Hydrogen Bonding. 47. Characterization of the Ethylene Glycol–Heptane Partition System: Hydrogen Bond Acidity and Basicity of Peptides

MICHAEL H. ABRAHAM,*,† FILOMENA MARTINS,‡ ROBERT C. MITCHELL,§ AND COLIN J. SALTER§

Contribution from *Department of Chemistry, University College London, 20 Gordon Street, London, WC1H OAJ, UK, Grupo de Estrutura e Reactividade Quimica, Department of Chemistry, University of Lisbon, Cc. Bento da Rocha Cabral 14, 1250 Lisboa, Portugal, and SB Pharmaceuticals, New Frontiers Science Park (North), Third Avenue, Harlow, Essex, CM19 5AW, UK.*

Received June 9, 1998. Accepted for publication October 19, 1998.

Abstract □ Twelve measured ethylene glycol–heptane partition coefficients, P_{eh}, have been combined with 20 measured literature values and 44 indirectly determined values to give a set of 76 values. Excluding one value for benzamide, the log P_{eh} values are correlated through our general solvation equation, log $P_{eh} = 0.336 - 0.075R_2 - 1.201\pi_2^H - 3.786\ \Sigma\alpha_2^H - 2.201\ \Sigma\beta_2^H + 2.085V_x$ with $r^2 = 0.966$, sd $= 0.28$, and $F = 386$. The solute descriptor R_2 is the excess molar refraction, π_2^H is the dipolarity/polarizability, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ are the overall hydrogen bond acidity and basicity, and V_x is the McGowan volume. The log P_{eh} equation has then been used to obtain descriptors for eleven peptides, all of which are end-protected. It is shown that for these end-protected peptides, hydrogen bond basicity makes a greater contribution to log P_{eh} than does hydrogen bond acidity.

Introduction

Karls et al.,[1] in 1991, showed that the hydrogen bond number of Stein[2] was useful in the correlation of absorption of a series of peptides in rats. The peptides, I–VII, are shown in Figure 1, and the hydrogen bond numbers, H#, are in Table 1. The latter are simply calculated from the total number of C=O and N–H groups, with NH_2 counted as two. Burton et al.[3] used the same seven peptides to study permeability across Caco-2 cell monolayers and showed graphically that there was a reasonable correlation between log k_{caco} and the Δlog P parameter of Seiler.[4] The latter is obtained from water–octanol and water–alkane partition coefficients through eq 1,

$$\Delta\log P = \log P_{oct} - \log P_{alk} \quad (1)$$

Burton et al.[3] did not explore the H# numbers, but investigated the use of a novel partitioning system, that of ethylene glycol–heptane:

$$P_{eh} = [\text{concn in heptane}]/[\text{concn in ethylene glycol}] \quad (2)$$

They showed that for the seven peptides there was a reasonable plot of log P_{eh} against Δlog P, and argued that since Δlog P could be taken as a measure of the hydrogen bond or desolvation potential of a solute, so could log P_{eh}. The various partition coefficients that were determined[3] for the seven peptides are given in Table 1; the alkane used was isooctane. In later work it was shown[5] that permeabilities of the seven peptides across an in vitro model of the blood–brain barrier (BBB) or from physiologic saline to rat brain could be correlated with H# or Δlog P or log P_{eh}. The hydrogen bonding capacity of the peptides was suggested to be the major factor governing both the in vitro and the perfusion in vivo permeabilities.

Table 1—Hydrogen Bond Numbers and Partition Coefficients for Peptides[a]

peptides	H#	log P_{oct}	log P_{iso}[b]	log P_{eh}
I	5	0.05	–4.92	–5.46 (–5.57)
II	7	1.19	–5.29	–6.52 (–6.28)
III	9	2.30	–5.02	–7.10 (–7.16)
IV	8	2.63	–4.20	–6.28
V	7	2.53	–3.10	–5.14
VI	6	2.92	–1.67	–4.20
VII	5	3.24	–0.69	–2.86
VIII	8			(–6.00)
IX	8			(–5.76)
X	7			(–4.35)
XI	5			(–3.44)

[a] log P values from ref 3; values in parentheses from ref 6. [b] P_{iso} refers to partition between water and isooctane.

To characterize the ethylene glycol–heptane partitioning system, Paterson et al.[6] measured log P_{eh} values for a set of 20 standard compounds, and correlated these values with the solvatochromic parameters of Kamlet et al..[7] The log P_{eh} values are in Table 2, nos. 1–20, and the correlation equation was:

$$\log P_{eh} = 0.30 - 1.53\pi^* - 4.41\alpha - 1.69\beta + 2.79V_I \quad (2)$$
$$n = 19,\ r^2 = 0.984,\ \text{sd} = 0.19,\ F = 223$$

In eq 2, π^* is the solute dipolarity/polarizability, α is the solute hydrogen bond acidity, β is the solute hydrogen bond basicity, and V_I is the solute intrinsic volume in (cm^3 mol^{-1})/100. Benzamide was left out of the correlation, hence n, the number of data points, is 19; the only other statistic given[6] was $r^2 = 0.980$, where r is the correlation coefficient. We have rerun the correlation and find $r^2 = 0.984$ (the value of 0.980 is the adjusted value), the standard deviation sd = 0.19 and the F-statistic $F = 223$, as shown above.

Paterson et al.[6] pointed out that eq 2 shows that the most important coefficient is the solute hydrogen bond acidity. However, this does not mean that the term in α always dominates; this will depend on the various parameter values for any given solute. Although the solvatochromic parameters of Kamlet et al.[7] are useful, there are difficulties in that parameters for many solutes have to be estimated, and that there is no protocol for the estimation of parameters for new structures such as peptides I–VII. In addition, although Paterson et al.[6] selected a reasonable set of compounds, the number is minimal for a four-

† University College London.
‡ University of Lisbon.
§ SB Pharmaceuticals.

10.1021/js980242l CCC: $18.00
Published on Web 12/10/1998

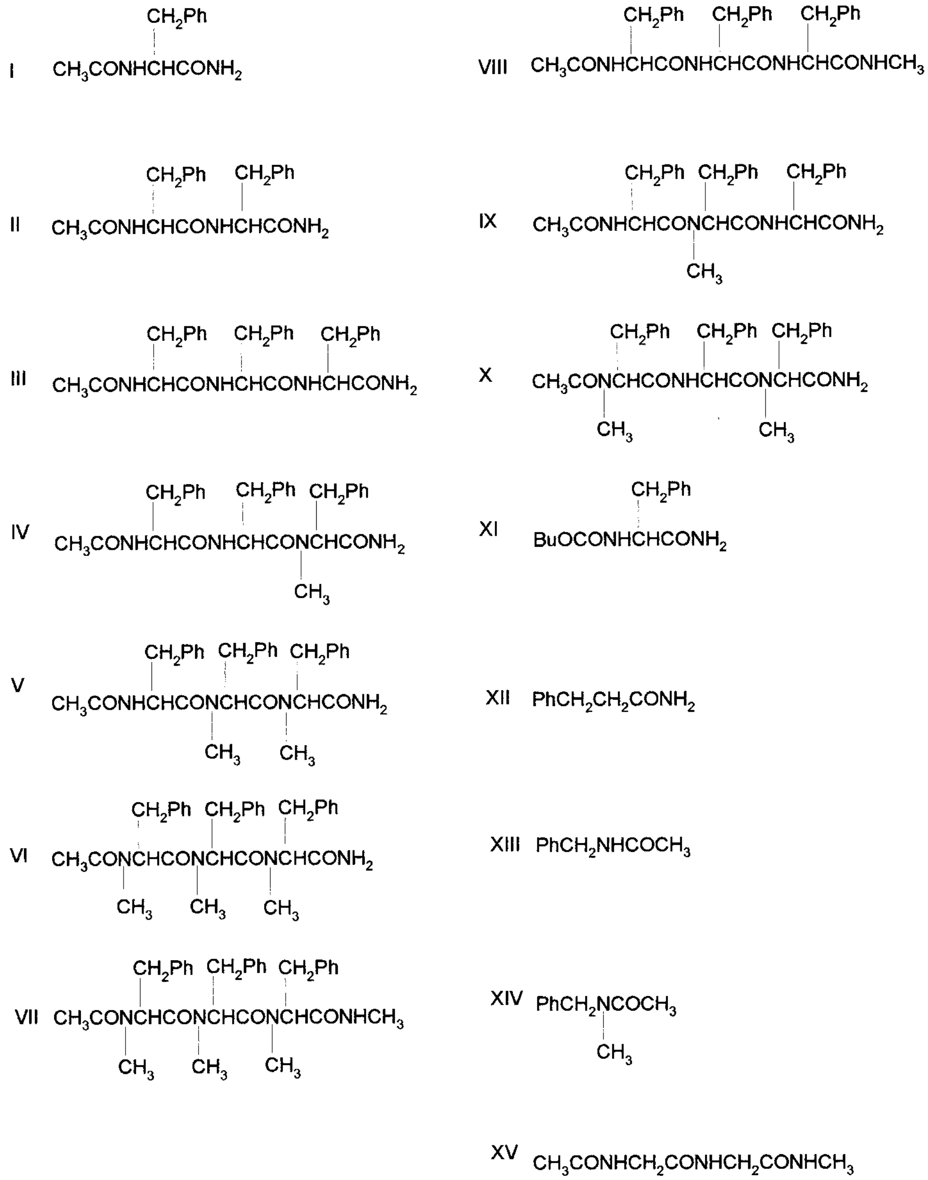

Figure 1—The structures of the peptides.

parameter equation, and the range of the various descriptors is not very large.

We therefore set out to expand the number and nature of the solutes, to characterize the ethylene glycol–heptane system through the solvation equation and the solvation parameters of Abraham,[8] and to compare coefficients in the correlation equation with those for many other partitioning systems that have been thus characterized.[9] Once this has been done, our aim is to determine solvation parameters for the peptides I–VII, and then to ascertain the factors that influence partitioning of peptides in the ethylene glycol–heptane system.

Experimental Section

Ethylene glycol was Fisons Analytical Reagent, and heptane was Fisher Scientific Analytical Reagent. Partition experiments were carried out with mutually saturated solvents. The various solutes were from Sigma or Aldrich, except for benzyl alcohol, zolantidine, and clonidine which were in-house specimens. Preliminary experiments showed that the maximum water content of the ethylene glycol was 0.11% by Karl Fischer titration, either with fresh solvent or after shaking the solvent in a separating funnel with a significant air space. Since the ethylene glycol did not appear to be particularly hydroscopic, no special precautions were taken to exclude water. A number of partition experiments were carried out on benzyl alcohol, with different combinations of sonication and tumbling. Results showed that 15 min sonication using a SEMAT ultrasonic bath followed by mixing for 1 h on a mechanical flask tumbler was sufficient to reach equilibrium.

Partition measurements were carried out by dissolving the solute in presaturated ethylene glycol. In the case of compounds obtained as salts (imipramine hydrochloride, mepyramine maleate, and clonidine hydrochloride) solid potassium hydroxide was first dissolved in the ethylene glycol at a concentration of 1.0 mg ml^{-1}. A known volume of the ethylene glycol solution was transferred to a Nalgene FEP bottle, and a known volume of preequilibrated heptane was added. The phases were mixed, as above, and allowed to separate, with centrifugation for 3 min at 3000 rpm, if necessary. The ethylene glycol solutions before and after partitioning were analyzed by UV spectrometry. In all case, determinations were carried out in duplicate, using different volume ratios of ethyene glygol:heptane. The volume ratios were chosen so that the initial and final absorbances were significantly different. In the case of pentachlorophenol and anthracene, solubility in ethylene glycol was very low. The compounds were therefore initially dissolved in heptane, and it was this phase that was analyzed before and after partitioning. Results are in Table 2 under compounds nos. 21–32.

Table 2—Solute Descriptors and Values of log P_{eh}

no.	solute	R_2	π_2^H	$\Sigma\alpha_2^H$	$\Sigma\beta_2^H$	V_x	log P_{eh}	no.	solute	R_2	π_2^H	$\Sigma\alpha_2^H$	$\Sigma\beta_2^H$	V_x	log P_{eh}
1	benzene	0.610	0.52	0.00	0.14	0.7164	0.475	39	carbon dioxide	0.000	0.28	0.05	0.10	0.2809	0.320
2	toluene	0.601	0.52	0.00	0.14	0.8573	0.994	40	methane	0.000	0.00	0.00	0.00	0.2495	1.080
3	bromobenzene	0.882	0.73	0.00	0.09	0.8914	0.970	41	ethane	0.000	0.00	0.00	0.00	0.3904	1.360
4	methyl phenyl ether	0.708	0.75	0.00	0.29	0.9160	0.673	42	propane	0.000	0.00	0.00	0.00	0.5313	1.270
5	benzaldehyde	0.820	1.00	0.00	0.39	0.8730	–0.131	43	butane	0.000	0.00	0.00	0.00	0.6722	1.750
6	acetophenone	0.818	1.01	0.00	0.48	1.0139	–0.051	44	2-methylpropane	0.000	0.00	0.00	0.00	0.6722	1.550
7	propyl phenyl ketone	0.797	0.95	0.00	0.51	1.2957	0.660	45	hexane	0.000	0.00	0.00	0.00	0.9540	2.370
8	butyl phenyl ketone	0.795	0.95	0.00	0.50	1.4366	0.950	46	heptane	0.000	0.00	0.00	0.00	1.0949	2.590
9	heptyl phenyl ketone	0.720	0.95	0.00	0.50	1.8593	1.720	47	octane	0.000	0.00	0.00	0.00	1.2358	2.820
10	benzonitrile	0.742	1.11	0.00	0.33	0.8711	–0.274	48	2,2,4-trimethylpentane	0.000	0.00	0.00	0.00	1.2358	2.580
11	benzamide	0.955	0.96	0.26	0.41	0.8162	–1.250	49	nonane	0.000	0.00	0.00	0.00	1.3767	2.990
12	nitrobenzene	0.871	1.11	0.00	0.28	0.8906	0.015	50	cyclohexane	0.305	0.10	0.00	0.00	0.8454	2.050
13	benzamide	0.990	1.50	0.49	0.67	0.9728	–3.690	51	methylcyclohexane	0.244	0.06	0.00	0.00	0.9863	2.310
14	acetanilide	0.870	1.40	0.50	0.67	1.1133	–2.740	52	ethylcyclohexane	0.263	0.10	0.00	0.00	1.1272	2.540
15	phenol	0.805	0.89	0.60	0.30	0.7751	–2.460	53	ethene	0.107	0.10	0.00	0.07	0.3474	0.890
16	4-ethylphenol	0.800	0.90	0.55	0.36	1.0569	–1.800	54	propene	0.103	0.08	0.00	0.07	0.4883	0.950
17	4-chlorophenol	0.915	1.08	0.67	0.20	0.8975	–2.750	55	2-methylbut-2-ene	0.159	0.08	0.00	0.07	0.7701	1.890
18	2-nitrophenol	1.015	1.05	0.05	0.37	0.9493	–0.258	56	hex-1-ene	0.078	0.08	0.00	0.07	0.9110	2.020
19	benzyl alcohol	0.803	0.87	0.39	0.56	0.9160	–1.780	57	hept-1-ene	0.092	0.08	0.00	0.07	1.0519	2.230
20	pyridine	0.631	0.84	0.00	0.52	0.6753	–1.070	58	oct-1-ene	0.094	0.08	0.00	0.07	1.1928	2.460
21	propanone	0.179	0.70	0.04	0.49	0.5470	–0.490	59	buta-1,3-diene	0.320	0.23	0.00	0.10	0.5862	0.970
22	butanone	0.166	0.70	0.00	0.51	0.6879	–0.050	60	2-methylbuta-1,3-diene	0.313	0.23	0.00	0.10	0.7271	1.540
23	anthracene	2.290	1.34	0.00	0.28	1.4544	0.960	61	cyclohexene	0.395	0.20	0.00	0.10	0.8024	1.800
24	hexafluorobenzene	0.088	0.66	0.00	0.00	0.8226	1.550	62	ethyne	0.190	0.25	0.21	0.15	0.3044	–0.110
25	3,5-dichlorophenol	1.020	1.00	0.91	0.00	1.0199	–2.070	63	1,4-dioxane	0.329	0.75	0.00	0.64	0.6810	–0.320
26	pentachlorophenol	1.220	0.87	0.96	0.01	1.3871	–0.920	64	butanone	0.166	0.70	0.00	0.51	0.6879	–0.120
27	antipyrine	1.320	1.50	0.00	1.48	1.5502	–2.280	65	ammonia	0.139	0.35	0.14	0.62	0.2084	–1.320
28	clonidine	1.847	1.83	0.35	1.08	1.5317	–1.800	66	trimethylamine	0.140	0.20	0.00	0.67	0.6311	–0.450
29	mepyramine	1.819	1.92	0.00	1.59	2.3870	0.000	67	nitromethane	0.313	0.95	0.06	0.31	0.4237	–1.170
30	imipramine	1.480	1.75	0.00	1.19	2.4020	0.910	68	ethanol	0.246	0.42	0.37	0.48	0.4491	–1.840
31	zolantidine	2.689	2.64	0.40	1.38	2.9946	–1.470	69	benzene	0.610	0.52	0.00	0.14	0.7164	0.910
32	deoxycorticosterone	1.740	3.50	0.14	1.31	2.6802	–1.520	70	toluene	0.601	0.52	0.00	0.14	0.8573	1.190
33	neon	0.000	0.00	0.00	0.00	0.0850	0.780	71	ethylbenzene	0.613	0.51	0.00	0.15	0.9982	1.400
34	argon	0.000	0.00	0.00	0.00	0.1900	1.030	72	*o*-xylene	0.663	0.56	0.00	0.16	0.9982	1.370
35	xenon	0.000	0.00	0.00	0.00	0.3290	1.120	73	*m*-xylene	0.623	0.52	0.00	0.16	0.9982	1.440
36	hydrogen	0.000	0.00	0.00	0.00	0.1086	0.810	74	*p*-xylene	0.613	0.52	0.00	0.16	0.9982	1.460
37	nitrogen	0.000	0.00	0.00	0.00	0.2222	1.160	75	propylbenzene	0.604	0.50	0.00	0.15	1.1391	1.660
38	nitrous oxide	0.068	0.35	0.00	0.10	0.2810	0.610	76	isopropylbenzene	0.602	0.49	0.00	0.16	1.1391	1.600

Results and Discussion

Characterization of the Ethylene Glycol–Heptane System—We use the solvation equation of Abraham,[8] as shown in eq 3:

$$\log SP = c + rR_2 + s\pi_2^H + a\Sigma\alpha_2^H + b\Sigma\beta_2^H + vV_x \quad (3)$$

The solute descriptor R_2 is excess molar refraction, π_2^H is the polarizability/dipolarity, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ are the overall or effective hydrogen bond acidity and basicity, and V_x is the characteristic volume of McGowan[10] in units of (cm^3 mol^{-1})/100. As solutes we used the 20 compounds studied by Paterson et al.,[6] nos. 1–20 in Table 2, and we also measured log P_{eh} for another 12 compounds, nos. 21–32 in Table 2, by the standard shake-flask method, as described above. The 12 compounds were chosen specifically to increase the range of the descriptors; inclusion of the drugs nos. 28–32, and of deoxycortisone, very considerably extends the range. However, there is still room for improvement because over solutes nos. 1–32, the lowest value of π_2^H is 0.52 for benzene, and the lowest value of V_x is 0.5470 for propanone. Since the solubility of ethylene glycol in heptane is very small,[6] we felt that the partition coefficient between these equilibrated phases could be obtained through solubility of gases and vapors in the two pure solvents, eq 4:

$$\log P_{eh} = \log L_h - \log L_e \quad (4)$$

where log L_h and log L_e are the Ostwald solubility coefficients (or gas–liquid partition coefficients) in heptane[11–23] and ethylene glycol[13–16,22–29] at 298 K. Values of log P_{eh} could thus be obtained for 44 extra solutes (nos. 33–76 in Table 2) some with zero values of π_2^H, and with V_x as low as 0.0850 for neon; details are in Table 3. The range of descriptors over the total 72 solutes is now very large indeed; values are collected in Table 2. We note that three of the solutes in set 1–32 are duplicated in set 33–76 (benzene, toluene, and butanone) so that we have 76 data points for 73 solutes. Application of eq 3 to the 76 data points in Table 2 showed that again benzamide was a marked outlier. It is hard to believe that both the Kamlet and the Abraham set of descriptors for benzamide are wrong, and so it is possible that the log P_{eh} value is in error. We did not carry out a determination of log P_{eh} for benzamide ourselves, because the value quoted (–3.69) is lower than we can measure by our shake flask and UV analytical method. Omission of benzamide leads to the correlation eq 5; the sd values for the coefficients are also given.

$$\log P_{eh} = 0.336(0.067) - 0.075(0.134)R_2 - 1.201(0.140)\pi_2^H - 3.786(0.177)\Sigma\alpha_2^H - 2.201(0.163)\Sigma\beta_2^H + 2.085(0.097)V_x \quad (5)$$

$$n = 75,\ r^2 = 0.966,\ sd = 0.28,\ F = 386$$

Table 3—Indirect Calculation of log P_{eh}

no.	solute	log L_e[13−16,22−29]	log L_h[11−23]	log P_{eh}
33	neon	−0.45	−1.23	0.78
34	argon	−1.41	−0.38	1.03
35	xenon	−0.47	0.65	1.12
36	hydrogen	−1.75	−0.94	0.81
37	nitrogen	−1.81	−0.65	1.16
38	nitrous oxide	−0.13	0.48	0.61
39	carbon dioxide	−0.02	0.30	0.32
40	methane	−1.14	−0.06	1.08
41	ethane	−0.63	0.73	1.36
42	propane	0.03	1.30	1.27
43	butane	0.21	1.96	1.75
44	2-methylpropane	0.14	1.69	1.55
45	hexane	0.54	2.91	2.37
46	heptane	0.85	3.44	2.59
47	octane	1.13	3.95	2.82
48	2,2,4-trimethylpentane	0.77	1.81	2.58
49	nonane	1.43	1.56	2.99
50	cyclohexane	1.07	3.12	2.05
51	methylcyclohexane	1.20	3.51	2.31
52	ethylcyclohexane	1.52	4.06	2.54
53	ethene	−0.37	0.52	0.89
54	propene	0.33	1.28	0.95
55	2-methylbut-2-ene	0.56	2.45	1.89
56	hex-1-ene	0.79	2.81	2.02
57	hept-1-ene	1.07	3.30	2.23
58	oct-1-ene	1.33	3.79	2.46
59	buta-1,3-diene	0.99	1.96	0.97
60	2-methylbuta-1,3-diene	0.82	2.36	1.54
61	cyclohexene	1.39	3.19	1.80
62	ethyne	0.50	0.39	−0.11
63	1,4-dioxane	3.27	2.95	−0.32
64	butanone	2.64	2.52	−0.12
65	ammonia	2.25	0.93	−1.32
66	trimethylamine	2.24	1.79	−0.45
67	nitromethane	3.09	1.92	−1.17
68	ethanol	3.48	1.64	−1.84
69	benzene	2.02	2.93	0.91
70	toluene	2.28	3.47	1.19
71	ethylbenzene	2.49	3.89	1.40
72	*o*-xylene	2.67	4.04	1.37
73	*m*-xylene	2.51	3.95	1.44
74	*p*-xylene	2.49	3.95	1.46
75	propylbenzene	2.68	4.34	1.66
76	isopropylbenzene	2.59	4.19	1.60

The rR_2 term is statistically not significant and can be omitted to give:

$$\log P_{eh} = 0.343(0.066) - 1.247(0.112)\pi_2^{H} - 3.807(0.172)\Sigma\alpha_2^{H} - 2.194(0.162)\Sigma\beta_2^{H} + 2.065(0.089)V_x \quad (6)$$

$$n = 75,\ r^2 = 0.966,\ \mathrm{sd} = 0.28,\ F = 488$$

These equations confirm the finding of Paterson et al.[6] that the *a*-coefficient is numerically the largest coefficient. However, as pointed out above, this does not necessarily mean that solute hydrogen bond acidity is the major factor that influences log P_{eh}. In any case, solute hydrogen bond acidity is not the only factor, as shown by the large *s*-, *b*-, and *v*-coefficients. Indeed, it is chemically unreasonable to suggest that solute hydrogen bond basicity, for example, will have little effect on log P_{eh}. The *b*-coefficient is related to the difference in solvent hydrogen bond acidity of ethylene glycol and heptane. Since ethylene glycol is a reasonably strong hydrogen bond acid, and heptane has no acidity, there must be a considerable difference in solvent hydrogen bond acidity, and this must give rise to a substantial *b*-coefficient in the solvation equation.

Table 4—Coefficients in Eq 3 for Various Systems

log SP	*c*	*r*	*s*	*a*	*b*	*v*
log P_{oct}	0.088	0.562	−1.054	0.034	−3.460	3.814
log P_{cyc}	0.127	0.816	−1.731	−3.778	−4.905	4.646
log P_{iso}	0.288	0.382	−1.668	−3.639	−5.000	4.561
log P_{eh}	0.336	−0.075	−1.201	−3.786	−2.201	2.085
-Δlog *P*	0.039	0.254	−0.677	−3.822	−1.445	0.832

Table 5—Values of R_2 and V_x, and Calculated Descriptors for Peptides I–VII from Three Simultaneous Equations

peptide	R_2	V_x	π_2^{H}	$\Sigma\alpha_2^{H}$	$\Sigma\beta_2^{H}$
I	1.453	1.6519	4.01	0.64	0.85
II	2.466	2.7979	5.18	0.73	1.60
III	3.479	3.9439	6.41	0.68	2.33
IV	3.441	4.0848	6.46	0.50	2.36
V	3.403	4.2257	6.91	0.12	2.40
VI	3.365	4.3666	7.68	−0.18	2.20
VII	3.302	4.5075	6.33	−0.30	2.66

We can compare the coefficients in eq 5 with those for a number of partitions,[9,30] as shown in Table 4. For convenience we give coefficients for −Δlog *P*. There is certainly a connection between the coefficients for log P_{eh} and −Δlog *P*, so it is not surprising that log P_{eh} and −Δlog *P* are linearly related for a series of similar solutes, e.g., peptides I−VII, as shown by Burton et al.[3] However, there are considerable differences between the *b*- and *v*-coefficients in the two partition systems. Hence a relationship between log P_{eh} and Δlog *P* for peptides I−VII may not be general, especially as these peptides have particularly large volumes and large basicities (as shown later).

Analysis of the Hydrogen Bonding Capacity of Peptides—To our knowledge, there are no published studies on quantitative determinations of either the hydrogen bond acidity or the hydrogen bond basicity of peptides. It is therefore of some interest to attempt to determine the solvation descriptors, including $\Sigma\alpha_2^{H}$ and $\Sigma\beta_2^{H}$, for peptides I−VII. Once the solvation descriptors are known, it is possible to determine quantitatively the factors that influence the distribution of the peptides.

For the peptides I−VII, both R_2 and V_x can be calculated from structure. The former is obtained by the summation of fragments of known R_2 value,[9] and the latter by McGowan's method of atomic fragments.[10] There remain three descriptors that have to be determined, π_2^{H}, $\Sigma\alpha_2^{H}$, and $\Sigma\beta_2^{H}$. Now if there are available for a given peptide log *P* values in three different partition systems for which the coefficients in eq 3 are known, all three descriptors can be calculated from a set of three simultaneous equations. In Table 1 are log *P* values for the peptides in three partition systems, and in Table 4 are the required coefficients. The calculated descriptors are in Table 5, together with the necessary R_2 and V_x values. There is no point in giving any calculated log *P* values, because the descriptors reproduce the three log *P* values exactly. The obtained descriptors can be compared with descriptors calculated by simple addition of fragment values. We take π_2^{H}, $\Sigma\alpha_2^{H}$, and $\Sigma\beta_2^{H}$ as follows: for a primary amide (1.30, 0.55, and 0.70), for a secondary amide (1.30, 0.40, and 0.70), and for the CH_2Ph group (0.51, 0.00, and 0.15). The calculated descriptors for the mono-, di-, and tripeptides I, II, and III are in Table 6. Values of π_2^{H} from the fragment addition and the simultaneous equations agree quite well, compare Tables 5 and 6, but $\Sigma\alpha_2^{H}$ and $\Sigma\beta_2^{H}$ from the simultaneous equations are much lower than those from fragment addition.

Table 6—Calculated Descriptors from Fragment Values

peptide	π_2^H	$\Sigma\alpha_2^H$	$\Sigma\beta_2^H$
I	3.11	0.95	1.95
II	4.92	1.35	2.82
III	6.73	1.75	3.69

Table 7–Revised Descriptors for Peptides I–VII

				calculated log P[a]		
peptide	π_2^H	$\Sigma\alpha_2^H$	$\Sigma\beta_2^H$	oct	iso	eh
I	3.90	0.65	0.89	0.04	–4.94	–5.43
II	5.20	0.67	1.63	1.05	–5.27	–6.38
III	6.60	0.64	2.27	2.30	–5.08	–7.05
IV	6.45	0.50	2.37	2.62	–4.20	–6.26
V	6.60	0.16	2.48	2.59	–3.13	–5.10
VI	6.50	0.00	2.50	3.13	–1.85	–4.12
VII	6.10	0.00	2.55	3.88	–0.82	–3.45

[a] Oct is water–octanol, iso is water–isooctane, eh is ethylene glycol–heptane.

This might be due to peptide conformations in which intramolecular hydrogen bonding takes place, or in which there is restricted access of solvent molecules to hydrogen bond sites. In either case, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ will be lower than expected.

Even taking these effects into account, some of the obtained descriptors by the simultaneous equation method are not very reasonable. The negative $\Sigma\alpha_2^H$ values for VI and VII have no physical meaning, and π_2^H for VI seems too large. There are a number of possible reasons. First, the determination of very negative log P values is difficult, and the observed log P values may have a larger error than usual. Second, there are numerous possible combinations of the three descriptors that will all give calculated log P values in reasonable accord with the observed values. Third, although we have considerably extended the range of descriptors by incorporation of compounds nos. 21−32 (Table 2), some of the peptides still lie outside this range. Hence small variations in the coefficients of the log P_{eh} equation could lead to large differences in calculated descriptors. We therefore take the descriptors in Table 5 as a first estimate only and then calculate a revised set that (i) will reproduce quite well the observed log P values, and (ii) will be chemically more reasonable. In particular we have assigned $\Sigma\alpha_2^H$ as zero for VI and VII, instead of the negative quantities in Table 5. From the fragment values in Table 6, the difference in π_2^H between I and II and between II and III would be around 1.80 units. However, we find that such an adjustment leads to poorer fits of calculated and experimental log P values, and as a compromise we take the difference as 1.5 units. Also, we expect π_2^H for the tripeptides III−VII to be nearly the same. The revised descriptors are in Table 7, together with the calculated log P values using the revised descriptors. Results are very good for peptides I−VI, with observed and calculated log P values being in good agreement, but for peptide VII, the agreement is not so good. The general trend of the descriptors for I−VII, Table 7, now seems more reasonable than those from the simultaneous equations, Table 5, especially regarding π_2^H and $\Sigma\beta_2^H$.

We can test whether the assigned descriptors for I−VII are compatible with those in Table 2 by incorporating the peptides into the log P_{eh} regression. The resulting equations are very close to eq 5 and eq 6. If only peptides I−VI are used, the regression equation coefficients are almost exactly the same as those in eq 5, or in eq 6:

$$\log P_{eh} = 0.336(0.061) - 0.063(0.109)R_2 - 1.218(0.070)\pi_2^H - 3.788(0.157)\Sigma\alpha_2^H - 2.192(0.144)\,\Sigma\beta_2^H + 2.087(0.092)\,V_x \quad (7)$$

$$n = 81,\ r^2 = 0.986,\ \text{sd} = 0.27,\ F = 1043$$

$$\log P_{eh} = 0.339(0.061) - 1.231(0.066)\pi_2^H - 3.816(0.149)\Sigma\alpha_2^H - 2.206(0.141)\Sigma\beta_2^H + 2.063(0.082)V_x \quad (8)$$

$$n = 81,\ r^2 = 0.986,\ \text{sd} = 0.27,\ F = 1315$$

The range of descriptors in eq 7 and eq 8 is now so large that it is possible to estimate new values of log P_{eh} for a very large number of compounds; we suggest eq 7 or eq 8 rather than eq 5 or eq 6 be used for this purpose.

Paterson et al.[6] also gave values of log P_{eh} (Table 1) and log P_{oct} (the latter graphically) for peptides VIII−XI. We can use the same procedure as that used for peptides I−VII to assign descriptors. Thus for VIII, π_2^H is expected to be around 6.5−6.6 and $\Sigma\beta_2^H$ to be near to 2.4, by analogy with III, VI, and VII. For IX there is a direct analogy with IV ($\pi_2^H = 6.45$ and $\Sigma\beta_2^H = 2.37$), and for X there is an analogy with V ($\pi_2^H = 6.6$ and $\Sigma\beta_2^H = 2.48$). Our assigned descriptors are in Table 8, together with log P values calculated from the descriptors; there is quite good agreement with the expected values. For XI, it is more difficult to estimate descriptors, but by comparison of *N*-alkylamides and alkylcarbamates we expect that π_2^H for XI should be somewhat less than that for I ($\pi_2^H = 3.9$). If we take π_2^H as 3.5, then the descriptors shown in Table 8 for XI are obtained.

Table 8—Calculated Descriptors for Peptides VIII–XIV

						calculated log P		
peptide	R_2	V_x	π_2^H	$\Sigma\alpha_2^H$	$\Sigma\beta_2^H$	oct	iso	eh
VIII	3.441	4.0848	6.55	0.41	2.35	2.58	–3.93	–6.00
IX	3.441	4.0848	6.40	0.40	2.34	2.77	–3.60	–5.76
X	3.403	4.2257	6.45	0.03	2.45	2.84	–2.26	–4.36
XI	1.263	2.1333	3.50	0.45	1.01	1.77	–2.02	–3.44
XII	1.060	1.2546	1.65	(0.40)	0.82	0.91	–1.89	–2.43
XIII	1.000	1.2546	1.54	(0.30)	0.83	0.95	–1.42	–1.94
XIV	0.960	1.3955	1.60	0.00	0.89	1.18	–0.10	–0.71

Conradi et al.[31] also investigated peptides XII−XIV in terms of hydrogen bond numbers, H#, but gave no partition coefficients. We can use our usual method of addition of fragments, together with known values of log P_{oct} for XII (0.91),[32,33] and XIII (0.95),[33] to estimate the descriptors shown in Table 8. There is not enough data to obtain reliable $\Sigma\alpha_2^H$ values for XII and XIII, and so the calculated values for log P_{iso} and log P_{eh} are provisional only.

Inspection of Table 7 reveals surprising trends especially with the descriptor $\Sigma\alpha_2^H$. For the mono-, di-, and tripeptides I, II, and III, π_2^H and $\Sigma\beta_2^H$ increase regularly with the number of amide groups, but $\Sigma\alpha_2^H$ remains constant even though the number of acidic (CO)N−H bonds is increased. This is not at all due to our assignment of descriptors, because the same trends are shown in Table 5. *N*-Methylation of III decreases the hydrogen bond acidity, as expected, but VI and VII would still be predicted to have some hydrogen bond acidity; yet from the results in Table 5 we have had to assign zero values of $\Sigma\alpha_2^H$. Our calculated descriptors for VIII−X show the same trend: $\Sigma\alpha_2^H$ decreases with increasing *N*-methylation, but the actual values are smaller than expected. Borchardt et al.[34,35] have shown how the secondary β-turn structure of peptides can influence their lipophilicity, as log P_{oct}, but

Table 9—An Analysis of the Factors That Influence Log P_{eh} for Peptides, Based on Eq 5

peptide	rR_2	$s\pi_2^H$	$a\Sigma\alpha_2^H$	$b\Sigma\beta_2^H$	vV_x	log P_{eh} calcd[a]	log P_{eh} obsd
I	–0.11	–4.68	–2.46	–1.96	3.44	–5.43	–5.46
III	–0.26	–7.93	–2.42	–5.00	8.22	–7.05	–7.10
VI	–0.26	–7.80	0.00	–5.50	9.10	–4.12	–4.20
XI	–0.07	–4.29	–1.66	–2.21	4.44	–3.45	–3.44

[a] With c = 0.336 in eq 5, and the descriptors in Table 7.

Table 10—Values of log k for Permeability of Peptides

peptide	log k_{mono}[a]	log k_{mono}[b]	log k_{mono}[c]	log k_{situ}[d]
I	–5.04	–5.05	–5.43	–5.74
II	–5.72	–5.65	–5.67	–6.77
III	–6.23	–6.71	–5.89	–7.06
IV	–5.47		–5.66	–6.68
V	–5.20		–5.51	–6.16
VI	–4.69		–5.12	–5.96
VII	–4.47		–4.89	–4.94
VIII	–6.10			
IX	–5.39			
X	–4.82			
XI	–4.05			
XII		–3.37		
XIII		–3.68		
XIV		–3.39		

[a] k cm/s for permeability across Caco-2 cell monolayers, ref 6. [b] k cm/s for permeability across Caco-2 cell monolayers, ref 14. [c] k cm/s for permeability across BBMEC monolayers, ref 5. [d] k cm/s for in situ rat perfusion from saline, ref 5.

the most direct evidence of the effect of structure on hydrogen-bonding potential is provided by Price et al.[36] These workers examined the peptide XV and calculated the electrostatic maxima and minima around the van der Waals surface, with the peptide in both an extended conformation and a helical conformation. In the extended conformation, only one substantial electrostatic maximum was found, and in the helical conformation only two maxima were found, yet the peptide possesses three acidic (CO)N−H bonds. The lack of maxima in the extended conformation was attributed to N−H bonds being parallel and coplanar with neighboring C═O groups, resulting in cancelation of potentials of opposite magnitude. In the helical conformation, the N−H in the terminal CONHMe group is in close proximity to the oxygen in the MeCONH group, again producing a cancelation of potentials.[36] Our findings can possibly be explained by similar effects that result in the loss of electrostatic maxima.

Interestingly, in the extended conformation of XV there are three substantial electrostatic minima, corresponding to the three C═O groups, and in the helical conformation there are two very large electrostatic minima corresponding to C═O(1)/C═O(2) and C═O(2)/C═O(3).[36] Hence cancellation of electrostatic potentials can result in almost complete loss of electrostatic maxima while still retaining substantial electrostatic minima. Our calculated descriptors seem to follow this trend; VI and VII have no hydrogen bond acidity, yet $\Sigma\beta_2^H$ gradually increases along the series III − VII, with increasing *N*-methylation.

Once descriptors for the peptides are available, it is possible to assess quantitatively the factors that influence values of log P_{eh} for peptides. In Table 9 is a term-by-term breakdown of eq 5 for a representative selection of peptides I−XIV. The most interesting finding is that solute hydrogen bond acidity is NOT the main factor that influences the value of log P_{eh}. Even for peptides I and III, with large values of $\Sigma\alpha_2^H$, the $a\Sigma\alpha_2^H$ term is numerically smaller than the dipolarity/polarizability and volume terms. And with the sole exception of peptide I, all the peptides I−XIV have a smaller $a\Sigma\alpha_2^H$ term than the hydrogen bond basicity term $b\Sigma\beta_2^H$. This illustrates that examination of coefficients is not enough to assess the contribution of the different terms in eq 5, or in eq 3, generally. Only a term-by-term analysis, such as that in Table 9, allows the contribution of each term to be found.

Paterson et al.[6] and Conradi et al.[31,37] determined permeability of Caco-2 cell monolayers for the fourteen peptides I−XIV, Table 10. For 10 of these peptides, there was a good correlation (r^2 = 0.943) of log k_{mono} against the hydrogen bond number H#.[34] We have repeated the correlation for all fourteen peptides and find r^2 = 0.857, still quite reasonable. However, such a correlation does not distinguish between effects of hydrogen bond acidity and hydrogen bond basicity. To establish these effects, a correlation equation and descriptors for the solutes concerned are required. Unfortunately, the number and variety of solutes is too small to yield a definitive correlation equation through the solvation equation, eq 3.

References and Notes

1. Karls, M. S.; Rush, B. D.; Wilkinson, K. F.; Vidmar, T. J.; Burton, P. S.; Ruwart, M. J. Desolvation Energy; a Major Determinant of Absorption but not Clearage of Peptides in Rats. *Pharm. Res.* **1991**, *8*, 1477−1481.
2. Stein, W. D. *The Movement of Molecules Across Cell Membranes*; Academic Press: New York, 1967; pp 65−91.
3. Burton, P. S.; Conradi, R. A.; Hilgers, A. R.; Ho, N. F. H.; Maggiora, L. L. The Relationship between Peptide Structure and Transport Across Epithelial Membranes. *J. Controlled Release* **1992**, *19*, 87−92.
4. Seiler, P. Interconversion of Lipophilicities from Hydrocarbon/water Systems into the Octanol/Water System. *Eur. J. Med. Chem.* **1974**, *9*, 473−479.
5. Chikale, E. G.; Ng, K.-Y.; Burton, P. S.; Borchardt, R. T. Hydrogen Bonding Potential as a Determinant of the in vitro and in situ Blood-brain Barrier Permeability of Peptides. *Pharm. Res.* **1994**, *11*, 412−419.
6. Paterson, D. A.; Conradi, R. A.; Hilgers, A. R.; Vidmar, T. J.; Burton, P. S. A Nonaqueous Partition System for Predicting the Oral Absorption Potential of Peptides. *Quant. Struct.-Act. Relat.* **1994**, *13*, 4−10.
7. Kamlet, M. J.; Doherty, R. M.; Abboud, J.-L. M.; Abraham, M. H.; Taft, R. W. Solubility − A New Look. *CHEMTECH* **1986**, 566−576.
8. Abraham, M. H. Scales of Solute Hydrogen-bonding: Their Construction and Application to Physicochemical and Biochemical Processes. *Chem. Soc. Rev.* **1993**, *22*, 73−83.
9. Abraham, M. H.; Chadha, H. S. *Lipophilicity in drug action and toxicology*; Pliska, V., Testa, B., van de Waterbeemd, H., Eds.; VCH: Weinheim, 1996; pp 311−337.
10. Abraham, M. H.; McGowan, J. C. The Use of Characteristic Volumes to Measure Cavity Terms in Reversed Phase Liquid Chromatography. *Chromatographia* **1987**, *664*, 243−246.
11. Clever, H. L.; Battino,; Saylor, J. H.; Gross, P. M. The Solubility of Helium, Neon, Argon and Krypton in some Hydrocarbon Solvents. *J. Phys. Chem.* **1957**, *61*, 1078−1082.
12. Brunner, E. Solubility of Hydrogen in 10 Organic Solvents at 298.15, 323.15 and 373.15 K. *J. Chem. Eng. Data* **1985**, *30*, 269−273.
13. Battino, R.; Rettich, T. R.; Tominaga, T. The Solubility of Nitrogen and Air in Liquids. *J. Phys. Chem. Ref. Data* **1984**, *13*, 563−600.
14. *Oxides of Nitrogen. Solubility Data series Vol. 8*; Young, C. L., Ed.; Pergamon Press: Oxford, 1981.
15. *Ammonia, Amines, Phosphines, Arsine, Stibine, Silane, Germane and Stannane in Organic Solvents. Solubility Data Series Vol. 21*; Young, C. L., Fogg, P. G. T., Eds.; Pergamon Press: Oxford, 1985.
16. Wilhelm, E.; Battino, R. Thermodynamic Functions of the Solubilities of Gases in Liquids at 25 °C. *Chem. Rev.* **1973**, *73*, 1−9.
17. Milanova, E; Cave, G. C. B. Limiting Activity Coefficients in Dilute Solutions of Nonelectrolytes. I. Determination for Polar-Nonpolar Binary Mixtures by a Novel Apparatus, and a Solubility Parameter Treatment of Results. *Can. J. Chem.* **1982**, *60*, 2697−2706.

18. Thomas, E. R.; Newman. B. A.; Long, T.C.; Wood, D.A.; Eckert, C.A. Limiting Activity Coefficients of Nonpolar and Polar Solutes in both Volatile and Nonvolatile Solvents by Gas Chromatography. *J. Chem. Eng. Data* **1982**, *27*, 399−405.
19. Rytting, J. H.; Huston, L. P.; Higuchi, T.. Thermodynamic Group Contributions for Hydroxyl, Amino and Methylene Groups. *J. Pharm. Sci.* **1978**, *67*, 615−618.
20. Jadot, R. Determination de Constants de Henry par Chromatographie. *J. Chim. Phys.* **1972,** 1036−1040.
21. Hayduk, W.; Minhas, B. S. Diffusivity and Solubility of 1,3-Butadiene in Heptane and Other Properties of Heptane-Styrene and Aqueous Potassium Laurate Solutions. *J. Chem. Eng. Data* **1987**, *32*, 285−290.
22. Bruckl, N.; Kim, J. I. Gibbs Free Energies of Solute Solvent Interactions for He, Ne, Ar, Kr, Xe, H_2, CH_4, SF_6, C_2H_4, CO_2, and C_2H_2 in Various Solvents: Comparison of Theoretical Prediction with Experiment. *Z. Phys. Chem. Neue Folge* **1981**, *126*, 133−150.
23. Park, J. H.; Hussam, A.; Couasnon, P.; Fritz, D.; Carr, P. W. Experimental Reexamination of Selected Partition Coefficients from Rohrschneider's Data Set. *Anal. Chem.*, **1987**, *59*, 1970−1976.
24. Fernandez-Prini, R.; Crovetto, R.; Gentili, N. Solubilities of Inert Gases in Ethylene Glycol. *J. Chem. Thermodyn.* **1987**, *19*, 1293−1298.
25. Brunner, E. Solubility of Hydrogen in Alcohols. *Ber. Bunsen-Ges. Phys. Chem.* **1979**, *83*, 715−721.
26. Hayduk, W.; Laudie, H. Solubilities of Gases in Water and Other Associated Solvents. *AIChE J.* **1973**, *19*, 1233−1238.
27. Vernier, P.; Raimbault, C.; Renon, H. Proprietes Thermodynamiques des Solutions Infiniment Diluees D′Hydrocarbures dans les Solvants Polaires. *J. Chim. Phys.* **1969**, *66*, 429−436.
28. Lenoir, J.-Y.; Renault, P.; Renon, H. Gas Chromatographic Determination of Henry's Constants of 12 Gases in 19 Solvents. *J. Chem. Eng. Data* **1971**, *16*, 340−342.
29. Arancibia, E. L.; Catoggio, J. A. Gas Chromatographic Study of Solution and Adsorption of Hydrocarbons on Glycols. *J. Chromatogr.* **1982**, *238*, 281−290.
30. Abraham, M. H.; Chadha, H. S.; Whiting, G. S.; Mitchell, R. C. Hydrogen Bonding. 32. An Analysis of Water-Octanol and Water-Alkane Partitioning and the Δlog *P* Parameter of Seiler. *J. Pharm. Sci.* **1994,** *83*, 1085−1100.
31. Conradi, R.A.; Hilgers, A.R.; Ho, N.F. H.; Burton, P.S.. The Influence of Peptide Structure on Transport Across Caco-2 Cells. *Pharm. Res.* **1991**, *8*, 1453−1460.
32. Leo, A. *The Medicinal Chemistry Project*; Pomona College: Claremont, CA 91711.
33. Sangster, J. A *Databank of Evaluated Octanol−Water Partition Coefficients*; Sangster Research Laboratories: Montreal, Quebec, Canada H3G 2A4, 1994.
34. Knipp, G. T.; Vander Velde, D. G.; Siahaan, T. J.; Borchardt, R. T. The Effect of β-Turn Structure on the Passive Diffusion of Peptides Across Caco-2 Cell Monolayers. *Pharm. Res.* **1997**, *14*, 1332−1340.
35. Sorensen, M.; Steenberg, B.; Knipp, G. T.; Wang, W.; Steffansen, B.; Frokjaer, S.; Borchardt, R.T. The Effect of β-Turn Structure on the Permeation of Peptides Across Monolayers of Bovine Brain Microvessel Endothelial cells. *Pharm. Res.* **1997**, *14*, 1341−1348.
36. Apaya, R. P.; Bondi, M.; Price, S. L. The Orientation of N−H···O═C and N−H···N Hydrogen Bonds in Biological Systems: How Good is a Point Charge as a Model for a Hydrogen Bonding Atom *J. Comput.-Aided Mol. Des.* **1997**, *11*, 479−490.
37. Conradi, R. A.; Hilgers, A. R.; Ho, N. F. H.; Burton, P. S. The Influence of Peptide Structure on Transport Across Caco-2 Cells. II. Peptide Bond Modification Which Results in Improved Permeability. *Pharm. Res.* **1992**, *9*, 435−439.

Acknowledgments

F. Martins thanks the North Atlantic Treaty Organization for a postdoctoral fellowship, the Department of Chemistry, University of Lisbon, for leave of absence, and Junta Nacional Investigaceo Cientisica Technologica for financial support under project PBIC/P/QUI/2199/95. We are grateful to Dr. Sarah L. Price for her helpful comments on electrostatic maxima and minima of peptides.

JS980242L