Hydrophobic Contribution Constants of Amino Acid Residues to the Hydrophobicities of Oligopeptides

Hua Gao,[1] Fengzhen Wang,[1] and Eric J. Lien[1,2]

Received January 27, 1995; accepted March 20, 1995

Purpose. The main purpose of this study is to explore the additive-constitutive nature of common amino acids in their contribution to the partition coefficients of small peptides. ***Methods.*** The Log P values and other physico-chemical parameters of the peptides studied are taken from the literature. The frequency of appearance (n_i) of each individual amino acid is calculated as the number of the amino acids in a given peptide. ***Results.*** The partition coefficients (Log $P_{(oct./buff.)}$ at pH 7) of 87 N-acetyl-peptide-amides have been correlated with the frequency of appearance of amino acids. From the correlation obtained, the *de novo* hydrophobic contribution constants of 19 amino acid residues are derived for the first time. The contribution constants are extended to 59 unmodified regular peptides with the inclusion of the pk_a values of both N-terminal and C-terminal amino acids. The models thus obtained have been validated with additional 27 peptides (both N-acetyl-peptide-amides and unmodified). ***Conclusions.*** The Log P of oligopeptides is very well correlated with the *de novo* hydrophobic contribution constants of amino acids. The models we have derived are reasonably accurate in predicting the hydrophobicities of new oligopeptides (≤tetrapeptides) at a fixed pH (*e.g.*, 7).

KEY WORDS: hydrophobicity; oligopeptides; hydrophobic contribution constant; amino acids.

INTRODUCTION

The hydrophobicity of a peptide is an important parameter in peptide drug delivery (1,2) and dosage formulation. It is also an important factor in determining the tertiary structures and biological activities of proteins. Fauchère (3) has found that the biological activity of peptides can be correlated with π value as well as other amino acid side chain parameters. We have found that the hydrophobicities of oligopeptides (up to pentapeptides) can be correlated with the hydrophobicity of the individual amino acid, the molecular weight, the frequency of β-turn formation, and the calculated dipole moments of peptides (4). To further simplify the calculation of hydrophobicities of oligopeptides, we propose that the hydrophobicities of oligopeptides can be correlated with the frequency of appearance and the hydrophobic contribution constants of individual amino acids, and the pk_a values of both N-terminal and C-terminal amino acids, and β-turn frequency of the peptide. The simplified model(s) should be very useful in estimating the hydrophobicities of any new oligopeptides (up to tetrapeptides).

[1] Department of Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, CA 90033.
[2] To whom correspondence should be addressed.

MATERIALS AND METHODS

The Log P values of 57 N-acetyl-C-amidopeptides or N-acetyl-peptide-amides as used by the authors in the original paper are taken from the excellent paper of Akamatsu *et al.* (5). Briefly, the apparent hydrophobicities (Log P) at pH 7 of oligopeptides were measured by using aqueous $C_3H_7NH_2/C_3H_7NH_2 \cdot Cl$ buffer for peptides with acidic side chains and aqueous C_2H_5COOH/C_2H_5COONa buffer for peptides with basic side chains. The Log P, pk_a, and β-turn formation frequency (F_β) values of unmodified peptides are taken from our previous report (4). The terms pk_1 and pk_2 refer to pk_a of the N-amino and C-carboxy groups, respectively.

The frequency of appearance (n_i) of each individual amino acid is calculated as the number of the amino acids in a given peptide. For example, for Phe-Ala-Gly, the n_i for Phe is 1; for Phe-Ala-Phe, the n_i for Phe is 2.

RESULTS

Correlation of the Hydrophobicities of N-acetyl-peptide-amides

We have correlated the hydrophobicities (see Table I) of 87 N-acetyl-peptide-amides with n_i according to the following model:

$$\text{Log P} = \sum_{i=1}^{n}(n_i \times f\text{aa}) + f_{CH_3C(O)\text{-}} + f\text{-}NH_2 \quad \text{(1a)}$$

$$\text{Log P} = \sum_{i=1}^{n}(n_i \times f\text{aa}) + \text{Log P}_{(CH_3CONH_2)} \quad \text{(1b)}$$

$$\text{Log P} = \sum_{i=1}^{n}(n_i \times f\text{aa}) - 1.26 \quad \text{(1c)}$$

Where n_i is the frequency of the appearance of each individual amino acid in the peptide, and *f*aa is the hydrophobic contribution constant of each amino acid, $f_{CH_3C(O)\text{-}}$ is the fragment constant of $CH_3C(O)$- with a value of −1.01 in the acetylated N-terminal, and $f\text{-}NH_2$ is the fragment constant of amido-NH_2 with a value of −0.23 (6).

Substitute $f_{CH_3C(O)\text{-}}$ and $f\text{-}NH_2$ with their numeric values, Eq. (1d) is derived.

$$\text{Log P} = \sum_{i=1}^{n}(\text{ni} \times f\text{aa}) - 1.24 \quad \text{(1d)}$$

Eq. (2) is obtained from the regression of the reported data. The *f*aa's of 19 amino acids are summarized in Table II. Because the lack of data of pepetides containing proline, cysteine, asparagine, and glutamine, we can not get the *f*aa's for these amino acids at this time. The correlation is statistically very significant. The constant term (−1.20) is very close to the summation of $f_{CH_3C(O)\text{-}}$ and $f\text{-}NH_2$ (−1.24) (Eq.

0724-8741/95/0900-1279$07.50/0

Table I. Hydrophobicities of N-acetyl-peptide-amides

Peptide	Log $P_{(cal)}$[a]	Log $P_{(obs)}$[b]	Peptide	Log $P_{(cal)}$[a]	Log $P_{(obs)}$[b]
Ac-Gly-Val-NH_2	−1.27	−1.33	Ac-Ala-Gly-Phe-NH_2	−0.82	−0.71
Ac-Ala-Val-NH_2	−1.10	−1.13	Ac-Ile-Ala-Val-NH_2	−0.21	−0.21
Ac-Leu-Val-NH_2	0.20	0.26	Ac-Phe-Gly-Leu-NH_2	0.48	0.60
Ac-Gly-Phe-NH_2	−0.48	−0.56	Ac-Phe-Ile-Gly-NH_2	0.38	0.34
Ac-Ile-Val-NH_2	0.10	0.16	Ac-Val-Val-Ile-NH_2	0.54	0.49
Ac-Val-Val-NH_2	−0.32	−0.32	Ac-Gly-Leu-Gly-NH_2	−1.25	−1.23
Ac-Phe-Val-NH_2	0.46	0.43	Ac-Ala-Tyr-Leu-NH_2	−0.03	−0.41
Ac-Ala-Leu-NH_2	−0.57	−0.54	Ac-Ala-Tyr-Phe-NH_2	0.24	0.26
Ac-Ala-Ala-NH_2	−1.88	−2.00	Ac-Trp-Ala-Ala-NH_2	−0.41	−0.38
Ac-Gly-Leu-NH_2	−0.75	−0.78	Ac-Trp-Ile-Gly-NH_2	0.62	0.62
Ac-Leu-Ile-NH_2	0.63	0.68	Ac-Trp-Gly-Phe-NH_2	0.98	0.99
Ac-Phe-Gly-NH_2	−0.48	−0.50	Ac-Trp-Ala-Val-NH_2	0.37	0.36
Ac-Val-Ala-NH_2	−1.10	−1.14	Ac-Ala-Met-Val-NH_2	−0.67	−0.63
Ac-Tyr-Val-NH_2	−0.21	−0.20	Ac-Ile-Met-Phe-NH_2	1.33	1.28
Ac-Tyr-Leu-NH_2	0.31	0.32	Ac-Leu-Ser-Phe-NH_2	0.21	0.23
Ac-Tyr-Phe-NH_2	0.57	0.54	Ac-Leu-Thr-Leu-NH_2	0.23	0.24
Ac-Trp-Val-NH_2	0.70	0.73	Ac-Lys-Phe-Val-NH_2	−1.96	−2.13
Ac-Met-Val-NH_2	−0.32	−0.28	Ac-Lys-Ile-Phe-NH_2	−1.54	−1.46
Ac-Met-Phe-NH_2	0.46	0.42	Ac-Lys-Phe-Leu-NH_2	−1.44	−1.51
Ac-Ser-Val-NH_2	−1.54	−1.53	Ac-Leu-Lys-Phe-NH_2	−1.44	−1.41
Ac-Ser-Phe-NH_2	−0.76	−0.79	Ac-Orn-Phe-Leu-NH_2	−1.34	−1.37
Ac-Thr-Val-NH_2	−1.26	−1.25	Ac-Leu-Orn-Phe-NH_2	−1.34	−1.38
Ac-Thr-Ile-NH_2	−0.83	−0.86	Ac-Arg-Ile-Phe-NH_2	−0.97	−0.90
Ac-Asn-Val-NH_2	−1.86	−1.85	Ac-Arg-Phe-Leu-NH_2	−0.87	−1.04
Ac-Asn-Ile-NH_2	−1.43	−1.43	Ac-Leu-Arg-Phe-NH_2	−0.87	−0.76
Ac-Asn-Phe-NH_2	−1.07	−1.14	Ac-Leu-Phe-Arg-NH_2	−0.87	−0.93
Ac-Leu-Asn-NH_2	−1.33	−1.30	Ac-Ile-Phe-Arg-NH_2	−0.97	−0.93
Ac-Ile-Asn-NH_2	−1.43	−1.41	Ac-His-Ile-Phe-NH_2	0.35	0.36
Ac-Gln-Val-NH_2	−1.85	−1.85	Ac-Phe-His-Leu-NH_2	0.45	0.46
Ac-Gln-Leu-NH_2	−1.33	−1.32	Ac-Ile-His-Val-NH_2	−0.44	−0.33
Ac-Gln-Phe-NH_2	−1.07	−1.14	Ac-Gly-Phe-His-NH_2	−1.03	−1.09
Ac-Phe-Gln-NH_2	−1.07	−1.03	Ac-Trp-His-Val-NH_2	0.16	0.16
Ac-Val-Gln-NH_2	−1.85	−1.82	Ac-Phe-Trp-His-NH_2	0.95	0.89
Ac-Lys-Phe-NH_2	−2.40	−2.43	Ac-Asp-Phe-Leu-NH_2	−1.33	−1.39
Ac-Phe-Lys-NH_2	−2.40	−2.23	Ac-Asp-Ile-Phe-NH_2	−1.43	−1.32
Ac-Orn-Phe-NH_2	−2.30	−2.23	Ac-Phe-Asp-Leu-NH_2	−1.33	−1.19
Ac-Val-Ala-Ala-NH_2	−1.44	−1.40	Ac-Leu-Asp-Leu-NH_2	−1.59	−1.55
Ac-Val-Ala-Val-NH_2	−0.66	−0.67	Ac-Ile-Leu-Asp-NH_2	−1.69	−1.90
Ac-Val-Ile-Gly-NH_2	−0.41	−0.45	Ac-Glu-Phe-Leu-NH_2	−1.42	−1.52
Ac-Ala-Leu-Val-NH_2	−0.13	−0.14	Ac-Glu-Ile-Phe-NH_2	−1.52	−1.57
Ac-Val-Phe-Ala-NH_2	0.13	0.06	Ac-Phe-Glu-Phe-NH_2	−1.16	−1.08
Ac-Ala-Gly-Ile-NH_2	−0.23	−0.20	Ac-Leu-Glu-Phe-NH_2	−1.42	−1.25
Ac-Ile-Phe-Ala-NH_2	0.55	0.52	Ac-Leu-Ile-Glu-NH_2	−1.78	−1.87
Ac-Gly-Ala-Val-NH_2	−1.61	−1.56			

[a] Calculated from Eq. (2).
[b] Taken from Ref. (5).

1d) and the Log P of CH_3CONH_2(− 1.26) (Eq. 1c). A plot of the experimental Log P values v the calculated Log P values from Eq. (2) is presented in Figure 1. The correlation is statistically highly significant as indicated by the F-test as well as the r and s values.

$$\text{Log P} = \sum_{i=1}^{n} (n_i \times f_{aa}) - 1.20 \qquad (2)$$

$$(n = 87,\ r = 1.00,\ s = 0.08,\ F_{19.67} = 596.89)$$

Extension of the Model to Unmodified Regular Peptides

We propose that the hydrophobicities of peptides without N-acetyl and C-amido groups should follow the model presented in Eq. (3).

$$\text{Log P} = a\left[\sum_{i=1}^{n}(n_i \times f_{aa})\right] + b\text{pk}_1 + c\text{pk}_2 + d \qquad (3)$$

Where a, b, and c are coefficients, d is a constant.

We have correlated the hydrophobicities (see Table III) of 59 di- and tripeptides according to Eq. (3). Eq. (4) is derived with an r value of 0.95 (r^2 = 0.90, 90% of data can be accounted for). The hydrophobicities of the peptides are well correlated with the summation of the hydrophobic contribu-

Table II. The Hydrophobic Contribution Constants of Amino Acid Residues

Amino Acid	Contribution Constant[a] (f_{aa})	Amino Acid	Contribution Constant (f_{aa})
Ala	−0.34	Lys	−2.43
Asn	−1.10	Met	0.43
Asp	−2.32	Orn	−2.33
Arg	−1.86	Phe	1.23
Gln	−1.09	Ser	−0.78
Glu	−2.41	Thr	−0.50
Gly	−0.51	Trp	1.47
His	−0.54	Tyr	0.55
Ile	0.87	Val	0.44
Leu	0.97		

[a] Obtained from Eq. (2).

tion constants of the amino acids obtained from Eq. (2) (see Table II), the pk_1, and the pk_2 values. A plot of the experimental Log P values *v.* the calculated Log P values from Eq. (4) is illustrated in Figure 2.

$$\text{Log P} = 0.81\left[\sum_{i=1}^{n} (n_i \times f_{aa})\right] - 0.31\ pk_1 - 0.12\ pk_2 + 0.11 \quad (4)$$

$$(n = 59,\ r = 0.95,\ s = 0.22,\ F_{3,55} = 165.38)$$

Validation of the Models

To be useful, the models obtained above should be capable of making reasonably accurate estimations of hydrophobicities for peptides not included in the correlation sets. We have tested a total of 27 additional peptides, 24 tetrapeptides and three N-acetyl-peptide-amides (1) (see Table IV) by using Eq. (2) and Eq. (4). From the results, one can see that the models worked very well even for tetrapeptides. We did try to include the β-turn frequency (F_β) of the tetrapeptides into the correlation, but the contribution of F_β turned out to

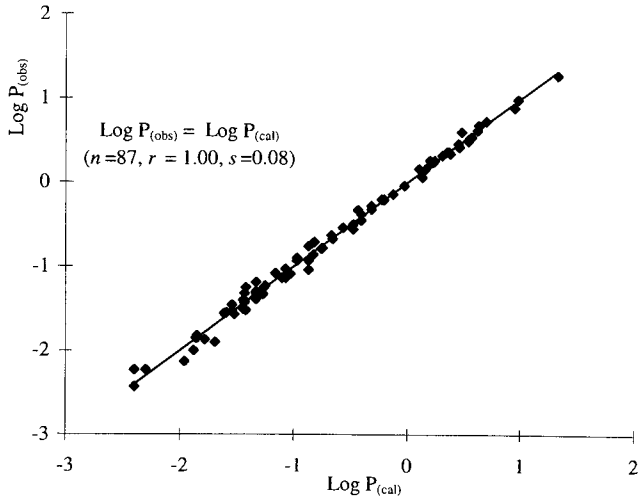

Fig. 1. The plot of Log $P_{(obs)}$ *v.* Log $P_{(cal)}$ of N-acetyl-peptide-amides.

Table III. Hydrophobicities and pk_a's of Peptides

Peptide[b]	Log $P_{(cal)}$[a]	Log $P_{(obs)}$[b]	pk1[b]	pk2[b]
Phe-Leu	−1.15	−1.17	9.10	1.64
Leu-Phe	−1.33	−1.15	9.60	1.80
Phe-Phe	−0.96	−0.85	9.10	1.80
Leu-Leu	−1.52	−1.46	9.60	1.64
Leu-Val	−2.02	−2.05	9.60	2.30
Val-Leu	−1.95	−2.07	9.60	1.64
Ala-Ile	−2.83	−2.60	9.90	2.32
Ile-Ile	−1.81	−1.82	9.76	2.32
Leu-Ile	−1.68	−1.64	9.60	2.32
Val-Val	−2.45	−2.82	9.60	2.30
Trp-Trp	−0.73	−0.27	9.39	2.38
Trp-Ala	−2.18	−1.98	9.39	2.30
Trp-Phe	−0.86	−0.47	9.39	1.80
Trp-Leu	−1.05	−0.73	9.39	1.64
Trp-Tyr	−1.45	−1.13	9.39	2.20
Leu-Tyr	−1.92	−1.94	9.60	2.20
Tyr-Leu	−1.70	−1.75	9.10	1.64
Val-Tyr	−2.35	−2.52	9.60	2.20
Phe-Tyr	−1.56	−1.68	9.10	2.20
Tyr-Tyr	−2.11	−1.87	9.10	2.20
Leu-Met	−2.03	−1.87	9.60	2.28
Met-Leu	−1.83	−1.84	9.21	1.64
Met-Val	−2.33	−2.53	9.21	2.30
Phe-Met	−1.66	−1.59	9.10	2.28
Ser-Leu	−2.81	−2.49	9.20	1.64
Phe-Phe-Phe	0.04	−0.02	9.10	1.80
Gly-Phe-Phe	−1.59	−1.33	9.80	1.80
Phe-Val-Phe	−0.60	−0.76	9.10	1.80
Phe-Val-Ala	−1.93	−2.19	9.10	2.30
Phe-Val-Gly	−2.08	−2.33	9.10	2.40
Leu-Val-Val	−1.67	−2.10	9.60	2.30
Leu-Ile-Ile	−0.98	−1.11	9.60	2.32
Leu-Val-Leu	−1.16	−1.57	9.60	1.64
Leu-Ala-Leu	−1.79	−2.03	9.60	1.64
Leu-Leu-Leu	−0.74	−0.94	9.60	1.64
Trp-Gly-Gly	−2.75	−2.72	9.39	2.40
Trp-Phe-Ala	−1.19	−1.00	9.39	2.30
Trp-Trp-Leu	0.14	0.36	9.39	1.64
Leu-Leu-Tyr	−1.14	−1.34	9.60	2.20
Val-Phe-Tyr	−1.36	−1.50	9.60	2.20
Gly-Phe-Tyr	−2.19	−1.96	9.80	2.20
Tyr-Leu-Val	−1.42	−1.45	9.10	2.30
Tyr-Val-Phe	−1.15	−1.37	9.10	1.80
Tyr-Gly-Phe	−1.92	−1.86	9.10	1.80
Tyr-Tyr-Leu	−1.26	−1.38	9.10	1.64
Ala-Tyr-Ile	−2.39	−2.04	9.90	2.32
Ile-Tyr-Val	−1.71	−1.77	9.76	2.30
Met-Leu-Phe	−0.85	−1.03	9.21	1.80
Leu-Ser-Leu	−2.15	−2.35	9.60	1.64
Ile-Ser-Leu	−2.28	−2.28	9.76	1.64
Ile-Ser-Ile	−2.44	−2.64	9.76	2.32
Ser-Leu-Ile	−2.19	−1.99	9.20	2.32
Ser-Leu-Leu	−2.03	−2.03	9.20	1.64
Phe-Ile-Thr	−1.70	−1.95	9.10	2.15
Leu-Ile-Thr	−2.07	−2.14	9.60	2.15
Ile-Ile-Thr	−2.20	−2.23	9.76	2.15
Leu-Thr-Ile	−2.09	−2.30	9.60	2.32
Thr-Leu-Ile	−1.94	−1.66	9.12	2.32
Thr-Val-Leu	−2.20	−1.97	9.12	1.64

[a] Calculated from Eq. (4).
[b] Taken from Ref. (4).

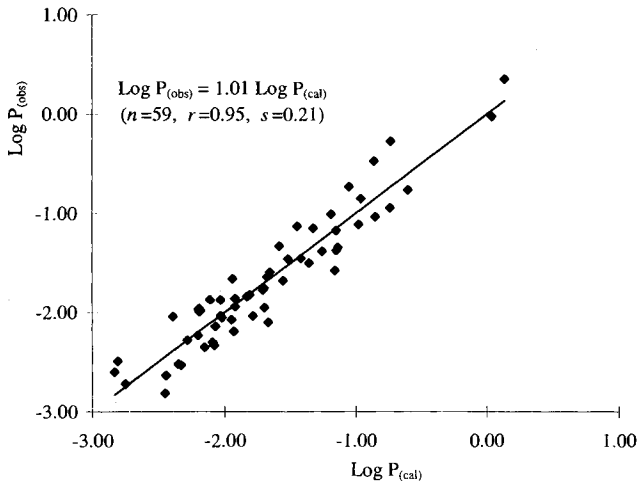

Fig. 2. The plot of Log $P_{(obs)}$ *v.* Log $P_{(cal)}$ of peptides listed in Table III.

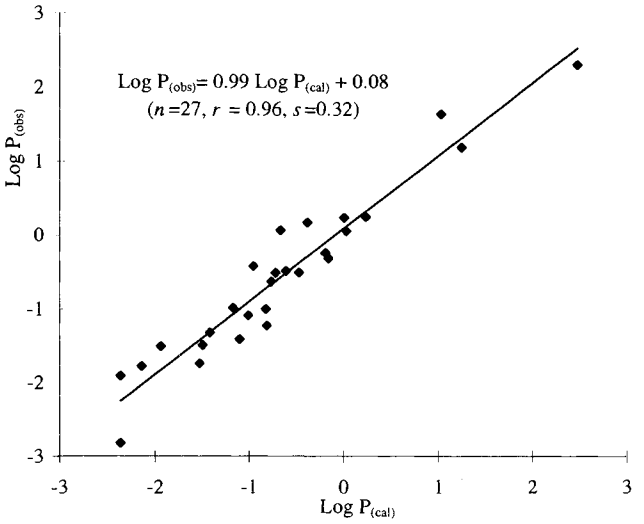

Fig. 3. A plot of Log $P_{(obs)}$ *v.* Log $P_{(cal)}$ of peptides listed in Table IV.

be not significant for tetrapeptides. Figure 3 shows the plot of experimental Log P values and calculated log P values from Eq. (2) and (4). Unfortunately we do not have data for the validation of peptides containing acidic or basic side chains. Since ionizable side chains of Arg, Lys, Orn, Glu, and Asp can form ion-pairs with counter ions in a buffer solution (5,7), the same buffer is recommended for the calculation of Log P values of oligopeptides containing these amino acids.

Table IV. Hydrophobicities and F_β of Peptides

Peptides	Log $P_{(cal)}$[a]	Log $P_{(obs)}$[b]	F_β[b]
Phe-Gly-Gly-Phe	−1.94	−1.51	0.34
Val-Ala-Ala-Phe	−2.36	−1.91	0.27
Leu-Leu-Val-Phe	−0.19	−0.25	0.34
Leu-Leu-Leu-Val	−0.47	−0.51	0.28
Val-Gly-Phe-Phe	−0.72	−0.51	0.27
Ala-Val-Leu-Leu	−1.53	−1.74	0.46
Ile-Ala-Gly-Phe	−2.14	−1.78	0.30
Phe-Phe-Phe-Phe	1.03	1.63	0.34
Leu-Leu-Gly-Phe	−0.95	−0.42	0.34
Leu-Leu-Ala-Phe	−0.82	−1.00	0.34
Leu-Leu-Leu-Phe	0.24	0.24	0.34
Ile-Ile-Val-Val	−1.17	−0.99	0.30
Ile-Ile-Gly-Phe	−1.10	−1.41	0.24
Ile-Ala-Ala-Ile	−2.36	−2.82	0.26
Phe-Phe-Gly-Phe	−0.38	0.17	0.34
Val-Leu-Val-Leu	−0.81	−1.23	0.28
Trp-Leu-Leu-Val	0.01	0.23	0.35
Trp-Gly-Leu-Leu	−0.66	0.06	0.44
Tyr-Ile-Leu-Gly	−1.50	−1.49	0.40
Phe-Val-Tyr-Phe	−0.16	−0.32	0.34
Ile-Tyr-Ile-Val	−1.01	−1.09	0.24
Val-Phe-Leu-Thr	−1.42	−1.32	0.48
Met-Ile-Leu-Ile	−0.61	−0.49	0.20
Val-Met-Phe-Ile	−0.76	−0.63	0.24
Ac-Phe-NH_2[c]	0.03	0.05	0.00
Ac-Phe-Phe-NH_2[c]	1.25	1.19	0.00
Ac-Phe-Phe-Phe-NH_2[c]	2.48	2.30	0.00

[a] Calculated from Eq. (2) for the modified peptides and Eq. (4) for the unmodified peptides, respectivley.
[b] Taken from Ref. (4).
[c] Taken from Ref. (1).

CONCLUSION

The *de novo* hydrophobic contribution constants of 19 amino acids have been derived. The hydrophobicities of di- to tetrapeptides are correlated very well with the summation of f_{aa} of individual amino acids, the pk_1, and the pk_2 of unmodified terminal amino acids. The models have been validated with a new set of peptides. A similar *de novo* constant approach has been successfully used by Meylan *et al.* (8) to estimate the octanol/water partition coefficients of many miscellaneous chemical compounds. This approach has not been used in estimation of hydrophobicities of peptides before. Our models should be very useful in estimating hydrophobicities for new di- to tetrapeptides.

To what extent the additive-constitutive models developed for small peptides can be extended to larger peptides (pentapeptides or higher) and proteins with modified amino acids, unnatural amino acids or D-amino acids remains to be studied. It is likely that additional parameters need to be introduced to account for differences in secondary and tertiary structures (*e.g.*, β-sheet, α-helix, and random coil forms) in the macromolecules.

REFERENCES

1. R. A. Conradi, A. R. Hilgers, N. F. H. Ho, and P. S. Burton. The influence of peptide structure on transport across Caco-2 Cells. *Pharm. Res.* 8(12):1453–1460 (1991).
2. V. H. L. Lee. *Peptide and protein drug delivery* Marcel Dekker, Inc., New York, 1991.
3. J. L. Fauchère. Amino acid side chain parameters and their use in QSAR studies of polypeptide hormones. In J. K. Seydel (eds), *QSAR and strategies in the design of bioactive compounds*. VCH Verlagsgesellschaft mbH, Weinheim, Germany, 1985, pp. 79–82.

4. H. Gao, E. J. Lein, and F. Z. Wang. Hydrophobicity of oligopeptides having unionizable side chains. *J. Drug Target.* **1**:59–66 (1993).
5. M. Akamatsu, T. Katayama, D. Kishimoto, Y. Kurokawa, H. Shibata, T. Ueno, and T. Fujita. Quantitative analysis of the structure-hydrophobicity relationship for N-acetyl di- and tripeptide amides. *J. Pharm. Sci.* **83**(7):1026–1033 (1994).
6. C. Hansch and A. Leo. *Substituent constants for correlation analysis in chemistry and biology*. John Wiley & Sons, New York, 1979, pp. 18–43.
7. P. H. Wang and E. J. Lien. Effect of different buffer species on partition coefficients of drugs used in quantitative structure-activity relationships. *J. Pharm. Sci.* **69**(6):662–669 (1980).
8. W. M. Meylan and P. H. Howard. Atom/fragment contribution method for estimating octanol-water partition coefficients. *J. Pharm. Sci.* **84**(1):83–92 (1995).