Control of Peptide Conformational Preferences by Intramolecular Ion-pairing: a NMR and Molecular Mechanics Analysis of Glycyl-containing Dipeptides in Aqueous Solutions

Craig Beeson [a] **and Thomas A. Dix** *[a,b]
[a] *Department of Chemistry, The University of California, Irvine, California 92717, USA*
[b] *Department of Biological Chemistry, The University of California, Irvine, California 92717, USA*

A NMR and molecular mechanics analysis of isomeric N- and C-terminal glycyl dipeptides in aqueous solutions is described, with the overall objective of understanding the basis of their different rotameric preferences. The C-terminal glycyl dipeptides are significantly more conformationally homogeneous than their N-terminal isomers in all water-containing solvents, as probed by the magnitude of the chemical shift non-equivalence of the diastereotopic glycyl methylene protons in the ^{1}H NMR spectra of both isomers. Further NMR experiments demonstrated that, with both sets of isomers, the degree of conformational homogeneity is under control of the extent of intramolecular ion pairing; the ion pairing is thus a greater conformational determinant in C-terminal glycyl dipeptides. Molecular mechanics studies failed to demonstrate an inherent structural basis for the differences in rotameric preferences between isomers. To explain the differences, it is hypothesized that alkyl substitution of the α-carbon in C-terminal glycyl dipeptides destabilizes the ammonium ion to a greater extent through steric inhibition of solvation than in N-terminal dipeptides, which forces intramolecular ion pairing to occur to a greater extent. The significance of these results towards the *de novo* prediction of polypeptide tertiary structure in solution is discussed briefly.

A satisfactory understanding of the factors controlling polypeptide conformational preferences in aqueous solutions remains elusive, despite decades of study employing a battery of spectroscopic and physical techniques. NMR spectroscopy, for example, has been applied extensively to the study of the conformational preferences of small peptides, with the identification of spectroscopic probes applicable to structurally more complex molecules as a primary focus. The intrinsically diastereotopic glycyl methylene protons are one such probe: a large difference in chemical shift between these protons can be indicative of preferred conformations of the peptide backbone.[1] However, the magnitude of the chemical shift difference is not directly related to conformer populations; the difference arises both from the predominance of a small subset of rotamers *and* anisotropic electrical fields, the magnitude of each effect being highly structure-dependent.[2] A classic example of how small changes in structure drastically affect rotamer preferences is the behaviour of isomeric sets of N-terminal glycyl (gly-L-aa)† and C-terminal glycyl (L-aa-gly) dipeptides in aqueous solutions. The large difference in chemical shift between the glycyl protons of L-aa-gly has been interpreted[3] to indicate that, in D_2O, selective population of only a few rotamers is occurring whereas with gly-L-aa, a large number of rotamers are significantly populated. Despite almost three decades of study by NMR spectroscopy,[3] the structural basis of these differences remains controversial.

In this manuscript, we describe a combined NMR and molecular mechanics analysis that re-evaluates the origin of the chemical shift differences of gly-L-aa and L-aa-gly dipeptides. We find that formation of an intramolecular, solvent-separated, ion pair can occur with both sets of dipeptides at low concentrations to selectively populate certain rotamers. In a given solvent, however, the ion-pairing rotamers are more favoured with L-aa-gly than with gly-L-aa because the alkyl-substituted α-carbon in L-aa-gly disturbs the ammonium ion solvation shell to a greater extent than it disturbs the carboxylate ion solvent shell in gly-L-aa. Thus, L-aa-gly dipeptide ammonium ions are forced to 'solvate' by intramolecular ion pairing to a greater extent than the ammonium ions of gly-L-aa dipeptides. Ultimately, the greater ability of the L-aa-gly ion pair to control rotamer preferences results from 'steric hindrance to solvation'[4] modulation of the free energy of the electrostatic interaction, which we have also defined as a factor in ion-pair formation in aqueous solutions using conformationally homogeneous molecules.[5] The key conformational determinant for the dipeptide isomers in water is thus the interaction of structure and solvent, as is certainly the case with larger peptides of biological interest.

Results

The ^{1}H NMR spectra of isomers glycyl-L-alanine (gly-L-ala) and L-alanylglycine (L-ala-gly) in D_2O, shown in Fig. 1(*a*) and 1(*b*), illustrate the previously observed[3] differences between dipeptide isomers. In D_2O, all methylene proton resonances of N-terminal glycyl dipeptides (gly-L-aas) exhibited A_2 spin systems and thus shift equivalence ($\Delta\delta = 0$) whereas all C-terminal glycyl dipeptides (L-aa-glys) exhibited AB spin systems (shift non-equivalence, $\Delta\delta > 0$)– the spectra of gly-L-ala and L-ala-gly are representative. In less polar solutions such as aqueous CD_3CN, the methylene proton resonances of gly-L-ala exhibited a $\Delta\delta$; the magnitude of $\Delta\delta$ for L-ala-gly increased in the same solution. The methylene proton resonances of other N-terminal glycyl dipeptides also reverted from an A_2 to an AB spin system in aqueous CD_3CN solutions; representative $\Delta\delta$ values are listed in Table 1. Clearly, the magnitude of $\Delta\delta$ is highly structure-dependent; the largest values are observed for dipeptides containing aromatic side chains.

† The abbreviations used are: gly-L-aa, a dipeptide with an N-terminal glycyl residue; L-aa-gly, a dipeptide with a C-terminal glycyl residue; gly-L-ala, glycyl-L-alanine; L-ala-gly, L-alanylglycine; $\Delta\delta$, proton chemical shift non-equivalence; CD_3CN, deuteriated acetonitrile; [$^2H_6$]DMSO, deuteriated dimethyl sulphoxide; DSS, 2,2-dimethyl-2-silapentane-3-sulphonate.

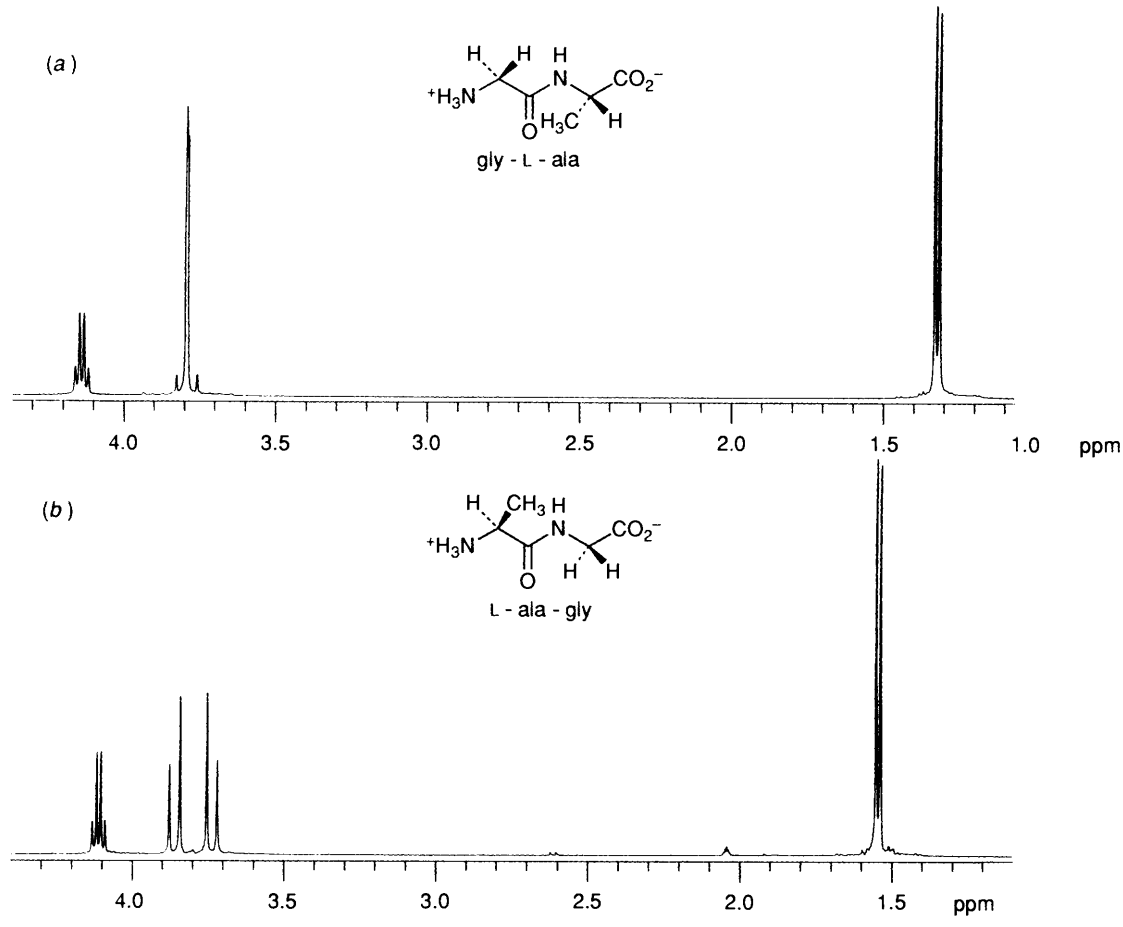

Fig. 1 Proton NMR spectra of: 5 mmol dm^{-3} gly-L-ala in D_2O (*a*) and 5 mmol dm^{-3} L-ala-gly in D_2O (*b*). For both spectra, the alanyl methyl resonance (1.32 or 1.55 ppm) is a doublet, the alanyl methine resonance (4.14 and 4.11) is a quartet. In (*a*), the glycyl methylene resonances (3.8 ppm) are an A_2 spin system; in (*b*) (3.8 ppm), an AB spin system.

Table 1 Methylene proton resonance chemical shifts for different dipeptides

	δ_1 CH_2 (ppm)	δ_2 CH_2 (ppm)	J_{ab}/Hz	$\Delta\delta$ (ppm)
L-aa-Gly				
Ala-gly	3.851	3.733	17.2	0.118
Val-gly	3.897	3.687	17.2	0.210
Phe-gly	3.838	3.527	17.2	0.311
Tyr-gly[a]	3.95	3.75	*b*	0.20
Leu-gly[a]	3.89	3.70	*b*	0.19
Ile-gly[a]	3.90	3.70	*b*	0.20
Gly-L-aa in D_2O–CD_3CN[c]				
Gly-ala	3.746	3.695	15.9	0.051
Gly-val	3.748	3.697	15.9	0.051
Gly-phe	3.587	3.454	16.0	0.133
Gly-tyr	3.698	3.558	16.1	0.140
Gly-leu	3.738	3.680	15.9	0.058
Gly-ile	3.772	3.714	15.9	0.058

[a] Values from ref. 3(f). [b] Range of reported values of J_{ab}, 17.0–17.6 Hz. [c] 46 mol% CD_3CN–54 mol% D_2O.

The failure of the glycyl residues of the N-terminal dipeptides to exhibit an observed $\Delta\delta$ in D_2O could be due to coincidentally similar magnetic environments for the two protons induced by the presence of the C-terminal α-carbon alkyl group. However, the $\Delta\delta$ of N-terminal glycyl dipeptides exhibited the same degree of sensitivity to the side chain in aqueous CD_3CN as did the C-terminal glycyl dipeptides in D_2O. The large difference in $\Delta\delta$ between L-ala-gly and L-phe-gly (Table 1) argued against possible screening of aromatic ring currents by D_2O. The lack of a temperature dependence for $\Delta\delta$ eliminated the possibility of hindered rotation.* These arguments support the initial claim[3] that the observed $\Delta\delta$ of methylene protons in glycyl dipeptides reflects the selective population of a small subset of rotamers. In addition, the similar sensitivity to side-chain substituents of all dipeptides demonstrated that the magnitude of $\Delta\delta$, within a pair of isomeric dipeptides, reflects the degree of conformational heterogeneity. Thus, the consistently greater $\Delta\delta$ observed with L-aa-gly *versus* gly-L-aa isomers in a given solvent indicated that N-terminal α-carbon substitution of the dipeptide is acting as a conformational determinant.

Acid titration of the sodium salt of L-ala-gly in D_2O (Fig. 2) demonstrated that the $\Delta\delta$ is most pronounced when the dipeptide is zwitterionic; equivalent behaviour was observed for gly-L-ala in aqueous CD_3CN solutions (data not shown). Previous studies[3] had also noted that only dipeptide zwitterions produced a significant $\Delta\delta$; it was proposed that the dipeptides were selectively populating rotamers in which the ammonium and carboxylate ions could interact. Further support for intramolecular ion pairing as a conformational determinant was obtained from the study of gly-L-ala in a series

* A small decrease in $\Delta\delta$ was observed at the higher temperature, which may indicate the population of rotamers that were thermodynamically less accessible at the lower temperature

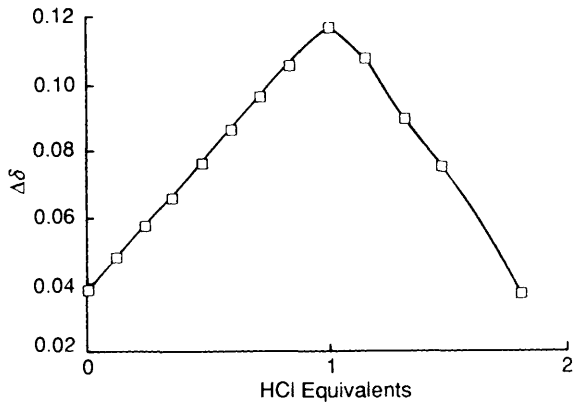

Fig. 2 The magnitude of chemical shift non-equivalence (Δδ) for the two methylene protons during the acid titration of 10 mmol dm^{-3} sodium L-alanylglycinate in D_2O

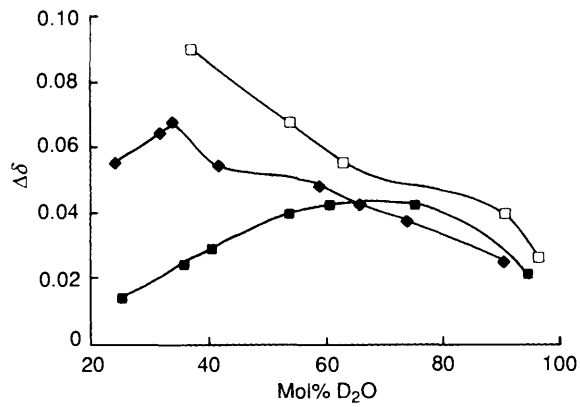

Fig. 3 The magnitude of chemical shift non-equivalence (Δδ) for the two methylene protons of 5 mmol dm^{-3} gly-L-ala in [$^2H_6$]DMSO–D_2O (■). CD_3CN–D_2O (◆) and [$^2H_8$]-1,4-dioxane–D_2O (□) solutions

of mixed aqueous solvents. (Gly-L-ala was chosen for detailed study because it was most amenable to computer simulations, see below.) Changes in Δδ as a function of solvent composition for gly-L-ala in aqueous [$^2H_8$]-1,4-dioxane, CD_3CN, and [$^2H_6$]DMSO solvent mixtures are shown in Fig. 3. For all three solvent mixtures, starting at high mol% D_2O, a decrease in mol% D_2O increased the stability of ion-paired rotamers, as reflected in the observed increase in Δδ. The ion-pairing maxima, with a decrease at lower mol% D_2O in CD_3CN and [$^2H_6$]DMSO, solutions had also been implicated from conductivity determinations of intermolecular ion-pair formation in the same solutions.[6] Experimental problems* prevented assessment of this effect in very low mol% D_2O in [$^2H_8$]-1,4-dioxane solvents.

A comparison of the behaviour of gly-L-ala amongst the different cosolvents provided further support for ion pairing as a conformational determinant. The two relevant cosolvent factors are bulk permittivity and hydrogen bond acceptor capability; the former affects the stability of both ammonium and carboxylate ions while the latter affects the stability of only ammonium ions. The relative permittivity values of DMSO and acetonitrile are similar (46.7 and 37.5) while the relative permittivity of 1,4-dioxane is much lower (2.2); the hydrogen-bond-acceptor values of DMSO and 1,4-dioxane are similar while that of acetonitrile is much lower.[7] The highest degree of solvent charge stabilization (as evidenced by lower Δδ values) was observed in the aqueous [$^2H_6$]DMSO solutions, which have high bulk permittivities and good hydrogen-bond-acceptor capabilities. In contrast, aqueous CD_3CN solutions exhibited larger magnitude Δδs and a lower mol% D_2O value at which the largest Δδ was observed; both results reflect ammonium ion destabilization and increased ion pairing due to the lack of hydrogen-bond-accepting ability of CD_3CN *versus* [$^2H_6$]DMSO. (The relative permittivities of each solvent at a given mol% D_2O are almost identical.[8]) Results in aqueous [$^2H_8$]-1,4-dioxane and CD_3CN amplified the importance of the hydrogen-bond-accepting ability of the cosolvent. For each solvent, the magnitude of Δδ was a comparable function of the mol% D_2O in the cosolvent– for [$^2H_8$]-1,4-dioxane, its hydrogen-bond-accepting ability compensated for its much lower permittivity compared to CD_3CN. The effect of a lower dielectric for [$^2H_8$]-1,4-dioxane was manifested in the greater magnitude of Δδ at comparable mol% D_2O in [$^2H_8$]-1,4-dioxane *versus* [$^2H_6$]DMSO solvents.

Both intra- and inter-molecular ion pairing will be enhanced by solvents poor in ion-solvating properties. Thus, the concentration dependence of Δδ for gly-L-ala in aqueous [$^2H_8$]-1,4-dioxane, CD_3CN, and [$^2H_6$]DMSO solutions, under conditions of significant ion-pair formation (as defined by the data presented in Fig. 3), was evaluated (Fig. 4). The decreased Δδs noted at higher dipeptide concentrations in [$^2H_8$]-1,4-dioxane- and CD_3CN-containing solvents were diagnostic† of the shift from intra- to inter-molecular associations; the abrupt transition in the [$^2H_8$]-1,4-dioxane experiments indicated that, at higher concentrations, intermolecular associations were dominant whereas at lower concentrations, intramolecular associations were dominant. (For the aqueous [$^2H_6$]DMSO solution, dipeptide saturation was reached before attaining the concentration at which intermolecular interactions became significant.) All of the data reported in Figs. 1–3 were accumulated at dipeptide concentrations of approximately 5 mmol dm^{-3}, and thus reflect primarily or exclusively intramolecular interactions. The chemical shift values for all of the data presented in Figs. 1–3 appear in Table 2.‡

Molecular mechanics[9] calculations were performed to evaluate the possibility that inherent structural differences in the dipeptides, independent of their relative intramolecular ion-pairing abilities, were the basis of the greater tendency of C- *versus* N-terminal glycyl residue dipeptides to adopt a smaller subset of preferred rotamers. Utilizing the CHARMM force field (initially implemented for the study of polypeptides)[10] low-energy conformations of gly-L-ala and L-ala-gly were identified and evaluated. The two torsion angles evaluated are defined in Fig. 5; the assumption of *trans* peptide configuration is strongly supported by both computational[11] and experimental[12] investigations. The definition of low-energy rotamers and energy minimizations were performed as follows. First, the neutral dipeptides were evaluated to avoid any artificial bias imposed by ammonium–carboxylate ion interactions; global conformational searches identified nine equivalent rotamers for each dipeptide. Once defined, the dipolar form of the minimum and maximum energy gly-L-ala conformations were evaluated using explicit water molecules and periodic boundary conditions. The resulting energy differences were then used to locate a distance-dependent permittivity (RDE = 4) which produced a similar energy range in the absence of explicit solvation. Each of the nine gly-L-ala and L-ala-gly zwitterions were then energy minimized using this permittivity; the nine rotameric structures with their relative energies are listed in Fig. 6. The absence of

* The concentration of dipeptide at low mol% water in [$^2H_8$]-1,4-dioxane solutions was insufficient to overcome the shim humps resulting from presaturation of the HOD resonance, which coincides with the glycyl proton resonances in these solutions.

† A decrease in Δδ may reflect a larger number of rotamers that can form intermolecular ion pairs.

‡ Changes in the grminal coupling constant J_{AB} were consistent with changes in rotameric populations; however it is difficult to extract useful geometric relations owing to the small dispersion (15–18 Hz).

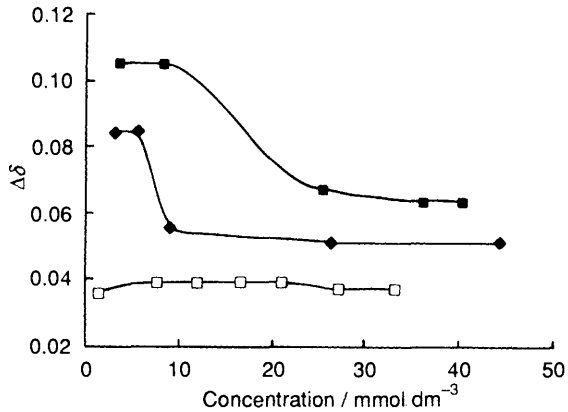

Fig. 4 The magnitude of chemical shift non-equivalence ($\Delta\delta$) for the two methylene protons of gly-L-ala at different dipeptide concentrations in 52 mol% [$^2H_6$]DMSO–48 mol% D_2O (□), 37 mol% [$^2H_8$]-1,4-dioxane–63 mol% D_2O (◆) and 58 mol% CD_3CN–42 mol% D_2O (■)

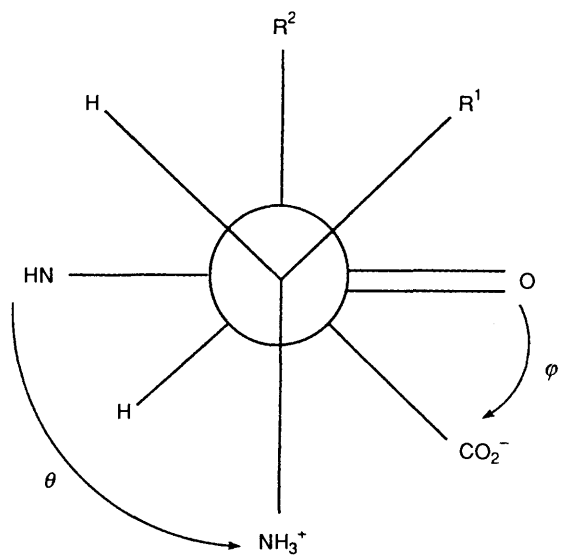

Fig. 5 Definition of the two torsions (θ, φ = positive degrees in direction of arrow) evaluated during conformational searches of gly-L-ala (R^1 = H, R^2 = CH_3) and L-ala-gly (R^1 = CH_3, R^2 = H). The perspective shown is a view from the N-terminal α-carbon towards the peptidic carbonyl carbon.

artificial ammonium–carboxylate ion interactions in the calculations was supported with the data from two conformers: the calculated lowest energy rotamer of gly-L-ala (conformation **1**) was all-*trans*, with the ions on opposite faces, while one of the highest energy rotamers (conformation **7**) was predicted to have the shortest distance between the ions. The lowest energy rotamer for gly-L-ala was essentially the same as previously predicted by *ab initio* calculations.[11a] Calculation of the solvent-accessible surface areas for each of the nine rotamers did not identify rotamers with significantly lower hydrophobic surface areas for either gly-L-ala or L-ala-gly that might be favoured in water (data not shown). The close energy correspondence between the nine rotamers of gly-L-ala and L-ala-gly support the contention that there is no purely structural basis for L-ala-gly to prefer a smaller rotamer population subset than gly-L-ala in identical solvents.

Discussion

In previous NMR studies,[3] the observation of a $\Delta\delta$ between the glycyl methylene proton resonances of glycyl-containing dipeptides was attributed to the molecule populating a preferred subset of an energetically accessible set of rotamers. Our ^1H NMR evaluation of glycyl-containing dipeptides in water and aqueous aprotic solutions fully supports this attribution.

Table 2 Methylene proton resonance chemical shifts and coupling constants for glycyl-alanyl dipeptides

	δ_1 CH_2 (ppm)	δ_2 CH_2 (ppm)	J_{AB}/Hz	$\Delta\delta$ (ppm)
L-Ala-gly (equiv. HCl)				
0.00	3.776	3.738	17.3	0.038
0.12	3.784	3.736	17.3	0.048
0.24	3.792	3.735	17.3	0.057
0.36	3.796	3.730	17.3	0.066
0.48	3.804	3.728	17.3	0.076
0.60	3.812	3.726	17.3	0.086
0.72	3.810	3.714	17.2	0.096
0.84	3.816	3.711	17.2	0.105
1.00	3.827	3.711	17.0	0.116
1.16	3.849	3.742	17.2	0.107
1.32	3.879	3.790	17.2	0.089
1.48	3.903	3.828	17.5	0.075
1.80	3.958	3.921	18.1	0.037
Gly-L-ala (mol% D_2O in DMSO)				
25	3.478	3.464	16.2	0.014
36	3.514	3.490	16.2	0.024
41	3.525	3.496	16.2	0.029
54	3.574	3.535	16.0	0.039
61	3.601	3.559	15.9	0.042
75	3.648	3.606	15.9	0.042
94	3.688	3.667	15.3	0.021
Gly-L-ala (mol% D_2O in dioxane)				
37	3.726	3.636	15.8	0.090
54	3.694	3.629	15.9	0.065
63	3.684	3.632	16.0	0.052
90	3.674	3.635	16.0	0.039
96	3.805	3.779	16.0	0.026
Gly-L-ala (mol% D_2O in CD_3CN)				
24	3.574	3.519	15.4	0.055
32	3.696	3.632	15.8	0.064
34	3.696	3.629	15.8	0.067
42	3.710	3.656	15.8	0.054
59	3.700	3.652	15.9	0.048
66	3.680	3.638	15.9	0.042
74	3.714	3.677	16.0	0.037
90	3.728	3.703	16.1	0.025
Gly-L-ala (mmol dm^{-3} in $(CD_3CN)^a$				
3.45	3.737	3.632	15.8	0.105
8.27	3.743	3.638	15.8	0.105
25.3	3.738	3.671	15.8	0.067
36.1	3.729	3.666	15.8	0.063
40.1	3.729	3.666	15.9	0.063
Gly-L-ala (mmol dm^{-3} in DMSO)^b				
1.44	3.540	3.504	16.5	0.036
7.57	3.558	3.519	15.9	0.039
11.9	3.560	3.522	15.9	0.039
16.6	3.566	3.527	15.8	0.039
20.9	3.566	3.527	15.8	0.039
27.2	3.570	3.533	15.8	0.037
33.2	3.570	3.533	15.8	0.037
Gly-L-ala (mmol dm^{-3} in dioxane)^c				
2.97	3.719	3.635	15.8	0.084
5.48	3.687	3.602	15.9	0.085
8.90	3.682	3.627	15.9	0.055
26.3	3.656	3.605	15.9	0.051
44.0	3.658	3.607	15.9	0.051

[a] 58 mol% CD_3CN–42 mol% D_2O. [b] 52 mol% DMSO–48 mol% D_2O. [c] 37 mol% [$^2H_8$]dioxane–63 mol% D_2O.

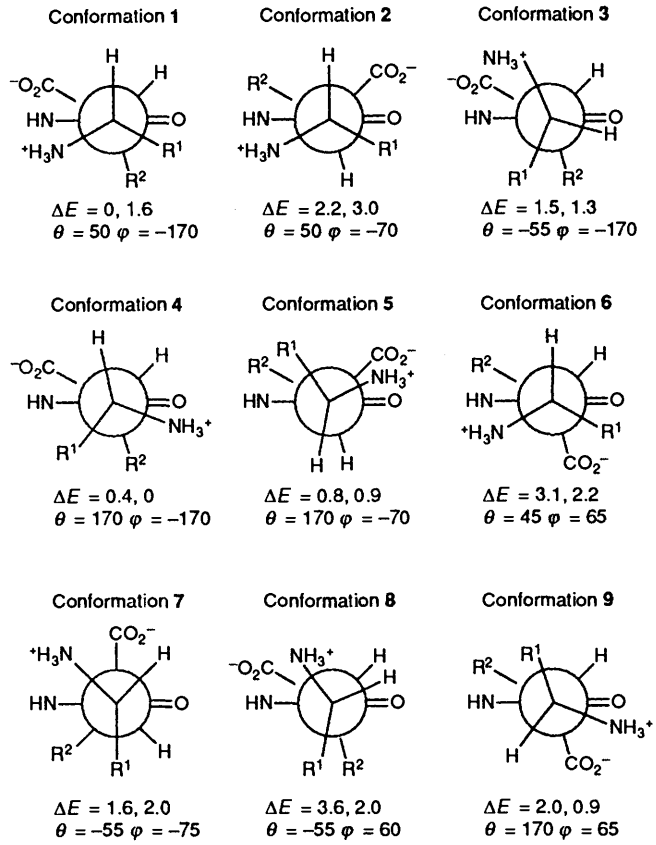

Fig. 6 The nine low energy rotamers of gly-L-ala (R^1 = H, R^2 = CH_3) and L-ala-gly (R^1 = CH_3, R^2 = H) and their relative energies (ΔE in kcal mol^{-1}) determined from molecular mechanics calculations as described in the text. (1 cal = 4.184 J.)

Furthermore, the similar sensitivity of $\Delta\delta$ to the side-chain substituent in gly-L-aa and L-aa-gly, as the solvating powers of the solution were varied, demonstrated that, within the isomeric pair, the magnitude of $\Delta\delta$ was indicative of the degree of conformational homogeneity. Since significant $\Delta\delta$ values were observed only when the dipeptides were zwitterionic, the intramolecular ion pair must act as a conformational determinant. To do so, the preferred rotamers must maximize the interaction by minimizing the distance between the ammonium and carboxylate ions; thus, the most likely candidates for selective population are lowest energy conformers **3** and **5** (Table 2) in which the distance between the ions is minimized. The nitrogen–oxygen distance predicted for **3** and **5** (5 Å) is an acceptable distance for a solvent-separated ion pair.* Others have used ^3J coupling constants determined for L-phenylalanylglycine to support the preponderance of the same conformations.[3e,f]

The relative strength of the ion pair in the dipeptides, as experimentally probed through the magnitude of $\Delta\delta$ for glycyl protons, provided further support for ion pair formation as a conformational determinant in the dipeptides. As the mol% water in a given solvent mixture was lowered, the strength of the electrostatic interaction increased and the distribution of rotamers narrowed to the conformations which maximized ion pairing to lower their free energy. This explains the increase in $\Delta\delta$ values for the methylene protons of gly-L-ala in the three solvent mixtures from 100–40 mol% water in dioxane, 100–35 mol% water in acetonitrile and 100–65 mol% water in DMSO (see Fig. 3). However, for each solvent mixture (and presumably, each dipeptide) there existed a certain composition that maximizes the extent of ion-pairing; at lower mol% water the extent of ion-pairing decreases.† Conductivity measurements were used previously to demonstrate that the reversal of the stability of an intermolecular ion-pair at lower mol% water occurs in aqueous acetonitrile, DMSO and 1,4-dioxane solutions.[6] The approximately co-incidental mol% water in co-solvent at which the ion pair stabilization was reversed, as seen in the previous[6] and current study, strongly argued for the proposed role of ion pairing as a dipeptide conformational determinant.

The substantially greater magnitude of $\Delta\delta$ observed for all L-aa-glys relative to gly-L-aas in all solvent mixtures evaluated was diagnostic of the population of a smaller number of rotamers in the C-terminal glycyl dipeptides; presumably, the ion-pairing conformers **3** and **5** (Fig. 6) were populated to a greater extent. For example, the electrostatic interaction was strong enough to function as a conformational determinant for L-ala-gly but not for gly-L-ala in water; in two-component aqueous solvent mixtures, this interaction was always a stronger conformational determinant for L-ala-gly. Previous workers attributed the conformational differences to a reduced conformational bias about the N-terminal Cα–CO torsion for the N-terminal glycyl dipeptides;[3c] however, our computational and experimental results argue against this attribution. Examination of the calculated energies for their various conformers of gly-L-ala and L-ala-gly (Fig. 6), in the context of the somewhat limited reliability of molecular mechanics calculations, provided little support for a more selective population of rotamers in L-ala-gly compared to gly-L-ala. Accordingly, any difference between the two dipeptides must arise from differential solvation of the free ions and ion pairs in various conformers. The key structural determinant of conformation appears to be the methyl substituent on the α-carbon of the N-terminus of L-ala-gly; and, to generalize, an alkyl substituent on the α-carbon of L-aa-glys *versus* gly-L-aas.

We propose that the N-terminal α-carbon alkyl substitution of L-aa-glys destabilizes the solvation sphere of the ammonium ion, thereby increasing the strength of its interaction with the carboxylate ion. This is supported by our studies[5] in which an *N*-hexyl substituent proved to be a greater than 2 kcal conformational determinant for formation of an intramolecular ion pair in a conformationally defined *cis*-decalin amino acid derivative. The sensitivity of ammonium ion solvation to adjacent alkyl groups has been noted previously. It has been well established in similar studies that the ammonium ion causes a greater constriction of solvent, and is more sensitive to substitution, than a carboxylate ion.[13] For example, the anomalous solution basicity of alkyl amines was initially attributed to solvation effects;[14] recent computational results have demonstrated that reduction in solvation due to an adjacent alkyl group diminished the transfer of charge from the ammonium ion to solvent water which caused net destabilization.[15] Experimental analyses of the molal volumes of dipeptides also demonstrated this phenomena; the presence of the methyl substituent at the N-terminal α-carbon in glycyl di- and tri-peptides perturbed the solvation of the ammonium ion which created non-additivity problems for group molal volume approximations.[16] Finally, a recent molecular dynamics analysis of a solvated enkephalin zwitterion also demonstrated the tighter solvation of the ammonium ion.[17] All of these effects

* The nitrogen–oxygen distance is adequate to allow for a water molecule of radius 1.4 Å to bridge the two ions.

† A general explanation for specific solvation effects in such solutions is lacking; however, in the case of the dipeptides, some of the observed $\Delta\delta$ decrease may be due to a shift of the zwitterion concentration towards the neutral species at very low mol% water. A corresponding upfield shift of the glycyl proton resonances is observed but it is partly due to solvent changes as well. A thorough analysis of this effect has been investigated with a more structurally defined *cis*-decalin amino acid and will be reported in another paper.

are manifestations of the concept of steric inhibition of solvation',[4] a central determinant of the nature and strength of electrostatic phenomena in water-containing solutions.

In conclusion, the extremely sensitive interaction of molecular structure and solvent is a major conformational determinant of dipeptide, and certainly of polypeptide, tertiary structure in aqueous solutions. Thus, approaches to the *de novo* prediction of polypeptide solution structure must incorporate the role of solvent; unfortunately, the role of solvent is the least understood aspect of the process. Ultimately, it is clear that the understanding of conformational preferences of polypeptides of biological interest must come from proportionally as detailed an analysis of structure–solvent interactions as was necessary to understand the preferences of the dipeptide isomers gly-L-ala and L-ala-gly. The orders-of-magnitude greater complexity of larger peptides is thus both daunting and challenging.

Experimental

Dipeptides obtained from Sigma Chemical Co. and United States Biochemical were dried *in vacuo* at 0 °C prior to use; deuteriated solvents were obtained from Aldrich Chemical Co. and Cambridge Isotopes Laboratories. All samples were prepared directly in the NMR tube under an argon atmosphere. The ^1H (300 or 500 MHz) spectra were recorded with a General Electric QE300, GN500 or Ω500 at 295 K. Chemical shifts in D_2O were measured from 2,2-dimethyl-2-silapentane-3-sulphonate (DSS); chemical shifts in D_2O–organic solvent mixtures were measured from the residual solvent resonance for the organic solvent. In a separate set of experiments, residual solvent resonances were measured relative to DSS in the D_2O–organic solvent mixtures and a calibration curve generated to enable correlation of all shift values relative to DSS. Spectra were sampled in a 32 K memory which was zero filled to 64 K. Apodization by exponential multiplication was done with a line broadening of 0.2 Hz. The intensity of the HOD and/or residual solvent resonances was often diminished with low power presaturation pulses. Molecular mechanics calculations were done on a Silicon Graphics 320VGX workstation using QUANTA from the Polygen Corporation.

Acknowledgements

We thank the Petroleum Research Fund, administered by the ACS (grant 21426-G4), and the University of California Cancer Coordination Research Committee, for financial support of this research (to T. A. D.). C. B. is grateful for support as a National Institutes of Health predoctoral trainee (GM07311). The computer hardware was purchased with instrument funds from the National Institutes of Health (RR-05690) to T. A. D.

References

1 H. Kessler, *Angew. Chem. Int. Ed. Engl.*, 1982, **21**, 512.
2 K. Mislow and M. Raban, *Top. Stereochem.*, 1967, **1**, 1.
3 (*a*) M. Mandel, *J. Biol. Chem.*, 1965, **240**, 1586; (*b*) V. J. Morlino and R. B. Martin, *J. Am. Chem. Soc.*, 1967, **89**, 3107; (*c*) A. Nakamura and O. Jardetzky, *Proc. Natl. Acad. Sci. USA*, 1967, **58**, 2212; (*d*) J. W. Westley and B. Weinstein, *J. Chem. Soc., Chem. Commun.*, 1967, 1232; (*e*) M. Kainosho, K. Ajisaka, M. Kamisaku and A. Murai, *Biochem. Biophys. Res. Commun.*, 1975, **64**, 425; (*f*) M. J. Anteunis, C. Becu, A. K. Lala, G. Verhegge and K. Narayan-Lala, *Bull. Soc. Chim. Belg.*, 1977, **86**, 161.
4 J. I. Brauman and L. K. Blair, *J. Am. Chem. Soc.*, 1970, **92**, 5986.
5 C. Beeson and T. A. Dix, submitted for publication in *J. Am. Chem. Soc.*
6 D. B. Rorabacher, W. J. MacKeller, F. R. Shu and (Sister) M. Bonavita, *Anal. Chem.*, 1971, **43**, 561; K. Mui, M. A. E. McBryde and E. Nieboer, *Can. J. Chem.*, 1974, **52**, 1821; E. Hawlicha and R. Grabowski, *Z. Naturforsch., Teil A*, 1991, **46**, 122.
7 M. H. Abraham, P. L. Grellier, D. V. Prior, R. W. Taft, J. J. Morris, P. J. Taylor, C. Larrence, M. Berthelot, R. M. Doherty, M. J. Kamlet, J. M. Abboud, K. Sraidi and G. Guiheneuf, *J. Am. Chem. Soc.*, 1988, **110**, 8534.
8 Y. Y. Akhador, *Dielectric Properties of Binary Solutions*, Pergamon Press, Oxford, 1981.
9 QUANTA 3.0, The Polygen Corp., Waltham, MA.
10 R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187.
11 (*a*) L. R. Wright and R. F. Borkman, *J. Phys. Chem.*, 1982, **86**, 3956; (*b*) W. L. Jorgensen and J. Gao, *J. Am. Chem. Soc.*, 1988, **110**, 4212.
12 D. B. Davies and M. A. Khaled, *J. Chem. Soc., Perkin Trans. 2*, 1973, 1651; J. Dorie, J. P. Gouesnard, B. Mechin, N. Naulet and G. J. Martin, *Org. Magn. Reson.*, 1980, **13**, 126.
13 R. Zana, *J. Phys. Chem.*, 1977, **81**, 1817; A. K. Mishra and J. C. Ahluwalia, *J. Phys. Chem.*, 1984, **88**, 86.
14 E. M. Arnett, F. M. Jones III, M. Taagepera, W. G. Henderson, J. L. Beauchamp, D. Holtz and R. W. Taft, *J. Am. Chem. Soc.*, 1972, **94**, 4724.
15 A. Mucci, R. Domain and R. L. Benoit, *Can. J. Chem.*, 1980, **58**, 953; P. Nagy, *J. Mol. Struct.*, 1989, **201**, 271.
16 M. K. Kumaran, G. R. Hedwig and I. D. Watson, *J. Chem. Thermodynamics*, 1982, **14**, 93; G. R. Hedwig, *J. Solution Chem.*, 1988, **17**, 383.
17 P. E. Smith, L. X. Dang and B. M. Pettitt, *J. Am. Chem. Soc.*, 1991, **113**, 67.

Paper 1/02826J
Received 12*th June* 1991
Accepted 8*th August* 1991