J|A|C|S ARTICLES

Published on Web 07/22/2009

Structures of Protonated Dipeptides: The Role of Arginine in Stabilizing Salt Bridges

James S. Prell,[†] Jeremy T. O'Brien,[†] Jeffrey D. Steill,[‡] Jos Oomens,[‡] and Evan R. Williams*,[†]

Department of Chemistry, University of California, Berkeley, California 94720-1460, and FOM Institute for Plasma Physics "Rijnhuizen", Edisonbaan 14, 3439 MN Nieuwegein, The Netherlands

Received March 10, 2009; E-mail: williams@cchem.berkeley.edu

Abstract: Structures of protonated dipeptides containing N-terminal Gly, Val, Pro, Lys, His, or Arg and C-terminal Arg are investigated with infrared multiple photon dissociation (IRMPD) spectroscopy between 900 and 1850 cm^{-1} and theory. The IRMPD spectra clearly indicate that, when Gly, Val, Pro, Lys, or His are N-terminal to Arg, these protonated dipeptides adopt gas-phase structures with a single formal charge site (SCS), whereas ArgArg·H^+ has a salt-bridge (SB) structure in which the C-terminus is deprotonated and two basic sites are protonated. There are only subtle differences in the IRMPD spectra for dipeptides containing Gly, Val, Pro, and Lys. A sharp, intense peak at 1080 cm^{-1} is observed for HisArg·H^+ that is attributed to the neutral histidine side chain, an assignment that is confirmed by comparison to the IRMPD spectrum of $(HisArg·H_2)^{2+}$. Lowest-energy B3LYP/6-31+G(d,p) structures and energies for the SCS and SB forms of these protonated dipeptides indicate that stability of the SB form relative to the SCS form generally increases with increasing gas-phase basicity of the N-terminal amino acid, but only ArgArg·H^+ is calculated to have a SB ground state at 298 K, in agreement with the results from IRMPD spectroscopy. This is the first direct experimental evidence for a salt-bridge structure in a gaseous protonated peptide, and ArgArg·H^+ is the smallest protonated peptide for which a SB structure has been reported. These results suggest that SB structures should be common for protonated peptides containing at least two arginine residues and may also occur for large protonated peptides or proteins with at least one arginine residue and other basic residues, such as lysine or histidine.

Introduction

Ionic interactions play a central role in the chemistry of amino acids, peptides, and proteins. Amino acids are zwitterionic in aqueous solution, and in peptides and proteins, basic and acidic residues can form salt bridges in which the protonated side chain of a basic residue interacts with a deprotonated side chain of an acidic residue. Salt bridge formation can either stabilize or destabilize the folded structure of proteins[1−6] and often occurs in protein−protein interfaces.[7−10] While common in aqueous solution, salt bridges can also occur in the absence of solvent. Although the naturally occurring amino acids are nonzwitterionic in isolation, attachment of a metal cation,[11−33] an electron,[34,35] or water molecules[36−38] can make the zwitterionic forms of some amino acids more stable. Intramolecular salt bridges have

[†] University of California.
[‡] FOM Institute for Plasma Physics "Rijnhuizen".

(1) Anderson, D. E.; Becktel, W. J.; Dahlquist, F. W. *Biochemistry* **1990**, *29*, 2403–2408.
(2) Errington, N.; Doig, A. J. *Biochemistry* **2005**, *44*, 7553–7558.
(3) Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1994**, *3*, 211–226.
(4) Horovitz, A.; Serrano, L.; Avron, B.; Bycroft, M.; Fersht, A. R. *J. Mol. Biol.* **1990**, *216*, 1031–1044.
(5) Lee, W. S.; Park, C. H.; Byun, S. M. *J. Biochem.* **2004**, *135*, 93–99.
(6) Marqusee, S.; Sauer, R. T. *Protein Sci.* **1994**, *3*, 2217–2225.
(7) Li, Y. L.; Urrutia, M.; Smith-Gill, S. J.; Mariuzza, R. A. *Biochemistry* **2003**, *42*, 11–22.
(8) Shimaoka, M.; Xiao, T.; Liu, J. H.; Yang, Y. T.; Dong, Y. C.; Jun, C. D.; McCormack, A.; Zhang, R. G.; Joachimiak, A.; Takagi, J.; Wang, J. H.; Springer, T. A. *Cell* **2003**, *112*, 99–111.
(9) Xu, D.; Lin, S. L.; Nussinov, R. *J. Mol. Biol.* **1997**, *265*, 68–84.
(10) Xu, D.; Tsai, C. J.; Nussinov, R. *Protein Eng.* **1997**, *10*, 999–1012.
(11) Jockusch, R. A.; Price, W. D.; Williams, E. R. *J. Phys. Chem. A* **1999**, *103*, 9266–9274.
(12) Talley, J. M.; Cerda, B. A.; Ohanessian, G.; Wesdemiotis, C. *Chem.−Eur. J.* **2002**, *8*, 1377–1388.
(13) Constantino, E.; Rodriguez-Santiago, L.; Sodupe, M.; Tortajada, J. *J. Phys. Chem. A* **2005**, *109*, 224–230.
(14) Hoyau, S.; Ohanessian, G. *Chem.−Eur. J.* **1998**, *4*, 1561–1569.
(15) Wyttenbach, T.; Witt, M.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 3458–3464.
(16) Strittmatter, E. F.; Lemoff, A. S.; Williams, E. R. *J. Phys. Chem. A* **2000**, *104*, 9793–9796.
(17) Lemoff, A. S.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1014–1024.
(18) Lemoff, A. S.; Bush, M. F.; Williams, E. R. *J. Phys. Chem. A* **2005**, *109*, 1903–1910.
(19) Lemoff, A. S.; Bush, M. F.; Wu, C. C.; Williams, E. R. *J. Am. Chem. Soc.* **2005**, *127*, 10276–10286.
(20) Lemoff, A. S.; Wu, C. C.; Bush, M. F.; Williams, E. R. *J. Phys. Chem. A* **2006**, *110*, 3662–3669.
(21) Wyttenbach, T.; Witt, M.; Bowers, M. T. *Int. J. Mass Spectrom.* **1999**, *182−183*, 243–252.
(22) Kapota, C.; Lemaire, J.; Maitre, P.; Ohanessian, G. *J. Am. Chem. Soc.* **2004**, *126*, 1836–1842.
(23) Dunbar, R. C.; Polfer, N. C.; Oomens, J. *J. Am. Chem. Soc.* **2007**, *129*, 14562–14563.
(24) Forbes, M. W.; Bush, M. F.; Polfer, N. C.; Oomens, J.; Dunbar, R. C.; Williams, E. R.; Jockusch, R. A. *J. Phys. Chem. A* **2007**, *111*, 11759–11770.

10.1021/ja901870d CCC: $40.75 © **2009 American Chemical Society**

also been implicated in the structures of gaseous peptides and proteins,[39−42] as well as in their dissociation,[42−46] and H/D exchange reactivity.[47−49]

Because of the important role that salt bridges play in the reactivity and structures of peptides and proteins both in solution and in the gas phase, much effort has gone into investigating various factors that contribute to salt-bridge stability, including gas-phase basicity of the proton accepting site,[15,21,24−27,50−54] gas-phase acidity of the deprotonation site,[29] molecular orientation,[55] and charge solvation.[25−33,36,55−57] Structural information about cationized amino acids has been inferred from various dissociation energy,[18,54,58,59] ion mobility,[15,21] and H/D exchange[60] measurements. Because arginine is the most basic amino acid, its propensity to form salt bridges is especially important to understand. Attachment of an electron,[34,35] a larger metal cation (Na, K, or Cs, but not Li or Ag),[11,24,25] divalent metal cations (Ba or Sr),[28] other arginine molecules,[61] or a protonated arginine cation[41,58] can result in the zwitterionic form of arginine being most stable. Because zwitterions typically have large dipole moments, they can be stabilized by an electric field from a nearby charge. The nonzwitterionic forms can also be stabilized in charge-solvated structures in which heteroatoms interact with and donate electron density to the nearby charge.

Despite a significant and growing body of evidence that salt-bridge structures can be stable in complexes without any solvent, evidence for the role of stable salt bridges in the structures and reactivities of protonated peptides and proteins is indirect. Salt-bridge intermediates have been proposed to explain backbone cleavages at acidic residues,[46,62] intermolecular proton transfer,[50,51] and in the H/D exchange reactivity at acidic sites in peptides and proteins.[47,48] A stable salt-bridge structure for singly protonated bradykinin, a nonapeptide with arginine residues at each terminus, has been inferred based on blackbody infrared radiative dissociation (BIRD),[42] H/D exchange,[49,63,64] ion mobility measurements,[65] loss of small molecules in ultraviolet photodissociation experiments,[66] and theory.[41,67] A stable salt-bridge structure for higher charge states of multiply protonated ubiquitin is indicated from BIRD measurements to explain facile and entropically favored cleavage of the backbone amide bond near acidic residues.[62]

Because IRMPD spectroscopy has provided some of the most detailed information about salt-bridge formation in a variety of ionic clusters with attached cations,[22−30,36,55,56,68−72] its use for investigating salt bridges in protonated peptides, for which only indirect evidence has been reported, appears to be particularly promising. IRMPD spectroscopy using light generated by a free electron laser was used to study the protonated pentapeptide leucine enkephalin and its dissociation products to elucidate the role of mobile protons.[73] Several small, singly protonated

(25) Bush, M. F.; O'Brien, J. T.; Prell, J. S.; Saykally, R. J.; Williams, E. R. *J. Am. Chem. Soc.* **2007**, *129*, 1612–1622.
(26) Bush, M. F.; Forbes, M. W.; Jockusch, R. A.; Oomens, J.; Polfer, N. C.; Saykally, R. J.; Williams, E. R. *J. Phys. Chem. A* **2007**, *111*, 7753–7760.
(27) Bush, M. F.; Oomens, J.; Saykally, R. J.; Williams, E. R. *J. Phys. Chem. A* **2008**, *112*, 8578–8584.
(28) Bush, M. F.; Oomens, J.; Saykally, R. J.; Williams, E. R. *J. Am. Chem. Soc.* **2008**, *130*, 6463–6471.
(29) O'Brien, J. T.; Prell, J. S.; Steill, J. D.; Oomens, J.; Williams, E. R. *J. Phys. Chem. A* **2008**, *112*, 10823–10830.
(30) Armentrout, P. B.; Rodgers, M. T.; Oomens, J.; Steill, J. D. *J. Phys. Chem. A* **2008**, *112*, 2248–2257.
(31) Cerda, B. A.; Wesdemiotis, C. *Analyst* **2000**, *125*, 657–660.
(32) Remko, M.; Rode, B. M. *J. Phys. Chem. A* **2006**, *110*, 1960–1967.
(33) Remko, M.; Fitz, D.; Rode, B. M. *J. Phys. Chem. A* **2008**, *112*, 7652–7661.
(34) Skurski, P.; Rak, J.; Simons, J.; Gutowski, M. *J. Am. Chem. Soc.* **2001**, *123*, 11073–11074.
(35) Xu, S. J.; Zheng, W. J.; Radisic, D.; Bowen, K. H. *J. Chem. Phys.* **2005**, *122*, 091103.
(36) Bush, M. F.; Prell, J. S.; Saykally, R. J.; Williams, E. R. *J. Am. Chem. Soc.* **2007**, *129*, 13544–13553.
(37) Blom, M. N.; Compagnon, I.; Polfer, N. C.; von Helden, G.; Meijer, G.; Suhai, S.; Paizs, B.; Oomens, J. *J. Phys. Chem. A* **2007**, *111*, 7309–7316.
(38) Bachrach, S. M. *J. Phys. Chem. A* **2008**, *112*, 3722–3730.
(39) Cox, H. A.; Gaskell, S. J.; Morris, M.; Whiting, A. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 522–531.
(40) McLafferty, F. W.; Guan, Z. Q.; Haupts, U.; Wood, T. D.; Kelleher, N. L. *J. Am. Chem. Soc.* **1998**, *120*, 4732–4740.
(41) Strittmatter, E. F.; Williams, E. R. *J. Phys. Chem. A* **2000**, *104*, 6069–6076.
(42) Schnier, P. D.; Price, W. D.; Jockusch, R. A.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 7178–7189.
(43) Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640–10644.
(44) Deery, M. J.; Summerfield, S. G.; Buzy, A.; Jennings, K. R. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 253–261.
(45) Summerfield, S. G.; Whiting, A.; Gaskell, S. J. *Int. J. Mass Spectrom. Ion Processes* **1997**, *162*, 149–161.
(46) Lee, S. W.; Kim, H. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1998**, *120*, 3188–3195.
(47) Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1995**, *117*, 12840–12854.
(48) Jurchen, J. C.; Cooper, R. E.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 1477–1487.
(49) Freitas, M. A.; Marshall, A. G. *Int. J. Mass Spectrom.* **1999**, *182*, 221–231.
(50) Strittmatter, E. F.; Williams, E. R. *Int. J. Mass Spectrom.* **2001**, *212*, 287–300.
(51) Strittmatter, E. F.; Wong, R. L.; Williams, E. R. *J. Phys. Chem. A* **2000**, *104*, 10271–10279.
(52) Julian, R. R.; Jarrold, M. F. *J. Phys. Chem. A* **2004**, *108*, 10861–10864.
(53) Lemoff, A. S.; Bush, M. F.; O'Brien, J. T.; Williams, E. R. *J. Phys. Chem. A* **2006**, *110*, 8433–8442.
(54) Lemoff, A. S.; Bush, M. F.; Williams, E. R. *J. Am. Chem. Soc.* **2003**, *125*, 13576–13584.
(55) Prell, J. S.; Demireva, M.; Williams, E. R. *J. Am. Chem. Soc.* **2009**, *131*, 1232–1242.
(56) Polfer, N. C.; Paizs, B.; Snoek, L. C.; Compagnon, I.; Suhai, S.; Meijer, G.; von Helden, G.; Oomens, J. *J. Am. Chem. Soc.* **2005**, *127*, 8571–8579.
(57) Rodgers, M. T.; Armentrout, P. B.; Oomens, J.; Steill, J. D. *J. Phys. Chem. A* **2008**, *112*, 2258–2267.
(58) Price, W. D.; Jockusch, R. A.; Williams, E. R. *J. Am. Chem. Soc.* **1997**, *119*, 11988–11989.
(59) Moision, R. M.; Armentrout, P. B. *J. Phys. Chem. A* **2006**, *110*, 3933–3946.
(60) Cox, H. A.; Julian, R. R.; Lee, S. W.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2004**, *126*, 6485–6490.
(61) Julian, R. R.; Beauchamp, J. L.; Goddard, W. A. *J. Phys. Chem. A* **2002**, *106*, 32–34.
(62) Jockusch, R. A.; Schnier, P. D.; Price, W. D.; Strittmatter, E. F.; Demirev, P. A.; Williams, E. R. *Anal. Chem.* **1997**, *69*, 1119–1126.
(63) Lifshitz, C. *Int. J. Mass Spectrom.* **2004**, *234*, 63–70.
(64) Wyttenbach, T.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 9–14.
(65) Wyttenbach, T.; Kemper, P. R.; Bowers, M. T. *Int. J. Mass Spectrom.* **2001**, *212*, 13–23.
(66) Kjeldsen, F.; Silivra, O. A.; Zubarev, R. A. *Chem.−Eur. J.* **2006**, *12*, 7920–7928.
(67) Rodriquez, C. F.; Orlova, G.; Guo, Y. Z.; Li, X. M.; Siu, C. K.; Hopkinson, A. C.; Siu, K. W. M. *J. Phys. Chem. B* **2006**, *110*, 7528–7537.
(68) Drayss, M. K.; Blunk, D.; Oomens, J.; Schäfer, M. *J. Phys. Chem. A* **2008**, *112*, 11972–11974.
(69) Rajabi, K.; Fridgen, T. D. *J. Phys. Chem. A* **2008**, *112*, 23–30.
(70) Wu, R.; McMahon, T. B. *J. Mass Spectrom.* **2008**, *130*, 3065–3078.
(71) Wu, R. H.; McMahon, T. B. *J. Am. Chem. Soc.* **2007**, *129*, 4864–4865.
(72) Kong, X. L.; Tsai, I. A.; Sabu, S.; Han, C. C.; Lee, Y. T.; Chang, H. C.; Tu, S. Y.; Kung, A. H.; Wu, C. C. *Angew. Chem., Int. Ed.* **2006**, *45*, 4130–4134.
(73) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2007**, *129*, 5887–5897.

peptides have also been investigated with IRMPD spectroscopy[74−78] and IR-UV double-resonance spectroscopy.[79−83] No evidence for zwitterionic or salt-bridge structures was found for any of these ions.

Previous IRMPD results[55] with ArgGly·M^+ and GlyArg·M^+, M = Li, Na, K, and Cs, indicate that the sequence of the amino acids in these two dipeptides is important for the relative stabilities of the salt-bridge structures. The propensity of ArgGly·M^+ to form a salt-bridge structure is lower than that of the corresponding Arg·M^+, whereas that for GlyArg·M^+ is greater. This was attributed to a more favorable linear orientation of formal charge sites in the salt-bridge form possible for the latter and not to differences in gas-phase basicity. Here, we investigate the propensity of the protonated dipeptides XxxArg·H^+, Xxx = Gly, Val, Pro, Lys, His, and Arg, to form salt-bridge structures where a favorable quasilinear arrangement of formal charge sites is possible. These results show that salt-bridge stability is related to the basicity of the N-terminal amino acid and provide unambiguous evidence that the most stable form of ArgArg·H^+ is a salt-bridge structure in which both arginine residues are protonated and the C-terminus is deprotonated. This is the first direct evidence for a salt bridge in a protonated peptide and it is also the smallest protonated peptide for which a stable salt-bridge form has been reported. These results strongly indicate that stable salt bridges should be common in larger protonated peptides and proteins containing multiple arginine residues.

Experimental Section

IRMPD Spectroscopy. A 4.7 Fourier-transform ion cyclotron resonance mass spectrometer coupled with a free electron laser (FELIX), which produces tunable infrared radiation, was used to collect IRMPD spectra. Detailed descriptions of the instrument[84] and experimental methods[85] can be found elsewhere. The 60/40 water/methanol solutions containing XxxArg, Xxx = Val, Pro, Lys, His, or Arg (BaChem Distribution Services GmbH, Weil am Rhein, Germany; ProArg, LysArg, and ArgArg obtained as acetate salts) were prepared at a 2 mM concentration. Acetic acid (2.5 mM) was added to solutions used to generate protonated dipeptides. Ions were introduced into the mass spectrometer using electrospray ionization at an infusion rate of 10−15 μL/min. Precursor ions were isolated using stored waveform inverse Fourier transforms and subsequently irradiated for 3 s with tunable radiation from FELIX at a macropulse repetition rate of 5 Hz. The same irradiation time was used over the full frequency range scanned for each IRMPD spectrum, and irradiation times were optimized to produce extensive, but not complete fragmentation of the precursor ion at the frequency of maximum absorption. Fragmentation of ArgArg·H^+ in the region ~1350−1700 cm^{-1} was very extensive, so this spectrum was reacquired with 5 dB attenuation of the laser power to reduce peak broadening in this region due to saturation effects.

Computations. A detailed description of GlyArgH^+ calculated structures, energies, and absorption spectra were reported previously.[55] Low-energy structures for XxxArg·H^+, Xxx = Val, Pro, Lys, His, Arg, were generated with Monte Carlo conformational searching using the MMFFs force field as implemented in Macromodel 9.1 (Schrödinger, Inc., Portland, OR) in combination with chemical intuition. At least 30 000 structures were identified for the single charge site (SCS) and salt-bridge (SB) forms of each of these protonated dipeptides. The resulting low-energy structures were separated into families based on similarities in conformation. Geometries of representative structures from these families were optimized at the B3LYP/6-31G* level of theory in Jaguar v. 7.5 (Schrödinger, Inc., Portland, OR). All B3LYP/6-31G* structures were then further optimized and harmonic oscillator vibrational frequencies and intensities calculated at the B3LYP/6-31+G(d,p) level of theory in Q-Chem v. 3.1.[86] All B3LYP/6-31+G(d,p) structures were found to have all real vibrational frequencies, indicating that they are minima on the potential energy surface. Zero-point and temperature corrections at 298 K to electronic energies as well as entropies were calculated using unscaled vibrational frequencies, and Gibbs free energies at 298 K were obtained from these 298 K enthalpies and entropies. All calculated spectra are plotted with frequencies scaled by 0.975 (a scaling factor found previously to give reasonable agreement with experimental spectra in similar systems[26,27,29,30,55,57,87]), and peak shapes are approximated by convolution with a 30 cm^{-1} fwhm Lorentzian profile.

Results and Discussion

IRMPD Spectroscopy. IRMPD of XxxArg·H^+, Xxx = Gly,[55] Val, Pro, Lys, His, and Arg results in the formation of a variety of product ions, and the major IRMPD product ions (>10% of the total observed product ion population) are listed in Table 1. Loss of a water molecule was observed for all of the protonated dipeptides investigated and was the dominant product ion for Xxx = Pro, His, and Arg. Major products for Xxx = Val were largely analogous to those previously reported for Xxx = Gly.[55] Loss of the amino acid Xxx was observed for Xxx = Gly, Val, and Pro, and the y_1 ion (formed by cleavage of the amide C−N bond) was a major product for Xxx = Val, Pro, Lys, and His.

An IRMPD yield was determined from the precursor (I_p) intensity and the sum of the fragment ion intensities (I_f) after laser irradiation at each frequency:

$$\text{IRMPD yield} = \left(\sum I_f\right)/\left(I_p + \sum I_f\right)$$

and this was normalized linearly with laser power to roughly account for the change in laser power as a function of photon energy. IRMPD spectra were obtained for XxxArg·H^+, Xxx = Gly,[55] Val, Pro, Lys, His, and Arg, and $(YyyArg\cdot H_2)^{2+}$, Yyy = His and Arg, over the range ~900−1850 cm^{-1}, and these data are shown in Figures 1−3.

The singly protonated dipeptides can adopt structures in which there is a single formal charge site (SCS) where protonation occurs, or they can adopt a SB structure in which the C-terminus

(74) Fujihara, A.; Matsumoto, H.; Shibata, Y.; Ishikawa, H.; Fuke, K. *J. Phys. Chem. A* **2008**, *112*, 1457–1463.
(75) Lucas, B.; Grégoire, G.; Lemaire, J.; Maitre, P.; Glotin, F.; Schermann, J. P.; Desfrançois, C. *Int. J. Mass Spectrom.* **2005**, *243*, 105–113.
(76) Lucas, B.; Grégoire, G.; Lemaire, J.; Maitre, P.; Ortega, J. M.; Rupenyan, A.; Reimann, B.; Schermann, J. P.; Desfrançois, C. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2659–2663.
(77) Vaden, T. D.; de Boer, T. S. J. A.; Simons, J. P.; Snoek, L. C. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1443–1447.
(78) Vaden, T. D.; de Boer, T. S. J. A.; Simons, J. P.; Snoek, L. C.; Suhai, S.; Paizs, B. *J. Phys. Chem. A* **2008**, *112*, 4608–4616.
(79) Guidi, M.; Lorenz, U. J.; Papadopoulos, G.; Boyarkin, O. V.; Rizzo, T. R. *J. Phys. Chem. A* **2009**, *113*, 797–799.
(80) Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R. *J. Am. Chem. Soc.* **2007**, *129*, 13820–13821.
(81) Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R. *Chimia* **2008**, *62*, 240–243.
(82) Stearns, J. A.; Guidi, M.; Boyarkin, O. V.; Rizzo, T. R. *J. Chem. Phys.* **2007**, *127*, 154322.
(83) Stearns, J. A.; Seaiby, C.; Boyarkin, O. V.; Rizzo, T. R. *Phys. Chem. Chem. Phys.* **2009**, *11*, 125–132.
(84) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. *Rev. Sci. Instrum.* **2005**, *76*, 023103.
(85) Polfer, N. C.; Oomens, J.; Moore, D. T.; von Helden, G.; Meijer, G.; Dunbar, R. C. *J. Am. Chem. Soc.* **2006**, *128*, 517–525.
(86) Shao, Y.; et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.
(87) Polfer, N. C.; Oomens, J.; Dunbar, R. C. *ChemPhysChem* **2008**, *9*, 579–589.

Table 1. Precursor and Major Product Ions[a] from IRMPD Spectroscopy

complex studied	product ion *m/z*	assignment
GlyArg·H^+	157	Loss of glycine
(*m/z* 232)	158	Loss of glycinamide
	172	Loss of 60 Da
	214	− H_2O
	215	− NH_3
ValArg·H^+	157	Loss of valine
(*m/z* 274)	158	Loss of valinamide
	175	y_1
	256	− H_2O
	257	− NH_3
ProArg·H^+	70	$C_4H_8N^+$
(*m/z* 272)	157	Loss of proline
	175	y_1
	254[b]	− H_2O
LysArg·H^+	175[b]	y_1
(*m/z* 303)	273	− CH_2O
	285	− H_2O
HisArg·H^+	175	y_1
(*m/z* 312)	294[b]	− H_2O
ArgArg·H^+	253	Loss of 60 Da and H_2O
(*m/z* 331)	271	Loss of 60 Da
	313[b]	− H_2O

[a] Major products are defined as those that constitute at least 10% of the observed product ion population for a given complex. [b] Indicates product ions that form at least 40% of the observed product ion population.

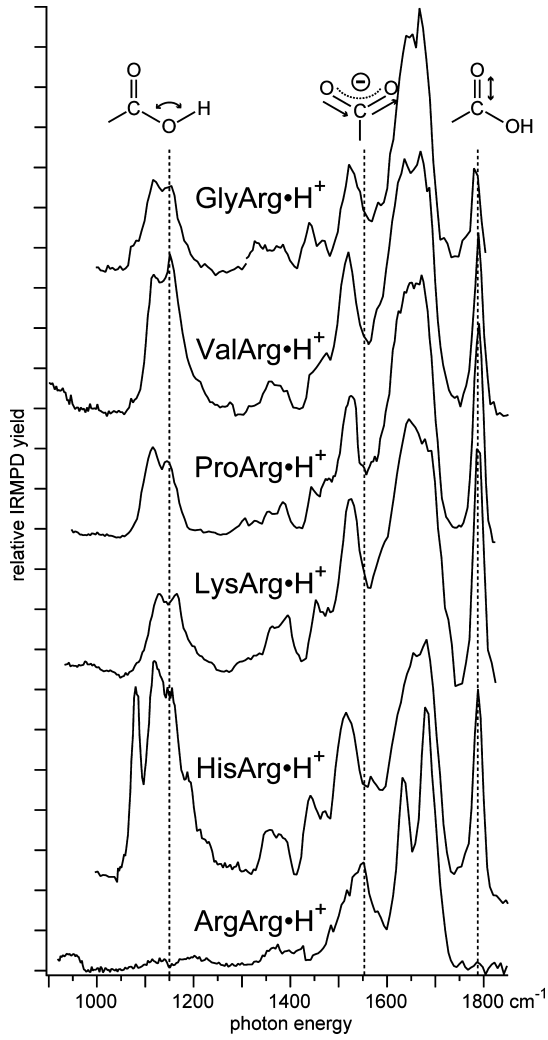


Figure 1. IRMPD spectra of protonated peptides at 298 K.

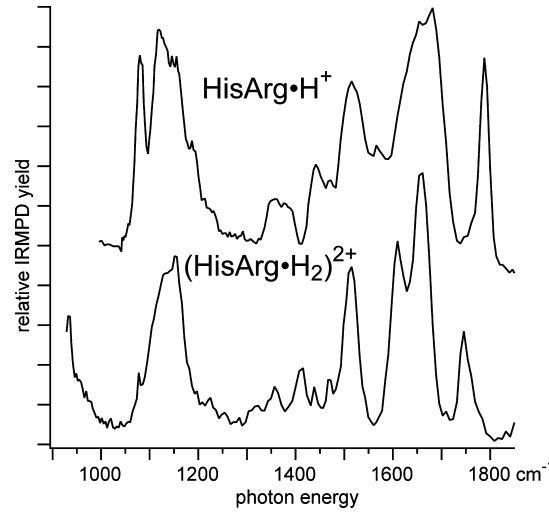


Figure 2. IRMPD spectra of singly and doubly protonated HisArg at 298 K.

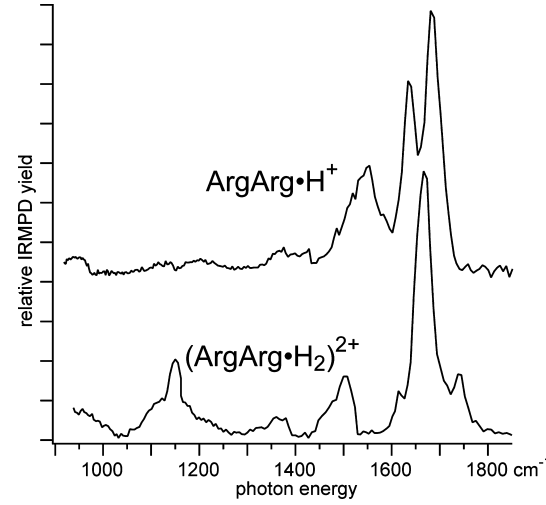
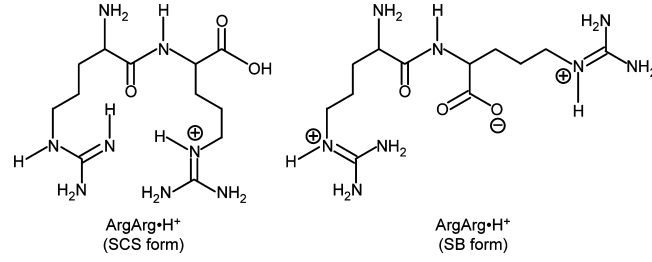


Figure 3. IRMPD spectra of singly and doubly protonated ArgArg at 298 K.

Scheme 1



is deprotonated and two basic sites are protonated. These two structures are illustrated schematically above for ArgArg·H^+ (Scheme 1):

Because the arginine side chain is significantly more basic than other basic functional groups in all of the dipeptides investigated, protonation likely occurs on the arginine side chain in both the SCS and SB structures.

IRMPD Spectra of XxxArg·H^+, Xxx = Gly, Val, Pro, and Lys. The IRMPD spectra of XxxArg·H^+, Xxx = Gly, Val, Pro, and Lys (Figure 1) are remarkably similar to each other over the region investigated (950−1800 cm^{-1}). Each spectrum has

intense bands centered around 1150 and 1780 cm^{-1}, which are attributed to the hydroxyl in-plane bend and carbonyl stretch of an *endo* carboxylic acid group, respectively, based on previous results for other protonated and cationized amino acids[24,26,29,30,57] and dipeptides.[55,87] These two bands unambiguously show that these four protonated dipeptides adopt SCS structures.

The most intense feature in these spectra occurs from 1600−1700 cm^{-1}, a region associated with arginine side chain NH bends, a *trans* amide carbonyl, and amine group NH bends, all of which can be strong IR chromophores.[22,26,27,29,55,87,88] Because of the spectral congestion, assigning contributions of individual modes to this broad band is challenging. Strong photodissociation also occurs around 1530 cm^{-1} and is assigned to the in-plane NH bend of a *trans* amide group. A weaker feature between 1300 and 1400 cm^{-1} previously attributed[55] to α-CH bending modes for GlyArg·H^+ is observed for each of these dipeptides.

The IRMPD spectra of XxxArg·H^+, Xxx = Gly, Val, Pro, and Lys, are nearly the same even though the C-terminal amino acids in these ions have significantly different structures. Because the spectral resolution obtained is insufficient to distinguish these ions, a detailed structural analysis of the conformation of the Xxx side chains based on these IRMPD spectra is not highly informative. There are, however, some minor differences between these spectra. Notably, the relative intensities of the in-plane hydroxyl bend (~1150 cm^{-1}) and the carboxylic acid carbonyl stretch (~1780 cm^{-1}) differ. This variation is especially surprising between GlyArg·H^+ and ValArg·H^+, because these two protonated dipeptides differ only by an aliphatic side chain, which is unlikely to couple strongly to carboxylic acid vibrational modes. The ~1150 cm^{-1} band is also blue-shifted by ~15 cm^{-1} for LysArg·H^+ relative to that in XxxArg·H^+, Xxx = Gly, Val, and Pro, but the carboxylic acid carbonyl group does not blue-shift, suggesting that the geometry or chemical environment of the carboxylic acid group in LysArg·H^+ is slightly different from that of the other three protonated peptides. A band at ~1600 cm^{-1} corresponding to lysine side chain NH bending modes, as observed for Lys·H^+ and protonated *ε-N*-methyllysine,[26] is not present as a distinct feature in the spectrum of LysArg·H^+ and is likely buried beneath the very intense feature from 1600−1700 cm^{-1} (vide supra). Although an intense, sharp feature at 1656 cm^{-1} corresponding to the NH_2 scissoring mode has been reported for the SB form of sodiated proline[22] (although not for potassiated proline,[68] which has a similar structure), no sharp, resolved features attributable to the proline ring NH bending modes are observed from 1600−1700 cm^{-1} for ProArg·H^+.

IRMPD Spectra of HisArg·H^+ and (HisArg·H_2)$^{2+}$. As is the case for Gly, Val, Pro, and Lys C-terminal to Arg, the intense *endo* carboxylic acid carbonyl stretch (1788 cm^{-1}) and in-plane hydroxyl bend (~1150 cm^{-1}) in the IRMPD spectrum of HisArg·H^+ clearly indicate that this protonated dipeptide adopts a SCS structure. A strikingly sharp feature at 1080 cm^{-1} is too low in frequency to be the hydroxyl in-plane bend of a second conformer. This region of the infrared spectrum of condensed-phase imidazole,[89] an analogue for the histidine side chain, contains bands due to CN stretches and in-plane CH and NH bends. This similarity and the narrowness of the 1080 cm^{-1} band suggest that it arises from a histidine side chain bending or stretching mode, which is much less "floppy" than those in the side chains of the other dipeptides due to the rigid, planar structure of the aromatic ring.

Further support for the assignment of the 1080 cm^{-1} band as a signature of the histidine side chain comes in comparing this spectrum to the IRMPD spectrum of (HisArg·H_2)$^{2+}$ (Figure 2), where this band is substantially weaker. Because 5-methylimidazole, a model of the histidine side chain, has a substantially higher gas-phase basicity (~920 kJ/mol) than glycine (848 kJ/mol),[90] it is very likely that the histidine side chain of (HisArg·H_2)$^{2+}$ is protonated, which should alter the frequencies and intensities of the histidine side chain CN stretches and in-plane CH and NH bends from those of neutral histidine side chain. The carboxylic acid carbonyl stretch is red-shifted in this spectrum to 1746 cm^{-1}, likely due to hydrogen bonding to one of the protonated side chains. This is consistent with the electrostatic repulsion between the protonated arginine and histidine side chains, which should make interaction of either side chain with the carboxylic acid carbonyl oxygen more favorable than direct interaction between the two side chains.

IRMPD Spectra of ArgArg·H^+ and (ArgArg·H_2)$^{2+}$. In striking contrast to the other singly protonated dipeptides, there are no bands corresponding to a carboxylic acid carbonyl stretch or in-plane hydroxyl bend in the spectrum of ArgArg·H^+ (Figure 1). A band at 1553 cm^{-1} in the spectrum of ArgArg·H^+ is not present in the spectra of the other singly protonated peptides investigated. Photodissociation in this region is consistent with a carboxylate asymmetric stretch.[24,26,30,55] These results strongly indicate that ArgArg·H^+ adopts a SB structure in which both arginine side chains are protonated and the C-terminus is deprotonated.

The assignment of a SB structure for ArgArg·H^+ is supported by comparison to the IRMPD spectrum of (ArgArg·H_2)$^{2+}$. In contrast to ArgArg·H^+, the presence of a carboxylic acid carbonyl stretch band at ~1740 cm^{-1} and an intense in-plane hydroxyl bend at ~1150 cm^{-1} in the spectrum of (ArgArg·H_2)$^{2+}$ (Figure 3), as well as the absence of the band at 1553 cm^{-1} in the spectrum (ArgArg·H_2)$^{2+}$, indicate that the C-terminus is not deprotonated in the doubly protonated peptide. Because the guanidino group of arginine is substantially more basic (>90 kJ/mol) than glycine, a model system for protonation of the N-terminus, both side chains of (ArgArg·H_2)$^{2+}$ are likely to be protonated, as in the SB form of ArgArg·H^+. The significant differences in the relative intensities of the two features between 1600 and 1700 cm^{-1} in the IRMPD spectra of ArgArg·H^+ and (ArgArg·H_2)$^{2+}$ may be due to a difference in hydrogen-bonding structure between the arginine side chains and the C-terminal oxygen atom in these two ions.

Neutral Water Molecule Loss. One of the major fragmentation pathways for all the protonated dipeptides upon IRMPD is the loss of a neutral water molecule (Table 1). This is the dominant fragmentation pathway observed for ArgArg·H^+ (~45%), despite the clear indication from the IRMPD spectra that ArgArg·H^+ adopts a SB structure. For the loss of a neutral water molecule from a SB structure to occur, a proton and a hydrogen atom must be transferred to the carboxylate group. Thus, a SCS structure must be an intermediate in this fragmentation pathway. The absence of a carboxylic acid carbonyl stretch in the spectrum of ArgArg·H^+ indicates that this intermediate structure is not reflected in the IRMPD spectrum, that is, the IRMPD

(88) Polfer, N. C.; Oomens, J.; Dunbar, R. C. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2744–2751.
(89) Marcia Cordes, d. N. D.; Walter, J. L. *Spectrochim. Acta, Part A* **1968**, *24*, 237–252.
(90) Harrison, A. G. *Mass Spectrom. Rev.* **1997**, *16*, 201–217.

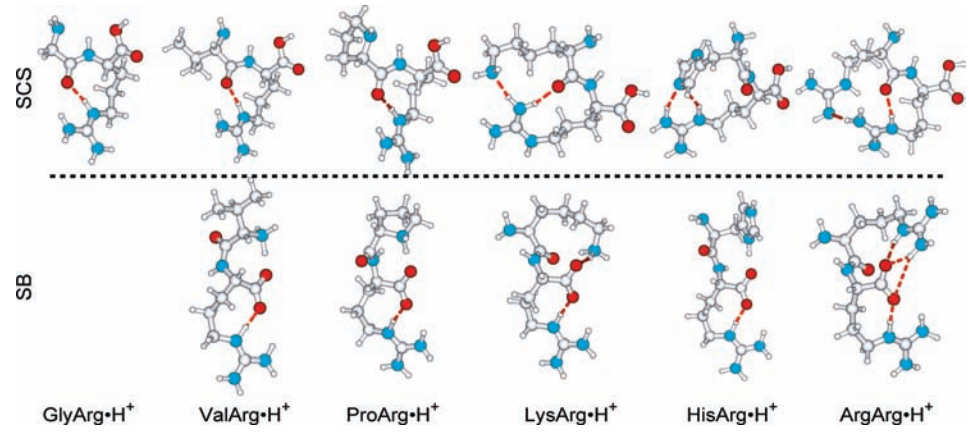


Figure 4. Lowest-energy B3LYP/6-31+G(d,p) SCS and SB structures at 298 K for singly protonated dipeptides. No stable SB form was identified at this level of theory for GlyArg•H^+.

spectrum must more closely resemble the absorption spectrum of the ground state SB structure. These results are consistent with previous IRMPD spectroscopy experiments on alkali metal cationized ArgGly and GlyArg, for which SB structures were found to dissociate by loss of a water molecule as a major fragmentation pathway as well.[55] Loss of water has also been observed for lithiated azetidine-2-carboxylic acid (a proline analogue) even though binding energy measurements and theory indicate a zwitterionic structure.[59] These results provide further evidence that the loss of small neutral molecules upon ion activation, which has been used to infer structural information,[11,66,91] is not a reliable indicator of the presence of salt bridges.

Calculated SCS and SB Structures. Lowest-energy B3LYP/6-31+G(d,p) structures of the SCS and SB forms of XxxArg•H^+, Xxx = Val, Pro, Lys, His, and Arg, are shown in Figure 4. The conformational space for each of the protonated dipeptides investigated is very large, so it is possible that there are even lower-energy conformers that were not identified in this search. The only candidate SB structure for GlyArg•H^+ identified at the B3LYP/6-31G* level of theory minimized to a SCS structure at the B3LYP/6-31+G(d,p) level of theory. In the lowest-energy SCS structures for each of these protonated dipeptides, solvation of the protonated arginine side chain by the amide carbonyl oxygen occurs. An *endo* conformation of the carboxylic acid is most stable for the SCS structure for all but HisArg•H^+, for which the lowest-energy *exo* structure (Figure 3) is calculated to be 19 kJ/mol more stable than the lowest-energy *endo* conformation. Stability is gained for this *exo* SCS structure of HisArg•H^+ by the donation of a hydrogen bond from the hydroxyl group to the histidine side chain and from the arginine side chain to the hydroxyl oxygen. Such an *exo* structure is less favorable for LysArg•H^+ and ArgArg•H^+ likely due to their longer side chains, which create less strain than the shorter histidine side chain when interacting directly with the charge site.

The lowest-energy SB structures calculated for these ions also share some common features. Each of the two protonation sites in these SB structures interacts primarily with one carboxylate oxygen to form a quasilinear arrangement of formal charge sites. The only exception is LysArg•H^+, for which the protonated lysine side chain donates a second ionic hydrogen bond to the amide carbonyl oxygen, resulting in a somewhat less linear arrangement of formal charge sites. Solvation of positive charge by an amide carbonyl oxygen, which is present for metal cationized glutamine[27] and dipeptides[55,87] as well as all of the lowest-energy SCS forms of the singly protonated dipeptides investigated here, occurs only for the lowest-energy SB structure of LysArg•H^+. In each of these SB structure, interaction of the Xxx protonation site with the carboxylate group induces ring strain. Unlike Pro and His, in which the protonation sites are comparatively much closer to the carboxylate group, the longer, floppier Lys side chain can accommodate the additional ring strain induced upon charge solvation by the amide carbonyl oxygen.

Comparison to Calculated Spectra. A detailed analysis of GlyArg•H^+ comparing its experimental IRMPD spectrum and calculated IR absorption spectra was reported elsewhere.[55] Figure 5 shows the experimental spectra of ProArg•H^+ and LysArg•H^+ along with the IR absorption spectra calculated for the lowest-energy SCS and SB structures discussed above, as well as a second low-energy single-charge-site structure (SCS*). The experimental spectra of both of these dipeptides are remarkably similar to each other and to that of GlyArg•H^+, indicating that the extent of structural information that can be gained by detailed comparisons to calculated spectra is limited. The calculated spectrum of the SCS structure for ProArg•H^+ is consistent with all of the major features in the experimental spectrum, indicating that this or similar structures are likely dominant in the experimental ion population. The SCS* structure for ProArg•H^+ differs from the lowest-energy SCS structure in that the carboxylic acid carbonyl oxygen is hydrogen-bonded to the Arg side chain and the hydroxyl hydrogen is hydrogen-bonded to the amide oxygen. These two structural differences are reflected in the calculated spectra, where the carboxylic acid carbonyl stretch is substantially red-shifted from that of the SCS structure and the in-plane hydroxyl bend is considerably weaker. This suggests that structures with hydrogen-bonded carboxylic acid oxygen atoms do not contribute substantially to the ion population. The calculated spectrum of the SB structure for ProArg•H^+ is clearly a poor match to experiment, having no band from 1750−1800 cm^{-1} nor any strong features from 1100−1150 cm^{-1}.

For LysArg•H^+, the SCS* structure differs from the SCS structure (Figure 5) in that different Arg side chain NH groups are coordinated to the amide oxygen and Lys side chain nitrogen atom. This structure is calculated to be slightly higher in energy (by 1.8 kJ/mol) than the SB structure described above, but this

(91) Schäfer, M.; Schmuck, C.; Geiger, L.; Chalmers, M. J.; Hendrickson, C. L.; Marshall, A. G. *Int. J. Mass Spectrom.* **2004**, *237*, 33–45.

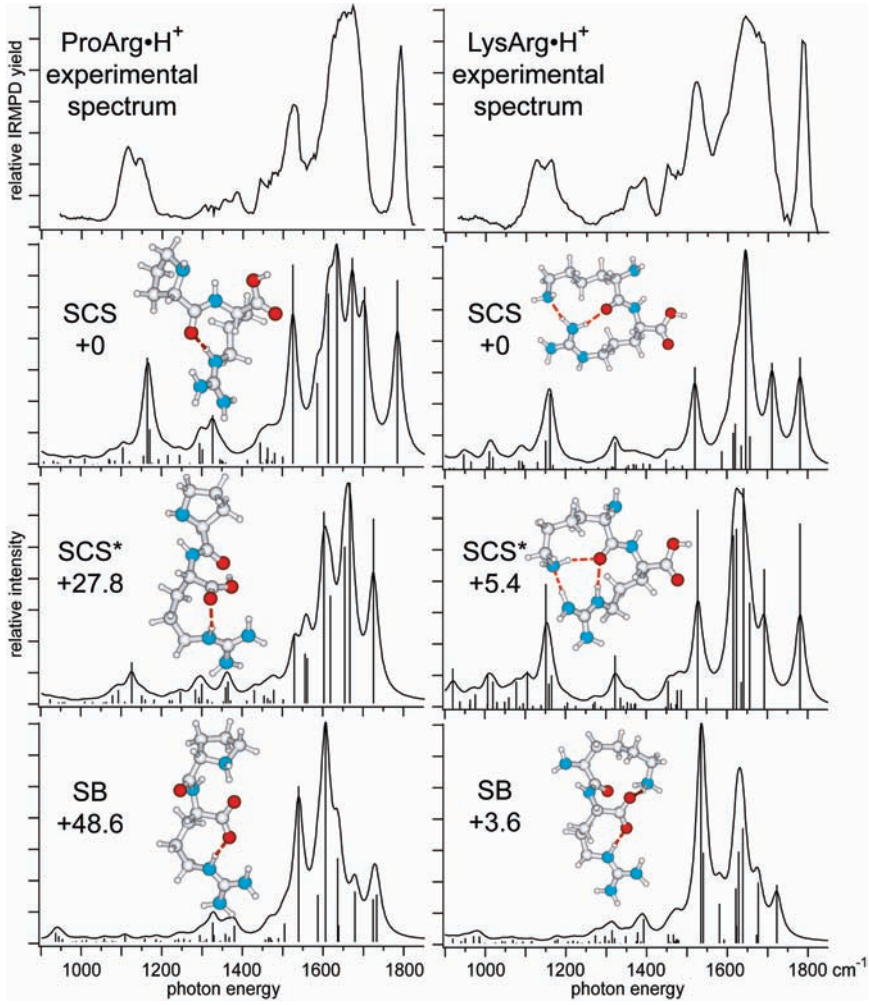


Figure 5. IRMPD spectra and calculated low-energy structures and spectra of ProArg•H^+ and LysArg•H^+, with relative 298 K Gibbs free energies (kJ/mol) calculated at the B3LYP/6-31+G(d,p) level of theory.

energy difference is likely within the uncertainy of the calculations. Although some minor differences in the spectra of these two structures are calculated, both the SCS and SCS* structures are consistent with the experimental spectrum. Thus, substantial populations of both structures, or other similar structures, may be present in the experimental ion population. The SB structure for LysArg•H^+ can clearly be eliminated based on the absence of calculated absorption features for these structures from 1750−1800 and 1100−1150 cm^{-1}.

Figure 6 shows the experimental spectrum of ArgArg•H^+ along with the calculated lowest-energy SCS and SB structures (described above), as well as a somewhat higher-energy salt-bridge structure (SB*) and its calculated spectrum. The SCS structure can be eliminated as a major component of the experimental ion population based on the predicted carboxylic acid carbonyl stretch feature at ~1780 cm^{-1}, which is clearly absent in the experimental spectrum. The calculated spectra of both the SB and SB* structures are consistent with the experimental spectrum. Both of these structures, as well as other similar structures, therefore likely contribute to the ion population.

These comparisons illustrate that the lowest-energy structures calculated for these protonated dipeptides are consistent with experiment and with structural assignments based on previous results for amino acids and peptides. In particular, the calculated spectra clearly show differences between SCS and SB structures and between different conformations of carboxylic acid groups, which can aid in the assignment of experimental spectra. Although other spectral differences due to alternative side chain conformations are calculated for these ions, the experimental spectra can be consistent with several different structures, so that a more detailed structural assignment based on these data alone is difficult.

Effects of Gas-Phase Basicity. The differences in Gibbs free energy at 298 K between the SCS and SB forms of XxxArg•H^+, Xxx = Val, Pro, Lys, His, and Arg, as a function of the basicity differences between the N-terminal amino acid are shown in Figure 7. Only for ArgArg•H^+ is the SB form calculated to be more stable, although the difference in energy for the SCS and SB forms of LysArg•H^+ is only 1.4 (3.6) kJ/mol at 0 (298) K. The computational results are in good agreement with the experimental results which show that only ArgArg•H^+ adopts a SB structure. This is consistent with calculations that predict a ground state SB structure for a related ion, protonated arginine with a guanidino group in place of an N-terminal amine group.[52] There is a general trend between the stability of the SB form and the basicity of the N-terminal amino acid, although this relationship is clearly not direct. The apparently unusual relative stability of the SB form of LysArg•H^+ compared to that of ProArg•H^+ and HisArg•H^+ may be attributable in part to solvation of the Lys side chain protonation site by the amide carbonyl oxygen, which does not occur for ProArg•H^+ and HisArg•H^+ likely due to their shorter side chains (vide supra).

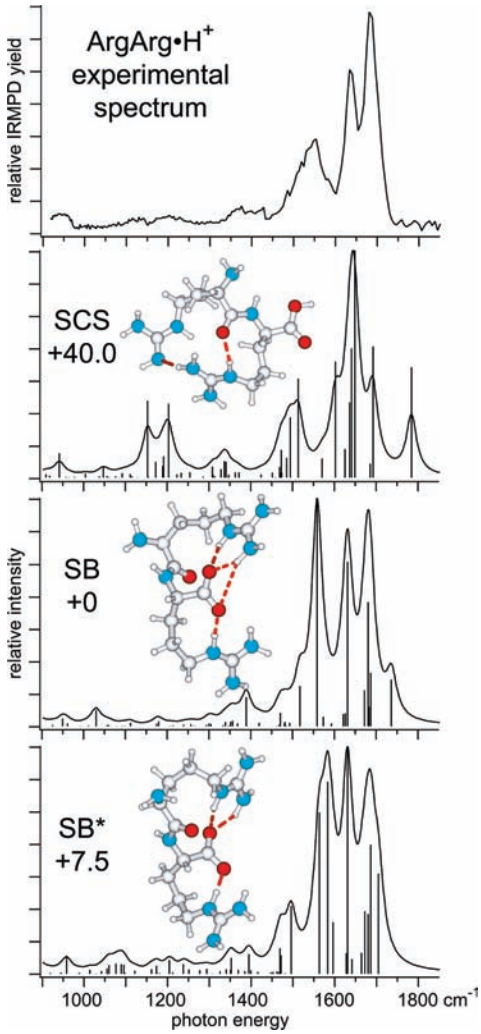


Figure 6. IRMPD spectra and calculated low-energy structures and spectra of ArgArg•H^+, with relative 298 K Gibbs free energies (kJ/mol) calculated at the B3LYP/6-31+G(d,p) level of theory.

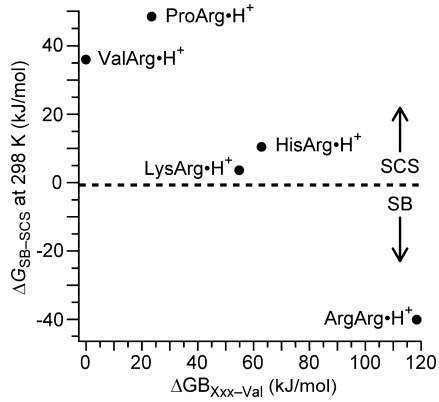


Figure 7. Differences in Gibbs free energy at 298 K (kJ/mol) between the lowest-energy SB and SCS forms of singly protonated dipeptides XxxArg•H^+ as a function of the gas-phase basicity of the amino acid Xxx relative to that of Val. Positive values of $\Delta G_{SB\text{-}SCS}$ correspond to a SCS ground state, and positive values of $\Delta GB_{Xxx\text{-}Val}$ correspond to Xxx being more basic than Val.

The lowest-energy SB forms for XxxArg•H^+, Xxx = Val, Pro, His, and Arg all have a nearly linear arrangement of formal charge sites. Previous results with GlyArg•M^+ and ArgGly•M^+, M = H, Li, Na, K, and Cs, showed that GlyArg•M^+ has a greater propensity to form stable SB structures despite the fact that these two peptides have essentially identical gas-phase basicities, a result attributed to a more linear orientation of formal charge sites possible in the SB form of GlyArg•M^+.[55] Because a C-terminal arginine makes possible a quasilinear formal charge site geometry in the SB structure of all these dipeptides, the change in the relative stability of the SB and SCS structures is most directly related to the gas-phase basicity of the N-terminal residue.

Conclusions

The role of the N-terminal amino acid in stabilizing the salt-bridge form of protonated dipeptides, XxxArg•H^+, Xxx = Gly,[55] Val, Pro, Lys, His, and Arg, was investigated with IRMPD spectroscopy and hybrid density functional theory. The spectra for Xxx = Gly, Val, Pro, Lys, and His clearly indicate that these ions form structures in which protonation occurs at a single site. The spectra for Xxx = Gly, Val, Pro, and Lys differ only slightly, but a signature peak due to a neutral histidine side chain was identified for HisArg•H^+ at 1080 cm^{-1} and confirmed by comparison to the spectrum of $(HisArg•H_2)^{2+}$. In contrast to the sharp NH_2 scissoring mode reported for the protonated pyrrolidine ring in sodiated proline[22] (which is not observed for potassiated proline despite its nearly identical structure[68]), this intense, unusually narrow feature due to the histidine side chain occurs in a less congested region of the spectrum and may prove useful in the analysis of IRMPD spectra of other histidine-containing peptides.

In striking contrast, the IRMPD spectrum of ArgArg•H^+ clearly indicates that this ion adopts a salt-bridge structure in which two basic sites are protonated and the C-terminus is deprotonated, a result that is in agreement with theory. There is a general trend in increasing stability of the SB form with increasing gas-phase basicity of the N-terminal amino acid, although this trend is not direct. Calculations indicate that the dipeptides with histidine and lysine residues are not sufficiently basic to stabilize the SB form preferentially, in agreement with the experimental data. These results are the first direct experimental evidence for a stable salt-bridge structure in a protonated peptide, and ArgArg•H^+ is the smallest peptide for which a salt bridge has been reported. These results indicate that salt-bridge structures should be common in larger protonated peptides that have at least two arginine residues and a carboxylic acid group. Such structures may also be common for larger peptides containing just a single arginine residue and a basic residue, such as Lys or His, where significant charge solvation of the formal charge sites can occur. Investigation of larger peptides should provide useful information about the role of salt bridges in the structures and reactivity of gas-phase ions as well as the role of intramolecular solvation in salt-bridge stability that may provide useful insights into the stability of salt bridges in the interior of proteins and at protein−protein interfaces in solution.

Acknowledgment. IRMPD spectra were acquired at the FOM Institute for Plasma Physics "Rijnhuizen," which is financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). The authors thank Dr. Britta Redlich and the rest of the FELIX staff for excellent support and Drs. Kathleen A. Durkin and Jamin L. Krinsky for helpful conversations. Generous financial support was provided by the National Science Foundation (Grants CHE-0718790 and OISE-730072).

Supporting Information Available: Complete ref 86. This material is available free of charge via the Internet at http://pubs.acs.org.

JA901870D