioned above, one value for each of the following residues: Ada, Car, Cys, Trp and Pen. They were not considered in the subsequent evaluation.

Fractionation of the π-value

Using the π values in Table II we computed the contributions, π_f, of single groups and tested the validity of the additivity rule (eqn. 7). For this purpose we selected the 30 functional groups in Table III and fractionated each individual value of π_R according to eqn. 7. In this way, we have defined 532 equations and solved the 30×532 matrix by means of the least-squares method. Calculating back from the π_f

TABLE III

FRAGMENTAL CONTRIBUTIONS, π_f, OF THE CONSTITUENT GROUPS OF THE AMINO ACID SIDE-CHAINS

Fragment	π_f (aa)*	π_f (al)**	π_f (ar)**
CH$_3$	0.42 ± 0.03	0.50	0.56
CH$_2$	0.41 ± 0.02	0.51***	
CH	0.35 ± 0.06	0.41***	
C	0.39 ± 0.09	0.57***	
C$_6$H$_5$ (phenyl)	1.22 ± 0.04	2.15	1.96
C$_{10}$H$_{15}$CH$_2$ (adamantylmethane)	2.83 ± 0.05		
C$_2$H$_{11}$B$_{10}$CH$_2$ (o-carboranylmethane)	3.73 ± 0.05		
OH (aliphatic)	−0.46 ± 0.04	−1.12	
OH (aromatic)	−0.50 ± 0.09		−0.67
O$^-$ (aromatic)	−0.86 ± 0.08		
OH (on pyrrolidine)	−0.67 ± 0.13		
OCH$_3$ (aromatic)	0.24 ± 0.06		0.22 §
NH$_2$	−1.19 ± 0.19	−1.19	−1.23
NH$_3^+$	−2.12 ± 0.20	−4.19	
NHCH$_3$	−0.33 ± 0.09	−0.67	−0.47
NH$_2^+$CH$_3$	−1.85 ± 0.19		
NHCHO (formamide)	−0.70 ± 0.18		−0.96
NHCONH$_2$	−1.37 ± 0.08	−1.67***	−1.30
NHC(NH)NH$_3^+$ (guanidinium)	−2.13 ± 0.07		
C$_3$H$_3$N$_2$ (imidazole)	−0.29 ± 0.05		
C$_3$H$_4$N$_2^+$ (imidazole cation)	−1.26 ± 0.06		
C$_8$H$_6$N (indole)	1.43 ± 0.05		
C$_8$H$_7$N$^+$ (indole cation)	0.56 ± 0.13		
C$_4$H$_8$N (pyrrolidine) §§	0.77 ± 0.07		
COOH	−0.41 ± 0.06	−0.77***	−0.32
COO$^-$	−1.51 ± 0.05	−4.67	−4.63
CONH$_2$	−0.86 ± 0.04	−1.71	−1.49
SH	0.32 ± 0.05	0.28	0.39
S$^-$	−0.92 ± 0.05		
SCH$_3$	0.59 ± 0.05		0.61

* Fragmental contributions for amino acid side-chains. Standard deviations computed by means of regression analysis[28].

** Values of fragmental contributions for substitution in aliphatic (al) and aromatic (ar) backbones. Summarized by Hansch and Leo[11].

*** Computed from partition coefficients of various aliphatic compounds reported by Collander[16].

§ Another reported[11] value is −0.02 (both values for m- and p-substitution).

§§ Cf., footnotes §§§ in Table II.

values, "predicted (optimized)" values of π_R were obtained and are listed in Table II (last column). The fractional contributions and their standard deviations are indicated in Table III. The multiple correlation coefficient was highly significant (0.987), proving the validity of the additivity rule (eqn. 7) for the given set of π_R values.

The π_f values appear to be different from literature values[11] for aliphatic and aromatic fragments, as listed in Table III. Close correlations between π_f values for aromatic–amino acid and aliphatic–amino acid fragments prove that these differences are systematic. A clear outlier in the former correlation is the COO^- group. After its omission, the regression coefficients (eqn. 9) are 0.93 ± 0.06 and 0.48 ± 0.05 for correlation with aromatic and aliphatic fractional contributions, respectively. These values are least-squares fits to the following equation:

$$\pi_{f,1} = \kappa + \lambda\pi_{f,2} \tag{9}$$

(subscripts: 1 = amino acid backbone; 2 = aliphatic or aromatic substitution). In both instances, however, κ is not significantly different from zero (t-test: $\kappa = 0.07 \pm 0.21$ for 10 degrees of freedom in the former instance and 0.03 ± 0.34 for 12 degrees of freedom in the latter). It seems that the relationship between two sets of fragmental constants π_f (subscripts 1 and 2) follows a simple linear law:

$$\pi_{f,1} = \lambda\pi_{f,2} \tag{10}$$

Collander constants

In addition to the computations of π_R and $\log P_{OW}$, constants α and β in eqn. 3 for *n*-octanol–water (subscript a) and *n*-butanol–water (subscript b) were estimated by means of eqn. 6. In 26 experiments, their values were normally distributed, with arithmetic means and standard deviations of $\alpha = -0.48 \pm 0.24$ and $\beta = 1.53 \pm 0.20$. Literature values are $\alpha = -0.42$, $\beta = 1.24$ (Collander[16]) and $\alpha = -0.55$, $\beta = 1.44$ (Leo[17]). The differences are not appreciable.

DISCUSSION

Like other presently existing tools for assessing hydrophobicity, the π-scale derived from partition data represents only an arbitrary measure of this property. As for amino acid side-chains, several other measures, some of them probably better thermodynamically established than π values, have been suggested in the literature during the past decade[6–8,18,22]. However, the π values may possess several advantages in this specific case. They can be established with equal accuracy for both hydrophobic and hydrophilic residues, in contrast to other hydrophobicity parameters[6], and they are related to very dilute solutions where intramolecular interactions are reduced to a minimum. They have the further advantage, in our opinion, of being almost universally employed in pharmacology and in related fields for many types of substances. Another equally important advantage is their additive nature in the sense of eqn. 7. Indeed, it has been known for a long time that R_M values themselves may be fractionated into individual group contributions; the corresponding literature has been compiled by Lederer and Lederer[29] and the theory was later reviewed by Bush[30].

Knight[31] and Pardee[32] observed reported additivity even in peptide series: R_M values can be estimated from contributions of individual amino acids. The additivity of group fragments for amino acids was confirmed by our data, and Table III may serve for a simple estimation of π_R from fragmental contributions, π_f. However, such estimates have only a tentative character. We suggest that any definitive π value of an unlisted amino acid side-chain should be finally obtained experimentally, for which the chromatographic procedure described in this paper may become the method of choice.

The relationship between partition coefficient and R_F was derived long ago[14,15], but the uncertainty in measuring cross-sectional areas of solvent and aqueous phases has largely prevented its use in its original form. We have avoided this uncertain step by applying a calibration based on known partition coefficients of structurally similar compounds. For the amino acids, the data pairs were sufficiently well fitted by this empirical relationship. Also, an additional indirect proof, the computation of Collander constants by eqn. 6, showed a good agreement with the literature data[16,17] obtained in a different way.

However, if the hydrophobicity range becomes too broad, the theoretical relationship (eqn. 4) does not hold exactly. This failure may not be unexpected as eqn. 4 is dependent on an assumption that the composition of the mobile and stationary phases, and their volume ratio, are constant at any point on the chromatogram. Such conditions are not likely to be fulfilled; rather, these properties follow a gradient change which may be steep in the proximity of the solvent front. For computational purposes, we have replaced the linear eqn. 4 by a second-order polynomial in R_M (eqn. 8), which covers with sufficient accuracy a broad range of values. However, should the hydrophobicity range be extremely broad, it would be desirable to divide it into several narrower sections and carry out the interpolation within each one separately. Parameters of these relationships are purely empirical but the practical usefulness of eqn. 8 is not impaired by this fact.

In many respects, chromatography is to be preferred to alternative methods for the determination of partition coefficients and π values. First, no quantitative analytical method is needed, nor does the substance need to be extremely pure. Second, thin-layer chromatography is much simpler to carry out than the partitioning in n-octanol–water. This is probably valid for most substances but certainly for amino acids and peptides which, as mentioned, are generally only slightly soluble in n-octanol. The method is not bound to a particular chromatographic system and allows great freedom in the search for optimal chromatographic conditions. Finally, the pH or similar conditions are relatively easy to adjust and to check, which is not always the case in other partitioning procedures. This flexibility makes the chromatographic method particularly attractive.

In the procedure described above, the unknown values of log P_{OW} for amino acids and π values for their side-chains are obtained by means of interpolation within a set of partition data measured at pH 7. If a side-chain bears an electrical charge, its partition coefficient is pH dependent. One can, however, consider log P_{OW} for fully charged and completely uncharged compounds at pH 7, as it is in our case. Clearly, such values are not accessible to any measurement and can be obtained only by computations like the one presented here. These hypothetical values are physico-chemically admissible but have little biological or physical relevance. In almost any

biological system, a substance is partitioned at a pH close to 7, *i.e.*, as a population of various dissociation forms. An attempt to estimate such overall partition coefficients has been made in our experiments with approximately neutral solute systems. However, a more systematic study of the pH dependence of the chromatographic mobility may soon become necessary. The problem may also be solved by using amino acid derivatives with uncharged backbones. Experiments with N-acetylated amino acid amides are now being carried out in our laboratory[33].

As expected, the amino acids bearing charged side-chains display pH-dependent P and π values, whereas this phenomenon is generally absent in uncharged amino acids (Table II). The contribution of both positive and negative charges to a π value of an uncharged group amounts to approximately -1.0 and is lower only for the hydroxyl group on an aromatic ring (-0.36).

In the list of data in Table II, the log P and π values for arginine measured in alkaline solute systems (pH 13) have been omitted because of serious difficulties in obtaining reproducible R_F values (see footnote to Table II). These difficulties are understable, as the pK_a value of the guanidylo group (13.2) is close to the pH of the solvent system and its dissociation is therefore not fully suppressed. In fact, we have full confidence only in the values for uncharged side-chains, which show less scatter than those for electrically charged side-chains under physiological conditions. A few values marked in Table II will have to be confirmed in future experiments and should still be regarded as tentative.

ACKNOWLEDGEMENTS

REFERENCES

1 V. Pliška and J. L. Fauchère, in E. Gross and J. Meienhofer (Editors), *Peptides, Structure and Biological Function*, Pierce, Rockford, IL, 1979, p. 249.
2 T. Fujita, J. Isawa and C. Hansch, *J. Amer. Chem. Soc.*, 86 (1964) 5175.
3 A. Leo, C. Hansch and D. Elkins, *Chem. Rev.*, 71 (1971) 525.
4 E. C. Jorgensen and R. J. Weinkam, in H. Nesvadba (Editor), *Peptides 1971*, North-Holland, Amsterdam, 1973, p. 331.
5 G. Lindeberg, H. Vilhardt, L.-E. Larsson, P. Melin and V. Pliška, *J. Receptor Res.*, 1 (1980) 389.
6 Y. Nozaki and C. Tanford, *J. Biol. Chem.*, 246 (1971) 2211.
7 A. A. Aboderin, *Int. J. Biochem.*, 2 (1971) 537.
8 C. Tanford, *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, Wiley, New York, 1973, p. 120.
9 V. Pliška, *Experientia*, 34 (1978) 1190.
10 V. Pliška, in A. Eberle, R. Geiger and Th. Wieland (Editors), *Perspectives in Peptide Chemistry*, Karger, Basle, 1980, p. 221.
11 C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York, 1979, p. 13.
12 I. Molnár and Cs. Horváth, *J. Chromatogr.*, 142 (1977) 623.
13 J. L. Fauchère, K. Q. Do, P. Y. Jow and C. Hansch, *Experientia*, 36 (1980) 1203.
14 A. J. P. Martin and R. L. M. Synge, *Biochem. J.*, 35 (1941) 1358.
15 R. Consden, A. H. Gordon and A. J. P. Martin, *Biochem. J.*, 38 (1944) 224.
16 R. Collander, *Acta Chem. Scand.*, 5 (1951) 774.

17 A. J. Leo, in W. van Valkenburg (Editor), *Biological Correlations —the Hansch Approach*, American Chemical Society, Washington, DC, 1972, p. 51.
18 E. C. Bate-Smith and R. G. Westall, *Biochim. Biophys. Acta*, 4 (1950) 427.
19 F. E. Grubs, *Ann. Math. Stat.*, 21 (1950) 27.
20 W. J. Dixon, *Ann. Math. Stat.*, 22 (1951) 68.
21 B. L. Welch, *Biometrika*, 29 (1937) 350.
22 A. J. P. Martin, *Ann. N.Y. Acad. Sci.*, 49 (19r8) 249.
23 R. F. Rekker, *The Hydrophobic Fragmental Constant*, Elsevier, Amsterdam, 1977.
24 A. England and E. J. Cohn, *J. Amer. Chem. Soc.*, 57 (1935) 634.
25 R. A. Klein, M. J. Moore and M. W. Smith, *Biochim. Biophys. Acta*, 233 (1971) 420.
26 F. H. Carpenter, W. H. McGregor and J. A. Close, *J. Amer. Chem. Soc.*, 81 (1959) 849.
27 L. Sachs, *Statistische Auswertungsmethoden*, Springer, Berlin, 2nd ed., 1969, p. 446.
28 A. Hald, *Statistical Theory with Engeneering Applications*, Wiley, New York, 1959, p. 640.
29 E. Lederer and M. Lederer, *Chromatography, A Review of Principles and Applications*, Elsevier, Amsterdam, 2nd ed., 1957.
30 I. E. Bush, *Methods Biochem. Anal.*, 13 (1965) 357.
31 C. A. Knight, *J. Biol. Chem.*, 190 (1951) 753.
32 A. B. Pardee, *J. Biol. Chem.*, 190 (1951) 757.
33 J. L. Fauchère and V. Pliška, in K. Brunfeldt (Editor), *Peptides 1980*, Scriptor, Copenhagen, 1980, p. 637.