On the Stability of Amino Acid Zwitterions in the Gas Phase: The Influence of Derivatization, Proton Affinity, and Alkali Ion Addition

Thomas Wyttenbach, Matthias Witt, and Michael T. Bowers*

Contribution from the Department of Chemistry, University of California at Santa Barbara, Santa Barbara, California 93106

Received July 19, 1999. Revised Manuscript Received February 7, 2000

Abstract: Collision cross sections have been measured for a series of *N*- and *C*-methylated glycines cationized by alkali ions using ion mobility methods. In all cases the measured cross sections are in excellent agreement with model structures obtained from a number of different theoretical approaches. Unfortunately both charge solvation and zwitterion structures are predicted to have nearly identical cross sections. On the basis of a conformational search by molecular mechanics methods and density functional theory calculations at the B3LYP/DZVP level it is found that the lowest energy forms of alkali cationized glycine and alanine are charge solvation structures, whereas lowest energy singly and doubly *N*-methylated glycines are salt bridges independent of metal ion. α-Amino isobutyric acid forms a salt bridge when sodiated and a charge solvation structure when rubidiated. In the most stable charge solvation structures rubidium is bound to one or both carboxyl oxygens, while sodium is bound to both the N- and the C-terminus. The stability of salt bridge structures relative to charge solvation structures is found to be nearly proportional to the amino acid proton affinity (PA). For sodiated molecules a PA of >217 kcal/mol results in salt bridge formation, for rubidiated a PA of >219. Predictions are made for the structural preferences of all the common amino acids as a function of cationizing metal ion.

Introduction

Early on in the history of biochemical research it was realized that controlling the pH is of critical importance for the proper function of biological processes. These studies implied that basic sites are usually protonated and acidic sites usually deprotonated in solution, and hence molecules containing both basic and acidic functional groups, like peptides and proteins, typically end up in a zwitterionic form. Often the active site in an enzyme is of ionic nature or the correct folding of the enzyme is only achieved when certain specific residues are charged. For instance, it has been suggested that isoleucine-16 in the well-studied digestive enzyme chymotrypsin has to be protonated to hold the enzyme chain in the proper shape for it to act as a catalyst and to keep histidine-57 near serine-195.[1]

In a buffered solution enzyme zwitterions are stabilized by solvent molecules, counterions, and in some cases either the substrate or the enzyme itself. In the absence of solvent, zwitterions are far less stable. For instance the glycine zwitterion is calculated to be intrinsically unstable by 20 kcal/mol[2] and gas-phase experiments confirm that glycine is not a zwitterion[3,4] (although it is a zwitterion in solution[5]). The intrinsic stability of a zwitterion (i.e. without solvent) depends on the basicity and acidity of the functional groups involved: the more basic the base and the more acidic the acid the more stable the zwitterion. For example, for arginine with its extremely basic guanidine group calculations indicate the zwitterion form is expected to be intrinsically about as stable as the neutral form.[6]

The basis for the above arguments is primarily theoretical as, unfortunately, unambiguous experimental data are difficult to obtain. Kinetic methods frequently used for this purpose[7] suffer from the inherent problem that the reactions probed might be driven by kinetics rather than by thermodynamics, thus yielding limited information about the structure of the ground state. Collision cross section data for ionic systems, obtained from the ion mobility based ion chromatography technique,[8] are often not conclusive because charge solvation structures assume in many cases an overall shape very similar to that of salt bridge (zwitterionic) structures.[9,10] Spectroscopy is probably the method of choice and it has been successfully applied to

(1) See, for example: Morrison R. T.; Boyd R. N. *Organic Chemistry*; Allyn Bacon: Boston, 1973.

(2) See, for example: (a) Oegerle, W. R.; Sabin, J. R. *J. Mol. Struct.* **1973**, *15*, 131. (b) Tse, Y.-C.; Newton, M. D.; Vishveshwara, S.; Pople, J. A. *J. Am. Chem. Soc.* **1978**, *100*, 4329. (c) Yu, D.; Armstrong, D. A.; Rauk, A. *Can. J. Chem.* **1992**, *70*, 1762. (d) Ding, Y.; Krogh-Jespersen K. *Chem. Phys Lett.* **1992**, *199*, 261. (e) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1995**, *117*, 8159. (f) Truong, T. N.; Stefanovich, E. V. *J. Chem. Phys.* **1995**, *103*, 3709−3717. (g) Nagaoka, M.; Okuyama-Yoshida, N.; Yamabe, T. *J. Phys. Chem.* **1998**, *A102*, 8202−8208.

(3) (a) Suenham, R. D.; Lovas, F. J. *J. Mol. Spectrosc.* **1978**, *72*, 372. (b) Brown, R. D.; Godfrey, P. D.; Storey, J. W. V.; Bassez, M.-P. *J. Chem. Soc., Chem. Commun.* **1978**, 547. (c) Iijima, K.; Tanaka, K.; Onuma, S. *J. Mol. Struct.* **1991**, *246*, 257. (d) Lovas, F. J.; Kawashima, Y.; Grabow, J.-U.; Suenram, R. D.; Fraser, G. T.; Hirota E. *Astrophys. J.* **1995**, *455*, L201.

(4) Locke, M. J.; McIver, R. T., Jr. *J. Am. Chem. Soc.* **1983**, *105*, 4226.

(5) See, for example: (a) Reference 1. (b) McClellan, A. L. *Tables of Experimental Dipole Moments*; Freeman: San Francisco, 1963.

(6) Price, W. D.; Jockusch, R. A.; Williams, E. R. *J. Am. Chem. Soc.* **1997**, *119*, 11988−11989.

(7) (a) Cerda, B. A.; Hoyau, S.; Ohanessian, G.; Wesdemiotis, C. *J. Am. Chem. Soc.* **1998**, *120*, 2437−2448. (b) Price, W. D.; Jockusch, R. A.; Williams, E. R. *J. Am. Chem. Soc.* **1998**, *120*, 3474−3484. (c) Cooks, R. G.; Wong, P. S. H. *Acc. Chem. Res.* **1998**, *31*, 379−386. (d) Caprioli, R. M. *J. Mass Spectrom.* **1999**, *34*, 73. (e) Armentrout, P. B. *J. Mass Spectrom.* **1999**, *34*, 74−78. (f) Cooks, R. G.; Koskinen, J. T.; Thomas, P. D. *J. Mass Spectrom.* **1999**, *34*, 85−92.

(8) (a) Bowers, M. T.; Kemper, P. R.; von Helden, G.; van Koppen, P. A. M. *Science* **1993**, *260*, 1446. (b) Clemmer, D. E.; Jarrold, M. F. *J. Mass Spectrom.* **1997**, *32*, 577−592.

(9) Wyttenbach, T.; von Helden, G.; Bowers, M. T. *J. Am. Chem. Soc.* **1996**, *118*, 8355.

10.1021/ja992546v CCC: $19.00
Published on Web 03/25/2000

both glycine[3] and arginine[11] to show that they are both not zwitterions in the gas phase. Unfortunately spectroscopy cannot be applied to most systems due to low signal intensities, spectral congestion, and ambiguity about the isomeric stability of any observed species relative to those not observed.

Given this array of experimental problems electronic structure calculations become a valuable alternative. While the complexity of the systems makes calculating very accurate absolute stabilities difficult, relative calculated stabilities are much more reliable. Hence, we can ask important questions concerning the role of the size of the cationizing metal ion on isomeric stability as well as the role of the basicity of the N-terminus in determining the onset of zwitterionic character in the amino acid substrate.

In this paper we will report results of electronic structure calculations of sodiated and rubidiated glycine, alanine, and several methyl-substituted derivatives of the form $R_1R_2N-CR_3R_4-COOH$ (R_i = H, CH_3). We will primarily focus on relative energies of different structures as a function of choice of R_i and of alkali ion, but we will also discuss cross sections calculated for those structures and compare them with experimental values.

Experimental Methods

Experimental ion−helium collision cross sections were obtained by measuring the ion mobility in helium, which is in turn determined by the ion drift time in helium for a given drift length and electric field.[12] The technique has previously been described in detail,[13] as has the experimental setup[14] and sample preparation[15] employed here.

Briefly, a pulse of ions is formed by matrix-assisted laser desorption ionization (MALDI), mass selected and injected into a drift cell containing typically 3 Torr of helium. Ions drift through the cell (4 cm length) under the influence of a weak electric field (2.5−25 V/cm) and are subsequently counted in the detection system as a function of their drift time, yielding an ion arrival time distribution. From the arrival time a mobility is determined and from the mobility a cross section.[12] These values can then be compared to cross sections of structures obtained by various theoretical molecular structure calculation methods.

Computational Methods

For each system conformational space was thoroughly scanned with a simulated annealing procedure based on molecular mechanics (MM).[9] Stable geometries thus located were then further optimized by higher level methods. All cationized amino acids were fully geometry optimized at the density functional theory (DFT) B3LYP/DZVP level.[16,17] Many of the sodiated species were in addition optimized using a large 6-311++G** basis set for comparison with the DZVP[17] results. To verify acceptable performance of DFT using B3LYP functionals[16] sodiated glycine was additionally fully optimized at the MP2/6-311++G** level. Software used included the AMBER suite of programs with the standard AMBER force field[18] for MM calculations and Gaussian 94[19] for the MP2 and DFT calculations.

For all of the theoretical structures orientation averaged cross sections were calculated using a Monte Carlo algorithm previously described.[20]

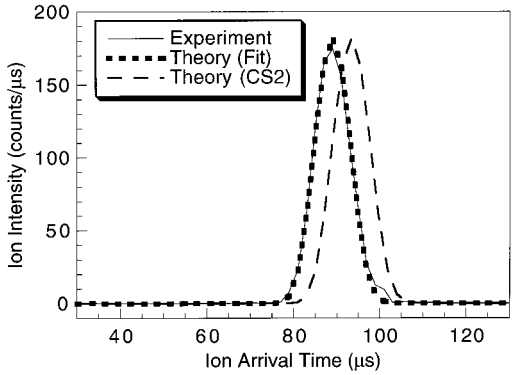

Figure 1. 300 K ion arrival time distribution (ATD) obtained by experiment for sodiated glycine (solid line) and a theoretical fit to the data, yielding a cross section of 49.1 Å^2 (dotted line). Assuming a 5% increased cross section, like that for structure CS2, an ATD shown as the dashed line is expected.

Atomic radii were determined from Lennard-Jones parameters obtained from detailed ion−helium collision studies.[21,22]

Results

All experimental ion arrival time distributions obtained for the systems studied here exhibit one narrow peak with a width that is determined by the original ion pulse width and the ion diffusion in the drift cell. A typical example is shown in Figure 1 for sodiated glycine (solid line). The dotted line indicates a theoretically expected[12] distribution under the assumption that all ions have the same collision cross section of 49.1 Å^2. Good agreement with experiment is obtained in all cases reported here. These results strongly suggest that each system is composed of one set of ions with a characteristic cross section. Those experimental cross sections are listed in Table 1. It can be seen that methylation increases the cross section by 3−6 Å^2 and that the cesiated species are 8−9 Å^2 larger than the corresponding sodiated forms.

A theoretical investigation regarding the geometry of the systems considered here reveals there are four relevant, distinctly different structures: CS1, CS2, CS3, and ZW shown in Figure 2. In agreement with previous studies,[23−26] charge solvation structure CS1 is energetically most favorable for sodiated glycine at all levels of theory used here. The salt bridge structure ZW is less stable by 3.0 (MP2/6-311++G**), 2.8 (B3LYP/6-311++G**), and 2.5 kcal/mol (B3LYP/DZVP), respectively.

(10) Wyttenbach, T.; Witt, M.; Bowers, M. T. *Int. J. Mass Spectrom.* **1999**, *183*, 243−252.

(11) Chapo, C. J.; Paul, J. B.; Provencal, R. A.; Roth, K.; Saykally, R. J. *J. Am. Chem. Soc.* **1998**, *120*, 12956−12957.

(12) Mason, E. A.; McDaniel, E. W. *Transport Properties of Ions in Gases*; Wiley: New York, 1988.

(13) Kemper, P. R.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 197.

(14) von Helden, G.; Wyttenbach, T.; Bowers, M. T. *Int. J. Mass Spectrom. Ion Proc.* **1995**, *146/147*, 349.

(15) Wyttenbach, T.; Bushnell, J. E.; Bowers, M. T. *J. Am. Chem. Soc.* **1998**, *120*, 5098−5103.

(16) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(17) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560.

(18) Pearlman, D. A.; Case D. A. Caldwell, J. C.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman P. A. *AMBER 4.0*; University of California: San Francisco.

(19) Gaussian 94, Revision C2; Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A.; Gaussian, Inc.: Pittsburgh, PA, 1995.

(20) von Helden, G.; Hsu, M.-T.; Gotts, N. G.; Bowers, M. T. *J. Phys. Chem.* **1993**, *97*, 8182.

(21) For an ion composed of 60 atoms the ion−helium Lennard-Jones parameters r_{60} and ϵ_{60} are those reported in ref 22. For smaller ions composed of n atoms the Lennard-Jones radius r_n is scaled down (analogous to ref 15) to account for the decreasing interaction with decreasing ion size using the empirical formula $r_n = r_{60}(0.86882 - 0.99427n + n^{0.99913})$ obtained by a three-parameter fit to hundreds of experimental data points of ions with 11 to 170 atoms. r_{60} = 2.38 Å (H), 3.02 (C, N, O), 2.89 (Na), 3.30 (Rb), 3.50 (Cs); potential well depth not scaled: $\epsilon_n = \epsilon_{60}$ = 0.340 kcal/mol (H), 0.370 (C, O, N), 0.364 (Na), 0.383 (Rb), 0.393 (Cs).

(22) Wyttenbach, T.; von Helden G.; Batka, J. J., Jr.; Carlat, D.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 275.

Table 1. Experimental and Calculated Cross Sections (Å^2) of Sodiated and Cesiated Species

		calcd[b]			
species[a]	exptl	CS1	CS2	CS3	ZW
(gly)Na^+	49.1	49.5	52.0	49.6	48.7
(ala)Na^+	54.1	54.7	57.6	55.1	54.1
(AIB)Na^+	58.8	59.1	61.9	59.9	58.6
(sar)Na^+	55.7	55.4	59.0	56.2	55.2
(DMG)Na^+	61.1	59.7	64.1	61.5	60.0
(betaine)Na^+	63.4				63.7
(gly)Cs^+	57.3	57.5	61.0	57.7	56.6
(ala)Cs^+	63.4	62.3	66.3	63.0	62.3
(AIB)Cs^+	67.6	67.1	70.3	67.8	66.7
(sar)Cs^+	64.6	63.2	67.9	64.5	63.2
(DMG)Cs^+	68.9	66.6	73.0	69.9	67.9
(betaine)Cs^+	71.6				71.6

[a] gly is glycine, ala alanine, AIB α-amino isobutyric acid, sar sarcosine, DMG *N*,*N*-dimethylglycine. [b] The various charge solvation structures CS1, CS2, and CS3 and the zwitterion structure ZW are given in Figure 2.

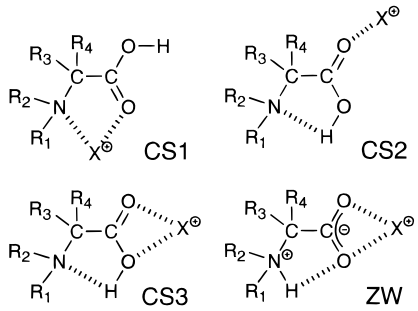

Figure 2. Schematic representations of the charge solvation structures CS1, CS2, CS3 and of the zwitterion structure ZW. R_i = H, CH_3; X^+ = Na^+, Rb^+.

Table 2. Relative Electronic[a] Energies, ΔE_{CS-ZW} (kcal/mol), of Charge Solvation Structures CS1, CS2, and CS3 (see Figure 2) to Salt Bridge Structure ZW Calculated on a B3LYP/DZVP Level for Sodiated and Rubidiated Species[b]

	Na^+					Rb^+		
species[c]	CS1		CS2		CS3	CS1	CS2	CS3
gly	−2.5	(−2.8)	+1.2		+1.4	−2.3	−3.6	−3.7
ala	−0.8	(−1.2)	+3.1	(+1.8)	+2.8	−0.6	−2.1	−2.3
AIB	+1.0		+3.8[d]		+3.8	+1.4	−1.3[d]	−1.3
sar	+3.6	(+3.2)	+6.4	(+5.1)	+6.5	+3.3	+1.1	+1.0
DMG	+7.7	(+7.0)	+10.1	(+8.9)	+10.0	+6.8	+4.2	+4.2

[a] Without zero-point energy correction which is ∼0.6 kcal/mol in favor of charge solvation structures. [b] Values in parentheses were calculated using B3LYP/6-311++G**. Positive values indicate the salt bridge is more stable. [c] gly is glycine, ala alanine, AIB α-amino isobutyric acid, sar sarcosine, and DMG *N*,*N*-dimethylglycine. [d] Converts to CS3.

The spread between ZW and CS1 energy increases slightly by 0.6 kcal/mol (B3LYP/6-311G) when corrected for zero-point energies. Structures CS2 and CS3 are least favorable with energies of ∼4 kcal/mol above the CS1 ground state (Table 2). Single-point calculations of structures between CS2 and CS3 indicate that there is essentially no barrier separating the two isomers. The potential is very flat and energy differences between CS2 and CS3 do not exceed 0.3 kcal/mol in any of the cases studied here. Thus, structures CS2 and CS3 can practically be treated as one and the same minimum on the potential surface with a low-frequency vibrational mode corresponding to an oscillation between CS2 and CS3.[27]

Replacing sodium by rubidium has essentially no effect on the CS1 and ZW stabilities. However, the relative energies of CS2 and CS3 drop by ∼5 kcal/mol for glyRb^+ compared to glyNa^+ and become the most stable geometries, ∼4 kcal/mol below the salt bridge structure ZW. Thus Rb^+ stabilizes charge solvation structures more than Na^+ does.

The situation is very similar for alanine. CS1 is the most stable structure for the sodiated species, while the CS2/CS3 structures are most favorable for the rubidiated form. However, the energy of the salt bridge structure ZW relative to CS1, CS2, and CS3 drops by 1−2 kcal/mol compared to glycine, a result consistent with the fact that the proton affinity (PA) of alanine is ∼3.6 kcal/mol[28] larger than that of glycine.[29] In α-amino isobutyric acid (glycine doubly methylated in the α-position) the effect observed for alanine is approximately doubled, yielding enough stabilization of ZW that it is now the lowest energy structure for the sodiated case. For the rubidiated form CS3 is still somewhat lower in energy than ZW. Thus, α-amino isobutyric acid is a molecule that forms a salt bridge when sodiated and a charge solvation structure when rubidiated.

N-Methylation of glycine has a larger effect on ZW relative stability than *C*-methylation. For both the sodiated and rubidiated form of sarcosine (*N*-methylglycine) ZW is the lowest energy structure in agreement with an older study.[24] The ZW energy drops by ∼6 kcal/mol relative to CS1, CS2,and CS3 upon *N*-methylation of glycine, reflecting the ∼8.3 kcal/mol increased PA of sarcosine over glycine.[28] In *N*,*N*-dimethylglycine the effects observed in sarcosine are amplified by a factor of ∼2. The ZW is now clearly the most stable structure by 7−8 kcal/mol for the sodiated and ∼4 kcal/mol for the rubidiated molecule (Table 2). Again in this example Na^+ is more effective at stabilizing the ZW isomer than Rb^+, an important trend that is observed for all the systems studied here.

Discussion

Potential minima of sodiated glycine have extensively been researched theoretically here and by others.[24−26] All studies agree that the sodiated zwitterion with structure ZW (see Figure 2) is fairly stable, but slightly higher in energy than the lowest energy charge solvation structure CS1 (Figure 2). The second most stable charge solvation structure was determined to be CS2 or CS3. Depending on the methods and basis sets used in the calculations either CS2 or CS3 or sometimes both were found to be stable minima on the potential surface. For glycine bound to the larger alkali ions K^+, Rb^+, and Cs^+ the charge solvation structure CS3 was located as the global minimum.[10,26] In summary, there is good agreement between different theoretical studies that alkali ion cationized glycine is a charge solvation

(23) Structures other than CS1, CS2, CS3, and ZW, that were considered in several theoretical studies on sodiated glycine (refs 24−26) and that turned out to be unfavorable high energy isomers are not located in our simulated annealing procedure, which is biased toward finding relevant low-energy isomers.

(24) Jensen, F. *J. Am. Chem. Soc.* **1992**, *114*, 9533.

(25) (a) Bouchonnet S.; Hoppilliard, Y. *Org. Mass Spectrom.* **1992**, *27*, 71. (b) Yu, D.; Rauk, A.; Armstrong, D. A. *J. Am. Chem. Soc.* **1995**, *117*, 1789.

(26) Hoyau, S.; Ohanessian, G. *Chem. Eur. J.* **1998**, *4*, 1561−1569.

(27) In view of the rapid interconversion between CS2 and CS3 at temperatures above 0 K, the exact location of the potential minimum (CS2 or CS3 or CS2/CS3 double well) is irrelevant in the context of this study; frequency calculations for CS2 and CS3 have not been performed here.

(28) Hunter, E. P.; Lias, S. G. Proton Affinity Evaluation. In *NIST Chemistry WebBook, NIST*; Standard Reference Database No. 69; Mallard, W. G., Linstrom, P. J., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, November 1998 (http://webbook.nist.gov).

(29) This result is in disagreement with an older study (ref 24) on the modest MP2/6-31G*//HF/6-31G* level, where the zwitterion structure was calculated to be more stable than CS1.

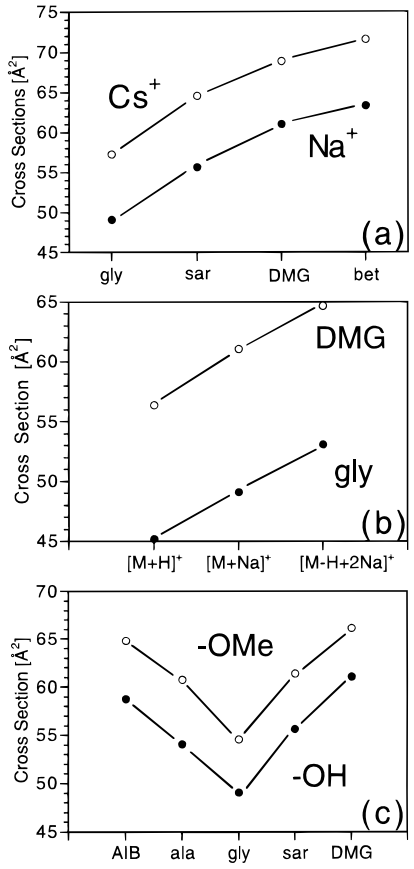

Figure 3. Experimental cross sections of sodiated and cesiated glycine derivatives. gly designates glycine, ala alanine, sar sarcosine, AIB α-amino isobutyric acid, DMG *N*,*N*-dimethylglycine, and bet betaine. (a) Cross section as a function of *N*-methylation (dots sodiated and circles cesiated species). (b) Cross sections for protonated $[M + H]^+$, sodiated $[M + Na]^+$, and deprotonated/doubly sodiated species $[M - H + 2Na]^+$ of glycine (dots) and *N*,*N*-dimethylglycine (circles). (c) Comparison of cross sections of sodiated amino acids (dots) with their methyl esters (circles).

structure and not a zwitterion. Thus, for the following discussion about collision cross sections the glycine system is used as a charge solvation reference.

The large body of experimental cross section data available here can be scanned for trends and abnormalities in an attempt to find a pattern that correlates with structure. An obvious series to examine is that of cationized glycine, sarcosine, *N*,*N*-dimethylglycine, and betaine. On the basis of the discussion in the previous paragraph alkali ion cationized glycine is expected to be nonzwitterionic, whereas betaine has to be a zwitterion. Thus, within this series of *N*-methylated glycines there has to be a transition from charge solvation form to salt bridge. However, cross sections measured for the sodiated and cesiated species increase fairly smoothly within this series (Figure 3a), not giving any clue about where the charge solvation → salt bridge transition might occur. Although this is somewhat discouraging, theory offers a sound explanation for this observation (see below).

Another series to consider is $[M + H]^+$ (charge solvation), $[M + Na]^+$, $[M - H + 2Na]^+$ (salt bridge). Using Tables 1 and 3 it becomes evident, that the $[M + Na]^+$ cross sections are always just about halfway between those of $[M + H]^+$ and $[M - H + 2Na]^+$ (see Figure 3b for glycine and *N*,*N*-dimethylglycine), again not indicating any unusual trend. Also a comparison of the amino acids with their methylesters (Table 4), for which CS1 is by far the most stable structure for all the

Table 3. Experimental and Calculated Cross Sections (Å²) of Protonated $(M + H)^+$ and Deprotonated/Doubly Sodiated $(M - H + 2Na)^+$ Species[a]

	$(M + H)^+$		$(M - H + 2Na)^+$	
species[b]	exptl	calcd	exptl	calcd
gly	45.2	44.5	53.1	53.0
ala	50.4	50.0	57.8	58.0
AIB	54.3	54.6	62.2	62.3
sar	51.3	51.6	59.3	59.2
DMG	56.4	56.5	64.7	63.4
betaine	60.1	60.5		

[a] The corresponding cesiated $(M - H + 2Cs)^+$ species is experimentally not observed. [b] gly is glycine, ala alanine, AIB α-amino isobutyric acid, sar sarcosine, and DMG *N*,*N*-dimethylglycine.

Table 4. Experimental and Calculated Cross Sections (Å²) of Protonated, Sodiated, and Cesiated Methyl Esters

	$(M + H)^+$		$(M + Na)^+$		$(M + Cs)^+$	
species[a]	exptl	calcd	exptl	calcd[b]	exptl	calcd[b]
gly	51.1	51.1	54.6	55.9	62.2	63.7
ala	57.2	56.3	60.8	60.8	69.1	68.4
AIB	61.8	60.6	64.8	65.1	73.8	72.8
sar	57.2	58.0	61.4	61.7	—[c]	—
DMG	62.6	62.7	66.2	66.0	73.6	72.6

[a] Methyl esters of glycine (gly), alanine (ala), α-amino isobutyric acid (AIB), sarcosine (sar), and *N*,*N*-dimethylglycine (DMG). [b] Structure CS1 is the only stable geometry for alkali cationized methyl esters.[10] [c] Low signal level.

systems studied here,[10] shows nothing else but the expected cross section increase with additional methylation (Figure 3c).

At this point it appears that the experimental data alone cannot be used to answer the question whether a particular species forms a salt bridge or not. A comparison with theoretical structures is indispensable. Collision cross sections calculated for the CS1, CS2, CS3, and ZW structures are compiled in Table 1 for comparison with experimental values. It can be seen that there is perfect agreement with experiment for all of the CS1, CS3, and ZW structures within the error limits of theory and experiment. The CS2 structures have a cross section, which is ∼5% larger than experiment and the other isomers. Ions with a 5% larger cross section would give rise to an ion arrival time distribution (ATD) shown as the dashed line in Figure 1 with a clearly shifted peak. From this comparison it is obvious that the CS2 structures are not a major component in the experiment. However, rapid interconversion between CS2 and any of the other more compact structures would yield a narrow ATD shifted by <5%, a possibility that cannot be completely ruled out, especially when excursions to CS2 are of short duration. Nevertheless, CS2 is not a major component in the experiment of any of the sodiated and cesiated glycine derivatives and will therefore be excluded in the following discussion.

The good agreement between experiment and the CS1, CS3, and ZW structures confirms that any of these isomers are good model candidates, validating the theory employed here. However, these comparisons also indicate that the cross section measurements cannot differentiate between the three structures and hence more detailed conclusions have to be based on theory.

The theoretical data presented in the results section indicate that zwitterion stability depends strongly on the proton affinity (PA) of the amino acid in question and to some degree on the choice of alkali ion bound to it. Figure 4 can be used to investigate the PA effect systematically. Plotted is the calculated zwitterionic salt bridge stability as a function of PA, where PA values for glycine, alanine, and sarcosine are taken from the literature[28] and those for α-amino isobutyric acid and *N*,*N*-

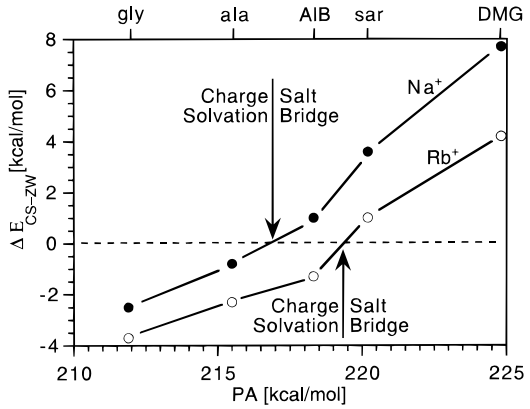

Figure 4. Energy difference ΔE_{CS-ZW} between most stable charge solvation and the salt bridge structure plotted vs proton affinity[28,30] (PA) of the substrate molecule. Solid dots indicate data for sodiated species, circles for rubidiated molecules: gly is glycine, ala alanine, AIB α-amino isobutyric acid, sar sarcosine, DMG *N*,*N*-dimethylglycine.

dimethylglycine are estimated.[30] It can be seen that there exists a fairly linear relationship between the two quantities both for $glyNa^+$ and $glyRb^+$. From this correlation it becomes evident that systems with a PA of >217 kcal/mol form zwitterions when sodiated, but a PA of >219 kcal/mol is required to form rubidiated zwitterions.

This type of correlation can be used to predict the most stable forms of other amino acids. For example, a safe prediction would be that the alkali ion cationized forms of the most basic amino acids histidine, lysine, and arginine are present as zwitterions. A caveat is that amino acids with heteroatoms in the side chain might act to preferentially stabilize charge solvation structures. Although this possibility should not affect our conclusions on arg, lys, and his, we cannot make accurate predictions for amino acids with side chain heteroatoms and a PA in the ~218 kcal/mol range. The model does predict that proline (PA = 220 kcal/mol[28]), phenylalanine (220.6), and isoleucine (219.3) are zwitterions when cationized by sodium and larger alkali ions, whereas valine (217.6) and leucine (218.6) are only zwitterions when cationized by the smaller alkali ions. In Table 5 our structural predictions are summarized for all 20 common naturally occurring amino acids.

The predictions above are in agreement with other calculations available for sodiated proline and cationized arginine.[10,31,32] Blackbody infrared radiative dissociation and collisionally activated dissociation experiments of alkali ion cationized arginine indicate salt bridge structures for all species except the lithiated one.[32] However, metastable ion (MI) mass spectrometry experiments of alkali ion−amino acid−amino acid methylester heterodimers[33] are generally not in good agreement with our predictions. Those experiments indicate as a general trend that larger alkali ions stabilize zwitterions better than smaller alkali ions. Thus, based on the MI data many cesiated species such as alanine, valine, leucine, and isoleucine are thought to be zwitterions and the corresponding species with smaller alkali ions charge solvation structures. Arginine is thought to be in a charge solvation form when lithiated and sodiated and in a salt bridge form when potassiated and cesiated. It should be emphasized, though, that results of kinetic experiments have to be interpreted with caution. The reactions probed may be kinetically driven rather than thermodynamically, therefore yielding limited information about ground-state structures. In the dissociation experiments of the heterodimers mentioned above,[33] the metal ion stays either with the amino acid or with the methyl ester. Since the methyl ester is intrinsically a better charge solvation agent, the metal ion tends to stay with the ester, unless the amino acid−metal ion complex is particularly stable as is the case in the ion−zwitterion complex. However, the fact that the metal ion stays with the amino acid does not necessarily indicate a zwitterion structure. Our calculations indicate that glycine-like compounds form particularly stable charge solvation structures with larger alkali ions (CS3 compared to CS1) making it hard for the methyl ester to compete for the metal ion. Note, that methyl esters cannot form stable CS3 structures and are always in CS1-like conformations.[10] In view of this uncertainty of interpreting the heterodimer MI data[33] we omitted any reference to it in Table 5.

Table 5. Most Stable Structures[a] of Alkali Ion Cationized Amino Acids

amino acid				
side chain characteristics		PA[b] (kcal/mol)	this work[c]	lit[e]
hydrocarbon	gly	211.9	CS	CS[f]
	ala	215.5	CS	ZW[g]
	val	217.6	ZW/CS[d]	
	leu	218.6	ZW/CS[d]	
	ile	219.3	ZW	
	pro	220	ZW	ZW[h]
	phe	220.6	ZW	
heteroatoms	cys	215.9	unclear	CS[h]
	asp	217.2	unclear	
	glu	218.2	unclear	
	ser	218.6	unclear	CS[h]
	thr	220.5	(ZW)	
	tyr	221	(ZW)	
	asn	222	(ZW)	
	met	223.6	(ZW)	
	gln	224.1	(ZW)	
	trp	226.8	(ZW)	
basic	his	236	ZW	
	lys	238	ZW	
	arg	251.20	ZW	ZW[i]

[a] ZW is zwitterion, CS charge solvation structure. [b] Reference 28. [c] Assignments in brackets indicate predictions are less reliable on the bases of PA values and the nature of side chains. [d] ZW when sodiated, CS when cesiated. [e] Results of ref 33 were not included in table due to potential problems with data interpretation (see the text). [f] CS when sodiated (refs 10, 15, 24−26,and 31) through cesiated (ref 26). [g] ZW when sodiated, ref 24. [h] Assignment for sodiated form, ref 31. [i] ZW when sodiated through the cesiated form , refs 10 and 32.

N- and *C*-methylation of alkali ion cationized glycine substantially lowers the ZW energy relative to the charge solvation structures, but there is hardly any energy change between the different charge solvation structures (Table 2). This is demonstrated in Figure 5 for the most basic system considered here, *N*,*N*-dimethylglycine, in comparison with glycine. It can be seen that the energy levels indicated with a solid line (*N*,*N*-dimethylglycine) agree very well with the dotted levels (glycine) for all of the charge solvation structures for both the sodiated and rubidiated systems. The ZW stability, on the other hand, is largely dependent on the amino acid considered.

From both Figures 4 and 5 it is also evident that rubidiated amino acids are less likely to form zwitterions than sodiated amino acids, an effect that has previously been reported for

(30) PA is estimated for α-amino isobutyric acid as PA(alanine) + PA(isopropylamine) − PA(ethylamine) and for *N*,*N*-dimethylglycine as PA(sarcosine) + PA(trimethylamine) − PA(dimethylamine). See ref 28.

(31) Hoyau, S.; Norrman, K.; McMahon, T. B.; Ohanessian, G. *J. Am. Chem. Soc.* **1999**, *121*, 8864−8875.

(32) Jockusch, R. A.; Price, W. D.; Williams, E. R. *J. Phys. Chem. A* **1999**, *103*, 9266−9274.

(33) Wesdemiotis, C.; Cerda, B. A. Proceedings of the 47th ASMS Conference on Mass Spectrometry and Allied Topics, 1999, Dallas, TX.

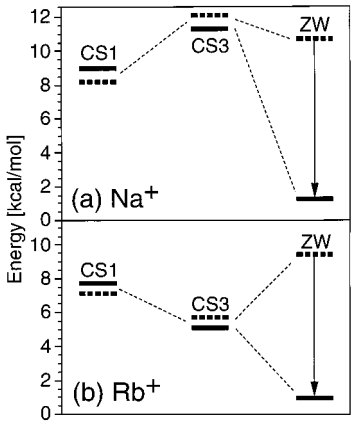

Figure 5. Energy level diagrams for (a) sodiated and (b) rubidiated glycine (dotted levels) and *N*,*N*-dimethylglycine (solid lines). The relative energies of charge solvation structures CS1 and CS3 are nearly independent of *N*-methylation, whereas the zwitterion stability increases dramatically upon *N*-methylation.

glycine[10,26] (but is in disagreement with conclusions drawn from results of kinetic methods experiments[33]). The reason is that the charge solvation structure CS3 becomes particularly stable in the rubidiated form, whereas CS1 and ZW are almost independent of the selection of alkali ion.

Why CS1 and ZW behave the same when the alkali ion is changed is curious. Their structures and hence their interactions with the alkali ions are very different. To investigate this observation, the following calculations were done. The idea is to break down the energy contribution to a certain structure into (a) glycine conformational strain and (b) alkali ion−glycine interaction. Thus, for $glyX^+$ the energy differences associated with the following processes are evaluated:

$$glyX^+ \rightarrow gly^* + X^+ \rightarrow gly + X^+ \qquad (1)$$

where "$glyX^+$" is a geometry optimized structure of alkali ion cationized glycine, "$gly^* + X^+$" is glycine in a conformation identical with that in the $glyX^+$ starting structure but infinitely separated from X^+, and "$gly + X^+$" is geometry optimized glycine infinitely separated from X^+. The energy difference for the first step in (1) is calculated by comparing the energies of the $glyX^+$ structure with that obtained from a single point calculation after removing the alkali ion. This energy difference can be regarded as the alkali ion−glycine binding energy for the corresponding frozen glycine conformation under the assumption that basis set superposition errors (BSSE) are negligible. This can be a rather poor assumption in absolute terms, but any errors should be similar for similar systems,[34−36] and hence relative binding energies should be accurate. But even ignoring correction for BSSE, our calculated sodium dissociation energy D_e for $glyNa^+$ of ∼43 kcal/mol compares very favorably with previous calculations[24−26] (38−45 kcal/mol) and with the experimental sodium−glycine binding enthalpy of $\Delta H^\circ_{298} = 37 \pm 3$ kcal/mol,[37] which is expected to be somewhat smaller than D_e.[31]

Table 6. Calculated Energy Contributions (kcal/mol)

	glycine−alkali ion binding energy[b]			glycine neutral conformational energy[c]		
isomer[a]	$glyNa^+$	$glyRb^+$	ΔE^{bond}_{Na-Rb}	$glyNa^+$	$glyRb^+$	$\Delta E^{strain}_{Na-Rb}$
CS1	−49.3	−12.1	−37.2	6.2	5.1	1.2
CS3	−41.4	−9.9	−31.5	2.2	1.5	0.7
ZW	−62.5	−24.8	−37.7	21.9	20.1	1.8

[a] See Figure 2. [b] Energy required to remove the alkali ion without changing the glycine geometry. BSSE not accounted for (see refs 34−36). [c] Assumes the neutral glycine structure is frozen in its $glyX^+$ conformation with X^+ at infinite separation and cannot relax to the global minimum. All energies are relative to the glycine neutral global minimum of energy 0.0 kcal/mol.

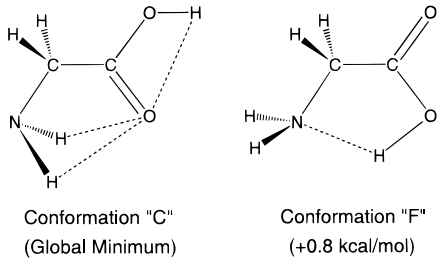

Figure 6. Glycine conformations corresponding to the levels C and F shown in Figure 7. "F" is a local and "C" the global minimum of neutral glycine.[2]

The energy difference for the second step in reaction 1 is determined by a comparison of the single-point energy of gly* with the energy of the global glycine minimum. These values are listed in Table 6 along with the glycine−alkali ion binding energies. First, it can be seen that addition of a rubidium ion perturbs the neutral glycine conformations less (e.g. by 5.1 kcal/mol for CS1) than addition of a sodium ion (6.2 kcal/mol) as expected. A second point to note regarding conformational energy is that the glycine zwitterion is very unstable (by >20 kcal/mol). In fact, on the B3LYP/DZVP level the zwitterion is not a minimum at all and converts into conformation F upon geometry optimization (see Figure 6). This lack of stability of the glycine zwitterion is well-known and has been the subject of several studies in the literature.[2]

The results of our calculations are shown in Figure 7, where three partial reaction coordinate diagrams are shown. The top trace (a) shows energies for glycine in specific conformations of interest with Na^+ and Rb^+ at infinite separation. Diagram b gives energies for $glyRb^+$ for CS1, CS3, and ZW structures whereas diagram c gives the same information for $glyNa^+$. In addition, in part c, energies for two isomerization transition states calculated by Hoyau and Ohanessian[26] are shown.

The labeling (A through H) refers only to the glycine conformation. For example, in part c the energy level label A corresponds to the calculated energy of $glyNa^+$ in the CS1 conformation. In part a the label A corresponds to the calculated energy of glycine frozen in the conformation it has as part of the CS1 $glyNa^+$ complex. The difference in the energies of these two levels corresponds to the first step in reaction 1 for X = Na. The label B has exactly the same meaning for X = Rb. The same interpretation is given for D and E for CS3 and G and H for ZW conformations.

Two other energies are given in part a. These correspond to the global minimum labeled C, which is arbitrarily set as the

(34) The calculated glycine−metal ion binding energies are expected to be too large due to BSSE. However, for DFT methods BSSE are generally considerably smaller than for electron-correlation methods such as MP2 and may be less than 1 kcal/mol for sodiated systems (ref 35) and probably even smaller for larger alkali ions (ref 36). In any case, BSSE are expected to be approximately the same for all of the CS1, CS3, and ZW structures for a given metal ion.

(35) Daza, M. C.; Dobado, J. A.; Molina, J. M.; Salvador, P.; Duran, M.; Villaveces, J. L. *J. Chem. Phys.* **1999**, *110*, 11806−11813.

(36) Hill, S. E.; Glendening, E. D.; Feller, D. *J. Phys. Chem. A* **1997**, *101*, 6125−6131.

(37) Klassen, J. S.; Anderson, S. G.; Blades, A. T.; Kebarle, P. *J. Phys. Chem.* **1996**, *100*, 14218.

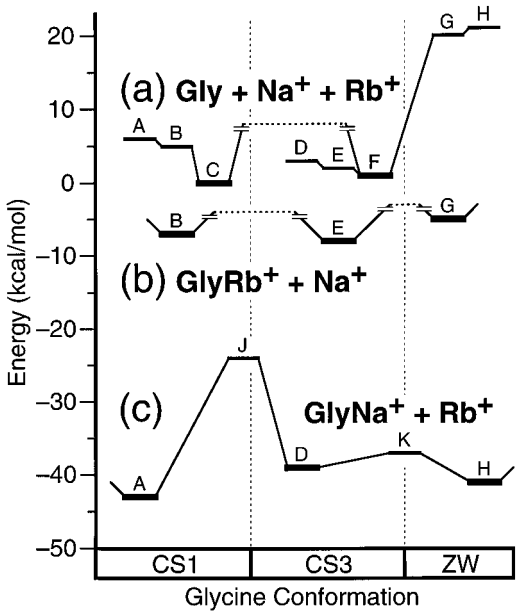

Figure 7. Energy level diagram of (a) neutral, (b) rubidiated, and (c) sodiated glycine. Glycine conformation changes from left to right with CS1-like conformations on the left, CS3 in the middle, and zwitterion (ZW) structures on the right. Bold levels (e.g. C and F) indicate potential minima. Energies are calculated at the B3LYP/DZVP level and are relative to the global minimum C of neutral glycine. Energies for the transition states J and K are from ref 26.

zero of energy, and a second conformation labeled F nearly as low in energy (see Figure 6). The energy difference between A and C in part a is just the energy of the second step in reaction 1, or the so-called glycine strain energy for CS1.

So why do CS1 and ZW behave the same as the metal ion changes even though their conformations are very different? It is the cancellation of two large effects. When you add a Na^+ ion to glycine in conformation CS1, the energy drops by 49.3 kcal/mol. When the same Na^+ ion is added to glycine in the ZW conformation the energy drops by 62.5 kcal/mol. This large difference in stabilization energies is nearly exactly offset by the difference in strain energies of the CS1 and ZW conformations relative to the global minimum: 6.2 and 21.9 kcal/mol, respectively. The net effect is CS1 is stabilized 43.1 kcal/mol by Na^+ and ZW 40.6 kcal/mol making CS1 the global minimum for $glyNa^+$ by 2.5 kcal/mol.

When Rb^+ is substituted for Na^+ the absolute stabilization numbers change dramatically (Table 6) but the *relative* differences are almost identical. The net result is CS1 is 2.3 kcal/mol more stable than ZW for $glyRb^+$, essentially identical with $glyNa^+$. This result appears to be coincidental since the nature of the interaction of the alkali ions with the two structures is so different.

A comparison of the energy changes for steps 1 and 2 in reaction 1 as Na^+ is replaced with Rb^+ is also interesting (ΔE^{bond}_{Na-Rb} vs $\Delta E^{strain}_{Na-Rb}$, Table 6). Values for $\Delta E^{strain}_{Na-Rb}$ are small (but consistently positive) and barely dependent on the structure, indicating that both alkali ions dislocate glycine in a similar way from conformation C with sodium having a slightly larger effect. On the other hand, values for ΔE^{bond}_{Na-Rb} are large and dependent on the structure. Na^+ binds more strongly to glycine than Rb^+, by 37−38 kcal/mol for CS1 and ZW and by 31−32 kcal/mol for CS3. Hence, the relative binding energies are responsible for the fact that the CS3 energy drops relative to CS1 and ZW when replacing Na^+ by a larger alkali ion.

The values for the two isomerization barriers in Figure 7, trace c, for $glyNa^+$ are very different. The isomerization CS3 ⇌ ZW, i.e., D ⇌ H, has a very small barrier[26] consistent with the fact that the only structural change is the migration of the bridging proton (−N···H−O−) ⇌ (−N−H^+···$^-$O−). Hence the value of the barrier *K* is essentially the difference in the stability of the two isomers. This situation mirrors one we previously reported[15] for the analogous isomerization in $(gly)_3Na^+$ where the zwitterion spontaneously reverted to the 5 kcal/mol more stable CS from.

On the other hand, the isomerization CS1 ⇌ CS3 requires substantial rearrangement (Figure 2). If starting from CS1, the X^+−N bond must be broken (forming a CS2 like intermediate) followed by a 180° rotation about the C−C bond. This concerted process takes ∼19 kcal/mol according to the calculations of Hoyau and Ohanessian[26] for $glyNa^+$. Hence glycine and its alkali-cationized congeners has two structural motifs that only very weakly communicate at thermal energies centered about the two low-energy structures given in Figure 6 and designated C and F in Figure 7.

Conclusions

On the basis of B3LYP/DZVP calculations, alkali cationized glycine and alanine form charge solvation structures, whereas singly and doubly *N*-methylated glycine form salt bridges independent of metal ion. Doubly *C*-methylated glycine (α-amino isobutyric acid) forms a charge solvation structure when rubidiated and a salt bridge when sodiated. Good agreement between experimental and theoretical cross sections for all systems reported here generates confidence in the theoretical results.

In all cases the relative energy between the salt bridge structure (ZW) and the charge solvation structure CS1 (where the metal ion is bound to the nitrogen and the carbonyl oxygen) is independent of the choice of alkali ion, which appears to be coincidental. However, the energy of CS3 (where the metal ion is directly bound to the oxygens of the C-terminus) significantly decreases as the alkali ion size increases. The neutral glycine conformation is least distorted in CS3 compared to CS1 and ZW, but the metal ion is also least strongly bound in CS3, an effect that is stronger for smaller alkali ions. As a consequence, sodiated glycine derivatives tend to form more stable salt bridges than the rubidiated molecules.

Finally, the stability of salt bridge structures relative to charge solvation structures increases nearly proportional to proton affinity for the systems studied here. For sodiated molecules a PA of ∼217 kcal/mol is required to make a salt bridge equally stable to a charge solvation structure; for rubidiated amino acids a PA of ∼219 kcal/mol is required. This correlation allows us to predict the structural preference for all 20 alkali ion cationized common amino acids.

Acknowledgment. The support of the National Science Foundation under grant CHE-9729146 is gratefully acknowledged.

JA992546V