J. Chem. Soc., Faraday Trans. 1, 1982, **78**, 1641-1665

Aqueous Solutions Containing Amino Acids and Peptides

Part 13.—Enthalpy of Dilution and Osmotic Coefficients of some *N*-Acetyl Amino Acid Amides and some *N*-Acetyl Peptide Amides at 298.15 K

BY G. MICHAEL BLACKBURN, TERENCE H. LILLEY* AND ELIZABETH WALMSLEY

Chemistry Department, The University, Sheffield S3 7HF

Received 17*th August*, 1981

The energetics of the interactions occurring between some *N*-acetyl amino acid amides and some *N*-acetyl peptide amides in aqueous solutions at 298.15 K have been investigated. Osmotic coefficients of solutions containing *N*-acetylglycinamide (G), *N*-acetyl-L-alaninamide (A) and *N*-acetyl-L-leucinamide (L) and equimolal solutions of G+A, G+L and A+L have been obtained using the isopiestic vapour pressure technique. Enthalpies of dilution of *N*-acetylglycylglycinamide (G_2), *N*-acetyl-L-alanyl-L-alaninamide (A_2), *N*-acetylglycylglycylglycinamide (G_3), *N*-acetyl-L-alanyl-L-alanyl-L-alaninamide (A_3) and *N*-acetyl-L-alanylglycinamide (AG) and equimolal solutions of $G+G_2$, $G+G_3$ and $A+A_2$ were obtained using microcalorimetry. The results obtained were used to calculate the pairwise free energy and enthalpy parameters for like–like and like–unlike solute interactions. The effects of molecular structure and substitution on these parameters are considered and the efficiency of a group interaction approach is investigated. The group interaction idea works well for both the rather limited free energy data set considered and the more extensive enthalpy data set, with the exception of the most hydrophobic molecule A_3. It is suggested that the results for A_3 indicate a degree of intramolecular folding which perturbs the intermolecular interactions.

There is at present considerable general interest in the so-called 'non-bonding' interactions which occur between atoms and molecules. Much of this interest stems from an appreciation of the important role which such interactions play in biological systems.[1-4] The work here is a continuation of earlier work[5] on the non-bonding interactions between some substituted peptides and amino acids in aqueous solutions and has been partly described in a preliminary communication.[6] The systems investigated embody certain features which must contribute, in some measure at least, to the behaviour of oligopeptides and proteins in aqueous environments, and our hope is that such studies will give insight into the factors which affect the conformational stability of proteins and enzyme–substrate interactions.

The available evidence indicates that globular protein and enzymes are stabilised in a relatively narrow distribution of conformations in aqueous systems[7-9] by various intramolecular solute–solute interactions and by a range of solute–solvent interactions, and it would seem that these manifold interactions are individually energetically weak but collectively of great importance with regard to conformational stability. The nature and arrangement of the amino acid side-chain along the protein backbone are responsible for the individual characteristics of the macromolecule and it has been recognised for some time that all of the information pertaining to the protein is implicit in the amino acid sequence.[10-15]

Because of the difficulty of studying protein systems *per se* it is of particular benefit to investigate the properties of systems containing small molecules[16] which incorporate

functional groups present in proteins. Indeed, most of our, albeit sketchy, knowledge of the magnitudes of non-bonding interactions has been obtained from studies on well-defined small molecules.[17] The objective of the present work is to study the energetics of the interactions occurring in solutions of amino acid and peptide derivatives which combine both biological relevance and structural simplicity. For the moment we have directed our attention to a study directed at substituted amino acids and peptides containing only glycine, L-alanine, L-valine and L-leucine. This particular set was chosen to demonstrate the effect of increasing side-chain hydrophobicity and is representative of the amino acids most commonly found in hydrophobic areas within proteins.

EXPERIMENTAL

The experimental procedures used for obtaining the heats of dilution[5] and the osmotic coefficients[18] have been described previously.

PREPARATION AND PURIFICATION OF MATERIALS

The preparation of the *N*-acetyl amino acid amides of glycine, L-alanine, L-leucine and L-valine has been described previously.[5]

N-ACETYL-L-ALANYL-L-ALANINAMIDE

The ethyl ester was obtained as follows. *N*-Benzyloxycarbonyl-L-alanyl-L-alanine ethyl ester (3.2 g, 10 mmol dm^{-3}) and *p*-toluenesulphonic acid monohydrate (1.9 g, 10 mmol dm^{-3}) were dissolved in dimethyl formamide (DMF) (20 cm^3). After purging with nitrogen, 10% Pd/C catalyst (0.5 g) was added and hydrogen bubbled through at atmospheric pressure overnight. The catalyst was filtered off and the solvent was lyophyllised to give an oil which was dissolved in dry pyridine at 0 °C and acetic anhydride (20 mmol dm^{-3}, 5 cm^3) added. After one hour stirring at 0 °C, excess solvent was lyophillised and the residue dissolved in ice/water. Excess Amberlite MB-3 ion-exchange resin was added and the solution stirred for 15 min at 0 °C. It was filtered, the solvent evaporated and the residue crystallised from ethyl acetate + petrol to give the product, m.p. 126-127 °C (74%), $[\alpha]_D^{22} - 54.7°$ (c 0.5, EtOH). (Found: C, 52.24; H, 7.76; N, 12.14; $C_{10}H_{18}N_2O_4$ requires C, 52.16; H, 7.88; N, 12.16%.) The ester was dissolved in anhydrous ethanol saturated with ammonia and set aside for one day at 20 °C. Solvent was removed *in vacuo* and the product repeatedly crystallised from ethanol + ether to constant melting point, 250-251 °C, $[\alpha]_D^{20} - 42°$ (c 1, MeOH). (Found: C, 47.81; H, 7.57; N, 20.82; $C_8H_{15}N_3O_3$ requires C, 47.75; H, 7.51; N, 20.88%.) $\delta(D_2O)$ 3.69-3.55 (2H, m, 2 αCH), 1.35 (3H, s, CH_3CO), 0.73 (3H, d, *J* 5 Hz, CH_3).

N-ACETYLGLYCYLGLYCINAMIDE

N-Benzyloxycarbonylglycylglycine ethyl ester (2.9 g, 10 mmol dm^{-3}) and *p*-toluenesulphonic acid (1.9 g, 10 mmol dm^{-3}) were dissolved in dry DMF (25 cm^3) and the solution purged with nitrogen and hydrogenated in the presence of 10% Pd/C catalyst (0.5 g) to yield the dipeptide ester as its *p*-toluenesulphonic acid salt. Solvent was removed *in vacuo* and the oily solid triturated with dry ether. Without further purification the residue was dissolved at 0 °C in pyridine (60 cm^3) and an excess (1.5 equiv.) of acetic anhydride added. The acetylation reaction was followed closely by t.l.c. and when judged to be complete the solvent was evaporated and the residue dissolved in ice/water. Excess Amberlite MB-3 ion-exchange resin was added. The solution was stirred for 15 min at 0 °C, then filtered, the solvent was evaporated and the residue crystallised from ethanol + ether, m.p. 148-149 °C, (68%), [lit.[19] m.p. 146-148 °C], $R_f = 0.39$ in 9:1 $CHCl_3$:i-PrOH. (Found: C, 47.7; H, 6.98; N, 13.9; $C_8H_{14}O_4N_2$ requires C, 47.5; H, 6.98; N, 13.86%.) δ ($CDCl_3$) 6.95 br (1H, m, NH), 6.63 br (1H, m, NH), 4.21 (2H, q, *J* 7 Hz, OCH_2), 4.02 (2H, d, *J* 5 Hz, αCH_2), 3.98 (2H, d, *J* 6 Hz, αCH_2), 2.03 (3H, s, CH_3CO), 1.27 (3H, t, *J* 7 Hz, CH_3).

This ethyl ester was dissolved in anhydrous ethanol saturated with ammonia and the solution allowed to stand in a pressure vessel at 20 °C for one day. Solvent was removed *in vacuo* and

the product repeatedly crystallised from ethanol+ether, m.p. 206-208 °C (82%). (Found: C, 41.8; H, 6.4; N, 24.3; $C_6H_{11}N_3O_3$ requires C, 41.6; H, 6.38; N, 24.26%.)

N-ACETYLGLYCYLGLYCYLGLYCINAMIDE

N-Acetylglycylglycylglycine ethyl ester[20] was dissolved in a large excess of anhydrous ethanol saturated with ammonia and the solution was allowed to stand at 20 °C for two days. Solvent was removed under reduced pressure and the product repeatedly crystallised from water+methanol, to constant melting point 253-255 °C. (Found: C, 41.66; H, 6.09; N, 24.40; $C_8H_{14}N_4O_4$ requires C, 41.74; H, 6.13; N, 24.34%.) $\delta(D_2O)$ 3.22 (2H, s, αCH_2), 3.30 (2H, s, αCH_2), 3.26 (2H, s, αCH_2), 1.38 (3H, s, CH_3).

N-ACETYL-L-ALANYL-L-ALANYL-L-ALANINAMIDE

N-Benzyloxycarbonyl-L-alanyl-L-alanyl-L-alanine ethyl ester[21] (3.93 g, 10 mmol dm^{-3}) was hydrogenated in dry DMF (20 cm^3) in the presence of *p*-toluenesulphonic acid (1.9 g, 10 mmol dm^{-3}) and Pd/C catalyst (0.6 g, 5% w/w) using the procedure described above. The tripeptide salt was then dissolved in an ice-cold mixture of pyridine (10 cm^{-3}) and excess acetic anhydride (2 cm^3). After stirring 1 h at 0 °C, solvent was evaporated and the residue dissolved in ice/water. Excess ion-exchange resin was added and the solution stirred at 0 °C for 15 min. After filtration, the residue was crystallised twice from ethyl acetate+ether to give the pure product, m.p. 246-246.5 °C (71%), $[\alpha]_D^{22} - 73.2°$ (c 0.5, EtOH), $R_f = 0.33$ in 9:1 $CHCl_3$: i-PrOH. (Found: C, 51.86; H, 7.58; N, 13.90; $C_{13}H_{23}N_3O_5$ requires C, 51.82; H, 7.69; N, 13.94%.)

This ester was dissolved in anhydrous ethanol saturated with ammonia and the solution allowed to stand at 20 °C for 36 h. Solvent was removed under reduced pressure and the product repeatedly crystallised from methanol ether to give the pure product, constant melting point 295-300 °C (64%), $[\alpha]_D^{22} - 61.5°$, (c 0.75, MeOH). (Found: C, 48.54; H, 7.5; N, 20.65; $C_{11}H_{20}N_4O_4$ requires C, 48.52; H, 7.40; N, 20.57%.) $\delta(D_2O)$ 3.77-3.60 (3H, m, 3 αCH), 1.42 (3H, s, CH_3CO), 0.83 (3H, d, *J* 6 Hz, CH_3), 0.79 (3H, d, *J* 7 Hz, CH_3), 0.76 (3H, d, *J* 7 Hz, CH_3).

N-ACETYL-L-ALANYLGLYCINAMIDE

N-Benzyloxycarbonyl-L-alanylglycine ethyl ester[22] was dissolved in dry DMF containing one equivalent of *p*-toluenesulphonic acid and hydrogenated overnight in the presence of 10% Pd/C catalyst. The solution was filtered and solvent evaporated *in vacuo* to give crude the dipeptide salt as an oil which without further purification was dissolved in ice-cold pyridine and acetic anhydride and stirred for 30 min at 0 °C. The solvent was evaporated and the residue dissolved in ice/water. Excess mixed-bed ion-exchange resin was added and the solution stirred at 0 °C for 15 min. After filtration the solvent was evaporated and the residue crystallised several times from ethyl acetate+petrol (60:80) to yield the product, m.p. 137-138 °C (63%). (Found: C, 50.36; H, 7.39; N, 12.99; $C_9H_{16}N_2O_4$ requires C, 49.99; H, 7.46; N, 12.96%.) $\delta(CDCl_3)$ 7.29 (1H, t, *J* 5 Hz, $NHCH_2$), 6.77 (1H, d, *J* 7 Hz, NH—CH), 4.65-4.59 (1H, m, αCH), 4.19 (2H, q, *J* 7 Hz, OCH_2), 3.99 (2H, d, *J* 5 Hz, αCH_2), 2.00 (3H, s, CH_3CO), 1.38 (3H, d, *J* 7 Hz, side chain CH_3), 1.26 (3H, t, *J* 7 Hz, CH_3). *N*-Acetyl-L-alanylglycine ethyl ester was dissolved in anhydrous ethanol saturated with ammonia and allowed to stand at 20 °C for 1 day. Solvent was removed *in vacuo* and the product repeatedly crystallised from ethanol+ether to constant melting point 148-149 °C. (Found: C, 44.83; H, 7.18; N, 22.43; $C_7H_{13}N_3O_3$ requires C, 44.90; H, 7.00; N, 22.45%.) $\delta(D_2O)$ 3.81 (1 H, q, *J* 7 Hz, CH), 3.42 (2H, s, CH_2), 1.54 (3H, s, CH_3CO), 0.88 (3H, d, *J* 7 Hz, CH_3).

RESULTS

The excess Gibbs free energy per kilogram of solvent may be represented[18] as a power series in solute molalities

$$G^{ex} = \sum_i \sum_j g_{ij} m_i m_j + \sum_i \sum_j \sum_k g_{ijk} m_i m_j m_k + \ldots \quad (1)$$

where the coefficients g_{ij}, g_{ijk} *etc.* are taken to represent interactions between the subscripted species. Using the relationships[23]

$$G_{\mathrm{w}}^{\mathrm{ex}} = [\partial(G^{\mathrm{ex}}/m)/\partial m^{-1}]_{y_i} \tag{2}$$

and

$$G_{\mathrm{w}}^{\mathrm{ex}} = \boldsymbol{RT}\,\mathbf{m}(1-\phi) \tag{3}$$

we obtain from eqn (1) the following expression for the osmotic coefficient

$$\phi = 1+(\mathbf{m}/\boldsymbol{RT})\,(\sum_i\sum_j g_{ij}\,y_i\,y_j + 2\mathbf{m}\sum_i\sum_j\sum_k g_{ijk}\,y_i\,y_j\,y_k + \ldots). \tag{4}$$

If we consider the application of eqn (4) to single-solute systems containing the solutes A and B, respectively, we obtain:

$$\phi_{\mathrm{A}} = 1+(g_{\mathrm{AA}}m_{\mathrm{A}} + g_{\mathrm{AAA}}m_{\mathrm{A}}^2 + \ldots)/\boldsymbol{RT} \tag{5}$$

$$\phi_{\mathrm{B}} = 1+(g_{\mathrm{BB}}m_{\mathrm{B}} + g_{\mathrm{BBB}}m_{\mathrm{B}}^2 + \ldots)/\boldsymbol{RT}. \tag{6}$$

TABLE 1.—ISOPIESTIC MOLALITIES FOR THE SYSTEMS INVESTIGATED

m(urea) /mol kg^{-1}	ϕ(urea)	**m**(solute)/mol kg^{-1}					
		G[a]	A	L	G+A[b]	G+L[b]	A+L[b]
0.5400	0.9781	0.5416	0.5441	0.6176	0.5461	0.5563	0.5631
0.5522	0.9777	0.5512	0.5556	0.6348	0.5800	0.5751	0.5812
0.5769	0.9768	0.5746	0.5798	0.6681	0.5786	0.6047	0.6113
0.6466	0.9742	0.6422	0.6479	0.7610	0.6480	0.6809	0.6882
0.6987	0.9723	0.6957	0.7023	0.8356	0.7004	0.7388	0.7467
1.0986	0.9588	1.0950	1.1045	1.4562	1.0959	1.1898	1.2041
1.1436	0.8577	1.1425	1.1513	1.5551	1.1418	1.2266	1.2609
1.1863	0.9559	1.1782	1.1875	1.6691	1.1788	1.2884	1.3101
1.2271	0.9547	1.2214	1.2803	1.7652	1.2228	1.3402	1.3635
1.2892	0.9528	1.2781	1.2917	1.8936	1.2870	1.4109	1.4320

[a] Abbreviations used: G = *N*-acetylglycinamide, A = *N*-acetyl-L-alaninamide, L = *N*-acetyl-L-leucinamide. [b] The binary solute systems were equimolal within 0.1 per cent.

The corresponding expression for the osmotic coefficient for a binary solute system containing the solutes A and B in equimolar amounts ($y_{\mathrm{A}} = y_{\mathrm{B}} = 0.5$) is

$$\phi_{\mathrm{AB}} = 1+[(g_{\mathrm{AA}}+2g_{\mathrm{AB}}+g_{\mathrm{BB}})\,\mathbf{m}/4+(g_{\mathrm{AAA}}+3g_{\mathrm{AAB}}+3g_{\mathrm{ABB}}+g_{\mathrm{BBB}})\,\mathbf{m}^2/4+\ldots]/\boldsymbol{RT}. \tag{7}$$

All three eqn (5)-(7) can be written in the form

$$\phi = 1+(g_2\mathbf{m}+g_3\mathbf{m}^2+\ldots)/\boldsymbol{RT} \tag{8}$$

where g_2 now embraces all pairwise interactions and g_3 all triplet interactions.

The experimental data obtained from the isopiestic experiments are presented in table 1. The osmotic coefficients of the various solutions were obtained from

$$\phi = [m(\mathrm{urea})\,\phi(\mathrm{urea})/\mathbf{m}] \tag{9}$$

the osmotic coefficients for urea being obtained from Ellerton and Dunlop's very precise measurements.[24]

Each of the data sets (ϕ, **m**) was fitted to a polynomial of the above form [eqn (8)] and the appropriate coefficients were determined by a least-squares procedure.

In the program used, if the 95% confidence limit of a particular coefficient was found to be more than 100% the data were reanalysed with that coefficient excluded from the fit. This procedure continued until the coefficients remaining all had errors less than 100% within a 95% confidence limit. Thus any coefficient excluded had an effectively zero value. Hence, only the minimum number of parameters was retained. The g_2 and higher-order interaction parameters yielded by linear regression for the systems studied are given in table 2.

TABLE 2.—EXCESS FREE ENERGY PARAMETERS FOR THE PEPTIDE SYSTEMS INVESTIGATED

solute		g_2	g_3	g_4	
A	B	/J kg mol^{-2}	/J kg^2 mol^{-3}	/J kg^3 mol^{-4}	$10^3\sigma^b$
G	G	−82.9 (4.2)[a]	—	—	2.2
A	A	−144.2 (11.7)	+38.5 (10.5)	—	1.3
L	L	−731.9 (62.9)	+279.5 (106.8)	−73.1 (41.4)	3.8
G	A	−149.5 (36.8)	+53.5 (33.5)	—	4.2
G	L	−312.4 (35.4)	+63.5 (29.1)	—	4.1
A	L	−358.8 (17.4)	+78.9 (14.0)	—	2.0

[a] The number in parentheses represents the 95% confidence range of the coefficient. [b] σ is the standard error of the least-squares fit.

TABLE 3.—PAIRWISE FREE ENERGY COEFFICIENTS FOR THE PEPTIDE SYSTEMS INVESTIGATED

solute		g_{AB}
A	B	/J kg mol^{-2}
G	G	−82.9 (4.2)[a]
G	A	−185.5 (81.5)
G	L	−217.3 (103.9)
A	A	−144.2 (11.7)
A	L	−279.6 (71.8)
L	L	−731.9 (63.0)

[a] Bracketed term represents the 95% confidence limit.

The expansions all converge rapidly in the sense that $|g_2| > |g_3| > |g_4|$; this in itself implies that the systems are not strongly interacting. Indeed for most systems the first two terms describe the behaviour quite adequately. Only the solution containing *N*-acetyl-L-leucyl amide requires a quartet term.

Comparing eqn (5) and (6) and (8) for a single-solute system, the g_2 term corresponds exactly to the pairwise interaction coefficient for like species. For binary-solute systems, the g_2 term embraces like and unlike solute pair interactions. These may be separated using eqn (7) in the form

$$g_2 = \tfrac{1}{4}(g_{AA} + 2g_{AB} + g_{BB}) \qquad (10)$$

and since we have values for the like terms g_{AA} and g_{BB}, the cross-interaction term g_{AB} can also be evaluated.

The pairwise like and unlike coefficients are presented in table 3. The corresponding

TABLE 4.—EXPERIMENTAL ENTHALPIES OF DILUTION AT 298.15 K

$\mathbf{m}$/mol kg^{-1}	$n/10^{-3}$ mol	$\mathbf{m}'$/mol kg^{-1}	$q/10^{-1}$ J	$\Delta/10^{-4}$ J
N-acetyl-L-alaninamide + *N*-acetyl-L-alanyl-L-alaninamide				
0.5077	0.8763	0.2538	−1.2571	0
0.5077	0.1457	0.3752	−1.0527	+10
0.5077	1.6196	0.3158	−1.7363	−6
0.1765	0.9076	0.0892	−0.1743	+3
0.1765	0.1536	0.0437	−0.1180	+6
0.1765	0.1341	0.0304	−0.1198	0
N-acetylglycylglycinamide + *N*-acetylglycylglycylglycinamide				
0.6057	1.1569	0.2889	1.1547	−25
0.6057	0.5105	0.1066	0.8462	+26
0.6057	2.2350	0.4718	0.9055	−61
0.6057	1.0297	0.1809	1.3736	−35
0.6057	0.6481	0.1569	0.0540	+12
0.2685	0.5048	0.1316	0.2349	+12
0.2685	0.2629	0.0664	0.1971	+26
0.2685	0.9263	0.2094	0.1732	−6
N-acetyl-L-alanyl-L-alaninamide				
0.3221	1.1685	0.2576	−0.7072	0
0.3221	0.6054	0.1523	−0.9750	−6
0.3221	0.9177	0.2168	−0.9025	+4
0.3221	0.2691	0.0659	−0.6482	0
0.1730	0.3425	0.0863	−0.2759	+3
0.1730	0.5553	0.1275	−0.2333	+3
0.1730	0.1765	0.0463	−0.2241	+2
0.1730	0.2457	0.0541	−0.2730	+2
N-acetyl-L-alanylglycinamide				
0.4686	0.9067	0.2252	−0.6348	−7
0.4686	1.6626	0.3619	−0.5162	−12
0.4686	0.4272	0.1057	−0.4244	+16
0.4686	0.8548	0.1469	−0.7807	+1
0.2404	0.4820	0.1171	−0.1808	−12
0.2404	0.2232	0.0452	−0.1218	+2
0.2404	0.8842	0.1829	−0.1305	+14
0.2404	0.2546	0.0601	−0.1225	+7
0.2404	0.4765	0.1128	−0.1614	0
0.2404	0.2758	0.0625	−0.1308	+8
N-acetylglycylglycinamide				
0.4542	0.8716	0.2142	1.0961	−10
0.4542	0.4323	0.1070	0.8280	+6
0.4542	0.4231	0.0845	0.8664	+4
0.4542	1.7096	0.3502	0.9009	+3
0.4542	1.1078	0.3040	0.8545	+1
0.2191	0.4275	0.1071	0.2918	+10
0.2191	0.2351	0.0554	0.2181	−12
0.2191	0.8411	0.1733	0.2248	+2
0.2191	0.1645	0.0357	0.1743	−7

TABLE 4.—(*continued*)

$\mathbf{m}$/mol kg^{-1}	$n/10^{-3}$ mol	$\mathbf{m}'$/mol kg^{-1}	$q/10^{-1}$ J	$\Delta/10^{-4}$ J
		N-acetylglycinamide + *N*-acetylglycylglycylglycinamide		
0.1109	0.2048	0.0544	0.0794	−2
0.1109	0.4284	0.0862	0.0764	+2
0.1109	0.2782	0.0783	0.0642	+1
0.0558	0.1031	0.0264	0.0202	−1
0.0558	0.1565	0.0402	0.0243	+7
0.0558	0.0507	0.0107	0.0634	−9
0.0558	0.0368	0.0084	0.0870	−3
		N-acetylglycylglycylglycinamide		
0.1145	0.1152	0.0287	0.1520	+4
0.1145	0.2061	0.0371	0.2336	−6
0.1446	0.2545	0.0459	0.3614	−15
0.1446	0.4205	0.1096	0.2191	−1
0.1446	0.1540	0.0493	0.2428	+23
0.0760	0.0790	0.0160	0.0772	+7
0.0760	0.0806	0.0201	0.0666	−1
0.0760	0.0425	0.0097	0.0444	+2
0.0760	0.1360	0.0240	0.1174	+11
		N-acetyl-L-alanyl-L-alanyl-L-alaninamide		
0.0304	0.0562	0.0147	−0.0173	0
0.0304	0.0535	0.0093	−0.0224	+3
0.0304	0.0118	0.0204	−0.0191	−1
0.0304	0.0328	0.0066	−0.0207	−2
0.0304	0.1125	0.0239	−0.0087	+1
0.0304	0.0319	0.0078	−0.0185	−1

expression to eqn (1) for the excess enthalpy per kilogram of solvent is

$$H^{ex} = \sum_i \sum_j h_{ij} m_i m_j + \sum_i \sum_j \sum_k h_{ijk} m_i m_j m_k + \ldots \quad (11)$$

where the coefficients h_{ij} *etc.* are the appropriate enthalpy parameters for interaction between the subscripted species. These are related to the free energy parameters by, for example,

$$h_{ij} = [\partial(g_{ij}/T)/\partial T^{-1}]_p. \quad (12)$$

If we consider a solution containing one solute species, A or B, then the expressions for the excess enthalpies are

$$H_A^{ex} = m_A^2(h_{AA} + h_{AAA} m_A + \ldots) \quad (13)$$

$$H_B^{ex} = m_B^2(h_{BB} + h_{BBB} m_B + \ldots). \quad (14)$$

The expression for the excess enthalpy of a solution containing solutes A and B in equimolar amounts is

$$H_{AB}^{ex} = \mathbf{m}^2[(h_{AA} + 2h_{BB} + h_{BB})/4 + (h_{AAA} + 3h_{AAB} + 3h_{ABB} + h_{BBB})\mathbf{m}/4 + \ldots]. \quad (15)$$

Consequently as with the osmotic coefficients we can write a generalised expression to include eqn (13)-(15) *viz.*

$$H^{ex} = \mathbf{m}^2(h_2 + \mathbf{m}h_3 + \ldots) \quad (16)$$

where the h_2 and h_3 terms represent all pairwise and all triplet interactions in the system under consideration.

The enthalpy coefficients were determined in the present investigation using dilution experiments. The experimental heat change, q, associated with a dilution is given by

$$q = n(\mathbf{m}' - \mathbf{m})\,[h_2 + (\mathbf{m}' + \mathbf{m})\,h_3 + \ldots]. \qquad (17)$$

The calorimetric data systems for the various systems (see table 4) were fitted to the above equation. The program used was a simple modification of that used to analyse the free energy data.

TABLE 5.—ENTHALPY OF INTERACTION PARAMETERS[a]

solute[b] A	solute[b] B	h_2 /J kg mol^{-2}	h_3 /J kg^2 mol^{-3}	$10^3\,\sigma$[d] /J mol^{-1}
G_2	G_2	−646.0 (26.6)[c]	175.4 (39.1)	0.8
A_2	A_2	939.5 (4.5)	—	0.3
G_3	G_3	−1499.1 (65.7)	—	1.2
A_3	A_3	4880.5 (1767.6)	−65.902 (3210)	0.2
G_2	G_2	−321.8 (29.1)	—	5.5
G_2	G	−701.7 (16.6)	—	0.4
A_2	A	622.7 (50.7)	—	0.7
AG	AG	284.3 (5.0)	—	1.0

[a] In the binary solute systems, the solutes were equimolal within 0.1%. [b] Abbreviations: G_2 = *N*-acetylglycylglycinamide; AG = *N*-acetyl-L-alanylglycinamide; A_2 = *N*-acetyl-L-alanyl-L-alaninamide; G_3 = *N*-acetylglycylglycylglycinamide; A_3 = *N*-acetyl-L-alanyl-L-alanyl-L-alaninamide. [c] Bracketed term represents the 95% confidence limit. [d] σ is the standard error of the least-squares fit.

Values for the parameters h_2 and h_3 determined from the least-squares analyses are given in table 5. The systems seem to be well described by rapidly converging series and $|h_2| > |h_3|$ for most systems. Frequently only the leading term is necessary to describe the concentration dependence. As can be seen from table 5 the coefficients vary considerably in both sign and magnitude.

A notable exception to this general pattern of behaviour is A_3. For this solute the pairwise interaction coefficient is very large and $|h_3| > |h_2|$. Comparison of these values with those for G_3 where the solute is approximately the same size suggests that the self-interaction behaviour of A_3 is qualitatively different. This will be commented on later.

The pairwise enthalpy interaction parameters are tabulated in table 6. For a single-solute system, h_2 corresponds to the pairwise interaction constant whereas for a binary-solute system with solutes present in equimolal quantities we have [see eqn (15) and (16)]

$$h_2 = \tfrac{1}{4}(h_{AA} + 2h_{AB} + h_{BB}). \qquad (18)$$

Using the relationship $g_{AB} = h_{AB} - Ts_{AB}$

[see eqn (12)] the entropy parameters for the solutions containing amino acid amides were determined. These values, together with the corresponding free energy and enthalpy parameters, are given in table 7.

TABLE 6.—PAIRWISE ENTHALPY COEFFICIENTS FOR THE PEPTIDE SYSTEMS INVESTIGATED

solute		h_{AB}
A	B	/J kg mol^{-2}
G_2	G_2	−646.0 (26.6)[a]
A_2	A_2	939.5 (4.5)
G_2	G_3	−1499.1 (65.7)
A_2	A_3	4880 (1768)
G_2	G	−210.6 (16.5)
G_3	G	−543.7 (70.3)
A_2	A	641.4 (110.5)
AG	AG	284.3 (5.0)

[a] The parenthetical figure denotes the 95% confidence limit. The terms for the binary solute were calculated from the 95% confidence limits in h_2 (see text) and the component single-solute terms.

TABLE 7.—FREE ENERGY, ENTHALPY AND ENTROPY PAIRWISE INTERACTION COEFFICIENTS

solute		g_{AB}	h_{AB}	Ts_{AB}
A	B	/J kg mol^{-2}	/J kg mol^{-2}	/J kg mol^{-2}
G	G	−83	−220	−137
G	A	−186	86	272
G	L	−217	547	764
A	A	−144	269	413
A	L	−280	899	1179
L	L	−732	1714	2446

DISCUSSION

The excess thermodynamic functions have been analysed using a molality expansion to yield appropriate solute pairwise interaction coefficients. Although these coefficients are thermodynamic quantities and as such relate to properties of the solutions, the relationship between them and molecular events is complicated. A link between the free energy coefficients and solute interactions in solution can be made in an exact way using the McMillan–Mayer[25] theory of solutions. It is possible to transpose[26,27] such experimental interaction coefficients to the McMillan–Mayer state *via* a knowledge of the partial molar volumes of the solutes and the isothermal compressibility of the pure solvent. Whilst the latter is known the partial molar volumes of the solutes investigated are not. The transposition of the enthalpy coefficients to the McMillan–Mayer state is even more demanding of primary experimental data. In particular the temperature dependence of the solute partial molar volumes is required.

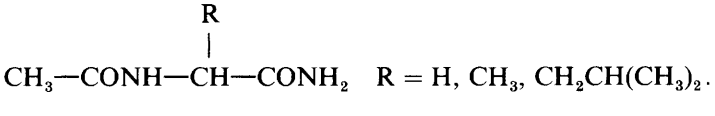
Consequently, given the lack of volumetric data for the present systems it has not been possible to pursue deconvolutions like those previously carried out and we have taken recourse to more qualitative and semi-empirical interpretations of the results.

We begin with a qualitative survey of the results obtained.

AMINO ACID AMIDES

We now have available (see table 7) the pairwise interaction coefficients for free energy, enthalpy and entropy for the amino acid derivatives G, A and L.

The negative value of the pairwise excess free energy parameters indicates that there is positive association or net attraction between these solutes in aqueous solution. The amino acid derivatives in this set all share the same general formula in which the 'backbone' of the molecule is constant, whereas the substituent on the α-carbon varies from one amino acid to another:



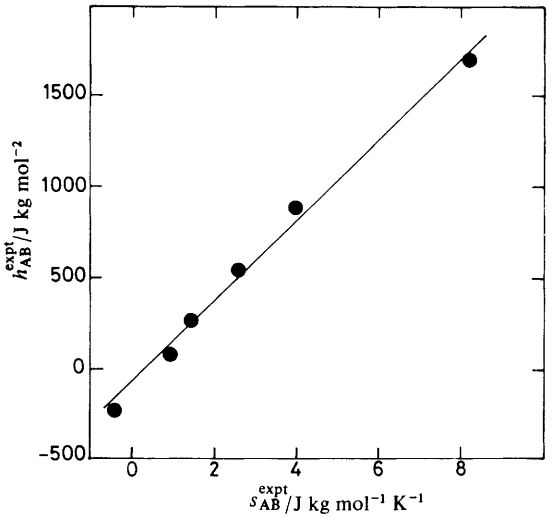


FIG. 1.—Plot of the experimental pairwise enthalpy coefficients (h_{AB}^{expt}) against the experimentally obtained pairwise entropy coefficient (s_{AB}^{expt}) for the *N*-acetyl amino acid amides investigated.

The pairwise interaction coefficients become increasingly more negative as the molecular size increases. This observed increase would therefore seem to be a result of greater association of the molecules through their side chains.

The g_{AB} values are the result of large but opposing h_{AB} and s_{AB} values. This compensatory effect is a widely-observed phenomenon[28] and the specific linear relationship between entropy and enthalpy change which is observed in a variety of processes involving small molecules in aqueous solution has become known as Lumry's law. This linear relationship is shown in fig. 1 for the present systems.

In fig. 2 g_{AB} is plotted against h_{AB}; this plot is qualitatively useful in that it reveals an approximately linear relationship between the pairwise free energy and enthalpy parameters for the systems studied. Although empirical, this could be useful in a predictive sense to evaluate order of magnitude values for the g_{AB} terms from a knowledge of the corresponding h_{AB} terms. As observed in this study the enthalpy terms are more easily determined experimentally than the corresponding free energy terms. From such a limited data set, however, it is not possible to judge fully the usefulness of this approach.

For molecular pair interactions which rely on solute hydrogen bonding or strong dipolar effects the enthalpy term[29, 30] should be negative. From table 7 we see that the only amino acid derivative from the set for which this is so is *N*-acetylglycinamide. The main functional groups of this molecule are the two amide residues in the backbone and unlike the other amino acids there is no alkyl substituent on the α-carbon atom. The molecule is predominantly polar and we can surmise from its solution

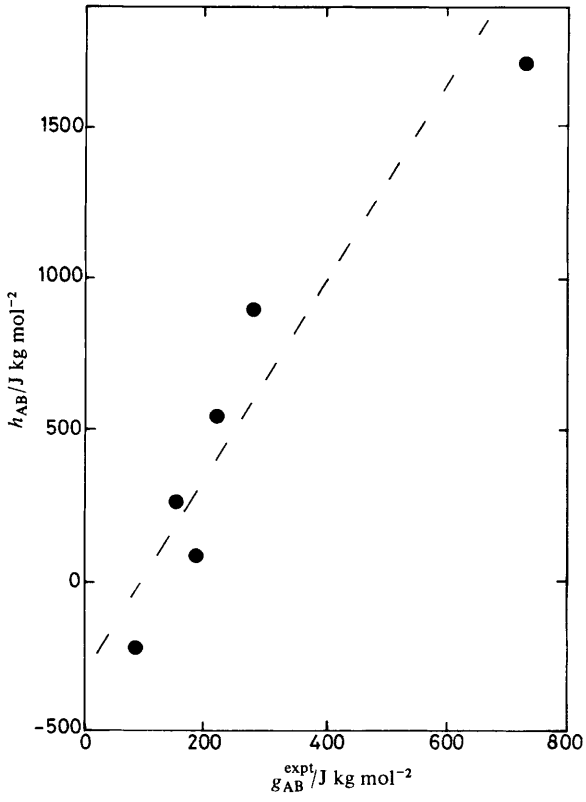


FIG. 2.—Plot of the experimental pairwise free energy coefficient (g_{AB}^{expt}) against h_{AB}^{expt} for the *N*-acetyl amino acid amides.

characteristics that the weak association of the solute in water is primarily through dipole interactions between the amide groups. The association of *G* is accompanied by an exothermic heat change. In general, the nature of the enthalpic change must depend on the relative enthalpic favourability of solute–solute and solvent–solvent interactions over the corresponding solute–solvent interactions. Since both water and acetylglycinamide have hydrogen-bonding characteristics the heat change can be rationalised[30] by assuming that breaking amide–water hydrogen bonds and forming amide–amide and water–water hydrogen bonds is a net exothermic process.

Spectroscopic studies on model amides[31, 32] in various solvents indicate that the strength of the intermolecular —C=O···H—N— bond varies inversely with the hydrogen-bonding potential of the solvent. Model compound studies, however, have been less than unanimous in their evaluation of the amide–amide interaction in aqueous solution, although it is generally agreed that the bond strength is weak.

Estimates[33-36] of the enthalpy of rupture of a peptide–peptide hydrogen bond range from 0 to 8 kJ mol^{-1}.

If we assume that the association of G is brought about entirely by intermolecular hydrogen bonding between amide groups and that the simple equilibrium is set up:

$$G + G \rightleftarrows G\cdots G$$

then it can be shown[36] that

$$h_{G-G} = \Delta H^{\ominus}_{G-G} K_{G-G} \tag{19}$$

$$g_{G-G} = -RT K_{G-G} \tag{20}$$

and

$$\Delta G^{\ominus} = -RT \ln K_{G-G} \tag{21}$$

where $\Delta H^{\ominus}_{G-G}$ is the standard enthalpy of association, K_{G-G} is the pairwise association constant of G with itself and $\Delta G^{\ominus}_{G-G}$ is the standard free energy of association. Combining the above equations we get

$$\Delta H^{\ominus}_{G-G} = -\frac{h_{G-G}}{g_{G-G}} RT \tag{22}$$

and substituting values from table 7 gives

$$\Delta H^{\ominus}_{G-G} = -6.6 \text{ kJ mol}^{-1}$$

$$K_{G-G} = 0.033 \text{ kg mol}^{-1}$$

$$\Delta G^{\ominus}_{G-G} = 8.4 \text{ kg mol}^{-1}.$$

TABLE 8.—THERMODYNAMIC ASSOCIATION CONSTANTS AND ASSOCIATED PARAMETERS FOR THE SYSTEMS INVESTIGATED

solute		K	$\Delta G^{\ominus}$	$\Delta H^{\ominus}$	$\Delta S^{\ominus}$
A	B	/mol^{-1} kg	/kJ mol^{-1}	/kJ mol^{-1}	/J K^{-1} mol^{-1}
G	G	0.033	8.4	−6.6	−50
G	A	0.075	6.4	1.2	−18
G	L	0.088	6.0	6.3	1
A	A	0.058	7.1	4.6	−8
A	L	0.113	5.4	8:0	+9
L	L	0.295	3.0	5.8	+9

Assuming further that on average four amide–amide hydrogen bonds are formed per dimer leads us to an estimate of −1.6 kJ mol^{-1} for the enthalpy of formation of an amide hydrogen bond. This value would clearly increase if less than four hydrogen bonds were formed and, although approximate, it is compatible with earlier estimates. Without justification we have completely ignored the possibility that methyl and methylene group effects are involved in the interaction; clearly if these effects are significant the estimate would need further adjustment.

Similar calculations can be performed for the other systems investigated and the results of such calculations are given in table 8.

Note that both dipole–dipole and hydrogen-bonding mechanisms are contributing in some way to the heat effects even though the enthalpy is dominated by the heat associated with the breaking of hydrophobic interactions. If we can assume that the contribution from polar effects is similar for all solutes in the peptide series, *i.e.*

equivalent to $\Delta H^{\ominus}_{G-G}$, it allows us to assign values to the alkyl group contributions by subtracting the polar (*i.e.* $\Delta H^{\ominus}_{G-G}$) contribution from the above values. Contributions from the methyl and isopropyl groups to the standard enthalpy of association of the peptides are thus evaluated as

$$\Delta H^{\ominus}_{Me-Me} = +11.1 \text{ kJ mol}^{-1}$$
$$\Delta H^{\ominus}_{Pr-Pr} = +12.3 \text{ kJ mol}^{-1}.$$

Although this type of group contribution approach is frequently adopted, especially with low molecular weight species,[33] the accuracy of the derived group values is difficult to assess. Several questionable assumptions have been made, in particular that the amide group contribution is invariable.

Studies[37, 38] on the affinities of salts to various alkyl-substituted amides and on the interaction of salts with α-amino acids indicate that the salt–dipole interaction is modulated by the alkyl groups adjacent to the amide dipole. Thus while methyl or, in general, apolar groups do not participate directly in the dipole interaction they are able to exert an influence through their hydration spheres. Part of the trend observed here might therefore be due to the varying influence of the hydrocarbon side chains on the amide interactions. It would therefore seem inappropriate to place much emphasis on the numerical values obtained from this analysis.

Of particular note from the present set of results is the significant difference in behaviour exhibited by G and A. These species differ by the addition of one methyl group to the backbone structure. The effect on the entropy of association is large and the hydrophobic effect of the methyl group rapidly outweighs the effect of the two polar amide groups.

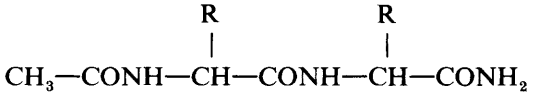
One feature of hydrophobic interactions which is of particular relevance to its biological role is cooperativity. This is also very marked in aggregation processes such as micellisation[39] but can be shown to contribute to the interactions of small alkyl derivatives such as alcohols and carboxylic acids.[30, 40] The triplet interactions included in table 2 demonstrate the cooperative element in the free energy of the peptide interaction. Another way of highlighting this cooperative effect is to compare contributions of pair and triplet effects at constant molality as a homologous series is ascended. A comparison of this sort is made in table 9.

TABLE 9.—COMPARISON OF PAIR AND TRIPLET CONTRIBUTIONS TO THE TOTAL SOLUTE INTERACTION AT 1 mol kg^{-1}

	G	A	L
$100\, g_3/g_2$	0	26	38

DIPEPTIDE AMIDES

The peptide amides used in this study have the general formula



where R represents either a methyl group or a hydrogen atom. Because of solubility limitations only enthalpy measurements were made. Comparison of the pair enthalpy

interaction coefficients for the dipeptides with the results from the corresponding monomer is shown in fig. 3.

The trend in the coefficients is more pronounced than in the corresponding monomer series but follows the same general pattern.

G_2 contains three amide groups in the backbone, which should result in the molecule being more hydrophilic than G. This is indeed reflected in the negative enthalpy

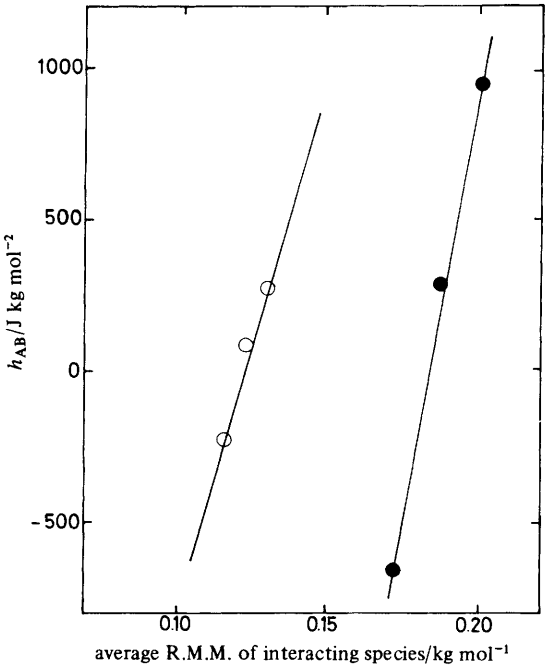


FIG. 3.—Plot of h_{AB}^{expt} against relative molar mass for the *N*-acetyl mono-amino acid amides (○) and the corresponding di-amino acid amides (●).

coefficient. If, however, the amide contribution was strictly additive the enthalpy of the dipeptide interaction would be 1.5 times that of the monomer; the experimental enthalpy term is actually nearer three times that of the monomer.

The pairwise enthalpy coefficient for the alanine dipeptide has a large positive value. While free energy data are not available it is reasonable to assume that the dipeptide is more strongly associated than the monomer and that hydrophobic interactions make large contributions to this. It would seem that the effect of additional methyl group far outweighs the effect of the additional amide group. The increased interaction over that of the monomer is larger than one would have predicted from the size alone.

The self-interaction of AG as expected from its constitution is intermediate in value between the interaction values for A_2 and G_2.

Thus while the pairwise enthalpy coefficients for the dipeptides are consistent with the corresponding values for the amino acid derivatives they could not have been extrapolated correctly from the monomer results.

TRENDS IN ENTHALPY COEFFICIENTS

By discussing the amino acid and peptide derivatives separately we were able to compare molecules with the same backbone structure differing only in side-chain functionality. If we now compare the pairwise enthalpy coefficients for the series amino acid, dipeptide, tripeptide, the combined effects on the solute interaction of increasing chain length and side-chain functionality may be considered.

Within these series of model compounds the structures have been varied systematically in an attempt to correlate the thermodynamic values obtained with the structural details of the molecules. These correlations are purely empirical and must be discussed with caution.

In the series G, G_2 and G_3 each glycine residue contributes the unit [NH—CH_2—CO] to the molecule. Throughout the series the pairwise enthalpy coefficients are negative and become increasingly so as the relative molar mass increases. The data are consistent with strong dipolar association and we have surmised (see above) that the attraction is essentially through the amide groups. The change in enthalpy is not, however, linear with the number of amide groups. Thus while the polar interaction is clearly important the total molecular interaction must be considered to be a more complex process.

In the series A, A_2 and A_3 each alanine contributes the unit [NH—$CH(CH_3)$—CO] to the molecules. The effect of replacing the hydrogen atom on the α-carbon by a methyl group has a most striking effect. Contrasting the solution behaviour of the glycine and alanine peptides highlights the importance of the side-chain on peptide association. The enthalpy coefficients all take positive values and these increase rapidly with increasing chain length. As in the glycine series, the trend is not monotonic and as the ratio of non-polar to polar groups is increased the solution behaviour of the solutes changes to that of a 'typically non-aqueous' solute.[41] The hydrophobic effect of the methyl groups seems to rapidly outweigh the hydrophilic effect of the backbone amide groups, *i.e.* the numerical values of the pairwise enthalpy terms increase more rapidly in the alanine than in the glycine series.

This difference in behaviour is most noticeable if we compare G_3 with A_3. Both solutes are similar in that they are approximately the same size and have the same backbone structure. While the pairwise enthalpy coefficient for A_3 is large and positive the triplet term is numerically even larger. If the free energy term is proportionately large then the interaction between these solutes is very strong and would indicate cooperative behaviour which is consistent with the relatively dominant hydrophobic character of the species.

It is apparent that the aqueous solution behaviour is determined by several features of the solute molecule, and as we anticipated the excess enthalpy of interaction is not a simple function of amide and hydrocarbon contributions or chain length. A complete description of the interaction must take into account the optimisation of all intermolecular effects, but without resort to a more sophisticated procedure this is not possible.

GROUP ADDITIVITY

The description of molecular interactions in terms of group interactions is potentially a useful approach. Whereas there is a vast number of chemical compounds of interest the number of functional groups which constitute these compounds is much smaller. If we presume that the physical properties of such compounds can be described in terms of the sum of contributions made by the molecules' functional groups, we obtain a method for correlating the properties of a large number of systems using a much smaller number of parameters. Each parameter would characterise some

property of an individual group. Any group contribution method is necessarily approximate since the contribution of a given group in one molecule is not necessarily the same in another molecule. It is similar in this regard to the concept of bond energies.

In order to implement a group contribution approach several assumptions must be made. The most fundamental of these is additivity: the contribution made by one group is assumed to be independent of that made by all other groups on the same molecule. Intermolecular forces acting on a group or whole molecule are thus determined by the average group composition of the system, *i.e.* are independent of how the groups are arranged in the molecule. This treatment is partially justified in that for non-electrolytic molecules, intermolecular forces are short range. The effect of distant groups in a molecule will be small unless they are brought together by conformational effects. However, nearest neighbour and steric effects are usually not insignificant and an approach which ignores molecular shape must remain approximate. Moreover, the local group composition adjacent to a particular molecule will not necessarily be equal to the average group composition of the whole solution. Random arrangement of the molecules is only strictly true for ideal solutions. Group additivity would thus seem to operate best within a group of molecules which are of related structure.

There is a growing collection of experimental data available for hydrophobic, hydrophilic and 'mixed' solutes in water and many attempts have been made at estimating additive contributions to thermodynamic quantities of certain functional groups.

Schrier and Schrier[42] have used a simple additivity scheme to describe the salting out behaviour of amides; the non-polar and polar residues on the amides were assumed to make independent contributions to the overall salting-out effect on the molecule. An alternative approach[43–45] using a 'solution of groups' model has been developed for estimating equilibrium properties of non-electrolyte mixtures. Functional groups are assumed to interact in solution in proportion to their mole fractions and are further assumed to be completely independent of how the groups are bonded together.

Using similar concepts to these Wood and Savage[46] developed a group additivity scheme which they applied to pair interactions between non-electrolytes in dilute aqueous solution. The approach was successfully applied to the description of the solution behaviour of over sixty solute species, including several amides. In view of this success, we felt it worthwhile to see how well such a procedure would represent the present results.

The group additivity approach assumes that when two solute molecules interact all groups on molecule A interact with all groups on molecule B and that each of these group interactions has a characteristic effect on the free energy or enthalpy, which is independent of the positions of the functional groups. The total pairwise interaction is then the sum of all the various group interactions that are present, *e.g.* for the enthalpy

$$h_{AB} = \sum_{i,j} n_i^A n_j^B H_{ij} \tag{23}$$

where n_i^A is the number of groups of type i on molecule A, n_j^B is the number of groups of type j on molecule B and H_{ij} is the characteristic contribution to the enthalpy of one i group interacting with one j group. The summation is taken over all groups i on molecule A and all groups j on molecule (B). An analogous equation can be used for the free energy term. If the magnitude of the interactions is small compared with the thermal energy of the molecules then all of the functional groups on one molecule

are able to interact with each of the groups on the other molecule. Neglecting steric and neighbouring group effects, the total free energy of interaction would consequently be group additive.

In applying this type of model it is necessary to assign a set of functional groups which when added together make up the set of molecules in the data set. The choice is arbitrary and the accuracy of any correlation improves with increasing distinction of groups. In the limit as more and more distinctions are made we recover the molecule itself, and in that event the advantage of the group additivity method is lost. Thus for practical purposes a compromise must be reached. The number of distinct groups must remain small but not so small as to neglect significant effects in molecular structure which affect the physical properties.

Following the previous work,[47] we considered that the peptide and amino acid derivatives studied here could be adequately described by two functional groups: the hydrocarbon and the peptide group.

The hydrocarbon function is based on the methylene group and as before it is further assumed that one CH_3 is equivalent to 1.5 CH_2 groups and that a CH group is equivalent to 0.5 CH_2 groups. By 'peptide group' we refer to the units —CONH or —$CONH_2$, *i.e.* no distinction is made between the peptide group proper and an unsubstituted amide group. These two primary groups give rise to three types of group interactions representing peptide–peptide, hydrocarbon–hydrocarbon and peptide–hydrocarbon contributions. Using the terminology more usually applied to describe peptide interactions, these correspond at least approximately to backbone–backbone, side-chain–side-chain and backbone–side-chain interactions.

In terms of the above group definitions, the pairwise enthalpy coefficients for the interaction between molecular species A and B can be expressed as

$$h_{AB} = (n^{A}_{CH_2} n^{B}_{CH_2})H_{CH_2\text{–}CH_2} + (n^{A}_{CH_2} n^{B}_{Pep} + n^{A}_{Pep} n^{B}_{CH_2})H_{CH_2\text{–}Pep} + (n^{A}_{Pep} n^{B}_{Pep})H_{Pep\text{–}Pep} \quad (24)$$

where, for example, $n^{A}_{CH_2}$ represents the number of CH_2 groups in species A and $H_{Pep\text{–}Pep}$ is a term representing the enthalpy of interaction of one peptide group with another. If an expression of the above type is used to fit a set of pairwise enthalpy data the group interactions can be obtained.

Initially[5] the group enthalpy interaction parameters generated by Savage and Wood were used to calculate the molecular interaction parameters, h_{AB}, for the present systems. The calculated values agree in sign and approximate magnitude with the experimentally determined enthalpies. However, notwithstanding the generally good agreement, the experimental results are usually greater than the calculated values and in particular the calculated coefficients for the solute A_2 and A_3 were considerably different from the experimental values.

In a later paper, Wood[36] applied the same additivity scheme to evaluate excess free energy group parameters. In adapting the approach to free energy data it is necessary to choose a concentration scale with which to define an ideal solution. On the molal scale

$$g_{AB} = \sum_{ij} n^{A}_{i} n^{B}_{j} G_{ij} \quad (25)$$

and the mole fraction scale

$$g_{AB} = -\frac{M_w RT}{2} + \sum_{ij} n^{A}_{i} n^{B}_{j} G_{ij}. \quad (26)$$

Wood found that there was a better fit between experimental and calculated values if the mole fraction scale was used. Using his group interaction terms and eqn (26)

the interaction terms were calculated for the various solutes and are compared with experimentally determined pair free energy terms in fig. 4.

The predicted values for the free energy terms do not agree will with the experimentally determined values. Although the correct sign is predicted the magnitudes of the values are not at all well reproduced.

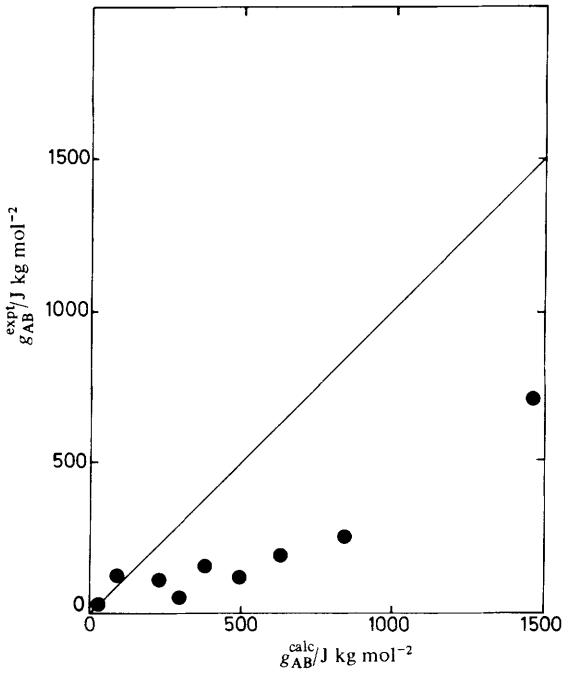


FIG. 4.—Plot of g_{AB}^{expt} against g_{AB}^{calc}. The calculated values were obtained using the earlier[36] parameters for group interactions.

The breakdown in group additivity for the free energy parameters could be a result of incorrect assumptions inherent in the model but it is also probable that the group interaction parameters are in error and need reassessment. In particular doubt are the interactions involving the peptide groups since in the original data set no molecule contained more than one such group.

In a group additivity approach it is assumed that the physical properties of a group may be transferred from one molecule environment to another and suffer relatively little change. This assumption is most likely to be valid if the group environments are similar in both molecules. Savage and Wood[46] used a very diverse set of non-electrolyte data to evaluate group interaction parameters. The data set included alcohols, sugars, amides and ureas. No distinction was made between unsubstituted, mono- or di-substituted amides and ureas.

Since the long-term aim of this work is to relate these model solute interactions to peptide and polypeptide systems it is important that the model solutes are good analogues for peptide residues. It is obvious considering the particular molecules investigated here that without reference to possible conformational effects the molecular environments for the functional groups are similar. Group interaction parameters derived from these model systems should therefore transpose well.

We therefore determined to select a more consistent data base by combining our substituted peptide and amino acid data with data on unsubstituted and monoalkyl-substituted amides from which we could extract a set of refined group interaction parameters which would better describe peptide interactions. Since few pairwise free energy coefficients have been reported only a limited data set could be assembled. The

TABLE 10.—REFINED GROUP FREE ENERGY PARAMETERS, G_{ij}(J kg mol^{-2})

	$G_{CH_2-CH_2}$	G_{CH_2-Pep}	$G_{Pep-Pep}$
molal scale	−29.8	+29.7	−59.1
mole fraction scale	−29.1	+27.8	−48.5

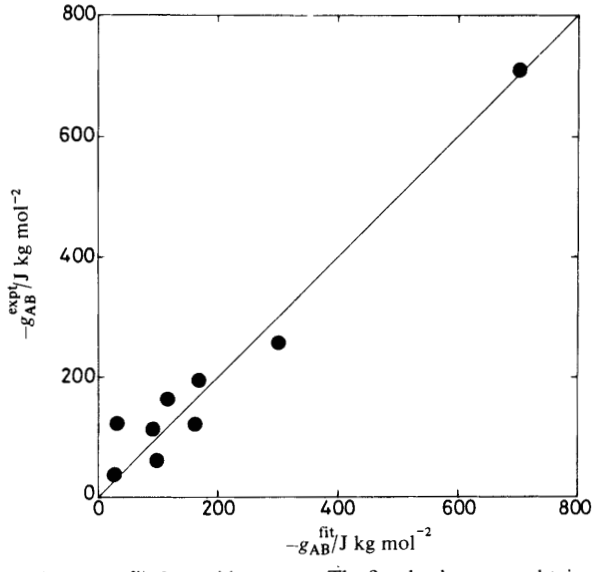


FIG. 5.—Plot of g_{AB}^{expt} against g_{AB}^{fit} for amide systems. The fitted values were obtained using the mole fraction group additivity parameters given in table 10.

amides selected for this data set were: *N*-acetylglycinamide, *N*-acetyl-L-alaninamide, *N*-acetyl-L-leucinamide, formamide, acetamide and propionamide.[48] By using a data set of greater coherence we avoid some of the unsatisfactory approximations used previously.

Refined group interaction parameters were evaluated using eqn (25) and (26) using a linear regression analysis. Pairwise free energy coefficients evaluated on the molal and mole fraction scales were analysed separately to test which gave the better correlation. The standard deviation of the fit on the mole fraction scale (53.2 J kg mol^{-2}) was slightly better than that on the molal scale (60.4 J kg mol^{-1}) but the values of the group parameters are not significantly different bearing in mind the crudity of the approximations made (see table 10).

The agreement between experimental values and values predicted using the mole fraction group parameters is shown in fig. 5.

The experimental data are reasonably reproduced by the refined group parameters.

The standard deviation is slightly higher than in the original correlation, largely as a result of the much smaller data base. However, we feel that the values are likely to be more realistic in the context of amide and peptide interactions.

To generate group enthalpy parameters, data from the present study were combined with results on unsubstituted and monoalkyl-substituted amides and all were fitted to eqn (26). When A_3 was included in the data set the fit was extremely poor with the standard deviation rising to 521.7 J kg mol^{-2} as compared with a standard

TABLE 11.—FUNCTIONAL GROUP PAIRWISE ENTHALPY OF INTERACTION PARAMETERS, H_{ij}(J kg mol^{-2})

	$H_{CH_2-CH_2}$	H_{CH_2-Pep}	$H_{Pep-Pep}$	σ^b
Savage and Wood parameters	+40 (8)[a]	+41 (34)	−252 (113)	220
refined parameters	+14 (13)	+95 (29)	−311 (57)	140

[a] The numbers in parentheses represent the 95% confidence limits. [b] σ is the standard deviation of the fit in J kg mol^{-2}.

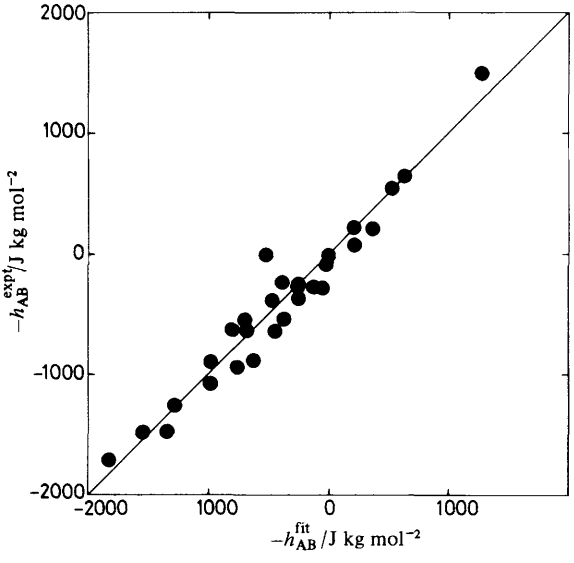


FIG. 6.—Plot of h_{AB}^{expt} against h_{AB}^{fit}. The fitted values were obtained using the group additivity parameters given in table 11.

deviation of 220 J mol^{-2} found in the original Savage and Wood[46] correlation. We felt justified in omitting this system from the parameterisation since it is our belief that additional specific effects are influencing the solution behaviour of this molecule.

With the omission of this system from the data base, enthalpy parameters were generated which gave a much improved correlation between experimental and calculated h_{AB} values. The new parameters are shown in table 11 as are the earlier[36] parameters. The results are also presented in a different form in fig. 6 where good agreement, both in sign and magnitude, is apparent between the fitted and experimental enthalpy coefficients.

The refined parameters while of similar sign are different in magnitude to those obtained earlier. The present data base includes molecules with one, two, three and four amide groups, whereas the previous data set was limited to molecules with only one amide group. We are therefore confident that the Pep–Pep and Pep–CH_2 interactions are better characterised by our refined parameters and that they will be more satisfactory for predicting peptide and amide interactions.

Using the above revised group free energy and enthalpy parameters we can evaluate the group entropy values from

$$G_{ij} = H_{ij} - TS_{ij}. \quad (27)$$

A list of all of the group parameters is presented in table 12.

TABLE 12.—FREE ENERGIES, ENTHALPIES AND ENTROPIES OF INTERACTION OF FUNCTIONAL GROUPS. UNITS ARE J kg mol^{-2}

i/j	CH_2	Pep
CH_2	G_{ij} −29	
	H_{ij} +14	
	TS_{ij} +43	
Pep	G_{ij} +28	G_{ij} −48
	H_{ij} +95	H_{ij} −311
	TS_{ij} +67	TS_{ij} −263

The CH_2–CH_2 pair interaction is qualitatively consistent with the established view[39, 47, 49, 50] of hydrocarbon interactions. A negative G_{ij} term is usually interpreted as a net attraction between the groups which arises from a large positive term. We also observe that $|TS_{CH_2-CH_2}| > |H_{CH_2-CH_2}|$, which is a recognised characteristic of the hydrophobic interaction.

The hydrophobic interaction has been widely studied[4, 51-55] both theoretically and experimentally, but unfortunately there are few published data directly comparable with the present work. It has been the custom to calculate hydrophobic side-chain contributions to protein conformation by considering the transfer[56-58] of amino acids and peptides from water to a non-aqueous solvent medium which might simulate the interior of the folded protein. One is thus able to calculate the difference in say the free energy between a side-chain completely surrounded by water and that of the side-chain surrounded by apolar groups without any contact with water molecules. The thermodynamic parameters thus estimated[50] do not correspond to the formation of pairwise hydrophobic bonds as described by our model. We have assumed a weak, non-specific association between hydrocarbon groups in an aqueous medium and anticipate only a partial shielding from water on association. Our pairwise association parameters are thus more in keeping with Némethy and Scheraga's minimum hydrophobic bond strength parameters.[49]

Comparison of the present results for the CH_2–CH_2 interaction with other quantitative estimates of the hydrophobic interaction is also difficult because of the different models used.[53-55] In our model systems the hydrocarbon parts are flanked by amide groups. We thus recognise that the polar group may interfere with the establishment of water structure around the non-polar group which would modify the hydrocarbon interaction. Our results are therefore most appropriately referred to peptide and amide systems and may only be extrapolated to alternative systems with caution.

Various values[33-36,59] have been attributed to the strength of the peptide–peptide bond and it is not clear to what extent these interactions affects the stability of protein structures. The present results indicate that the peptide group interactions make a large contribution to the overall peptide interaction and the magnitude of the Pep–Pep term shows that it has a larger effect than that for the CH_2–CH_2 interaction. The negative pairwise enthalpy term corresponds to an association dominated by hydrogen bonding or dipole effects. Using eqn (22) and values for $G_{\mathrm{Pep-Pep}}$ and $H_{\mathrm{Pep-Pep}}$ (from table 12) we estimate the standard enthalpy of association of two amide groups as

$$\Delta H^{\ominus}_{\mathrm{Pep-Pep}} = -16.1 \text{ kJ mol}^{-1}.$$

The present group estimate is much more negative than our preliminary estimate from the Gly data (see earlier). In this initial estimate it was assumed that *N*-acetylglycinamide self-interactions were dominated by the two amide groups and methyl and methylene group effects were ignored. Klotz and Franzen[33] reported a zero value for the enthalpy of rupture of the CO···HN bond from studies on *N*-methylacetamide but they too ignored the methyl group contribution. However, the positive enthalpy of formation of the hydrophobic bond will in part cancel the negative enthalpy contribution made by the hydrogen bond. Clearly if this destabilising contribution is ignored the amide interaction will be underestimated by an equivalent amount. Thus oversimplification of the model can lead to misleading conclusions.

The large decrease in entropy associated with Pep–Pep interaction indicates less freedom of motion when the two groups interact. This is consistent with the formation of a strong hydrogen bond or bonds which would tend to reduce the entropy of the dimer. Model studies[48] indicate that amide interactions probably do involve a hydrogen-bonding mechanism.

In previous studies on model compounds much emphasis was placed on evaluating the hydrocarbon and amide self-interactions and usually only passing reference[40] is made to the effect that polar and non-polar groups exert on each other. The results from this work indicate that these latter effects are by no means insignificant. The association of a peptide and a hydrocarbon group produces a strong destabilising effect which is approximately equal to the stabilising effect of the $CH_2\cdots CH_2$ interaction. Since there is no compatible mode of bonding between these two groups the enthalpic term is repulsive, but is compensated in part by the positive entropy term. This may be rationalised if it is assumed that the hydration spheres surrounding the polar and apolar groups are incompatible and that as the groups approach each other the hydration spheres overlap and the water structure is in part broken down. This effect is consistent with the view[37,38,60-63] that group effects can be modulated by adjacent groups *via* the solvent structure. While interpretation of model studies using a site-binding model would underestimate such cross-interaction terms, the present group additivity model[44-46] possibly overemphasises them. However, it seems likely that these interactions are real and their importance should be acknowledged.

CONCLUDING COMMENTS

The group additivity model has allowed us to evaluate free energy, enthalpy and entropy contributions for the peptide and hydrocarbon groups. In principle we could use these values to predict the molecular pair interaction coefficients for any molecule composed only of these groups. This model is recognised to be a first approximation and several questionable assumptions are inherent in this scheme.

Although the group expansion are empirical they can be given a rough interpretation. The molecular interaction parameters g_{AB} are related to the osmotic second virial coefficient *via* the integral of $[\exp(-W_{\mathrm{AB}}/kT)-1]$ where W_{AB} is the potential of

mean force for pairs of molecules. If W_{AB} is to be expressed in terms of additive group contributions, this implies that G_{ij} (the interaction between each pair of groups) must not depend on the presence of neighbouring groups on either molecule, *i.e.* all steric effects can be ignored, and that the effects of solvation are made up of additive contributions. Also the sum of all the group interactions must be small compared with kT to allow the factor $[\exp(W_{AB}/kT)-1]$ to be expanded into a linear sum of interaction terms. These assumptions, while providing a useful working model, represent a fairly extreme case and will not necessarily be applicable to all molecular interactions.

The model assumes that each functional group on one molecule is able to interact freely with each functional group on another molecule, independently of their positions. Steric effects could reduce interaction by preventing free access to all groups or conversely could enhance interactions. While such effects must cause deviations from the simple additivity rule, it should be possible to use these deviations to investigate the nature of such anomalous behaviour. As an example of this, we found that the additivity treatment was inadequate for solutions containing A_3. This discrepancy does not appear to arise from the behaviour of the peptide backbone since correspondence between experiment and calculation for G_3 is good. It thus almost certainly arise from intramolecular cooperativity of the alanine methyl groups, and we believe this intramolecular effect modifies the intermolecular behaviour.

While thermodynamic analysis in itself cannot reveal molecular features, it does establish constraints within which interpretation of molecular behaviour must operate. Structural molecular detail can be disclosed by a variety of spectroscopic probes and although such studies have not been reported for the solutes studied in this work, related results are enlightening. It has been shown that in aqueous solution small peptides adopt extended conformations which are not completely random.[64] Amino acid side-chains considerably effect the molecular conformation and free rotation is restricted for all peptide linkages except those involving glycine. Thus while glycine peptides are almost fully extended in aqueous solution, trialanine approaches a simple helix. There is also evidence which supports the existence of folding in short-chain alanine peptides and subsequent side-by-side aggregation of the folded species.[65, 66] We have used space-filling models of A_3 to investigate the possibility of folded structures in this species. Several conformations are accessible, in one of which the three methyl groups are drawn into a cluster arrangement giving the molecule one hydrophobic and one hydrophilic face with a consequent reduction in the overall extent of alkyl group's surface in contact with solvent. If this conformation was energetically favoured in aqueous solution it would be possible for the molecules to associate through their hydrophobic faces. Although unsubstantiated, this idea would explain the behaviour of A_3 and would be in accord with the enthalpy data.

In this study we have restricted our attention to molecules comprised of amino acids with aliphatic side-chains only. The relative simplicity of these molecules and the study of homologous series has helped to clarify the analysis. However, if the approach is to have predictive power for aqueous protein systems the group parameter analysis must be expanded to embrace other functions. If the analysis could be developed to handle longer peptides it might well highlight further examples of the A_3 type and would enter the zone in which intramolecular peptide interactions compete with intermolecular forces. This information should be particularly relevant to conformational features such as the β-turn and identification of polypeptide folding nucleation centres in proteins. One other area which might also be usefully explored using group additivity approaches is that of enzyme–substrate interactions.

These and other areas are currently being explored in this laboratory.

We acknowledge support from the A.R.C. for equipment and the S.R.C. for the award of a Research Studentship to E.W.

[1] G. Némethy and H. A. Scheraga, *Q. Rev. Biophys.*, 1977, **10**, 239.
[2] E. Clementi, *Lecture Notes in Chemistry* (Springer-Verlag, Berlin, to be published).
[3] See, *e.g. Water, A Comprehensive Treatise*, ed. F. Franks (Plenum Press, New York, 1973 and 1975), vol. 2 and 4.
[4] G. Némethy, W. J. Peer and H. A. Scheraga, *Annu. Rev. Biophys. Bioeng.*, 1981, **10**, 459.
[5] G. M. Blackburn, T. H. Lilley and E. Walmsley, *J. Chem. Soc., Faraday Trans. 1*, 1980, **76**, 915.
[6] G. M. Blackburn, T. H. Lilley and E. Walmsley, *J. Chem. Soc., Chem. Commun.*, 1980, 1091.
[7] F. Franks and D. Eagland, *CRC Crit. Rev. Biochem.*, 1975, **3**, 165.
[8] P. L. Privalov, *Adv. Protein Chem.*, 1980, **33**, 167.
[9] T. H. Lilley and I. R. Tasker, to be published.
[10] C. J. Epstein, R. F. Goldberger and C. B. Anfinsen, *Cold Spring Harbor Symp. Quant. Biol.*, 1963, **28**, 439.
[11] C. B. Anfinsen and H. A. Scheraga, *Adv. Protein Chem.*, 1975, **29**, 205.
[12] D. B. Wetlaufer and S. Ristow, *Annu. Rev. Biochem.*, 1973, **42**, 135.
[13] R. R. Hantgan, G. G. Hammes and H. A. Scheraga, *Biochemistry*, 1974, **13**, 3421.
[14] P. N. Lewis, F. A. Monamy and H. A. Scheraga, *Proc. Natl. Acad. Sci. USA*, 1971, **68**, 2293.
[15] E. Ralston and J-L. De Coen, *J. Mol. Biol.*, 1974, **83**, 393.
[16] M. N. Jones and H. A. Skinner, *Annu. Rep. Chem. Soc.*, 1979, **76C**, 253.
[17] F. Franks, in *Biochemical Thermodynamics*, ed. M. N. Jones (Elsevier, Amsterdam, 1979).
[18] T. H. Lilley and R. P. Scott, *J. Chem. Soc., Faraday Trans. 1*, 1976, **72**, 184.
[19] P. K. Nandi and D. R. Robinson, *J. Am. Chem. Soc.*, 1972, **94**, 1299.
[20] M. Bodansky, J. T. Sheehan, M. A. Ondetti and S. Lande, *J. Am. Chem. Soc.*, 1963, **85**, 991.
[21] M. Goodman, R. Rupp and F. Naider, *Bioorg. Chem.*, 1971, **1**, 294.
[22] P. G. Katsoyannis, K. Fukuda and A. Tometski, *J. Am. Chem. Soc.*, 1963, **85**, 1681.
[23] H. L. Friedman, *Ionic Solution Theory* (Interscience, New York, 1962).
[24] H. D. Ellerton and P. J. Dunlop, *J. Phys. Chem.*, 1966, **70**, 1831.
[25] W. G. McMillan and J. E. Mayer, *J. Chem. Phys.*, 1945, **13**, 276.
[26] See, *e.g.* B. P. Kelley and T. H. Lilley, *J. Chem. Soc., Faraday Trans. 1*, **74**, 2771.
[27] R. H. Wood, T. H. Lilley and P. T. Thompson, *J. Chem. Soc., Faraday Trans 1*, 1978, **74**, 1990.
[28] R. Lumry and S. Rajender, *Biopolymers*, 1970, **9**, 1125.
[29] R. B. Cassel and R. H. Wood, *J. Solution Chem.*, 1973, **2**, 119.
[30] F. Franks, M. Pedley and D. S. Reid, *J. Chem. Soc., Faraday Trans. 1*, 1975, **72**, 359.
[31] G. Rialdi and J. Hermans, *J. Am. Chem. Soc.*, 1966, **88**, 5719.
[32] I. M. Klotz, *J. Colloid Interface Sci.*, 1968, **27**, 84.
[33] I. M. Klotz and J. S. Franzen, *J. Am. Chem. Soc.*, 1962, **84**, 3461.
[34] S. J. Gill and L. Noll, *J. Phys. Chem.*, 1972, **76**, 3065.
[35] G. C. Kresheck, *J. Phys. Chem.*, 1969, **73**, 2441.
[36] B. Y. Okamoto, R. H. Wood and P. T. Thompson, *J. Chem. Soc., Faraday Trans. 1*, 1978, **74**, 1890.
[37] A. Hambata and P. H. von Hippel, *Biochemistry*, 1973, **12**, 1264.
[38] A. Hambata, S. Chang and P. H. von Hippel, *Biochemistry*, 1973, **12**, 1271.
[39] C. Tanford, *The Hydrophobic Effect* (John Wiley, New York, 1973).
[40] J. Konicek and I. Wadsö, *Acta. Chem. Scand.*, 1971, **25**, 1571.
[41] J. S. Rowlinson, *Liquids and Liquid Mixtures* (Butterworths, London, 1959).
[42] E. E. Schrier and E. B. Schrier, *J. Phys. Chem.*, 1967, **71**, 1851.
[43] G. M. Wilson and C. H. Deal, *Ind. Eng. Chem., Fundam.*, 1962, **1**, 20.
[44] G. A. Ratcliff and K. C. Chao, *Can. J. Chem. Eng.*, 1969, **47**, 148.
[45] A. Fredenslund, R. L. Jones and J. M. Prausnitz, *AIChE J.*, 1975, **21**, 1086.
[46] J. J. Savage and R. H. Wood, *J. Solution Chem.*, 1976, **5**, 733.
[47] C. Tanford, *J. Phys. Chem.*, 1962, **84**, 4240.
[48] T. H. Lilley and R. H. Wood, *J. Chem. Soc., Faraday Trans. 1*, 1980, **76**, 901.
[49] G. Némethy and H. A. Scheraga, *J. Phys. Chem.*, 1962, **66**, 1773.
[50] W. Kauzmann, *Adv. Protein Chem.*, 1959, **14**, 1.
[51] H. A. Scheraga, *Ann. N.Y. Acad. Sci.*, 1977, **2**, 303.
[52] H. A. Scheraga, *J. Am. Chem. Soc.*, 1979, **12**, 7.
[53] E. E. Schrier, M. Pottle and H. A. Scheraga, *J. Am. Chem. Soc.*, 1964, **86**, 3448.
[54] H. Schneider, G. C. Kresheck and H. A. Scheraga, *J. Phys. Chem.*, 1965, **69**, 1310.
[55] D. G. Oakenfull and D. E. Fenwick, *Aust. J. Chem.*, 1973, **26**, 2649.
[56] C. H. Spink and M. Auker, *J. Phys. Chem.*, 1970. **74**, 1742.

[57] P. K. Nandi, *Int. J. Pept. Protein Res.*, 1976, **255**, 253.
[58] G. C. Kresheck, H. Schneider and H. A. Scheraga, *J. Phys. Chem.*, 1965, **69**, 3132.
[59] H. Sasi, S. M. Timasheff and J. S. Ardi, *J. Biol. Chem.*, 1964, **239**, 3051.
[60] J. E. Desnoyers, G. Perron, L. Avedikian and J-P. Morel, *J. Solution Chem.*, 1976, **5**, 631.
[61] M. J. Mastroianni, M. J. Pikal and S. Lindenbaum, *J. Phys. Chem.*, 1972, **76**, 3050.
[62] G. Némethy, I. Z. Steinberg and H. A. Scheraga, *J. Am. Chem. Soc.*, 1963, **85**, 3866.
[63] T. H. Lilley and I. R. Tasker, *J. Chem. Soc., Faraday Trans. 1*, 1982, **78**, 1.
[64] B. A. Levine and R. J. P. Williams, *Jerusalem Symp. Quantum Chem. Biochem.*, 1976, 95.
[65] M. Goodman, N. Ueyama and F. Naider, *Biopolymers*, 1975, **14**, 901.
[66] M. Goodman, N. Ueyama, F. Naider and C. Gilon, *Biopolymers*, 1975, **14**, 915.

(PAPER 1/1324)