Experimental and Ab Initio Studies on Protonations of Alanine and Small Peptides of Alanine and Glycine

Carolyn J. Cassady,* Scott R. Carr, Kui Zhang, and Alice Chung-Phillips*

Department of Chemistry, Miami University, Oxford, Ohio 45056

Received November 7, 1994[⊗]

The gas-phase basicities of alanine (Ala) and di- and tripeptides of alanine and glycine were obtained by proton transfer reactions in a Fourier transform ion cyclotron resonance mass spectrometer. In addition, ab initio Hartree–Fock molecular orbital calculations were performed on the neutral and amino N-protonated species of glycine (Gly), alanine, and the four dipeptides (GlyGly, GlyAla, AlaGly, AlaAla). Minimum-energy structures were determined using both the 3-21G and 6-31G* basis sets with full geometry optimizations. Employing zero-point energies and thermal energies at 298.15 K and 1 atm calculated at the 3-21G//3-21G level and electronic energies at the 3-21G//3-21G, 6-31G*//3-21G, and 6-31G*//6-31G* levels, three sets of theoretical gas-phase basicity and proton affinity values were obtained. The relative basicities calculated at the highest level, 6-31G*//6-31G*, are in good agreement with the experimental values. Corrections that can be made to improve the calculated basicities are discussed in detail. The minimum-energy structures of the twelve species show consistent patterns of intramolecular hydrogen bonding in five-membered cyclic (C_5) forms; the presence of the alanyl methyl group has almost no effect on the structures. For the dipeptides and the tripeptides, the location of this methyl group at the N-terminus has the greatest impact on basicity, which shows the intrinsic ability of the methyl substituent near the protonation site to stabilize the ion.

Introduction

Recently there has been an increasing interest in the determination of fundamental thermodynamic and structural properties related to the proton transfer processes of biomolecules. Experimental studies have included measurements of gas-phase basicities (-ΔG of protonation) by mass spectrometry,[1-17] while ab initio calculations[1,2] have provided theoretical gas-phase basicities for the lowest energy protonated structures. These studies provide information about intrinsic properties in the gas-phase, in the absence of solvent effects that can play a large role in solution-phase measurements.[18] The protonation process is important because of its impact on hydrogen bonding, three-dimensional structure, and biological activity for peptides. Protonation is also of importance in gas-phase experiments employing mass spectrometry: the formation of protonated molecular ions is a dominant ionization pathway for peptides and the location of the proton may impact the ion's fragmentation patterns.

The experimental gas-phase basicities (GBs) of polyglycines (1–6 residues)[1,2] and serine and glycine dipeptides[3] have recently been determined in our laboratory. Due to the low vapor pressures of these biomolecules, the common technique of obtaining GBs by equilibrium measurements[19,20] could not be employed. Instead, protonated amino acid or peptide ions were produced by fast atom bombardment (FAB)[21] in the external source of a Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR). These ions were deprotonated by reactions with a series of reference compounds of known basicity, allowing the GB of the biomolecule to be bracketed. This same procedure has been used to determine the GBs of various biomolecules in the laboratory of Lebrilla,[12,13] and a similar procedure has been employed by Amster and co-workers[6,7] on ions generated by laser desorption.

The work in our laboratory on polyglycines included both mass spectral and ab initio studies.[1,2] In the ab initio research, Hartree–Fock (HF) calculations were carried out for glycine (Gly), diglycine (GlyGly), triglycine (GlyGlyGly), and all their stable N- and O-protonated species using either one or both of the smaller split-valence (3-21G) and the larger split-valence-plus-polarization (6-31G*) basis sets with full geometry optimizations. Zero-point energies and changes in enthalpy and entropy from 0 to 298 K were also obtained with either one or both basis sets for the subsequent gas-phase

[⊗] Abstract published in *Advance ACS Abstracts,* February 15, 1995.

(1) Zhang, K.; Zimmerman, D. M.; Chung-Phillips, A.; Cassady, C. J. *J. Am. Chem. Soc.* **1993**, *115*, 10812.
(2) Zhang, K.; Cassady, C. J.; Chung-Phillips, A. *J. Am. Chem. Soc.* **1994**, *116*, 11512.
(3) McKiernan, J. W.; Beltrame, C. E. A.; Cassady, C. J. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 718.
(4) Locke, M. J.; McIver, R. T., Jr. *J. Am. Chem. Soc.* **1983**, *105*, 4226.
(5) Meot-Ner, M.; Hunter, E. P.; Field, F. H. *J. Am. Chem. Soc.* **1979**, *101*, 686.
(6) Gorman, G. S.; Speir, J. P.; Turner, C. A.; Amster, I. J. *J. Am. Chem. Soc.* **1992**, *114*, 3986.
(7) Gorman, G. S.; Amster, I. J. *J. Am. Chem. Soc.* **1993**, *115*, 5729.
(8) Wu, Z.; Fenselau, C. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 863.
(9) Wu, Z.; Fenselau, C. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 403.
(10) Wu, Z.; Fenselau, C. *Tetrahedron* **1993**, *41*, 9197.
(11) Campbell, S.; Marzluff, E. M.; Rodgers, M. T.; Beauchamp, J. L.; Rempe, M. E.; Schwinck, K. F.; Lichtenberger, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 5257.
(12) Wu, J.; Lebrilla, C. B. *J. Am. Chem. Soc.* **1993**, *115*, 3270.
(13) Wu, J.; Lebrilla, C. B. *J. Am. Soc. Mass Spectrom.*, in press.
(14) Bliznyuk, A. A.; Schaefer, H. F., III; Amster, I. J. *J. Am. Chem. Soc.* **1993**, *115*, 5149.
(15) Bojesen, G. *J. Am. Chem. Soc.* **1987**, *109*, 5557.
(16) Li, X.; Harrison, A. G. *Org. Mass Spectrom.* **1993**, *28*, 366.
(17) Isa, K.; Omote, T.; Amaya, M. *Org. Mass Spectrom.* **1990**, *25*, 620.
(18) (a) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, 1987. (b) Tobias, D. J.; Brooks, C. L., III. *J. Phys. Chem.* **1992**, *96*, 3864.
(19) Kebarle, P. *Annu. Rev. Phys. Chem.* **1977**, *28*, 445.
(20) Aue, D. H.; Webb, H. M.; Bowers, M. T. *J. Am. Chem. Soc.* **1976**, *98*, 311.
(21) Barber, M.: Bodoli, R. S.; Sedwick, R. D.; Tyler, A. N. *J. Am. Chem. Soc., Chem. Commun.* **1981**, *7*, 325.

0022-3263/95/1960-1704$09.00/0

basicity (GB) calculations. The results were generally in good agreement with GBs obtained by experiments.

The polyglycine data provides a foundation from which other amino acids and smaller peptides can be examined because the polyglycine chain can be considered as the backbone of peptides. More complex peptides can be derived by the addition of functional groups at methylene sites. The simplest common substitution is addition of a methyl group at these sites to yield alanine residues. Such compounds are the focus of the present research. The GBs of alanine and dipeptides of alanine and/or glycine were determined experimentally and theoretically, while tripeptides of alanine and glycine were only studied experimentally.

In the past few years the ab initio literature relevant to the monomers[22] and dipeptides[23] of glycine and alanine has been surging rapidly. Recent interests on dipeptides center around model structures such as the blocked glycine and alanine dipeptides or the *N*-acetyl-*N*-methylglycinamide and -alaninamide which imitate the central regions of protein chains.[23] For the real dipeptides, we found practically no reports since the early SCF MO studies[24] on GlyGly, GlyAla, and AlaGly and their zwitterions using small basis sets and standard geometries. As for protonated peptides, there were a few studies[25,26] on the glycine systems prior to our work[1] on Gly and GlyGly.

In the present ab initio study, protonations of alanine (Ala), dialanine (AlaAla), and dipeptides of glycine and alanine (GlyAla and AlaGly) that lead to the most stable protonated species are investigated at the HF/3-21G and HF/6-31G* levels. On the basis of our prior experience, the most stable protonation site is the terminal nitrogen atom and the study is therefore limited to the amino N-protonated species in view of the large number of different peptide structures involved. In order to estimate the corrections on the calculated GBs due to deficiency in the basis set and lack of electron correlation, sufficiently high level calculations are performed on two model species, the neutral and protonated methylamine (CH_3NH_2 and $CH_3NH_3^+$). Thus, an objective of the ab initio research is to further examine the procedure for obtaining reasonable theoretical estimates of the GBs of peptides. The overall objective of this work is to increase our understanding of the protonation process for small peptides using structural and thermodynamic data obtained from experimental and ab initio research.

Experimental Methods

The experimental procedures used in obtaining gas-phase basicities by deprotonation reactions have been discussed previously.[1-3] All experiments were performed with a Bruker CMS-47X Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR).[27] Protonated peptide ions were produced in an external source using a Phrasor Scientific FAB gun[28] which employs a 6−10 kV beam of xenon atoms and ions. The peptides were dissolved in a glycerol matrix that contained trifluoroacetic acid to increase the abundance of protonated ions.

Electrostatic focusing was used to transfer ions from the external source to the FT-ICR cell. The ions were then collisionally cooled[29] by admitting a pulse of argon into the cell at a maximum pressure of 10^{-5} torr, followed by a 0.5 s collision period. The protonated peptide ions (MH^+) were mass-selected by resonant frequency ejection techniques[30] (with care taken not to cause translational excitation to the selected ions) and allowed to react with the reference compounds present at static pressures of $(2-20)\times10^{-8}$ torr. The overall rate constants were determined by observing the pseudo-first-order change in reactant ion intensity as a function of time at a constant pressure; these first-order decay plots were linear which suggests the predominance of ground state reactant ions. In cases where protonation was in competition with proton-bound dimer formation,[1] the rate constant for the protonation pathway (k_{proton}) was obtained from a plot of the relative protonated reference base intensity as a function of time using standard kinetics relationships.[31] Pressures were obtained with a calibrated ionization gauge.[1,3] Reported reaction efficiencies are the ratio of the experimental deprotonation rate constant (k_{proton}) to the collision rate constant that was obtained using the average dipole orientation model[32] (k_{ADO}). The FT-ICR cell was maintained at room temperature (*ca.* 298 K) during all experiments.

Computational Methods

Ab initio MO calculations[33] employed the Gaussian 92 program[34] on the IBM ES/9121/480 and DEC 4000-710 AXP computers at Miami University and the CRAY Y-MP8/864 computer at the Ohio Supercomputer Center. To find the most stable structures of the neutral and amino N-protonated peptide species, geometries were optimized with the 3-21G basis first and the 6-31G* basis next. The optimized 3-21G geometries were used to initiate searches for the optimized 6-31G* geometries. Each minimum-energy structure was verified by its all real harmonic vibrational frequencies in the case of HF/3-21G and its uniformly positive eigenvalues of the Hessian matrix in the case of HF/6-31G*.[34]

These geometry optimizations led to two sets of the twelve most stable structures of the peptide species and their corresponding electronic energies at the respective 3-21G//3-21G and 6-31G*//6-31G* levels. For purposes of comparative studies, single-point HF/6-31G* calculations at the HF/3-21G geometries were also performed to obtain the 6-31G*//3-21G electronic energies. In addition, zero-point energies and thermodynamic proper-

(22) (a) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J. L.; Marcoccia, J. F.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1991**, *113*, 6256. (b) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1991**, *113*, 7917. (c) Csaszar, A. G. *J. Am. Chem. Soc.* **1992**, *114*, 9568. (d) Hu, C. H.; Shen, M.; Schaefer, H. F., III. *J. Am. Chem. Soc.* **1993**, *115*, 2923.

(23) (a) Head-Gordon, T.; Head-Gordon, M.; Frisch. M. J.; Brooks II, C. L.; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989. (b) Gould, I. R.; Kollman, P. A. *J. Phys. Chem.* **1992**, *96*, 9255. (c) Böhm, H. J.; Brode, S. *J. Am. Chem. Soc.* **1993**, *115*, 6152. (d) Schäfer, L.; Newton, S. Q.; Cao, M.; Peeters, A.; van Alsenoy, C.; Wolinski, K.; Momany, F. A. *J. Am. Chem. Soc.* **1993**, *115*, 272. (e) Gould, I. R.; Cornell, W. D.; Hillier, I. H. *J. Am. Chem. Soc.* **1994**, *116*, 9250.

(24) Wright, L. R.; Borkman, R. F. *J. Phys. Chem.* **1982**, *86*, 3956.

(25) (a) Wright, L. R.; Borkman, R. F. *J. Phys. Chem.* **1982**, *86*, 3951. (b) Bouchonnet, S.; Hoppilliard, Y. *Org. Mass Spectrom.* **1992**, *27*, 71. (c) Jensen, F. *J. Am. Chem. Soc.* **1992**, *114*, 9533.

(26) Williams, R. W.; Kalasinsky, V. F.; Lowrey, A. H. *J. Mol. Struct. (THEOCHEM)* **1993**, *281*, 157.

(27) Kofel, P.; Allemann, M.; Kellerhals, H. P.; Wanczek, K. P. *Int. J. Mass Spectrom. Ion Proc.* **1985**, *65*, 97.

(28) Perel, J.; Faull, K.; Mahoney, J. F.; Tyler, A. N.; Barchas, J. D. *Am. Lab.* **1984**, Nov, 94.

(29) Ahmed, M. S.; Dunbar, R. C. *J. Am. Chem. Soc.* **1987**, *109*, 3215.

(30) Comisarow, M. B.; Grassi, V.; Parisod, G. *Chem. Phys. Lett.* **1978**, *57*, 413.

(31) Espenson, J. H. *Chemical Kinetics and Reaction Mechanisms*; McGraw-Hill: New York, 1981; p 55.

(32) Su, T.; Bowers, M. T. *Int. J. Mass Spectrom. Ion Phys.* **1973**, *12*, 347.

(33) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab initio Molecular Orbital Theory*; Wiley: New York, 1986.

(34) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. A.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 92*, Gaussian: Pittsburgh, 1992.

ties were obtained at the 3-21G//3-21G level for calculations of ideal-gas proton affinities and basicities at 298.15 K and 1 atm.

Following the recommended procedure from a previous protonation study of small molecules by Del Bene and Shavitt,[35] two model methylamine species were treated with geometry optimizations and frequency determinations at the MP2/6-31+G** level and correlation corrections at the MP4/6-31+G(2d,2p)//MP2/6-31+G** level. Full calculations in the Møller–Plesset (MP) perturbation approach were performed, i.e., MP2 = FULL and MP4 = FULL.[34]

Results and Discussion

Experimental Gas-Phase Basicities from Deprotonation Reactions. The gas-phase basicities of Ala, GlyAla, AlaGly, AlaAla, GlyGlyAla, GlyAlaGly, and AlaGlyGly were determined by the proton transfer reactions of protonated molecular ions (MH^+) with a series of neutral reference compounds (B) of known GB, reaction 1.

$$MH^+ + B \rightarrow M + BH^+ \qquad (1)$$

(The L-form of alanine was used in all studies.) The absence of reaction 1 gives a lower limit on the GB of M, indicating that the reference compound is less basic than the amino acid or peptide. The presence of reaction 1 gives an upper limit on the basicity of M. Typical mass spectra are shown in Figure 1 for the reaction of AlaGlyH$^+$ with *sec*-butylamine. By monitoring this reaction as a function of time, the proton transfer process was determined to have a rate constant of 9.3×10^{-10} cm^3 molecule^{-1} s^{-1} and a reaction efficiency of 0.73 (i.e., 73% of collisions with the reference compound result in proton transfer). This high efficiency indicates an exoergic process and brackets GB(AlaGly) < GB(*sec*-butylamine). The spectra also reveal a minor amount of proton bound dimer (MHB^+) formation. In the present study, dimer formation is the only process that was observed other than deprotonation; it was dependent on both reaction thermodynamics and pressure, as was noted in our previous study of polyglycines.[1]

In Table 1, the reference compounds, their GBs, and the measured reaction efficiencies for reaction 1 are shown. An in-depth discussion of the factors involved in these measurements has been presented elsewhere.[1] Gas-phase basicities were assigned at a 0.10 reaction efficiency level, which generally represents a sharp transition between the occurrence of slow endoergic reactions and fast exoergic reactions.[1,3,36] At least three compounds of higher and lower basicity were reacted with each analyte to insure proper determination of this break between slow and facile reactions.

A minor anomaly is seen in the GlyGlyAla and GlyAlaGly data, with pyridine (GB = 219.8 kcal/mol[37]) deprotonating these tripeptides with a lower efficiency than the less basic *sec*-butylamine (GB = 218.8 kcal/mol[38]). To check the basicity ordering of these two reference compounds, equilibrium proton transfer reactions were performed with gaseous *sec*-butylamine and pyridine both present in the FT-ICR cell. These experiments revealed that pyridine is definitely more basic than *sec*-butylamine; thus, the ordering in the literature[38] is correct. It is possible that the near thermoneutral tripeptide reactions with pyridine may have been slowed slightly by steric hindrance. Steric effects have been found to lower efficiencies in the proton transfer reactions of alkyl pyridines.[39] In addition, as we have noted previously for polyglycines,[1] the break between exoergic and endoergic reactions becomes less clear as the size of the peptide ion increases. This may be attributed to steric effects with the reference compounds experiencing decreased access to the protonation site as the size of the peptide increases. Our recent ab initio calculations[2] of GlyGlyGlyH$^+$ have shown that the most stable structure involves extensive intramolecular hydrogen bonding that essentially folds the peptide backbone around the N-terminal protonation site. It is highly probable that GlyGlyAla, GlyAlaGly, and AlaGlyGly have similar, bulky structures.

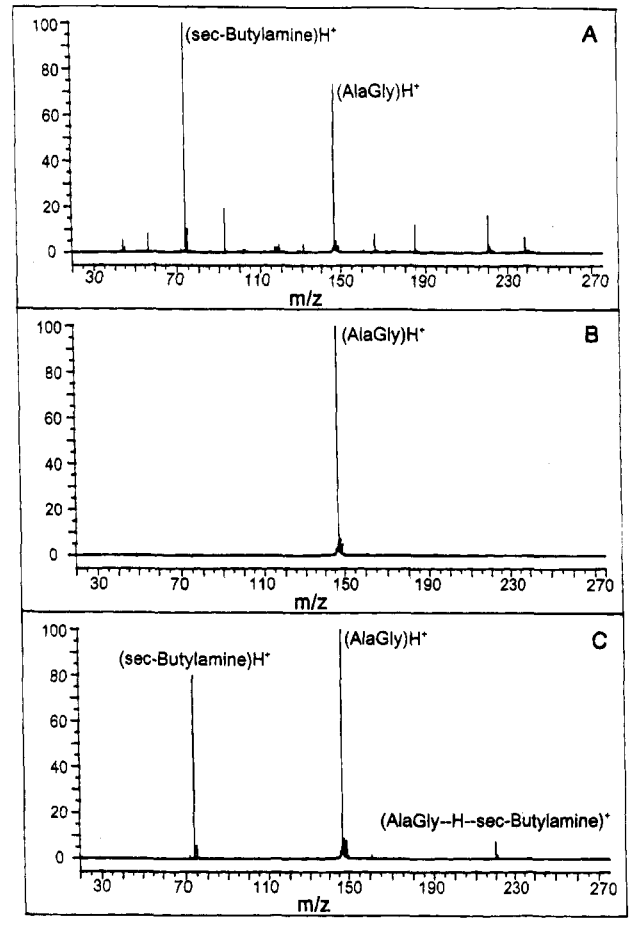

Figure 1. FT-ICR mass spectra for the reaction of (AlaGly)H$^+$ with *sec*-butylamine at a static pressure of 3.4×10^{-8} Torr. (a) All ions produced by FAB on a solution of AlaGly in glycerol. (b) Isolation of (AlaGly)H$^+$ by swept frequency ejection techniques. (c) Reaction of (AlaGly)H$^+$ and *sec*-butylamine for 0.6 s.

Table 2 contains a summary of the experimental and theoretical GBs and proton affinities (PAs) obtained in this study. "Experimental" PAs were obtained by adding the $-T\Delta S$ term for protonation obtained by ab initio

(35) Del Bene, J. E.; Shavitt, I. *J. Phys. Chem.* **1990**, *94*, 5514.
(36) Büker, H.-H.; Grützmacher, H.-F. *Int. J. Mass Spectrom. Ion Proc.* **1991**, *109*, 95.
(37) Moet-Ner, M.; Sieck, L. W. *J. Am. Chem. Soc.* **1991**, *113*, 4448.
(38) Lias, S. G.; Liebman, J. F.; Levin, R. D. *J. Phys. Chem. Ref. Data* **1984**, *13*, 695. All GBs from this compilation which are used in the present paper have been adjusted to the scale of ref 37.
(39) (a) Houriet, R.; Rolli, E. *New J. Chem.* **1987**, *11*, 221. (b) Meot-Ner, M.; Smith, J. C. *J. Am. Chem. Soc.* **1991**, *113*, 862.

Table 1. Reaction Efficiencies for the Proton Transfer Reactions of Protonated Biomolecules with Reference Compounds

reference compound	GB, kcal/mol[a]	reaction efficiency						
		Ala	GlyAla	AlaGly	AlaAla	GlyGlyAla	GlyAlaGly	AlaGlyGly
pyrrole	206.1	0.01	0.01	0.01	*b*	–	–	–
2-fluoropyridine	207.9	0.01	–	–	–	–	–	–
N,N-dimethylformamide	209.5	0.04 break[c]	0.00	0.00	–	–	–	–
3-methylaniline	211.5	0.17	0.02	0.01	–	–	–	–
3-fluoropyridine	212.9	0.81	0.03	0.02	–	–	–	–
allylamine	214.8	1.06	0.03	0.01	–	–	–	–
ethylamine	214.9	–	0.03	0.01	0.01	0.03	0.03	0.01
4-acetylpyridine	216.0	–	0.09 break	0.05	0.03	0.04	0.04	0.02
n-propylamine	217.0	–	0.24	0.08 break	0.03	0.07	0.07	0.07
isopropylamine	218.1	–	0.44	0.17	0.04 break	0.07 break	0.07 break	0.04
sec-butylamine	218.8	–	0.93	0.73	0.47	0.14	0.13	0.08
pyridine	219.8	–	1.09	0.95	0.51	0.07	0.08	0.06 break
cyclohexylamine	220.6	–	1.36	1.22	0.79	0.24	0.19	0.25
3-methylpyridine	223.7	–	–	–	1.16	0.49	0.47	0.49
diethylamine	225.3	–	–	–	–	0.56	0.54	0.67
1-methylpiperidine	229.5	–	–	–	–	0.85	0.55	0.68

[a] Reference compound GBs were obtained from refs 37 and 38. Those from ref 38 were adjusted to the scale of ref 37. [b] The "–" indicates that no experiment was performed. [c] The word "break" signifies the break between slow and facile proton transfer for each amino acid or peptide.

Table 2. Ab Initio and Experimental Gas-Phase Basicities and Proton Affinities of Monomers and Small Peptides[a]

	ab initio			
	3-21G// 3-21G	6-31G*// 3-21G	6-31G*// 6-31G*	experimental
A. Gas-Phase Basicity (GB in kcal/mol)				
Gly[b]	214.5	208.1	207.7	207.0 ± 3.1
Ala	220.1	211.5	212.2	210.4 ± 3.0
GlyGly[b]	224.6	214.3	214.7	214.8 ± 2.3
GlyAla	225.5	215.3	215.7	216.1 ± 2.5
AlaGly	227.7	217.6	218.1	217.2 ± 2.6
AlaAla	228.6	218.6	219.0	218.2 ± 2.4
GlyGlyGly[c]	229.3	218.8	219.6	218.5 ± 2.4
GlyGlyAla	–[d]	–	–	218.5 ± 2.4
GlyAlaGly	–	–	–	218.5 ± 2.4
AlaGlyGly	–	–	–	220.0 ± 2.4
B. Proton Affinity (PA in kcal/mol)[e]				
Gly	223.2	216.8	216.3	215.7
Ala	227.8	219.3	219.9	218.1
GlyGly	233.2	222.9	223.3	223.4
GlyAla	233.9	223.7	224.1	224.5
AlaGly	236.1	225.9	226.4	225.5
AlaAla	236.8	226.7	227.2	226.3
GlyGlyGly	237.7	227.1	227.9	226.8
GlyGlyAla	–[d]	–	–	226.8
GlyAlaGly	–	–	–	226.8
AlaGlyGly	–	–	–	228.3

[a] GB and PA are at 298.15 K and 1 atm. [b] Experimental Gly and GlyGly data are from ref 1. [c] Experimental and ab initio GlyGlyGly data are from ref 2. [d] Ab initio calculations were not performed on the tripeptides. [e] Experimental PA is deduced from eq 5 using the experimental GB term in part A and the ab initio $-T\Delta S$ term in Table 3. For all tripeptides, the ab initio $-T\Delta S$ term of triglycine from ref 2 (8.34 kcal/mol) was used.

calculations to the experimentally-derived GB value. The uncertainties listed with the experimental GB values include a standard 2.0 kcal/mol to account for errors in the literature basicity value of the reference compounds *plus* the difference in GB between the biomolecule and the reference compounds which bracket it. These uncertainties relate to the assignment of absolute GB values and not to the relative order of basicities for the compounds involved in this study.

The GB values of the reference compounds were obtained from the experimental gas-phase basicity and proton affinity scale of Meot–Ner and Sieck.[37] This scale assigns ammonia a GB of 199.7 kcal/mol and a PA of 208.3 kcal/mol. However, a recent study by Szulejko and McMahon[40] suggests that this scale is too high and that the earlier GB and PA ladder compiled by Lias *et al.*[38] is too low. Therefore, it is not clear as to which GB scale should be utilized in the present work. We have chosen to use the Meot-Ner[37] scale primarily because it was utilized in our earlier polyglycine work.[1,2] Our GB values from Tables 1 and 2 can be easily adjusted to any scale because their placement relative to the GBs of the reference compounds will remain the same.

Ab Initio Geometries and Energies. To obtain theoretical GBs the first step is to find the most stable structures for the neutral and protonated species of the molecules of interest. The HF/6-31G* minimum-energy structures for the monomers (Gly, $GlyH^+$, Ala, $AlaH^+$) are shown in Figure 2; those for the dipeptides (GlyGly, $GlyGlyH^+$, GlyAla, $GlyAlaH^+$, AlaGly, $AlaGlyH^+$, AlaAla, $AlaAlaH^+$) are presented in Figure 3. Intramolecular hydrogen bonding (XH···Y) with bonding distance (H···Y) below 2.3 Å is indicated by a dotted line and its distance given in angstroms . The atomic cartesian coordinates for the optimized HF/3-21G and HF/6-31G* geometries of the twelve species are provided as supplementary material in Tables S-I through S-IV. These ab initio geometries may be used for force-field parameter development in molecular modeling.[41]

Detailed results of the Hartree–Fock calculations for the molecular species are shown in Table S-V (supplementary material). These include SCF energies (E_{SCF} in hartrees) obtained from direct geometry optimizations at the HF/3-21G and HF/6-31G* levels and the single-point

(40) Szulejko, J. E.; McMahon, T. B. *J. Am. Chem. Soc.* **1993**, *115*, 7839.

(41) See, for example, Van Alsenoy, C.; Cao, M; Newton, S. Q.; Teppen, B.; Perczel, A.; Csizmadia, I. G.; Momany, F. A.; Schäfer, L. *J. Mol. Struct. (THEOCHEM)* **1993**, *286*, 149, and references cited therein.

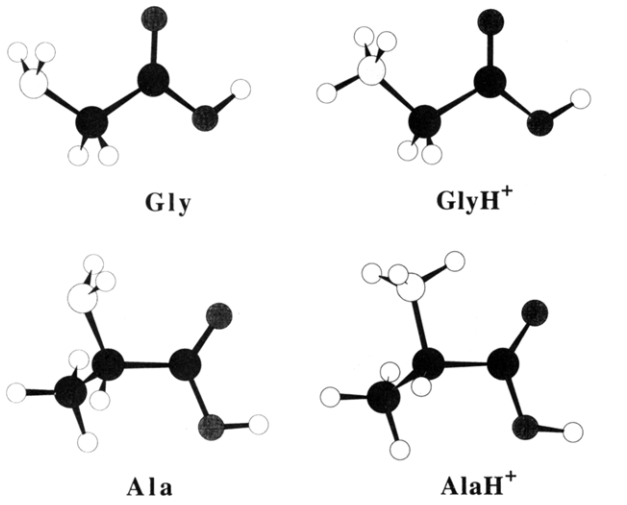

Figure 2. HF/6-31G* minimum-energy structures for the neutral and protonated glycine and alanine. Atoms are identified by shading (H, none; C, dark; N, light; O, medium) or size (C > N > O > H).

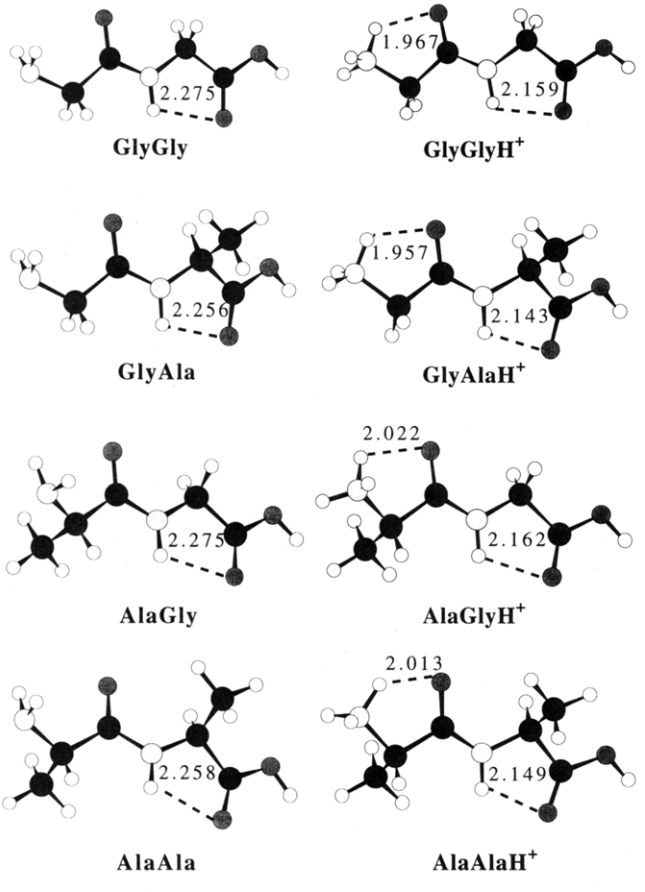

Figure 3. HF/6-31G* minimum-energy structures for the neutral and protonated dipeptides of glycine and alanine. Atoms are identified by shading (H, none; C, dark; N, light; O, medium) or size (C > N > O > H). Intramolecular hydrogen bonds are shown with dotted lines; bond distances are given in angstroms.

HF/6-31G* calculations over the HF/3-21G optimized geometries. The relative SCF energies (ΔE_{SCF} in kcal/mol), which refer to the energies of the protonated species relative to their respective neutral species, are listed in Table 3. Note the substantial differences between the HF/3-21G values (3-21G//3-21G) from the HF/6-31G* values over different geometries (6-31G*//3-21G and 6-31G*//6-31G*). Of particular importance is the great similarity between the two sets of HF/6-31G* values; this suggests that 3-21G geometries closely resemble the 6-31G* geometries.

Table 3. Ab Initio Relative Electronic and Thermal Energies for Monomers and Dipeptides[a]

	SCF energy			thermal energy	$-T\Delta S$
	3-21G// 3-21G	6-31G*// 3-21G	6-31G*// 6-31G*	3-21G// 3-21G	3-21G// 3-21G
Gly	−230.171	−223.758	−223.326	9.362	8.66
Ala	−235.200	−226.684	−227.307	8.862	7.79
GlyGly	−240.531	−230.265	−230.651	9.701	8.65
GlyAla	−241.179	−231.007	−231.398	9.427	8.43
AlaGly	−243.373	−233.194	−233.710	9.347	8.34
AlaAla	−243.985	−233.926	−234.413	9.085	8.14

[a] All quantities are in kcal/mol. The SCF and thermal energies are ΔE_{SCF} and ΔE_{therm} (scaled) for M → MH^+, and $-T\Delta S$ is for M + H^+ → MH^+, where M refers to the molecule listed. See text for details.

In the previous HF/6-31G* study on the diglycine system,[1] GlyGly and $GlyGlyH^+$ have conformations slightly different from those in Figure 3 and energies higher by 0.68 and 0.57 kcal/mol from those in Table S-V. Considering that the structures reported in this work are lower in energy, they replace the previous structures as representatives of the global energy minima. Existence of a large number of conformations with nearly equal energy to the global minimum energy is expected of the peptide systems.[1,2,22,23,42]

Hydrogen Bonding from Ab Initio Calculations. The hydrogen bonding in the four monomer species shown in Figure 2 is of the bifurcated type ($(NH_2$···O), i.e., double 1,5 interactions between two N-terminal H atoms at about equal distances from the carboxylic carbonyl O atom. The H···O bonding distance ranges from 2.5 in $AlaH^+$ to 2.8 in Gly. As discussed before,[1] these nonbonded interactions are weaker than those with shorter H···O separations, e.g., those below 2.3 Å shown explicitly on the diagrams. Although protonation at the terminal N atom in forming the ammonium group ($-NH_3^+$) increases the (Mulliken) charges on the ammonium H atoms (e.g., from +0.35 e in Ala to +0.48 e in $AlaH^+$), the attraction between the nonbonded H and O atoms still seems not strong enough to support significant torsions of the $-NH_3^+$ and C=O groups about the C−N bond to gain a shorter H···O distance and a stronger H-bond.

The eight dipeptide structures in Figure 3 display amazing similarity and regularity in their conformations. Hydrogen bonding is of the (NH···O) variety in a nearly planar five-membered cyclic (C_5) form. There are two types based on their locations: C_5−C, the bond between the C-terminal carboxylic carbonyl O and the amide H; C_5−N, the bond between the N-terminal ammonium H and the amide carbonyl O. Note in the neutral species only C_5−C exists, whereas in the protonated species both C_5−C and C_5−N are present. The H-bonding distance for C_5−C is around 2.26−2.28 Å in the neutral species but decreases to 2.14−2.16 Å in the protonated species. The bonding distance for C_5−N in the protonated species is considerable shorter, around 2.0 Å, inferring a stronger H-bond.

(42) Frey, R. F.; Coffin, J.; Newton, S. Q.; Ramek, M.; Cheng, V. K. W.; Momany, F. A.; Schäfer, L. *J. Am. Chem. Soc.* **1992**, *114*, 5369.
(43) Grev, R. S.; Janssen, C. L.; Schaefer, H. F., III. *J. Chem. Phys.* **1991**, *95*, 5128.
(44) Zhang, K.; Chung-Phillips, A. Unpublished results.

A comparison of the geometries of the neutral and protonated dipeptide species in Figure 3 suggests the following changes upon protonation. First, small torsions around the C–N bond allow one ammonium H atom to get closer to the amide carbonyl O to form the C_5–N ring (e.g., 2.022 for H···O in AlaGlyH$^+$). Second, the NH···O interaction disrupts the resonance in the amide group, O=C–NH ↔ $^-$O–C=N$^+$H, in favor of the $^-$O–C=N$^+$H structure to effect a shortening of the CN peptide bond (e.g., from 1.349 Å in AlaGly to 1.322 Å in AlaGlyH$^+$). Third, the above changes influence the geometry of the C_5–C ring, H–N–C–C=O, in gaining small increases of the C–N and C–C bondlengths and small decreases of the HNC and CCO angles to achieve an overall small decrease in the H···O distance (e.g., from 2.275 Å in AlaGly to 2.162 Å in AlaGlyH$^+$). These synergistic events bring about sufficient stabilization for the formation of the N-terminal H-bond (the C_5–N ring) against the steric resistance to the deformation of staggered conformations preferred originally by the three N–H bonds in $-NH_3^+$.

The repetitive nature of each type of H-bond among different dipeptide species implies that the presence of the methyl group as a side chain in the alanine residue has no significant effect on the pattern of hydrogen bonding. This is reasonable in view of the relatively short backbone of a dipeptide and the location of methyl group being noninterfering with the C_5 ring form of these specific H-bonds.

The H-bonds of the C_5 conformation shown here are characteristic of small peptides and protonated peptides in gas-phase. Yet, they can be of general interest in protein studies. For example, the C_5–C ring is roughly analogous to the C_5 bonding in a protein chain between an amide carbonyl O in one peptide unit and an amide H in the neighboring peptide unit. The general characteristics of the C_5–C rings here are comparable to those found in model dipeptides.[23] Likewise, the C_5–N interaction between the terminal ammonium group ($-NH_3^+$) and the first amide oxygen (C=O) is a plausible model of internal H-bonding in a zwitterion in solution.[26]

The above discussion on H-bonding (Figures 2 and 3) is based on conformations generated from Hartree–Fock calculations using medium-sized bases such as the 3-21G and 6-31G* sets in this work and the comparable sets in most previous works.[1,2,22–26] For an amino acid or peptide, it is well known that numerous stable conformations, differing slightly in energy, exist in the low-energy region of the potential energy surface. When geometry optimizations are carried out at a level more advanced than HF/6-31G*, for example, relative stabilities of different conformers may change.[42] Changes become more likely when correlation effects are partially accounted for and polarization functions on the hydrogen atoms are added.[42] Therefore, the conformations in Figures 2 and 3 could be slightly modified as new calculations at more advanced levels become available.

Theoretical Gas-Phase Basicities from Ab Initio Calculations. The protonation reaction of a compound M at 298 K and 1 atm corresponds to reaction 2.

$$M + H^+ \rightarrow MH^+ \quad (2)$$

For this reaction the Gibbs free energy may be expressed as[1,2]

$$\Delta G = \Delta E_{elec} + \Delta E_{therm} + 6.28 \text{ kcal/mol} \quad (3)$$

Table 4. Comparisons of Relative Gas-Phase Basicities of Monomers and Dipeptides[a]

	3-21G// 3-21G	6-31G*// 3-21G	6-31G*// 6-31G*	experimental
Gly → GlyGly	10.0	6.2	7.0	7.8
GlyGly → GlyAla	0.9	1.0	1.0	1.3
Ala → AlaGly	7.6	6.1	5.9	6.8
AlaGly → AlaAla	0.9	1.0	0.9	1.0
Gly → Ala	5.5	3.4	4.5	3.4
GlyGly → AlaGly	3.1	3.3	3.4	2.4
GlyAla → AlaAla	3.1	3.3	3.3	2.1

[a] For M1 → M2, relative GB = GB(M2) − GB(M1), in kcal/mol. The GB values are those from Table 2(A).

where ΔE_{elec} and ΔE_{therm} represent changes in the electronic and thermodynamic properties for M → MH$^+$. Specifically, $\Delta E_{elec} = \Delta E_{SCF}$ due to the omission of electron correlation; ΔE_{therm} accounts for the change in zero-point energy (ΔE_{ZP}) and changes associated with internal energy and entropy corrections between 0 and 298 K. The ΔE_{ZP} values are scaled by a factor of 0.91.[43]

Because of the generally similar geometries resulting from the HF/3-21G and HF/6-31G* optimizations, the corresponding ΔE_{therm} values calculated from these two levels are also expected to be similar. Separate studies on numerous compounds (CH_3NH_2, $CH_3CH_2NH_2$, Gly, GlyGly, and others) indicate that their differences are typically below 0.5 kcal/mol.[44] To cut the unnecessarily high cost in vibrational frequency calculations involving the 6-31G* basis,[34] thermal values are computed at the lower level of HF/3-21G. The resulting values for individual thermodynamic terms are presented in Table S-V (supplementary material). Values for the composite ΔE_{therm}, as well as the $-T\Delta S$ term for protonation, are included in Table 3.

Theoretical gas-phase basicity (GB) and proton affinity (PA) of compound M are related to the above quantities as follows:

$$GB = -\Delta G \quad (4)$$

$$PA = GB - T\Delta S \quad (5)$$

Applying the three different sets of electronic energy (3-21//3-21G, 6-31G*//3-21G, 6-31G*//6-31G*) and the same set of thermodynamic quantities (3-21G//3-21G) in Table 3 to eqs 3–5 yield the three sets of theoretical PAs and GBs summarized earlier in Table 2. Note the "experimental" PA values in Table 2(B) are those deduced from experimental GBs in Table 2(A) and the theoretical $-T\Delta S$ in Table 3.

Comparisons of the calculated GB values to the experimental values in Table 2 show that the HF/3-21G values are about 10 kcal/mol too high while the two sets of HF/6-31G* values generally agree to ±1 kcal/mol. As for comparisons of relative basicities in Table 4, i.e., the basicity of one peptide relative to another among the six peptides, all values bear the same positive sign, indicating an overall agreement on which of the two peptides in question is more basic.

Some patterns emerge upon close inspections of the relative GBs in Table 4. For the two sequences, Gly → GlyGly → GlyAla and Ala → AlaGly → AlaAla in the top four rows, where protonations occur uniformly on the same type of amino acid residue (Gly in the first sequence and Ala in the second), the agreement between the calculated and experimental values is generally good to within 1 kcal/mol. However, comparisons for protona-

Table 5. Ab Initio Calculations for the Protonation of Methylamine[a]

	HF/3-21G //HF/3-21G	HF/6-31G* //HF/6-31G*	MP4/6-31+G(2d,2p) //MP2/6-31+G**
electronic energy	−237.002	−228.201	−224.200
thermal energy[b]	10.045	9.986	9.538
$-T\Delta S$	7.552	7.525	7.570
GB	220.7	211.9	208.4
PA	228.2	219.5	216.0
GB (scaled)[b]	221.6	212.9	208.8
PA (scaled)	229.2	220.4	216.3

[a] All quantities are in units of kcal/mol. See Table 3 and text for definitions of symbols. Thermodynamic properties are calculated at the same level as for geometry optimization. [b] For the three respective models from left to right, changes in zero-point energies, ΔE_{ZP}, are 10.240, 10.177, and 9.705 kcal/mol without scaling. After scaling with the factors 0.91, 0.91, and 0.96, respectively, the ΔE_{therm} (scaled) values are 9.123, 9.070, and 9.150 kcal/mol.

tions of different types of residues (Gly vs Ala) in the last three rows, Gly → Ala, GlyGly → AlaGly, and GlyAla → AlaAla, show slightly larger discrepancies between calculated and empirical values in the absolute sense. The relative GBs of these three pairs, following the sequence as written, are 4.5, 3.4, and 3.3 kcal/mol according to 6-31G*//6-31G*, which parallel the values 3.4, 2.4, and 2.1 kcal/mol based on experiments, with a nearly constant gap of about 1 kcal/mol. This implies that both theory and experiment agree that the addition of a methyl group to a glycine residue at the N-terminus (to form an alanine residue) increases the basicity. This increase is more pronounced for the monomer and is attenuated by about 1 kcal/mol when the monomer expands into a dipeptide.

Several deductions may be drawn from comparisons of relative basicities calculated at the three different theoretical levels and those measured by mass spectrometry. At the highest level, 6-31G*//6-31G*, agreement between theory and experiment is very good. The performance of the lowest level, 3-21G//3-21G, is seen to be less consistent than the other two. The middle level, 6-31G*//3-21G, is generally comparable to 6-31G*//6-31G*.

Absolute GBs derived from the same theoretical level are expected to be reduced by roughly the same amount when the level is upgraded to a "top" level that yields more accurate GBs.[1,2] The MP4/6-31+G(2d,2p)//MP2/6-31+G** level has been shown before[35] to yield accurate proton affinities for small molecules and is taken to be the top level here. This theoretical model employs a high order of correction for electron correlation (MP4), a large and balanced basis set [6-31+G(2d,2p)], and a reliable molecular geometry (MP2/6-31+G**). To determine the approximate constants that would correct the GBs from the present HF/3-21G and HF/6-31G* levels to the GBs of the top level, calculations on numerous model compounds were carried out.[44] As protonation of CH_3NH_2 emulates the amino N-protonations of the chosen peptides, the relevant results are shown in Table 5. Indeed, the GB from the top level, 208.4 kcal/mol, is very accurate as it falls between the experimental values of 205.4 kcal/mol obtained with the presumably low Lias basicity scale[38] and 211.1 kcal/mol from the presumably high Meot-Ner basicity scale.[37] Against this "top" value of 208.4 kcal/mol, the correction constants are estimated to be −12.3 kcal/mol for HF/3-21G and −3.5 kcal/mol for HF/6-31G*. Further considerations on the overestimation of zero point energy as a result of neglecting anharmonicity in molecular vibration (for all three models in Table 5) and correlation in electronic motion (for the Hartree−Fock models in particular) lead to possible scaling of E_{ZP} with different factors, i.e., 0.91 for HF/3-21G and HF/6-31G* but 0.96 for MP2/6-31+G**.[33,43] With scaling, the correction constants become roughly −13 kcal/mol for HF/3-21G and −4 kcal/mol for HF/6-31G*; see GB (scaled) in Table 5.

Table 6. Experimental Gas-Phase Basicities from the Literature for Alanine and Its Small Peptides[a]

	Ala	AlaAla	GlyGlyAla	AlaGlyGly
this work	210.4	218.2	218.5	220.0
McIver[b]	213.0	–[c]	–	–
Moet-Ner[d]	209.5	–	–	–
Amster[e]	213.9	–	–	–
Harrison[f]	211.0	–	–	–
Lebrilla[g]	211.2	217.6	–	–
Fenselau[h]	–	221.8	223.9	224.7

[a] All values are in kcal/mol and have been adjusted to the scale of ref 37. In general, experimental errors for all studies are on the order of ±(2−3) kcal/mol. [b] Reference 4. [c] The "–" indicates that no experiment was noted in that reference. [d] Reference 5. [e] Reference 6. [f] Reference 16. [g] Reference 13. [h] Reference 10.

We therefore propose to adjust the GBs of 3-21G//3-21G, 6-31G*//3-21G, and 6-31G*//6-31G* in Table 2 downward by 13, 4, and 4 kcal/mol, respectively, to correct for deficiencies in the basis set, electron correlation, and vibrational anharmonicity inherent in each theoretical model. On the basis of this work, we suggest that the best theoretical estimates for the GBs correspond to those of 6-31G*//6-31G* in Table 2, reduced by 4 kcal/mol each. The GBs and PAs after this reduction will be referred to as the "corrected" ab initio values in following sections.

The corrected ab initio value considered here may be further improved by removing the basis set superposition error (BSSE) and by taking into account conformational equilibria between the most stable conformer and the conformers that are closely above in energy. A previous BSSE study[35] on the protonation of NH_3 showed that the BSSE error was small at the level, MP4/6-31+G(2d,2p)//MP2/6-31+G** in Table 5, which was used to derive the correction constants here. Although a significant population of low-energy conformers was found in an ab initio study for neutral glycine at 298 K,[22b] the overall impact on the theoretical GB based on the lowest-energy conformers of neutral and protonated peptides alone is expected to be small compared with experimental errors on the order of 2−3 kcal/mol (cf. Table 2).

Thermodynamic and Structural Aspects of the Protonation of Alanine. The GB of Ala was experimentally determined to be 210.4 ± 3.0 kcal/mol, which is in good agreement with the corrected ab initio value of 208 kcal/mol. In fact, if the Meot-Ner scale[37] of reference GBs is several kcal/mol too high, as has recently been suggested,[40] our experimental GB value for Ala will be adjusted downward by several kcal/mol and the agreement between experiment and theory may be even better. Table 6 gives experimental GB values from the literature for Ala and several other molecules involved in this study. Our experimental value agrees to within ±0.9 kcal/mol with values from three previous studies,[5,13,16] while the GB values from two other studies[4,6] are 1.6−3.5 kcal/mol higher. These basicity values are clearly more consistent with protonation of an amine rather than a carboxylic acid; the latter falls around 180 kcal/mol.[29] This observation that protonation of Ala involves the terminal amino nitrogen is also consistent

with the lowest energy structure of $AlaH^+$ in Figure 2.

Our experimental results indicate that the GB of Ala is 3.4 kcal/mol higher than that of Gly. For protonated amines, alkyl groups on the carbon α to the nitrogen provide stabilization of the positive ion; a proposed mechanism involves a through-bond charge-induced dipole interaction between the charged site and the polarizable alkyl group.[20,45] For example, this leads to a 1.8 kcal/mol increase in GB between n-propylamine and *sec*-butylamine.[38] The effect induced by the presence of a methyl group appears to occur with alanine. Our HF/6-31G* optimized geometry calculations show a decrease in the total (Mulliken) charge carried by the ammonium group ($-NH_3^+$) on going from $GlyH^+$ (+0.5934 e) to $AlaH^+$ (+0.5919 e). Although this change in the positive charge of the ammonium group is very small (0.0015 e), the direction of change is consistent with the proposed mechanism.[20,45,46] It is more difficult to explain why methyl substitution has a greater impact on basicity for these amino acids than it does on alkylamines without additional ab initio calculations on alkylamines for comparison.

Thermodynamic and Structural Aspects of the Protonation of Glycine and Alanine Dipeptides. The addition of one methyl group to GlyGly leads to a 1.3–2.4 kcal/mol increase in experimental basicity; while significant this increase is not as large as that in going from Gly to Ala. The experimental GBs of GlyAla and AlaGly are 216.1 and 217.2 kcal/mol, respectively, versus 214.8 kcal/mol for GlyGly. The corresponding corrected ab initio values are 212 and 214 versus 211 kcal/mol. The addition of the methyl group at the N-terminus (AlaGly) has a greater impact on basicity than its addition at the C-terminus (GlyAla). These results are consistent with increased stabilization of the positive ion involving an N-terminal methyl group, while a C-terminal methyl group provides lesser stabilization due to its remoteness from the protonation site. The HF/6-31G* charges on the $-NH_3^+$ group in $GlyGlyH^+$ (+0.5757 e), $AlaGlyH^+$ (+0.5673 e), and $GlyAlaH^+$ (+0.5736 e) again support this polarizability model:[20,45] there is a noticeable decrease of positive charge (0.0084 e) when the methyl group resides in close vicinity of the protonation site, whereas the decrease (0.0021 e) is significantly smaller when the methyl group is one residue away.

Adding the second methyl group to produce AlaAla yields another increase of 1.0–2.1 kcal/mol in experimental basicity. Again, the larger increase on going from GlyAla to AlaAla, as compared with a smaller increase from AlaGly to AlaAla, is totally analogous to the changes with the addition of the first methyl group. Also, the calculated charge for the ammonium group in $AlaAlaH^+$ (+0.5651 e) can be used in conjunction with those in $GlyAlaH^+$ and $AlaGlyH^+$ given earlier for similar explanations. Our experimental GB values are in agreement with the results of Gorman and Amster[7] in which divaline (ValVal) was found to be more basic than glycylvaline (GlyVal) and valylglycine (ValGly). However, their work revealed no difference in GB between GlyVal and ValGly,[7] while we have observed a difference in GB between AlaGly and GlyAla in the present study and between serylglycine (SerGly) and glycylserine (GlySer) previously.[3]

Ab initio calculations for the N-protonated dipeptides show intramolecular hydrogen bonding between the terminal nitrogen as the donor and the amide carbonyl oxygen as the acceptor in forming a C_5 ring. The basic geometries of GlyAla, AlaGly, and AlaAla are remarkably similar to that of GlyGly. As Figure 3 illustrates, methyl substitution at either of the two methylene carbons does not interfere with the intramolecular H-bonding inherent in the basic dipeptide geometry of GlyGly. In fact, it is possible that substitution of even larger functional groups will not appreciably impact on this conformation.

Thermodynamic and Structural Aspects of the Protonation of Glycine and Alanine Tripeptides. Due to their complexities, the tripeptides GlyGlyAla, GlyAlaGly, and AlaGlyGly were not subjected to ab initio calculations. The experimental deprotonation reaction results show that the addition of a third amino acid residue to a dipeptide has a less marked increase in basicity than going from a amino acid to a dipeptide. Following the reasoning described above, the basicity of AlaGlyGly fits the trend that methyl substitution near the N-terminus has a greater impact on the basicity than elsewhere on the triglycine chain: our experimental GB for AlaGlyGly is 220.0 kcal/mol while GBs for GlyAlaGly and GlyGlyAla are the same at 218.5 kcal/mol. In fact, our experiments do not reveal any basicity difference between GlyGlyGly, GlyGlyAla, and GlyAlaGly. Our experimental GBs for GlyGlyAla and AlaGlyGly agree to within a few kcal/mol to values recently obtained by Fenselau and co-workers using the kinetic method of metastable dissociation on a proton bound dimer containing the peptide and a reference compound.[10]

The recent ab initio study on triglycine from this laboratory indicates that protonation at the N-terminal amino nitrogen yields the most stable protonated species.[2] Both the neutral GlyGlyGly and the amino N-protonated $GlyGlyGlyH^+$ species have highly folded structures for their lowest energy conformations. Based on these calculated geometries, addition of a methyl group at any of the three methylene sites would in no way interfere with the three H-bonds (in C_5, C_7, and C_{11} ring forms) present in the two structures. (In fact, the addition of even larger functional groups should not interfere.) Therefore, it is reasonable to assume that the most stable structures of neutral and protonated forms of GlyGlyAla, GlyAlaGly, and AlaGlyGly are very similar to those found in the GlyGlyGly study. This conclusion is consistent with the experimental GBs. The experimental values support amino N-terminal protonation because the expected increase in GB is observed when the methyl group is near the charge site (AlaGlyGly). In contrast, a methyl group remote to the charge site (GlyAlaGly and GlyGlyAla) has little impact on stabilizing the ion relative to that provided by the remainder of the peptide chain. The presence of the same basic geometric structure for these tripeptides is supported by the fact that the experimental GBs for GlyGlyGly, GlyGlyAla, and GlyAlaGly are the same and that of AlaGlyGly is only 1.5 kcal/mol higher, indicating that no additional change in GB has occurred as a result of disruption of the H-bonding pattern found in the GlyGlyGly species.

(45) (a) Brauman, J. I.; Blair, L. K. *J. Am. Chem. Soc.* **1968**, *90*, 6561. (b) Brauman, J. I.; Riveros, J. M.; Blair, L. K. *J. Am. Chem. Soc.* **1971**, *93*, 3914.

(46) The use of Mulliken charges is generally adequate for qualitative interpretation [cf. ref 45(b)]. Here Mulliken charges are employed to show the direction of change (an increase or decrease) in the charge distribution as a result of methyl substitution, rather than to give a quantitative assessment of this change.

Concluding Remarks

Deprotonation reactions in an FT-ICR mass spectrometer have been used to determine the gas-phase basicities of alanine and di- and tripeptides of alanine and glycine. In addition, ab initio calculations at the 6-31G*//6-31G* level have confirmed the relative basicities of the monomers and dipeptides. After correcting for basis inadequacy and correlation neglect, the theoretical GBs fall within a few kcal/mol of the experimental values. The 6-31G*//3-21G approach, a far less expensive procedure than 6-31G*//6-31G*, is shown to be adequate for future GB calculations of larger peptides. Both the experimental and the theoretical results indicate that the addition of a methyl group to the polyglycines—to produce alanyl residues—has the greatest impact on basicity when it is placed at the N-terminal residue. In addition, the presence of the methyl group appears to have negligible impact on the lowest-energy structures and does not interfere with the extensive intramolecular hydrogen bonding which is present in the di- and tripeptide structures.

Acknowledgment. The computational work was supported by Miami University through the Miami Computing and Information Services (MCIS) and by a grant from the Ohio Supercomputer Center. We thank Mr. Stephen M. Moore and Mr. Kent A. Covert of MCIS for their valuable assistance with the Gaussian 92 program. Financial support of the mass spectrometry research by the National Institutes of Health (R15-GM47657-01A1) and the Ohio Board of Regents Academic Challenge Program is gratefully acknowledged.

Supplementary Material Available: For the twelve monomer and dipeptide species, Figures S-1 and S-2 on HF/6-31G* minimum-energy structures with atom labels, Tables S-I and S-II on HF/6-31G* atomic cartesian coordinates in bohr, Tables S-III and S-IV on HF/3-21G atomic cartesian coordinates in bohr, and Table S-V on ab initio energy and thermodynamic data (12 pages). This material is contained in libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

JO9418846