Correlation of Physicochemical Parameters to the Hydrophobic Contribution Constants of Amino Acid Residues in Small Peptides

Darshana Palekar,[1] Michael Shiue,[1] and Eric J. Lien[1,2]

Received February 29, 1996; accepted May 1, 1996

Purpose. This paper attempts to correlate the hydrophobic contribution constants (f_{aa}) of 21 amino acids in small peptides with commonly used physicochemical parameters. These f_{aa} constants can then be used to predict hydrophobicity change in peptides when any one of the amino acid residue is substituted with another.

Method. Non-weighted least squares method was used in deriving regression equations with a BMDP program. A Hyperchem program for Windows was used to calculate the group dipole moments of the side chain.

Results. A good correlation (r = 0.97) was obtained using a four parameter equation including molecular weight (log MW), hydrogen bond forming ability (HB), dipole moment (μ) and an indicator variable (I) to account for the presence of a free primary amine group in the side chain.

Conclusions. This proposed model should be useful in predicting the hydrophobic contribution constants of other uncommon amino acids and in the estimation of log P' values of numerous peptides containing different possible combinations of these amino acids, as well as log P' values resulting from amino acid substitution as is done in site-directed mutagenesis.

KEY WORDS: amino acids; hydrophobicity; hydrophobic contribution constants; partition coefficient.

INTRODUCTION

The hydrophobicity of a peptide is a function of its component amino acids and is an important parameter in determining its physicochemical and biological activities. It affects the three-dimensional structure of peptides in solution and influences their transport behavior across biomembranes as well as their interaction with receptor sites (1). Gao et al. (2), using the *de novo* approach have shown that the partition coefficients (log P') of oligopeptides (up to pentapeptides) can be correlated with their frequency of appearance and the hydrophobic contribution constants (f_{aa}) of the individual amino acids, as shown by the following equation:

$$\text{Log P'} = 0.81[\Sigma(n_i \times f_{aa})] - 0.31\ pK_1 - 0.12\ pK_2 + 0.11$$
$$(n = 59,\ r = 0.95,\ s = 0.22,\ F_{3,35} = 165.38) \quad (1)$$

where Log P' is the 1-octanol/buffer (pH = 7) partition coefficient of the peptide, n_i is the frequency of appearance of each individual amino acid in the peptide, f_{aa} is the hydrophobic contribution constant of each amino acid and pK_1 and pK_2 are the pK_a's of the N-amino and C-carboxyl groups of the peptide, respectively. In this paper, an attempt is made to correlate the hydrophobic contribution constants (f_{aa}) of the individual amino acids with commonly used physicochemical parameters such as molecular weight, hydrogen bond forming ability and dipole moment of the amino acid side chains (R). This model can then be used in conjunction with the above model for Log P' (Eq. 1) to predict the partition coefficients of uncommon amino acids and peptides containing them.

METHODS

The hydrophobic contribution constants for the first 19 essential amino acids were taken from the paper by Gao and Lien (2). These were obtained by correlating the Log P' of 57 oligopeptides reported by Akamatsu et al. (1) with the frequency of appearance of the individual amino acids.

The physicochemical parameters for the side chains (R) of the 19 essential amino acids are summarized in Table I. The maximum hydrogen bond number for each functional group was calculated based on the method described by Gao et al. (3), and was the sum of the total number of hydrogen donors and hydrogen acceptors. Further, the number of hydrogen donors equals the number of hydrogens which can form hydrogen bonds and the number of hydrogen acceptors equals the lone electron pairs of a given group. The values for commonly used functional groups are summarized in Table II.

The hydrogen bond forming ability (HB) of the side chain of any amino acid is the sum of the maximum hydrogen bond number of the various functional groups in the side chain. In the case of amino acids such as arginine, tyrosine and tryptophan, the HB is reduced by 1, in order to account for the delocalization of one lone pair of electrons into the π—electron network, thus making them unavailable to serve as H-bond acceptors.

The dipole moments (μ) for the side chains of the amino acids were calculated using a computer assisted program Hyperchem for Windows. For the calculations, the side chains were constructed as R-H, (where R- is the side chain of the amino acid) and energy minimized, to determine the optimum conformation for dipole moment determination.

An indicator variable (I = 1) was used for amino acids containing a free primary amine group in its side chain. In all other cases, a value of I = 0 was used. Stepwise multiple regression analysis was performed using the BMDP program (4) to derive the equations shown in Table III.

RESULTS AND DISCUSSION

Correlation of the Fragment Constants (f_{aa}) of Amino Acids with Molecular Weight (log MW), Hydrogen Bond Forming Ability (HB), Dipole Moment (μ), and Indicator Variable (I)

The values obtained for logMW, HB, μ and I were correlated to the (f_{aa}) constant via a stepwise multiple regression analysis using the BMDP program. The equations obtained are listed in Table III.

[1] Department of Pharmaceutical Sciences, School of Pharmacy, University of Southern California, 1985 Zonal Avenue, Los Angeles, California 90033.

[2] To whom correspondence should be addressed.

0724-8741/96/0800-1191$09.50/0

Table I. Physicochemical Parameters for the Amino Acids Side Chains (-R)

Amino Acid	R	f_{as} calc.[a]	f_{as} calc.[b]	MW	logMW	HB	μ	I
Ala (A)	$-CH_3$	−0.34	0.31	15.09	1.18	0	0.00	0
Arg(R)	$-(CH_2)_3$-NH-C(= NH)-NH_2	−1.86	−1.91	100.20	2.00	7(6)[f]	3.52	0
Asn(N)	$-CH_2$-C(= O)-NH_2	−1.10	−1.44	58.12	1.76	5	3.51	0
Asp(D)	$-CH_2$-COOH	−2.32	−2.07	59.10	1.77	5	1.65	0
Gln(Q)	$-CH_2-CH_2$-C(= O)-NH_2	−1.09	−1.27	72.15	1.86	5	3.75	0
Glu(E)	$-CH_2-CH_2$-COOH	−2.41	−1.96	73.13	1.86	5	1.74	0
Gly(G)	-H	−0.51	−0.76	1.07	0.03	0	0.00	0
His(H)	$-CH_2$(4-imidazolyl)	−0.54	0.25	81.16	1.91	3	3.93	0
Ile(I)	$-CH(CH_3)-CH_2-CH_3$	0.87	0.85	57.18	1.76	0	0.00	0
Leu(L)	$-CH_2-CH-(CH_3)_2$	0.97	0.85	57.18	1.76	0	0.00	0
Lys(K)	$-CH_2-CH_2-CH_2-CH_2-NH_2$	−2.43	−2.34	72.19	1.86	3	1.51	1
Met(M)	$-CH_2-CH_2-S-CH_3$	0.43	0.17	75.21	1.88	2	1.77	0
Orn(O)	$-CH_2-CH_2-CH_2-NH_2$	−2.33	−2.42	58.12	1.76	3	1.53	1
Phe(F)	$-CH_2-(C_6H_5)$	1.23	1.12	91.19	1.96	0	0.25	0
Ser(S)	$-CH_2$-OH	−0.78	−0.94	31.09	1.49	3	1.62	0
Thr(T)	$-CH(OH)-CH_3$	−0.50	−0.82	45.12	1.65	3	1.52	0
Trp(W)	$-CH_2$(3-indolyl)	1.47	1.23	130.23	2.11	2(1)[f]	2.18	0
Tyr(Y)	$-CH_2(C_6H_4OH)$	0.55	0.16	107.19	2.03	3(2)[f]	1.33	0
Val(V)	$-CH(CH_3)_2$	0.44	0.73	43.15	1.63	0	0.00	0
Pro(P)	$-(CH_2CH_2CH_2)-$	−0.19[c]	−0.17	41.10	1.61	2	1.49	0
Cys(C)	$-CH_2$-SH	−0.29[d]	−0.66	47.16	1.67	3	1.93	0

[a] Calculated from equation (2) of reference 1 (except Pro and Cys).
[b] Calculated from equation (4) of table III in this work.
[c] $f_{proline} = f_{leucine} - f_{CH3} + f_{cyclization} = 0.97 - 0.89 - 0.27 = 0.19$.
[d] $f_{cysteine} = f_{alanine} - f_{-H} + f_{-SH} = -0.34 - 0.23 + 0.28 = -0.29$.
[e] Represents an Indicator variable for the presence of an free primary amine group in the side chain of the amino acid.
[f] The HB value was reduced by one from the maximum number. See text for details.

Table II. Assignment of Hydrogen Bond Number to the Functional Groups Present in Amino Acid Side Chain

Function	Group in which present	HB
-OH	Alcohol	3
$-NH_2$	Primary amine Primary amide	3
-N(R)H	Secondary amine Secondary amide	2
-CO-	Amide Carboxylic acid	2

For these equations, ***r*** represents the correlation coefficient, ***s*** the standard deviation and ***n*** the number of data points (amino acids). Equation 4 (see Table III) was found to be the best equation according to the statistical F-test, with r = 0.97. All the variables in equation 4 were found to be statistically significant at the 95% significance level. Furthermore, by using this equation, 93% (r^2 = 0.932) of the variance in the data could be accounted for. The addition of other parameters such as pKa and an indicator variable for the presence of benzene ring in the side chain did not further improve the correlation significantly. A plot of f_{aa} (calculated) by Gao et al. (2) vs. f_{aa} (calculated) using equation 4 is shown in Figure 1.

These results indicate that physicochemical parameters can be used to estimate the hydrophobicities of amino acids and peptides containing these amino acids. Using equation 4, the f_{aa} constant for any uncommon amino acid can be derived. This calculated f_{aa} can then be used to predict the log P′ of peptides containing this amino acid, using the model described by Gao et al. (2).

Several investigators have described different means of quantifying the hydrogen bond forming abilities of various

Table III. The Stepwise Regression of f_{aa} with logMW, HB, μ, I

Equation	*n*	*s*	*r*	F statistic	Eq.
$f_{aa} = -0.474$ HB + 0.608	19	0.867	0.762	$F_{1,17} = 23.48$	1
$f_{aa} = -0.448$ HB − 1.767 I + 0.731	19	0.673	0.873	$F_{1,16} = 12.23$	2
$f_{aa} = -0.533$ HB − 1.852 I + 1.109 logMW − 0.939	19	0.466	0.945	$F_{1,15} = 18.30$	3
$f_{aa} = -0.702$ HB − 1.701 I + 0.931 logMW + 0.347μ − 0.787	19	0.383	0.966	$F_{1,14} = 8.21$	4
$f_{aa} = -0.645$ HB − 1.875 I + 1.094 logMW + 0.296μ − 1.003	19	0.532	0.932	$F_{1,14} = 3.03$	4a

Note: $F_{1,14,0.95} = 4.60$; $F_{1,15,0.95} = 6.20$.

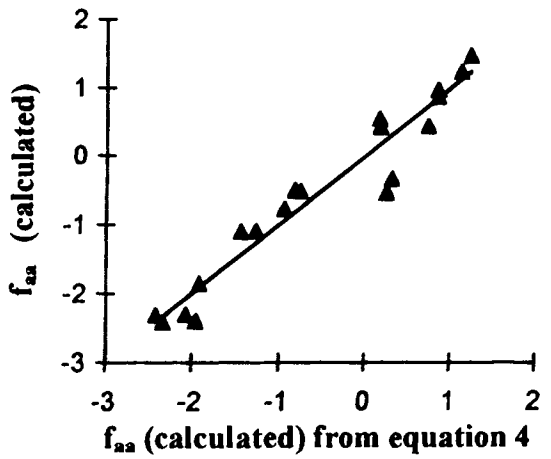

Fig. 1. Plot of f_{aa} (calculated) by Gao et al. (2) vs. f_{aa} (calculated) using Equation 4.

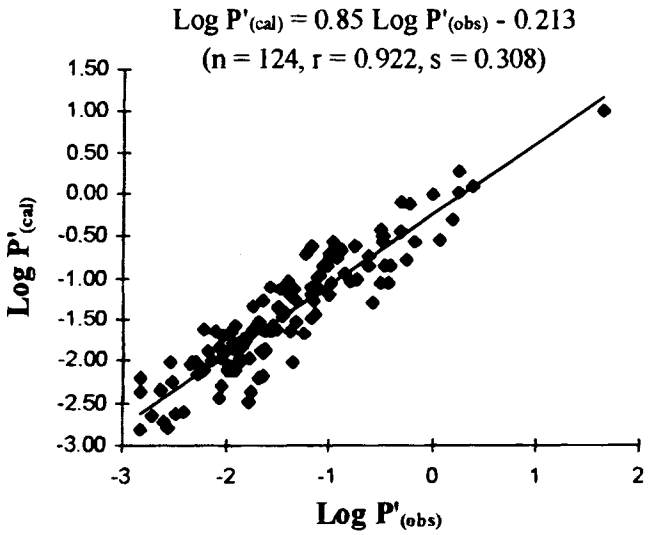

Fig. 2. A plot of Log $P'_{(cal)}$ vs Log $P'_{(obs)}$ of Peptides Listed in Table V.

donor/acceptor groups (5–8). Steric constraints, pH and nature of the solvent system and presence of ionic species are some of the factors affecting hydrogen bonding. In addition, amino acid side chains in large peptides are also involved in intramolecular hydrogen bonding. Since it would be difficult to parameterize all these effects, the sum of the maximum H-bond number for various donor/acceptor groups was chosen, as a measure of hydrogen bond forming ability for all the amino acid side chains. This approach has been shown to work reasonably well, especially for determining the hydrophobicity of peptides (3,9–10). In the case of arginine, tyrosine and tryptophan, the maximum hydrogen bond forming ability was reduced by 1, in order to account for the participation of the lone pair of electrons in resonance through the conjugated π system. As shown in Table III, the adjusted HB values increase the correlation slightly over the one using total HB numbers without correction. (Eq. 4 r = 0.966 vs. Eq. 4a r = 0.932). However, the coefficients associated with the parameters used were not drastically affected.

Multiple stepwise regression of f_{aa} using log MW, HB and μ yielded an equation with r = 0.87, with amino acids lysine and ornithine as possible outliers. There is a possibility that a three parameter equation was not able to completely explain the f_{aa} constants of these amino acids, due to an overestimation of the hydrogen bond forming ability of the amino acid side chains. Therefore, an indicator variable of I = 1 for lysine and ornithine and I = 0 otherwise, was included to improve the correlation. The inclusion of the indicator variable yielded a statistically significant four parameter equation as shown by equation 4, with r = 0.97 and no outliers and hence was selected as the 'best model' to describe the hydrophobic contribution constants (f_{aa}) of essential amino acids.

The positive contribution of increasing carbon chain length to the hydrophobic contribution constants (f_{aa}), is reflected in the f_{aa} values obtained for amino acids lysine and ornithine using equation 4 (Table III). In contrast, the values obtained for lysine and ornithine by Gao et al. (2) using the *de novo* approach, may reflect the reduced flexibility/ability of a shorter carbon chain (less than 4 carbon) to interact with its aqueous environment, thus making it less hydrophilic than expected.

The squared correlation matrix for the parameters used is given in Table IV. From this table, it is apparent that some

Table IV. The Squared Correlation Matrix of the Variables Used

	logMW	HB	μ	pKa	I
logMW	1.0000				
HB	0.1278	1.000			
μ	0.1902	0.6559	1.0000		
pKa	0.1082	0.1390	0.0988	1.0000	
I	0.0076	0.0095	0.0002	0.0061	1.0000

covariance exists between μ and HB. All the other parameters are practically independent of each other. Contrary to expectation, from equation 4, we observe that μ makes a small positive contribution to the f_{aa} constant. This may be a result of overestimation of the ability of the side chains to form hydrogen bonds as determined by the maximum H bond-forming ability method. Moreover, since HB and μ are correlated, an overestimation of HB may cause μ to make a positive compensation to the f_{aa} constant. Similar results have also been reported by Gao et al. (3).

Validation of the Model

To be useful, the model described in equation 4 must be capable of predicting the f_{aa} constant of any uncommon amino acid. In order to verify the utility of the equation obtained, we have calculated the f_{aa} constant of two additional amino acids namely proline to be −0.17 and cysteine to be −0.66 (see Table I). These appear to be in good agreement with the values calculated using the fragment method (see footnotes c and d under Table I). Based on the method described by Gao et al. (2), we further used the f_{aa} constants of the composition amino acids to calculate the log P′ values of 124 peptides containing them (see Table V). From these results, the model appears to work very well for tetrapeptides and pentapeptides. Introduction of the β-turn frequency (F_β) of tetrapeptides and pentapeptides into the correlation, turned out to be statistically significant. Figure 2 shows a plot of the calculated log P′ values vs. the observed log P′ values. These results indicate that the model

Table V. Hydrophobicities and pk_a's of Peptides

Peptide	$LogP'_{(obs)}$[b]	$LogP'_{(cal)}$[a]	pK_1[b]	pK_2[b]	F(β)
FL[c]	−1.17	−1.10	9.10	1.64	0
LF	−1.15	−1.28	9.60	1.80	0
FF	−0.85	−0.93	9.10	1.80	0
LL	−1.46	−1.44	9.60	1.64	0
LV	−2.05	−1.96	9.60	2.30	0
VL	−2.07	−1.84	9.60	1.64	0
AI	−2.60	−2.72	9.90	2.32	0
II	−1.82	−1.77	9.76	2.32	0
LI	−1.64	−1.64	9.60	2.32	0
VV	−2.82	−2.36	9.60	2.30	0
WW	−0.27	−0.77	9.39	2.38	0
WA	−1.98	−2.11	9.39	2.30	0
WF	−0.47	−0.84	9.39	1.80	0
WL	−0.73	−1.00	9.39	1.64	0
WY	−1.13	−1.42	9.39	2.20	0
LY	−1.94	−1.86	9.60	2.20	0
YL	−1.75	−1.61	9.10	1.64	0
VY	−2.52	−2.26	9.60	2.20	0
FY	−1.68	−1.52	9.10	2.20	0
YY	−1.87	−2.03	9.10	2.20	0
LM	−1.87	−1.97	9.60	2.28	0
ML	−1.84	−1.73	9.21	1.64	0
MV	−2.53	−2.25	9.21	2.30	0
FM	−1.59	−1.62	9.10	2.28	0
SL	−2.49	−2.64	9.20	1.64	0
PF	−2.07	−2.43	10.60	1.80	0
PL	−2.41	−2.60	10.60	1.64	0
PI	−2.56	−2.80	10.60	2.32	0
FP	−1.36	−2.02	9.10	1.99	0
LP	−1.76	−2.36	9.60	1.99	0
IP	−1.79	−2.49	9.76	1.99	0
FFF	−0.02	−0.01	9.10	1.80	0
GFF	−1.33	−1.52	9.80	1.80	0
FVG	−2.33	−2.02	9.10	2.40	0
FVF	−0.76	−0.60	9.10	1.80	0
FVA	−2.19	−1.87	9.10	2.30	0
LVV	−2.10	−1.63	9.60	2.30	0
LII	−1.11	−0.99	9.60	2.32	0
LVL	−1.57	−1.11	9.60	1.64	0
LAL	−2.03	−1.70	9.60	1.64	0
LLL	−0.94	−0.71	9.60	1.64	0
WGG	−2.72	−2.64	9.39	2.40	0
WFA	−1.00	−1.19	9.39	2.30	0
WWL	0.36	0.10	9.39	1.64	0
LLY	−1.34	−1.13	9.60	2.20	0
VFY	−1.50	−1.34	9.60	2.20	0
GFY	−1.96	−2.11	9.80	2.20	0
YLV	−1.45	−1.40	9.10	2.30	0
YVF	−1.37	−1.11	9.10	1.80	0
YGF	−1.86	−1.82	9.10	1.80	0
YYL	−1.38	−1.19	9.10	1.64	0
AYI	−2.04	−2.30	9.90	2.32	0
IYV	−1.77	−1.67	9.76	2.30	0
MLF	−1.03	−0.84	9.21	1.80	0
LSL	−2.35	−2.03	9.60	1.64	0
ISL	−2.28	−2.15	9.76	1.64	0
ISI	−2.64	−2.35	9.76	2.32	0
SLI	−1.99	−2.11	9.20	2.32	0
SLL	−2.03	−1.91	9.20	1.64	0
FIT	−1.95	−1.64	9.10	2.15	0
LIT	−2.14	−1.99	9.60	2.15	0
IIT	−2.23	−2.11	9.76	2.15	0
LTI	−2.30	−2.02	9.60	2.32	0
TLI	−1.66	−1.88	9.12	2.32	0
TVL	−1.97	−2.07	9.12	1.64	0
PLL	−1.64	−1.87	10.60	1.64	0
LPL	−1.56	−1.57	9.60	1.64	0
LLP	−1.58	−1.63	9.60	1.99	0
IPI	−1.65	−1.89	9.76	2.32	0
FGGF	−1.51	−1.61	9.10	1.80	0.34
VAAF	−1.91	−2.11	9.60	1.80	0.27
LLVF	−0.25	−0.13	9.60	1.80	0.34
LLLV	−0.51	−0.43	9.60	2.30	0.28
VGFF	−0.51	−1.06	9.60	1.80	0.27
AVLL	−1.74	−1.33	9.90	1.64	0.46
IAGF	−1.78	−1.96	9.76	1.80	0.3
FFFF	1.63	1.01	9.10	1.80	0.34
LLGF	−0.42	−0.84	9.60	1.80	0.34
LLAF	−1.00	−0.71	9.60	1.80	0.34
LLLF	0.24	0.27	9.60	1.80	0.34
IIVV	−1.41	−1.04	9.76	2.30	0.24
IIGF	−0.99	−1.05	9.76	1.80	0.3
IAAI	−2.82	−2.21	9.76	2.32	0.26
FFGF	0.17	−0.30	9.10	1.80	0.34
VLVL	−1.23	−0.71	9.60	1.64	0.28
WLLV	0.23	−0.02	9.39	2.30	0.35
WGLL	0.06	−0.54	9.39	1.64	0.44
YILG	−1.49	−1.36	9.10	2.40	0.44
FVYF	−0.32	−0.10	9.10	1.80	0.34
IYIV	−1.09	−0.96	9.76	2.30	0.24
VFLT	−1.32	−1.26	9.60	2.15	0.48
MILI	−0.49	−0.50	9.21	2.32	0.2
VMFI	−0.63	−0.73	9.60	2.32	0.24
PLLL	−1.06	−0.84	10.60	1.64	1.13
LPLL	−0.92	−0.75	9.60	1.64	0.35
LLPL	−1.00	−0.75	9.60	1.64	0.35
LLLP	−1.18	−0.61	9.60	1.99	1.13
IPGI	−1.69	−2.21	9.76	2.32	0.26
VPVL	−1.91	−1.56	9.60	1.64	0.28
VPGV	−2.83	−2.81	9.60	2.30	0.22
YPGW	−1.25	−1.67	9.10	2.38	0.79
YPGI	−1.65	−2.18	9.10	2.32	0.54
GGFVF	−1.40	−1.12	9.80	1.80	1.72
VFVGL	−0.97	−0.67	9.60	1.64	1.11
VGFVF	−0.50	−0.55	9.60	1.80	0.95
GAALL	−2.55	−2.02	9.80	1.64	1.43
AFGVF	−0.59	−1.29	9.90	1.80	0.7
AGFVF	−1.10	−1.13	9.90	1.80	1.32
LIIGA	−1.65	−1.26	9.60	2.30	1.37
GLLGF	−0.18	−0.57	9.80	1.80	3.03
ALLGF	−0.63	−0.85	9.90	1.80	1.62
IIIIG	−0.97	−0.57	9.76	2.40	1.10
IVVVI	−0.89	−0.65	9.76	2.32	0.48
FGAGI	−1.87	−1.85	9.10	2.32	1.79
FAAAL	−2.23	−1.62	9.10	1.64	0.91
WGGFV	−0.44	−1.05	9.39	2.30	1.21
WLFAA	−0.32	−0.44	9.39	2.30	1.05
IAYWG	−1.47	−1.12	9.76	2.40	1.66
GLSVL	−1.64	−1.84	9.80	1.64	1.12
SLAIV	−1.94	−1.78	9.20	2.30	0.96
YTGFI	−1.18	−1.19	9.10	1.64	1.22
LVGTF	−1.18	−1.47	9.60	1.80	0.88
YGGFL	−0.80	−1.03	9.10	1.64	1.58
YGGFM[d]	−1.39	−1.64	9.10	2.28	1.25

[a] Calc. using the equation $Log\ P'_{(cal)} = 0.751\Sigma f_{aa} - 0.300pK_1 - 0.187pK_2 + 0.265F(\beta) + 0.521$.

[b] Taken from reference (11).

[c] Calc. using $Log\ P'_{(cal)} = 0.751(1.23 + 0.97) - 0.300(9.1) - 0.187(1.64) + 0.265(0) + 0.521 = 1.36 - 3.54 - 0.30 + 1.36 = -1.10$.

[d] Calc. using $Log\ P'_{(cal)} = 0.751(0.55 - 0.51 - 0.51 + 1.23 + 0.43) - 0.300(9.1) - 0.187(2.28) + 0.265(1.25) + 0.521 = 0.63 - 3.54 - 0.42 + 0.33 + 1.36 = -1.64$.

described by equation 4 is practical and valid for estimating the fragment constants of any amino acids, including those containing an ionizable side chain. Thus, these f_{aa} values can be used to estimate the log P′ values of numerous small peptides (up to pentapeptides). Further work will be needed to predict the log P′ values of larger peptides with different secondary and tertiary structures.

CONCLUSIONS

In this paper, an attempt is made to correlate the hydrophobic contribution constants of 21 amino acids with various physicochemical parameters. A highly significant correlation (r = 0.966, Eq. 4, Table III) was obtained using molecular weight, hydrogen bond-forming ability and dipole moment of the side chain and an indicator variable to account for the presence of a primary amine group in the side chain of the amino acid. This model can now be used to predict the fragment constants of other uncommon amino acids and in the calculation of log P′ values of numerous small peptides containing them.

REFERENCES

1. M. Akamatsu, T. Katayama, D. Kshimoto, Y. Kurokawa, H. Shibata, T. Ueno and T. Fujita. Quantitative analyses of the structure-hydrophobicity relationship for N-acetyl di- and tripeptide amides. *J. Pharm. Sci.* **83**(7):1026–1033 (1994).
2. H. Gao, F. Wang and E. J. Lien. Hydrophobic contribution constants of amino acid residues to the hydrophobicities of oligopeptides. *Pharm. Res.* **12**:1279–1283 (1995).
3. H. Gao, E. J. Lien and F. Wang. Hydrophobicity of oligopeptides having unionizable side chains. *J. Drug Targeting* **1**:59–66 (1993).
4. W. J. Dixon, M. B. Brown, L. Engelman, J. W. Frane, M. A. Hill, R. I. Jennrich and J. D. Toporek. (eds.) Regression. *BMDP Statistical Software,* University of California Press, Berkeley, USA, 1985, pp. 235–283.
5. D. E. Leahy, J. J. Morris, P. J. Taylor and A. R. Wait. Model solvent systems for QSAR. Part 3. An LSER analysis of the 'critical quartet.' New light on hydrogen bond strength and directionality. *J. Chem. Soc. Perkin Trans.* **2**:705–722 (1992).
6. W. D. Stein. The molecular basis of diffusion across cell membranes. In W. D. Stein (eds.), *The Movement of Molecules Across Cell Membranes,* Academic Press, New York, 1967, pp. 65–125.
7. M. H. Abraham, P. L. Grellier, D. V. Prior and P. P. Duce. Hydrogen bonding: Part 7. A scale of solute hydrogen bond acidity based on log K values for complexation in tetrachloromethane. *J. Chem. Soc. Perkin Trans.* **2**:699–711 (1989).
8. G. Z. Yang, E. J. Lien, and Z. R. Guo. Physical factors contributing to hydrophobic constant π. *Q.S.A.R.* **5**:12–18 (1986).
9. E. J. Lien, H. Gao and H. Prabhakar. Physical factors contributing to the partition coefficient and the retention time of 2′, 3′-dideoxynucleotide analogues, *J. Pharm. Sci.* **80**:517–521 (1991).
10. E. J. Lien, H. Gao, F. Z. Wang and H. G. Shinouda. Partition behaviors of solvents, drugs and chemicals. 9th European Symposium On Structure-Activity Relationships: QSAR and Molecular Modeling, September 7th-11th, 1992, Strasbourg, France.
11. Akamatsu and T. Fujita. Quantitative analyses of hydrophobicity of di- to pentapeptides having un-ionizable side chains with substituent and structural parameters. *J. Pharm. Sci.* **81**:164–174 (1992).