Gaseous Bradykinin and Its Singly, Doubly, and Triply Protonated Forms: A First-Principles Study

Christopher F. Rodriquez,[†] Galina Orlova,[†,§] Yuzhu Guo,[†] Xiaomao Li,[‡] Chi-Kit Siu,[†] Alan C. Hopkinson,[†] and K. W. Michael Siu*[,†]

Department of Chemistry and Centre for Research in Mass Spectrometry, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada, and Biotechnology Centre for Applied Research and Training, Seneca@York, Seneca College, 70 The Pond Road, Toronto, Ontario M3J 3M6, Canada

Received: September 3, 2004; In Final Form: February 23, 2006

The conformers of gaseous bradykinin, BK, (Arg^1-Pro^2-Pro^3-Gly^4-Phe^5-Ser^6-Pro^7-Phe^8-Arg^9) and its protonated forms, $[BK + H]^+$, $[BK + 2H]^{2+}$, and $[BK + 3H]^{3+}$, were examined theoretically using a combination of the Merck molecular force field, Hartree−Fock, and density functional theory. Neutral BK, $[BK + H]^+$, and $[BK + 2H]^{2+}$ exist in zwitterionic forms that are stabilized by internal solvation and have compact structures; $[BK + 3H]^{3+}$ differs by the absence of a salt bridge and adopts an elongated form. The common structural feature in all four BK species is a β-turn in the Ser^6-Pro^7-Phe^8-Arg^9 sequence. The gas-phase basicity of $[BK + H]^+$ estimated from the calculated protonation energy is in accord with published experimental basicity; population-weighted collision cross-sections of the three ionic forms are in agreement with experimental cross-sections in the literature.

Introduction

The three-dimensional structures of gaseous polypeptide ions have received much attention in recent years.[1−24] The secondary and tertiary structures, which play crucial roles in the chemistries of biomolecules in the condensed phase, can affect the reactivities of gaseous ions, including their fragmentations.[5,20,21] Another important aspect that has been highlighted in a number of publications[23−27] is the difference between the structures of naked and solvated polypeptides[23] and amino acids[25−27] that reveals the influence of solvation on structure.

Arguably the most important role of a solvent is charge stabilization. An example of note is water's ability to stabilize the zwitterionic (charge-separated) form of amino acids, NH_3^+−CHR−COO^- or, when the side chain contains a basic group, NH_2−$CHRH^+$−COO^-. The zwitterionic form predominates in aqueous solutions. By contrast, naked, gaseous amino acids exist in the canonical form, NH_2−CHR−COOH. The zwitterionic form of glycine is not even at a minimum on the potential energy surface.[27d] However, recent theoretical studies[25−27] show that a charge-separated form of amino acids can exist in the gas phase; the presence of a few solvent molecules or a net charge can make a charge-separated form thermodynamically stable. Zwitterionic forms of hydrated,[25] protonated,[26a−c] alkali-cationized,[26b−d] and electron-anionized[27d] glycine are all at local minima. An amino acid may also act as a solvent; it has been shown that neutral arginine dimer exists in a zwitterionic form, with the two monomers bound in a cyclic array by two salt bridges.[27e] By extension, internal solvation may also offer sufficient stabilization to the zwitterionic form of a naked, neutral polypeptide.

Bradykinin (BK), a nonapeptide having the amino acid sequence of Arg^1-Pro^2-Pro^3-Gly^4-Phe^5-Ser^6-Pro^7-Phe^8-Arg^9, is one of the most thoroughly studied polypeptides in the gas phase. BK plays diverse functions in humans; it is a plasma kinin responsible for smooth muscle contraction, vasodilatation, and permeability of capillaries,[28] a selective precursor for the synthesis of nitric oxide,[29] an agent that plays a role in the cardioprotective effect of angiotensin-converting enzyme inhibitors,[30] and a growth factor for small-cell lung cancer and prostate cancer.[31] It is also a frequently used model peptide in analytical techniques such as mass spectrometry and ion mobility spectrometry.[1−24] Three protonated forms of bradykinin, $[BK + H]^+$, $[BK + 2H]^{2+}$, and $[BK + 3H]^{3+}$, are known and have been examined to varying degrees of comprehensiveness. Experimental techniques are frequently used in conjunction with theoretical calculations to obtain structural information. BK is relatively large, and until quite recently, the only viable means to compute its structure were molecular mechanical force field or semiempirical methods.

Wyttenbach et al.[4a] reported the structures of singly protonated BK, $[BK + H]^+$, and $[BK + Na]^+$, calculated using molecular mechanics/dynamics with the AMBER force field. These structures were then evaluated via comparison of their experimental collision cross-sections with helium and calculated cross-sections. For $[BK + H]^+$, three types of tautomers were examined: BK protonated on the side chain of Arg^1, BK protonated on the side chain of Arg^9, and BK protonated on both arginyl residues and deprotonated on the C-terminal carboxylic group (i.e., an ion−zwitterion form). For all tautomers, the lowest-energy structures are compact, with the charge effectively solvated by carbonyl oxygen atoms on the backbone; for the ion−zwitterion tautomer, the two positively charged guanidine groups were found to flank the carboxylate anion to form a salt-bridge structure. Subsequent studies, mainly employing molecular mechanics/dynamics simulation, supported the salt-bridge structure of protonated BK.[5,19,20] Strittmatter and

* Author to whom correspondence should be addressed. E-mail: kwmsiu@yorku.ca.
[†] York University.
[‡] Seneca College.
[§] Current address: Department of Chemistry, St. Francis Xavier University, Antigonish, NS B2G 2W5, Canada.

10.1021/jp046015r CCC: $33.50
Published on Web 03/23/2006

Williams[24] reported calculations of the structures of $[BK + H]^+$ using molecular mechanics with the Merck molecular force field (MMFF), followed by density functional theory (DFT) with the EDF1/6-31G(d) method. They determined that the salt-bridge structure is ca. 10 kcal/mol lower in energy than the ion−molecule structure in which protonation takes place only on Arg^1 or Arg^9. Thus theoretical studies support experimental findings of a salt-bridge structure for protonated BK.[4,5,8,19,20] It is noteworthy that a β-turn in the Ser^6-Pro^7-Phe^8-Arg^9 segment, observed in solution using NMR[32] was predicted by the DFT calculations.

There are a number of studies on $[BK + 2H]^{2+}$,[2,3,6,8,11−13,19−22] although comparatively fewer definitive studies. The blackbody infrared radiative dissociation (BIRD) study of Schnier et al.[2] found $[BK + 2H]^{2+}$ to have a more extended ion conformation than $[BK + H]^+$, thereby reducing charge−charge repulsion. Molecular modeling by Kaltashov et al.[3] using CHARMM supported elongated structures with maximum charge separation for $[BK + 2H]^{2+}$. Szilágyi et al.,[6] however, concluded from kinetic energy release experiments that $[BK + 2H]^{2+}$ is relatively compact and rigid. Using semiempirical MNDO calculations and assuming protonation on the Arg^1 and Arg^9 guanidine groups, these authors obtained a folded low-energy structure that is consistent with their experimental observations (although they admitted that there were probably hundreds of other possible conformers, some probably of lower energy). High-resolution ion mobility experiments of Counterman et al.,[8] however, failed to reveal the existence of more than one conformer; furthermore, the collision cross-section of $[BK + 2H]^{2+}$ is virtually identical to that of $[BK + H]^+$, thereby indicating a compact conformation. Freitas and Marshall[12] raised the possibility of a zwitterionic structure for $[BK + 2H]^{2+}$, which has a C-terminal carboxylate anion, two protonated guanidine groups on the two terminal arginine residues, and the third proton on the N-terminal amino group. Data from gas-phase reactions between BK and hydroiodic acid and thermal dissociation of BK ions led McLuckey and co-workers[11,13] to conclude that the dissociation of the $[BK + 2H]^{2+}$ ions probably originates from an interconverting mixture of precursor conformers. AMBER calculations of Gill et al.[19] showed that low-energy conformers of nonzwitterionic and zwitterionic $[BK + 2H]^{2+}$ (in the former, the guanidine groups of both Arg^1 and Arg^9 are protonated; in the latter, Arg^1, Arg^9, and the N-terminal amino group are protonated, and the C-terminal carboxylic group is deprotonated) have similar collision cross-sections. In the work of Cassady and co-workers,[20,21] molecular mechanics/dynamics modeling using CHARMM reveals a low-energy nonzwitterionic structure that has the protonated guanidine groups occupying opposite ends of a compact conformation with the guanidine groups folded back onto the peptide backbone and interacting with carbonyl oxygen atoms via hydrogen bonds. Hydrogen/deuterium exchange experiments of Levy-Seri et al.[22] confirmed compact conformation(s) for $[BK + 2H]^{2+}$; in addition, two non-interconverting populations were observed. In summary, conflicting conclusions abound for the structures of $[BK + 2H]^{2+}$; earlier studies favored extended conformations,[2,3] while later work implied or reported compact structures that were zwitterionic[9,12,22] or nonzwitterionic.[6,20,21]

Relatively little is known about $[BK + 3H]^{3+}$ ions.[8,11,13] The collision cross-section of $[BK + 3H]^{3+}$ is considerably larger than those of $[BK + 2H]^{2+}$ and $[BK + H]^+$, which are almost identical; this means that $[BK + 3H]^{3+}$ is considerably more extended than the lower-charged BK ions.[8,33] Hydrogen/deuterium exchange using deuterated hydroiodic acid suggested that there may be two reactive populations of $[BK + 3H]^{3+}$.[11] No suggestions of protonation sites have been made. To the best of our knowledge, no theoretical studies on $[BK + 3H]^{3+}$ have been reported, nor have there been any on neutral BK.

In this article, we present the first systematic theoretical study of neutral gaseous BK and its protonated forms, $[BK + H]^+$, $[BK + 2H]^{2+}$, and $[BK + 3H]^{3+}$. The questions that we are addressing are: (1) whether neutral BK and doubly charged $[BK + 2H]^{2+}$ exist in nonzwitterionic or zwitterionic forms; (2) why addition of a second proton to $[BK + H]^+$ does not increase the apparent ion size, as revealed by ion cross-sections, despite an increasing Coulombic repulsion; (3) what the structure of $[BK + 3H]^{3+}$ is; (4) how the structures of BK, $[BK+2H]^{2+}$ and $[BK + 3H]^{3+}$ relate to that of $[BK + H]^+$; and (5) how the proton affinity decreases with increased charge on BK? To answer these points, we performed thorough examinations of the lowest-energy isomers and their conformers, starting from neutral BK as a parent structure. All structures with the proton(s) added to all chemically reasonable sites for protonation, including the guanidine nitrogen atoms, the N-terminal amino nitrogen, the C-terminal carboxylate anion, the carbonyl oxygen atoms of the peptide bonds, as well as the proline nitrogen atoms, have been examined.

Methods

The conventional computational technique for predicting relative energies and geometries of many possible conformers on a relatively flat potential energy surface typically starts with an exhaustive search of candidate conformers using molecular mechanical force field methods. Among all commonly used force fields, MMFF has been shown to yield the most reliable set of conformers.[24,34] To obtain accurate energetics and geometries, the candidate conformers are further refined with high-level first-principles methods, e.g., Hartree−Fock (HF) or DFT. The combined approach of MMFF and DFT has been employed in the work of Strittmatter and Williams[24] for the study of $[BK + H]^+$. For tractability, they used relatively loose criteria for self-consistent field (SCF) convergence and geometry optimization. In this work, we employed the ab initio HF method as the basic first-principles method in tandem with MMFF. The HF approximation has been shown to predict reliable energetics and geometries and reproduce even subtle hydrogen-bond cooperative effects for the 3_{10}- and α-helices,[35a] chains and ribbons,[35b] and β-sheets.[35c] To include electron correlation effects, the HF geometries of selected isomers were then reoptimized using hybrid DFT. Standard criteria for SCF convergence and geometry optimization were used in all calculations.

The Merck molecular force field 94 (MMFF94) and Monte Carlo conformational search were employed to determine candidate low-energy structures using the SPARTAN code.[36] The Gaussian 98 program suite[37] was used for HF and DFT calculations with the doubly split-valence 3-21G[38] and 6-31G[39] basis sets. For each tautomer, ca. 12 000−15 000 conformers were generated with MMFF. The 100 lowest-energy conformers were displayed and examined visually; chemically improbable structures were discarded. The 10 lowest-energy conformers were then fully optimized with the HF/3-21G method. It should be noted that different MMFF conformers often converged to the same structure during HF optimization, thereby reducing the number of conformers to be recalculated at the next level. HF/3-21G conformers that are within 10 kcal/mol from the lowest-energy conformer were then reoptimized at the HF/6-31G level. Finally, a few selected HF/6-31G structures were

TABLE 1: Relative Electronic Energies (kcal/mol) of Canonical and Zwitterionic Arginine and Protonated Arginine

method	**c5**[a]	**c4**	**c2**	**c1**	**z3**[b]	**z2**	**z1**
HF/6-31G[c]	0.00	0.67	3.61	3.43	−6.35	−1.47	−1.03
B3LYP/6-31G[c]	0.00		3.61	6.65	−6.98	−3.63	−4.14
B3LYP/6-31++G(d,p)[d]	0.00	1.95	1.83	2.53	1.82	4.27	4.08
MP2/6-31++G(d,p)[d]	0.00	0.82	1.82	3.10	1.68	2.58	3.73
CCSD/6-31++G(d,p)[d]	0.00	1.32	2.00	1.99	3.97	4.11	4.65

method	**p1**[e]	**p2**	**p3**	**p4**	**p5**	**p6**
HF/6-31G[c]	0.00	3.03	4.09	5.16	3.35	9.58
B3LYP/6-31G[c]	0.00	3.40	4.52	6.34	3.70	9.92
B3LYP/6-31++G(d,p)[d]	0.00	2.73	3.21	4.74	3.03	9.61
MP2/6-31++G(d,p)[d]	0.00	5.56	5.58	6.41	6.89	8.19
CCSD/6-31++G(d,p)[d]	0.00	4.75	5.27	5.78	6.02	7.90

[a] Canonical arginine; labeling as in refs 42 and 43. [b] Zwitterionic arginine; labeling as in refs 42 and 43. [c] This study. [d] References 42 and 43. [e] Protonated arginine; labeling as in refs 42 and 43.

reoptimized using DFT with the B3LYP hybrid exchange−correlation functional[40,41] and the 6-31G and 6-31++G(d,p) basis sets.

Strittmatter and Williams[24] performed a detailed and critical analysis on performance of the MMFF(Monte Carlo)/DFT technique for predicting the structure of the $[BK + H]^+$ ion. They reported that the order of structures with respect to energy changed when molecular-mechanics-optimized structures were reoptimized with DFT. Similarly, in our study, the progression from MMFF94 to HF/3-21G and to HF/6-31G resulted in changes of the ordering of nearly degenerate conformers due to changes in relative energies. As a consequence, the strategy of full optimization on the 10 lowest-energy MM conformers with HF using increasingly larger basis sets cannot be considered as fail-safe in locating the lowest-energy conformation. Obviously, increasing the number of MM structures for full HF optimization would decrease the likelihood of missing the lowest-energy conformer; the decision to limit the number to 10 was based on a compromise of accuracy and tractability. To verify the MMFF/HF/DFT technique and the accuracies of our results, we compared them with available experimental data: structural information, collision cross-sections, and gas-phase basicities. We also compared our lowest-energy structures obtained using MMFF/HF/DFT with previously reported lowest-energy structures of $[BK + H]^+$ located with AMBER (molecular dynamics)[4a] and with MMFF/DFT.[24]

To assess the performance of the relatively small 6-31G basis set, we computed the lowest-energy conformers of canonical and zwitterionic arginine as well as those of canonical arginine protonated at the guanidine group. The arginine system has previously been examined by Simon and co-workers[42,43] using the B3LYP, MP2, and CCSD methods with the 6-31++G(d,p) basis set. These studies are the currently highest level theory work done on arginine. Using the structures optimized with the MP2/6-31++G(d,p) method[42] as initial geometries, we reoptimized these structures with the HF/6-31G and B3LYP/6-31G methods. The resulting relative electronic energies are shown in Table 1; structure labeling is identical to that used in Simon's work.[42,43] Optimized structures are detailed in Figures 1s and 2s in the Supporting Information; geometrical parameters for hydrogen bonds are in Table 1s. Additional calculations were performed on the lowest-energy conformers, **c5**, of canonical and, **z3**, of zwitterionic arginine using the 6-31G(d) and 6-31+G(d) basis sets. These results are shown in Table 2.

Both the HF/6-31G and the B3LYP/6-31G methods reproduce correctly the lowest-energy conformers of canonical and zwitterionic arginine and protonated arginine (Table 1). Ordering of the conformers in these methods is also in agreement with that in CCSD/6-31++G(d,p). The B3LYP/6-31G method did not locate the canonical conformer **c4**; upon optimization, the carboxylic proton transferred to the guanidine group yielding the zwitterionic conformer **z3**. By contrast, the HF/6-31G method predicted the **c4** minimum in accord with the higher-level methods.[42,43] The obvious drawback of the 6-31G basis set, irrespective of method, is overestimating the relative stability of zwitterionic structures versus the canonical forms partly because a smaller basis set usually gives a larger basis set superposition error (BSSE) in the calculated energies of biomolecules containing many intramolecular hydrogen bonds.[44] The relative performance of basis sets is emphasized in Table 2. Incorrect ordering is evident not only for 6-31G but also for 6-31G(d) and 6-31+G(d). Only the largest basis set, 6-31++G(d,p), correctly predicted that canonical arginine, **c5**, has a lower energy than the zwitterionic conformer **z3**, indicating that diffuse and polarization functions employed for all atoms are essential to obtaining a reasonably accurate potential energy surface for amino acids, similar to the findings for carbohydrates.[44] For the smallest basis set, 6-31G, the error is approximately 9 kcal/mol. This error might be even larger for polypeptides such as bradykinin. For this reason, the geometries of the lowest-energy structures of the canonical and zwitterionic $[BK + 2H]^{2+}$ were further optimized at the B3LYP/6-31++G(d,p) level to assess the adequacy in the use of the smaller 6-31G basis set for bradykinin.

TABLE 2: Relative Electronic Energies (kcal/mol) of Canonical and Zwitterionic Arginine Determined Using the B3LYP Exchange−Correlation Functional and Various Basis Sets

structure	6-31G[a]	6-31G(d)[a]	6-31+G(d)[a]	6-31++G(d,p)[b]
c5	0.00	0.00	0.00	0.00
z3	−6.98	−1.00	−0.94	1.81

[a] This study. [b] References 42 and 43.

The analysis of hydrogen bonding is necessary in understanding the folding of polypeptides. For systems with intramolecular hydrogen bonding, the energetics and geometries of hydrogen bonds are influenced by strain energy and steric effects. In this study, we assumed that strain and steric effects in the polypeptides that we examined are relatively insignificant, and intramolecular hydrogen bonds manifest the same trends as intermolecular hydrogen bonds. Thus we considered all A−H···B (where A = C, N, or O and B = O or N) interactions having H···B distances shorter than 2.5 Å; and A, H, and B atoms and the lone electron pair of the B atom in the same plane as a hydrogen bond. We also assumed that the strength of the A−H···B intramolecular hydrogen bond increases with decreasing H···B distance.[45] In addition, the intramolecular hydrogen bond in a charged system is stronger than that in a neutral system, similar to the intermolecular hydrogen bond.[46] For neutral and protonated arginine, both the HF/6-31G and the B3LYP/6-31G methods reproduced correctly the "short" and "long" hydrogen bonds predicted with the MP2/6-31++G(d,p) method[42] (Table 1s). The only exception (out of a total of 28) is the O2···HN9 distance in the highest-energy zwitterions **z1**, which is notably longer as determined with the MP2/6-31++G(d,p) method than with the B3LYP/6-31G and HF/6-31G methods (2.051 Å vs 1.583 and 1.709 Å, respectively). The theory level of the quantum chemical calculation has little effect on the hydrogen-bonding structure and is unlikely to change qualitatively the secondary folding structures.

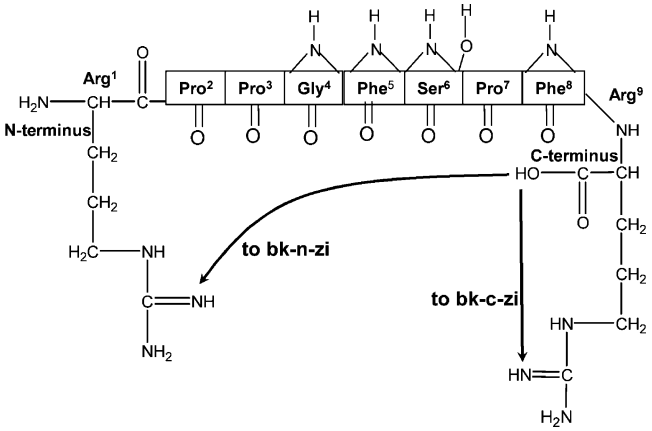


Figure 1. Idealized primary structure of bradykinin. For internal residues, the secondary amino groups, carbonyl oxygen atoms, and hydroxy group are shown. The arrows indicate two possible proton migrations resulting in zwitterionic structures.

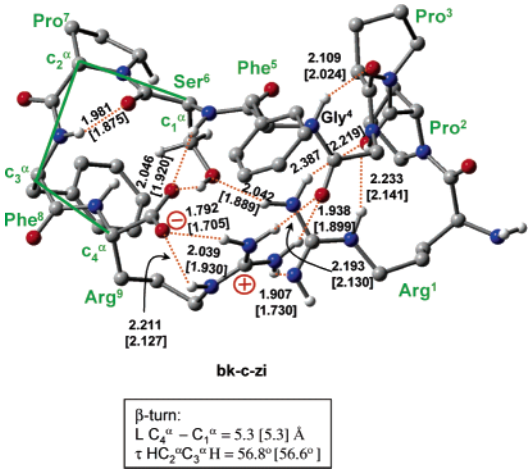


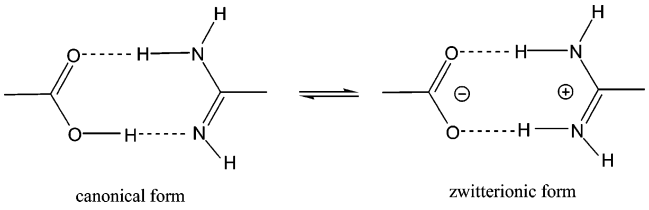
Figure 2. The lowest in energy structure of neutral BK, **bk-c-zi**, predicted with the HF/6-31G and B3LYP/6−31 (values in square brackets) methods. For clarity, only hydrogen atoms of importance to the structure are shown. Hydrogen bonds are illustrated by dashed orange lines. The distances are in angstroms. The β-turn is highlighted in green, and its α-carbon atoms are denoted.

Results and Discussion

BK. The primary structure of BK and the functional groups that could be involved in protonation, hydrogen bonding, and folding of the polypeptide are sketched in Figure 1. There are more potential proton acceptors than proton donors due to the presence of two basic Arg residues. Three possible tautomeric structures of neutral BK were examined. The first tautomer, **bk-non-zi**, has no formal charge centers and is nonzwitterionic (Figure 1). The second possible tautomer, **bk-c-zi**, is zwitterionic and is formed via transfer of the C-terminal carboxylic acid proton to the guanidine group of Arg[9]. The third tautomer, **bk-n-zi**, is also zwitterionic and is formed via transfer of the same carboxylic acid proton to the guanidine group of Arg[1].

According to MMFF/HF/DFT results, neutral BK energetically favors the zwitterionic form, **bk-c-zi**. The lowest-energy conformer of BK predicted with the HF/6-31G and B3LYP/6-31G methods is shown in Figure 2. For this zwitterion, the carboxylate oxygen atoms form bifurcated hydrogen bonds. A total of 12 hydrogen bonds results in a tightly packed structure. In the absence of a formal net charge, the hydrogen bonds are rather long, all calculated to be greater than 1.9 Å with the HF/6-31G method (Figure 2). The B3LYP/6-31G method yields shorter hydrogen bonds, typical of DFT.[47] A short hydrogen bond between the hydroxy group of Ser[6] and a carboxylate oxygen of 1.792 and 1.705 Å is predicted with the HF/6-31G and B3LYP/6-31G methods, respectively. The B3LYP/6-31G method also shortens notably (from 1.907 to 1.730 Å) the hydrogen bond between the guanidine groups.

SCHEME 1



A structural feature of note is the β-turn exhibited by the Ser[6]-Pro[7]-Phe[8]-Arg[9] sequence; this is stabilized by a hydrogen bond. The backbone geometry of the β-turn is somewhat twisted (τ $HC_2^{\alpha}C_3^{\alpha}H$ = ca. 57°), with the distance between the C_1^{α} and C_4^{α} atoms being 5.3 Å (note that the critical value for the β-turn is 7 Å).[48] Interestingly, the B3LYP/6-31G method, despite a significant change in the lengths of hydrogen bonds, predicts essentially the same characteristic values for the β-turn as does the HF/6-31G method.

Low-energy conformers of nonzwitterionic BK with the carboxylic acid group interacting "face-to-face" with the guanidine group of Arg[1] (Scheme 1) were located at the MMFF level. However, upon reoptimization with the Hartree−Fock method, the carboxylic proton was transferred without an energy barrier to the guanidine group in Arg[1], thereby producing a zwitterionic structure of the **bk-n-zi** type, which is only 2.1 kcal/mol higher in energy than the lowest-energy structure, **bk-c-zi**. The B3LYP/6-31G method yields a similar result, with the **bk-n-zi** type structure being 1.8 kcal/mol higher in energy than **bk-c-zi**. The fact that both the HF and B3LYP methods fail to locate nonzwitterionic structures suggests that increasing the size of the basis set will probably not change the conclusion (vide infra).

The **bk-n-zi** lowest conformer predicted with MMFF yields a tautomeric, zwitterionic, and compact structure (Figure 3s in the Supporting Information) upon optimization at higher levels of theory. This structure is 4.2 and 5.6 kcal/mol higher in energy than the **bk-c-zi** global minimum, with the HF/6-31G and B3LYP/6-31G methods, respectively.

$[BK + H]^+$. Conceptually, $[BK + H]^+$ can be formed by protonating **bk-c-zi** (Figure 2). There are two highly basic sites for protonation: the C-terminal carboxylate anion, which carries a formal negative charge, and the guanidine group of Arg[1]. Protonation on the carboxylate anion leads to nonzwitterionic structures that are relatively high in energies at MMFF and HF/3-21G levels. As the preference of an ion−zwitterion form for $[BK + H]^+$ has been proven both experimentally[4,5,8,19,20] and theoretically,[5,19,20,24] we did not examine nonzwitterionic structures further. Protonation on Arg[1] leads to salt-bridge tautomers, with the carboxylic group deprotonated and guanidine groups of Arg residues protonated, similar to those predicted earlier by Strittmatter and Williams[24] using the MMFF/DFT approach.

The global minimum structure predicted in this study for $[BK + H]^+$ with the MMFF/HF technique is shown in Figure 3. The same defining structural characteristics of extensive hydrogen-bonding network, compact structure, and the β-turn that are featured prominently in neutral BK are again apparent. As in the structure of BK, the oxygen atoms of the carboxylate anion form bifurcated hydrogen bonds. Details of the hydrogen-bond network, however, differ. The defining characteristic in $[BK + H]^+$ is the prominent salt bridge that spans the protonated

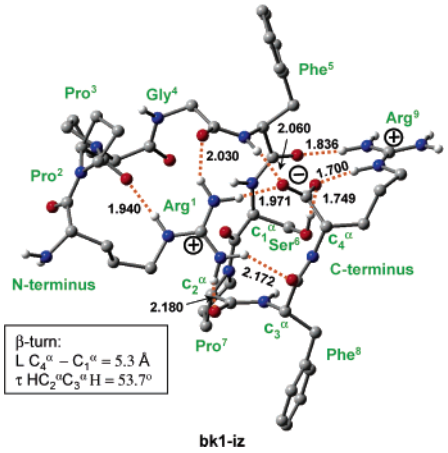


Figure 3. The lowest-energy structure of $[BK + H]^+$, **bk1-iz**, predicted with the HF/6-31G method. For clarity, only hydrogen atoms of importance to the structure are shown. Hydrogen bonds are illustrated by dashed orange lines. The distances are in angstroms. For the β-turn, α-carbon atoms are denoted.

guanidine group of Arg^1, the carboxylate anion, and the secondary amino group of protonated Arg^9 plus the hydroxy group of Ser^6 and the amide nitrogen of Phe^5. A total of nine hydrogen bonds provide a compact globelike shape. Protonation decreases the number of hydrogen bonds (there are 12 in BK) but clearly reduces the lengths of hydrogen bonds, thereby increasing their strengths.[45] Three short hydrogen bonds (1.700, 1.749, and 1.836 Å) in the vicinity of the carboxylate group are predicted with the HF/6-31G method.

The $[BK + H]^+$ structure exhibits a β-turn in the Ser^6-Pro^7-Phe^8-Arg^9 sequence with, however, no stabilizing hydrogen bond in this section of the conformer. Protonation does not change the distance between the C_1^α and C_4^α atoms (5.3 Å) and only slightly reduces the dihedral angle τ $HC_2^\alpha C_3^\alpha H$ (53.7°). The global minimum zwitterionic structure of $[BK + H]^+$ predicted earlier[24] at the MMFF/DFT level is reminiscent of our **bk1-iz** structure obtained with MMFF/HF. The main difference is that, in the former, the hydroxyl group of Ser^6 interacts with the carbonyl oxygen of Phe^5 rather than with a carboxylate oxygen. We reoptimized the MMFF/DFT zwitterionic structure[24] with the Hartree−Fock method. The fully optimized conformer is only 3.0 kcal/mol higher in energy than **bk1-iz** at the HF/6-31G level.

$[BK + 2H]^{2+}$. The protonation of $[BK + H]^+$ to form $[BK + 2H]^{2+}$ is a more challenging situation. The second positive net charge introduces Coulombic repulsion into the system and, at the same time, increases the strength of hydrogen bonds. For example, the intermolecular hydrogen bonding of charged systems in the $N-H^+-O$ and $O-H^+-O$ bridges spans the range of 18−32 kcal/mol versus 5−10 kcal/mol for neutral systems.[45]

To form $[BK + 2H]^{2+}$, several probable protonation sites on $[BK + H]^+$ were considered. The two most likely sites are the C-terminal carboxylate anion and the N-terminal amino nitrogen. Protonation of the carboxylate anion results in nonzwitterionic (ion−molecular) structures (vide infra). In this case, the two positive charges reside on the two termini; this arrangement provides maximum separation and, therefore, minimum Coulombic repulsion, while losing the extra binding due to Coulombic attraction in a salt bridge. By contrast, protonation of the N-terminal amino nitrogen produces ion−zwitterion structures, preserves the salt bridge, and provides more opportunities to form strong hydrogen bonding, while leaving the formal positive charge sites closer to each other. The MMFF/HF/DFT

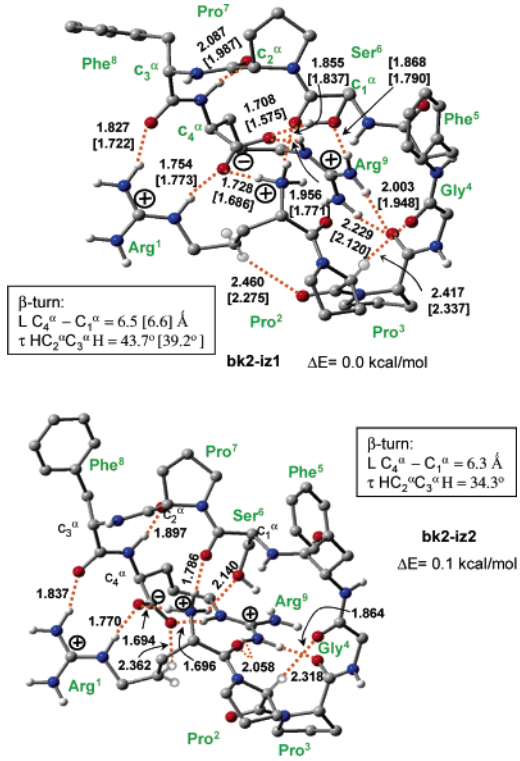


Figure 4. The lowest-energy ion−zwitterion structures of $[BK + 2H]^{2+}$, **bk2-iz1** and **bk2-iz2**, predicted with the HF/6-31G and B3LYP/6-31G (values in square brackets) methods. For clarity, only hydrogen atoms of importance to the structure are shown. Hydrogen bonds are illustrated by dashed orange lines. The distances are in angstroms. For the β-turn, α-carbon atoms are denoted.

approach favors the ion−zwitterion structures containing a salt bridge. The two lowest-energy ion−zwitterion conformers, **bk2-iz1** and **bk2-iz2**, predicted for $[BK + 2H]^{2+}$ are shown in Figure 4. The extensive hydrogen-bonding network features salt bridges connecting the C-terminal carboxylate anion and the secondary amino groups of protonated Arg^1 and Arg^9. This arrangement provides a separation between the formal positively charged sites (carbon atoms of the guanidine groups) of 7.7 Å with the B3LYP/6-31G method. Separate branches connect the carboxylate anion with the hydroxyl group of Ser^6 and also with the protonated N-terminal amino group, which in turn is hydrogen-bonded with the carbonyl oxygen of Ser^6. Other hydrogen bonds span between guanidine hydrogen atoms and carbonyl oxygen atoms on the peptide backbone. For the global minimum structure, **bk2-iz1**, a total of 12 hydrogen bonds affords a compact, globular structure that features a β-turn. It is of note that, in addition to conventional N−H−O and O−H−O hydrogen bonding, two $C^\alpha-H\cdots O{=}C$ interactions also participate in the folding of $[BK + 2H]^{2+}$. It has been shown independently that the $C^\alpha-H\cdots O{=}C$ interaction, with an association enthalpy of ca. 3.5 kcal/mol, is an important factor in the folding of biosystems.[49] The second protonation causes a further decrease in hydrogen bond lengths; in **bk2-iz1**, six of them are shorter than 1.9 Å with the HF/6-31G method, and seven with B3LYP/6-31G. The nearest-energy conformer, **bk2-iz2**, is only 0.1 kcal/mol higher in energy. As in **bk2-iz1**, it has a very compact structure that is stabilized by 11 hydrogen bonds and incorporates a β-turn. In comparison with the β-turns in BK and $[BK + H]^+$, the one in $[BK + 2H]^{2+}$ displays a minor decrease in dihedral angle (**bk2-iz1**, τ $HC_2^\alpha C_3^\alpha H = 43.7°$) but a major increase in the distance between the C_1^α and C_4^α atoms (L $C_1^\alpha - C_4^\alpha = 6.5$ Å).

TABLE 3: Harmonic Vibrational Frequency Analyses at the B3LYP Level for Some Conformers of Arginine and Protonated Arginine Using the 6-31G and 6-31++G(d,p) Basis Sets and for the Lowest-Energy Structures of Zwitterionic and Nonzwitterionic Forms of $[BK + 2H]^{2+}$, bk2-iz1 and bk2-im1, Respectively, Using the 6-31G Basis Set[a]

		arginine				protonated arginine		$[BK + 2H]^{2+}$	
basis set		**c5**	**c2**	**z3**	**z2**	**p1**	**p2**	**bz2-iz1**	**bk2-im1**
6-31G	ΔE	0.0	3.61	−6.98	−3.64	0.0	3.40	0.0	25.5
	ΔE_{ZPVE}	0.0	4.03	−5.88	−2.58	0.0	3.37	0.0	23.0
	ΔH_{298K}	0.0	4.08	−6.27	−2.75	0.0	3.55	0.0	24.0
	ΔG_{298K}	0.0	3.28	−4.97	−2.40	0.0	2.90	0.0	20.6
6-31++G(d,p)	ΔE	0.0	1.83	1.81	4.29	0.0	2.73	0.0	13.3
	ΔE_{ZPVE}	0.0	2.15	1.12	3.98	0.0	2.65		
	ΔH_{298K}	0.0	2.12	1.08	3.97	0.0	2.82		
	ΔG_{298K}	0.0	1.69	1.72	4.08	0.0	2.25		
	$\Delta ZPVE_{BS}$[b]	0.42	0.52	2.21	1.79	2.02	2.03		

[a] The energies are in kcal/mol. [b] $\Delta ZPVE_{BS}$ = ZPVE(6-31G) − ZPVE(6-31++G(d,p)), the difference of absolute zero-point vibrational energies obtained by the two basis sets.

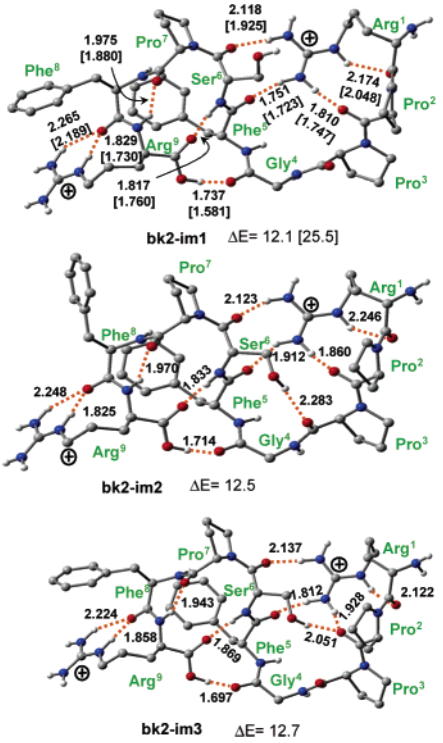


Figure 5. The lowest-energy ion−molecule structures of $[BK + 2H]^{2+}$, **bk2-im1**, **bk2-im2**, and **bk2-im3**, predicted with the HF/6-31G and B3LYP/6-31G (values in square brackets) methods. For clarity, only hydrogen atoms of importance to the structure are shown. Hydrogen bonds are illustrated by dashed orange lines. The distances are in angstroms. For the β-turn, α-carbon atoms are denoted.

As discussed above, protonating the C-terminal carboxylate anion of $[BK + H]^+$ results in nonzwitterionic structures. The three lowest-energy conformers, **bk2-im1**, **bk2-im2**, and **bk2-im3**, predicted with the MMFF/HF method all feature compact structures with extensive hydrogen-bonding networks and β-turns (Figure 5). For **bk2-im1**, the structure is stabilized by nine hydrogen bonds; five of them are shorter than 1.9 Å with HF and DFT methods. The formal sites of the positive charges (carbon atoms of the guanidine groups) are separated by 14.1 Å, at both the HF/6-31G and the B3LYP/6-31G levels of theory. The **bk2-im2** conformer has 10 hydrogen bonds; four of them are shorter than 1.9 Å. For **bk2-im3**, there are nine hydrogen bonds (four of them are short). These tautomers are ca. 12 kcal/mol higher in electronic energy than the lowest-energy conformer of the zwitterionic tautomer, **bk2-iz1**, at the HF/6-31G level. At the B3LYP/6-31G level, the energy preference for the **bk2-iz1** zwitterionic structure, relative to **bk2-im1**, increases to 25.5 kcal/mol. Adding diffuse and polarization functions to the 6-31G basis set should decrease the large energy gap (as in arginine) but is unlikely to change the relative order of the zwitterionic and nonzwitterionic forms because of the large difference in energy. To verify this, the geometries of the lowest-energy structures for **bk2-iz1** and **bk2-im1**, were further optimized at the B3LYP/6-31++G(d,p) level. As expected, the energy gap decreases, but **bk2-iz1** is still lower in energy than **bk2-im1** by 13.3 kcal/mol. The larger basis set does not change the structures significantly (comparing **bk2-iz1** in Figure 4 and **bk2-im1** in Figure 5 with those in Figure 4s in the Supporting Information). The short and long hydrogen bonds are also correctly reproduced, as they are in the case of arginine (Table 1s).

The thermal and entropic effects might also be significant factors in altering the relative energy. To appreciate these effects, harmonic vibrational frequency analyses at the B3LYP level were performed for some conformers of arginine and its protonated form using both the 6-31G and the 6-31++G(d,p) basis sets, and for **bk2-iz1** and **bk2-im1** using the 6-31G basis set. These results are shown in Table 3. In general, the zero-point vibrational energy (ZPVE) correction and thermal correction at 298 K have only a small effect on the relative stability for both arginine and protonated arginine. Evidently, the conformers have the same covalent backbones and differ only in the relatively weak van der Waals interactions and hydrogen bonding, both of which make relatively small contributions to the vibrational energy. Interestingly, ZPVE slightly disfavors the zwitterionic structures (**z3** and **z2**) using the smaller basis set, 6-31G; their energies without ZPVE (ΔE, relative to the canonical structure **c5**) are smaller than those with ZPVE (ΔE_{ZPVE}). However, it favors the zwitterionic structures using the larger basis set, 6-31++G(d,p), with ΔE being larger than ΔE_{ZPVE}. Again, this is a consequence of the fact that hydrogen bonding estimated with the 6-31G basis set is stronger than that with 6-31++G(d,p). This is evident from the difference in absolute ZPVE calculated with the 6-31G and 6-31++G(d,p) basis sets. Thus structures with stronger hydrogen bonds, zwitterionic and protonated, have larger differences in ZPVE obtained with the two basis sets. Similarly, the entropic effects, which manifest in the ΔG values, are also insignificant. Entropies calculated with 6-31++G(d,p) are slightly smaller than those with 6-31G. As expected, zwitterionic structures have lower entropies than canonical ones as the former structures are more compact as a result of stronger hydrogen bonding.

For $[BK + 2H]^{2+}$, frequency analyses were performed only with the 6-31G basis set. Similar to arginine, ZPVE and thermal correction are unimportant even for such a relatively large peptide. The entropic effect is considerably larger for **bz2-iz1** as it is more rigid with the charges being localized at the middle of the molecule. By contrast, **bk2-im1** is relatively flexible with the two formal charges being separated far away from each other. Results for arginine, the test model, strongly suggest that the combined ZPVE, thermal, and entropic effects for $[BK + 2H]^{2+}$ calculated with the 6-31G basis set at ~5 kcal/mol can be used as an upper limit for these same effects calculated with 6-31++G(d,p). Thus, the relative free energy of **bk2-im1** versus **bk2-iz1** at the B3LYP/6-31++G(d,p) level of theory is at a minimum of 8 kcal/mol (which gives an equilibrium constant, K, of 1×10^{-6} at 298 K). Evidently, the lowest-energy zwitterionic structure of $[BK + 2H]^{2+}$ is more favorable than its lowest-energy canonical structure.

The nitrogen atom of the proline residue has been proposed by some as a competitive site for protonation.[21] This possibility, however, is not borne out by our MMFF results. Visual analysis of the structures produced by proline protonation showed that the protonated proline ring is not involved in any hydrogen-bonding network, thereby significantly reducing structural stability. Consequently, further optimization at the ab initio level was not performed on the proline protonated tautomers.

$[BK + 3H]^{3+}$. Conceptually, additional protonation of $[BK + 2H]^{2+}$ can result in the third proton being on the C-terminal carboxylate anion, giving a nonzwitterionic structure, or on a formally neutral functional group, giving a zwitterionic structure that contains four formally positively charged centers and one negatively charged one. Optimization revealed that the nonzwitterionic tautomers (with a neutral C-terminal carboxylic acid group) are much lower in energy than the zwitterionic ones and that the tautomer with the lowest energy is not necessarily the one with protons on the guanidine nitrogens of Arg^1 and Arg^9, and the N-terminal amino group of Arg^1. There is little doubt that, because of their high gas-phase basicities, the two guanidine groups will be protonated. It is the location of the third proton that needs to be determined.

Conformers of the $[BK + 3H]^{3+}$ tautomers that are protonated at the N-terminal amino group or one of the carbonyl groups of Pro^3, Gly^4, Phe^5, Ser^6, and Pro^7 were studied with MMFF/HF. Among them, a conformer, **bk3-Phe^{5}H**, that bears the third proton on the carbonyl oxygen of Phe^5 is the lowest in energy; its structure is shown in Figure 6. This $[BK + 3H]^{3+}$ ion structure features an elongated polypeptide chain that is "kinked" with seven hydrogen bonds, three of them shorter than 1.9 Å. The protonated Phe^5 carbonyl oxygen forms an exceptionally short (1.542 Å) hydrogen bond with the hydroxyl group of Ser^6 that provides additional stabilization. It is of note that in this extended, nonzwitterionic structure the three formally positively charged groups reside at the two termini and in the middle of the molecular framework, thereby minimizing electrostatic repulsion. The nearest-in-energy tautomer, **bk3-Gly^{4}H**, which is 2.2 kcal/mol higher in energy than **bk3-Phe^{5}H**, bears the third proton on the carbonyl oxygen of Gly^4. The protonated carbonyl oxygen forms a short (1.623 Å) hydrogen bond with the carbonyl group of Pro^3. The elongated structure of **bk3-Gly^{4}H** is stabilized with five hydrogen bonds. These theoretical results are in accord with cross-section measurements that showed $[BK + 3H]^{3+}$ having a much larger cross-section than $[BK + 2H]^{2+}$ and $[BK + H]^{+}$.[11,33] These results indicate unfolding of the polypeptide with increasing charge from +2 to +3. Despite unfolding, a β-turn with a strongly twisted backbone (τ $HC_2^{\alpha}C_3^{\alpha}H = 168.1°$, L $C_1^{\alpha}-C_4^{\alpha}$ = 6.5 Å) still exists in the Ser^6-Pro^7-Phe^8-Arg^9 sequence.

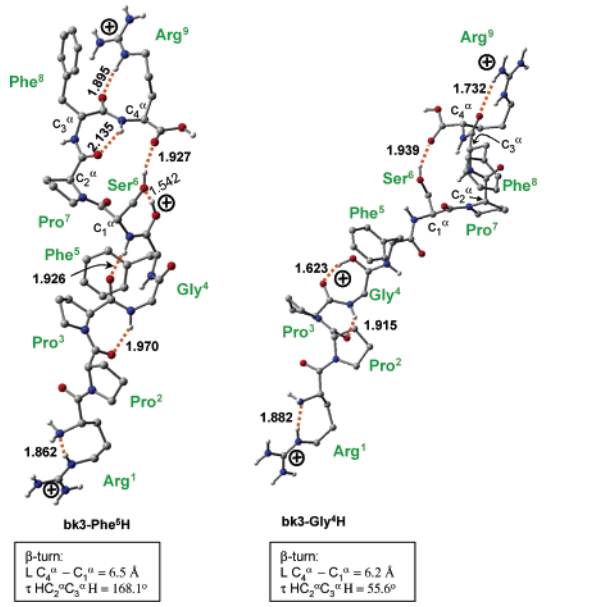


Figure 6. The lowest-energy structure of $[BK + 3H]^{3+}$, **bk3-Phe^{5}H**, and the nearest-in-energy tautomer, **bk3-Gly^{4}H**, predicted with the HF/6-31G method. For clarity, only hydrogen atoms of importance to the structure are shown. Hydrogen bonds are illustrated by dashed orange lines. The distances are in angstroms. For the β-turn, α-carbon atoms are denoted.

TABLE 4: Experimental and Population-Weighted Theoretical Cross-Sections (Å^2) of $[BK + H]^{+}$, $[BK + 2H]^{2+}$, and $[BK + 3H]^{3+}$ at the HF/6-31G Level Calculated Using Projection Approximation (PA), Exact Hard-Spheres Scattering (EHSS), and Trajectory Calculations (TC)[a]

isomer/ conformer	cross-sections			
	experimental	PA	EHSS	TC
$[BK + H]^{+}$	245 ± 3[b]; 239[c]; 248 ± 6[d]	241	274	264
$[BK + 2H]^{2+}$	255[b]; 240[c]; 242 ± 4[d]	244	278	266
	Zwitterionic			
bk2-iz1		245/*245*[e]	278/*279*	269/*272*
bk2-iz2		243	277	263
	Ion−Molecular			
bk2-im1		247/*250*[e]	279/*284*	271/*277*
bk2-im2		244	276	268
bk2-im3		242	275	266
$[BK + 3H]^{3+}$	284[c]; 293 ± 5[d]	281	316	312

[a] Theoretical cross-sections for individual isomers of $[BK + 2H]^{2+}$ are also shown. [b] From ref 11. [c] From ref 10. [d] From ref 33. [e] Italicized values were obtained from optimized B3LYP/6-31++G(d,p) structures.

Collision Cross-Sections. As collision cross-sections of BK ions have been measured and are available,[10,11,33] support of the structures predicted in this work with the HF method will be provided, if there is agreement between theoretical cross-sections calculated from the proposed structures and experimental cross-sections. The comparisons are shown in Table 4. The theoretical cross-sections are population-weighted averages of cross-sections of individual conformers at 298 K, calculated from the HF/6-31G geometries. Cross-sections of individual conformers for $[BK + 2H]^{2+}$ are also given as an illustration. For **bk2-iz1** and **bk2-im1**, cross-sections of the optimized structures calculated at the B3LYP/6-31++G(d,p) level were also given for comparison. At 298 K, only those conformers that have energies not more than 2 kcal/mol above that of the lowest-energy conformer will make significant contributions to

TABLE 5: Electronic Energies, E_{el}, (in hartree) Predicted for the Lowest-Energy Structures of BK, $[BK + H]^+$, $[BK + 2H]^{2+}$, and $[BK + 3H]^{3+}$ with the HF/6-31 and B3LYP/6-31G Methods[a]

		HF/6-31G		B3LYP/6-31G	
species	structure	E_{el}	PrE	E_{el}	PrE
BK	**bk-c-zi**	−3574.86513	282.3	−3597.22158	291.4
$[BK + H]^+$	**bk1-iz**	−3575.31494	247.9	−3597.68588	235.6
$[BK + 2H]^{2+}$	**bk2-iz1**	−3575.71006	169.8	−3598.06129[b]	n/a
$[BK + 3H]^{3+}$	**bk3-Phe5H**	−3575.98062		n/a	

[a] Protonation energies, PrE, ($PrE = E_{el}[BK + nH]^n + - E_{el}[BK + (n + 1)H]^{(n+1)+}$, $n = 0, 1, 2$) are reported in kcal/mol. [b] −3599.31328 at B3LYP/6-31++G(d,p).

the population. Furthermore, the cross-sections of the conformers within that 2 kcal/mol range are all comparable. Thus, calculating the theoretical cross-section from only the structure of the lowest-energy conformer would have introduced a negligible error. Table 4 shows three types of theoretical cross-sections, depending on the calculation methods, projection approximation (PA),[50] exact hard-spheres scattering (EHSS),[51] and trajectory calculations (TC).[52] The theoretical cross-sections are in general agreement with literature experimental values; the $[BK + 2H]^{2+}$ ion is comparable to $[BK + H]^+$ in cross-section, while $[BK + 3H]^{3+}$ is approximately 15% larger in cross-section. The difference in the calculated cross-sections of the zwitterionic and nonzwitterionic forms of $[BK + 2H]^{2+}$ is ca. 1%; thus they could not be distinguished even under the best experimental conditions. In addition, the calculated cross-sections from structures optimized using HF/6-31G and B3LYP/6-31++G(d,p) differ by ≤2%. Experimental cross-sections are typically between theoretical cross-sections calculated using the projection[50] and trajectory[52] methods. The second method is typically considered to be the most rigorous.[52]

Protonation Energies. We define the protonation energy (PrE) of BK and its protonated forms as the difference between electronic energies (E_{el}): $PrE = E_{el}[BK + nH]^{n+} - E_{el}[BK + (n + 1)H]^{(n+1)+}$, where $n = 0$, 1, or 2. PrE can be used as an estimate for proton affinity (PA) and gas-phase basicity (GB). PrE differs from PA only by the lack of the relative ZPVE, which is a systematic value reflecting principally the contribution of harmonic vibrational frequencies due to the attached proton. A previous theoretical study on the PAs of methyl esters of N-acetylated amino acids shows that the PAs are lower than the PrEs by ca. 7 kcal/mol.[53] GBs that include ZPVE, thermal, and entropy corrections are ca. 7−11 kcal/mol lower in energy than PAs.[54]

Electronic energies of the lowest-energy structures of BK and its protonated forms, **bk-c-zi**, **bk1-iz**, **bk2-iz1**, and **bk3-Phe^{5}H**, are listed in Table 5 along with the calculated protonation energies. The HF/6-31G results show a decrease in PrE from BK to $[BK + H]^+$ of 34.4 kcal/mol. This decrease in PrE approximately doubles to 78.1 kcal/mol from $[BK + H]^+$ to $[BK + 2H]^{2+}$. As the PrE of $[BK + 2H]^{2+}$ at 169.8 kcal/mol is already quite low, $[BK + 3H]^{3+}$ should be the highest observable charge state of BK, in accordance with observation.

The B3LYP/6-31G method yields a larger decrease of 55.7 kcal/mol in PrE from BK to $[BK + H]^+$. The PrE of $[BK + H]^+$ is 235.6 kcal/mol. Through the use of the relative values between PrE, PA, and GB in ref 53, the GB of $[BK + H]^+$ is approximately 218−222 kcal/mol, which is in agreement with the published experimental apparent GB of 225.8 ± 4.2 kcal/mol determined via bracketing[20] and GB of 217.8 ± 1.7 kcal/mol after incorporating the reverse activation energy barrier[3].

Conclusions

BK, $[BK + H]^+$, and $[BK + 2H]^{2+}$ prefer the zwitterionic form and have compact globular shapes. $[BK + 3H]^{3+}$ exists in an ion−molecule form and adopts an elongated shape. For neutral BK, the canonical form was only located at the MMFF level; at the HF and DFT levels, barrierless proton transfer resulted in zwitterionic forms. For $[BK + H]^+$ and $[BK + 2H]^{2+}$, the ion−molecule forms were located by the MMFF/HF/DFT technique as higher-energy tautomers. For $[BK + 2H]^{2+}$, the lowest ion−molecular tautomer is 12.1, 25.5, and 13.3 kcal/mol higher in energy than the zwitterionic global minimum structure, at the HF/6-31G, B3LYP/6-31G, and B3LYP/6-31++G(d,p) levels of theory, respectively. Including zero-point, thermal, and entropic corrections lowers the energy of the ion−molecule form by a maximum of only 5 kcal/mol relative to the zwitterionic form.

The lowest-energy zwitterionic form of BK has the carboxylic group deprotonated and the guanidine group of Arg^9 protonated. The Arg^9 of BK can be considered as a zwitterionic arginine "solvated" by the long polypeptide chain. The lowest-energy structures of $[BK + H]^+$ and $[BK + 2H]^{2+}$ contain salt bridges. In both ions, the oxygen atoms of the carboxylate anion, the central unit of the salt bridge, form bifurcated hydrogen bonds. In $[BK + H]^+$, the salt bridge spans the carboxylate anion, the amino group of protonated Arg^1, and the secondary amino group of protonated Arg^9, as well as the hydroxy group of Ser^6 and the amide nitrogen of Phe^5. For $[BK + 2H]^{2+}$, the carboxylate anion connects the secondary amino groups of protonated Arg^1 and Arg^9. This arrangement increases the distance between positively charged sites (formally, carbon atoms of the guanidine groups). Separate branches link the carboxylate anion with the hydroxyl group of Ser^6 and also with the protonated N-terminal amino group, which in turn is hydrogen-bonded with the carbonyl oxygen of Ser^6.

The obvious stabilizing effect of a salt bridge is the Coulombic attraction within the "+ − +" arrangement. The stronger hydrogen bonding bridged by the proton (A−H$^+$−B) relative to that bridged by the hydrogen atom (A−H···B) also contributes to stabilization of multiply protonated zwitterions. Despite significant difference in energies, ion−molecule and ion−zwitterion species possess surprisingly similar compact structures and sizes. The tight "knot" of the salt bridge with very short hydrogen bonds apparently is compensated by loose hydrogen binding for the rest of the polypeptide chain. This results in an ion−zwitterion species whose size is effectively the same as one provided by the hydrogen-bonding network of moderate-length bonds in the ion−molecule species. There is also not much difference between the sizes of $[BK + H]^+$ and $[BK + 2H]^{2+}$; the binding energies of the hydrogen-bonding network and the salt bridge are apparently sufficient to compensate for the increase in Coulombic repulsion upon the second protonation.

For $[BK + 3H]^{3+}$, no ion−zwitterionic forms were located. Electrostatic repulsion is the major factor that determines the unfolded structure. The protonation sites are the terminal guanidine groups and the carbonyl oxygen of Phe^5. This arrangement provides the maximum distances between the charges.

The common structural element in BK, $[BK + H]^+$, $[BK + 2H]^{2+}$, and $[BK + 3H]^{3+}$ is the β-turn in the Ser^6-Pro^7-Phe^8-Arg^9 sequence. Geometrical parameters of the β-turn predicted with the HF/6-31G method are: a dihedral angle, τ $HC_2^{\alpha}C_3^{\alpha}H$, of 56.8°, 53.7°, 39.2°, and 168.1°, and $C_1^{\alpha}-C_4^{\alpha}$ distance of

5.3, 5.3, 6.5, and 6.5 Å for BK, $[BK + H]^+$, $[BK + 2H]^{2+}$, and $[BK + 3H]^{3+}$, respectively.

The relative energies of the conformers depend on the level of theory as well as on the size and the flexibility of the basis set. The geometries appear to be less sensitive to the theoretical method. First-principles methods, HF and DFT, reveal similar trends. Our results for ion−zwitterion structures are supported by experimental results: a consensus exists for the $[BK + H]^+$ structure; the β-turn is supported by results in solution; the protonation energy of $[BK + H]^+$ is in accord with apparent gas-phase basicity measured via bracketing; and calculated collision cross-sections are in agreement with published cross-sections.

Acknowledgment. This article is dedicated to the memory and in honor of C.F.R., who passed away suddenly on December 11, 2005, in Lake Charles, Louisiana. Financial support of this research was provided by the Natural Sciences and Engineering Research Council of Canada. Funding for computational infrastructure was provided by the Canadian Foundation of Innovation and the Ontario Innovation Trust. We thank Professor Martin F. Jarrold for making available to us his collision cross-section programs and Dr. Eric F. Strittmatter for providing the MMFF/DFT coordinates of $[BK + H]^+$. In addition, we thank Professor Zhifeng Liu and Ms. Ka-Wai Chan for helpful discussions.

Supporting Information Available: Details of hydrogen bonding in neutral and protonated arginine; MMFF, HF/3-21G, and HF/6-31G energies of BK, $[BK + H]^+$, ion−molecule conformers of $[BK + 2H]^{2+}$, zwitterionic conformers of $[BK + 2H]^{2+}$, and $[BK + 3H]^{3+}$; optimized structures of neutral and protonated arginine, structures of BK; and structures of $[BK + 2H]^{2+}$ (**bk2-iz1** and **bk2-im1**) optimized at B3LYP/6-31++G(d,p). This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 9765−9766.

(2) Schnier, P. D.; Gross, D. S.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1086−1097.

(3) Kaltashov, I. A.; Fabris, D.; Fenselau, C. C. *J. Phys. Chem.* **1995**, *99*, 10046−10051.

(4) (a) Wyttenbach, T.; von Helden, G.; Bowers, M. T. *J. Am. Chem. Soc.* **1996**, *118*, 8355−8364. (b) von Helden, G.; Wyttenbach, T.; Bowers, M. T. *Science* **1995**, *267*, 1483−1485.

(5) (a) Schnier, P. D.; Price, W. D.; Jockusch, R. A.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 7178−7189. (b) Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859−866.

(6) Szilágyi, Z.; Drahos, L.; Vékey, K. *J. Mass Spectrom.* **1997**, *32*, 689−696.

(7) Zhang, Z.; Guan, S.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 659−670.

(8) Countermann, A. E.; Valentine, S. J.; Srebalus, C. A.; Henderson, S. C.; Hoaglund, C. S.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 743−759.

(9) Green, M. K.; Lebrilla, C. B. *Int. J. Mass Spectrom. Ion Processes* **1998**, *175*, 15−26.

(10) Wyttenbach, T.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 9−14.

(11) Schaaff, T. G.; Stephenson, J. L., Jr.; McLuckey, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 8907−8919.

(12) Freitas, M. A.; Marshall, A. G. *Int. J. Mass Spectrom.* **1999**, *182*/*183*, 221−231.

(13) Butcher, D. J.; Asano, K. G.; Goeringer, D. E.; McLuckey, S. A. *J. Phys. Chem. A* **1999**, *103*, 8664−8671.

(14) Freitas, M. A.; Hendrickson, C. L.; Marshall, A. G. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1639−1642.

(15) Freitas, M. A.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Chem. Soc.* **2000**, *122*, 7768−7775.

(16) Gimon-Kinsel, M. E.; Barbacci, D. C.; Russell, D. H. *J. Mass Spectrom.* **1999**, *34*, 124−136.

(17) Schaaff, T. G.; Stephenson, J. L., Jr.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1999**, *11*, 167−171.

(18) Schnier, P. D.; Jurchen, J. C.; Williams, E. R. *J. Phys. Chem. B* **1999**, *103*, 737−745.

(19) Gill, A. C.; Jennings, K. R.; Wyttenbach, T.; Bowers, M. T. *Int. J. Mass Spectrom.* **2000**, *195*/*196*, 685−697.

(20) Ewing, N. P.; Pallante, G. A.; Zhang, X.; Cassady, C. J. *J. Mass Spectrom.* **2001**, *36*, 875−881.

(21) Pallante, G. A.; Cassady, C. J. *Int. J. Mass Spectrom.* **2002**, *219*, 115−131.

(22) Levy-Seri, E.; Koster, G.; Kogan, A.; Gutman, K.; Reuben, B. G.; Lifshitz, C. *J. Phys. Chem. A* **2001**, *105*, 5552−5559.

(23) (a) Woenckhaus, J.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119*, 9586−9587. (b) Shelimov, K. B.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119*, 2987−2994. (c) Lee, S.-W.; Freivogel, P.; Schindler, T.; Beauchmap, J. L. *J. Am. Chem. Soc.* **1998**, *120*, 11758−11765. (d) Jarrold, M. F. *Acc. Chem. Res.* **1999**, *32*, 360. (e) Kohtani, M.; Jarrold, M. F. *J. Am. Chem. Soc.* **2002**, *124*, 11148−11158.

(24) Strittmatter, E. F.; Williams, E. R. *J. Phys. Chem. A* **2000**, *104*, 6069−6076.

(25) (a) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1995**, *117*, 8159−8170. (b) Kassab, E.; Langlet, J.; Evleth, E.; Akacem, Y. *J. Mol. Struct. (THEOCHEM)*, **2000**, *531*, 267−282. (c) Yamabe, S.; Ono, N.; Tsuchida, N. *J. Phys. Chem. A* **2003**, *107*, 7915−7922.

(26) (a) Price, W. D.; Jockusch, R. A.; Williams, E. R. *J. Am. Chem. Soc.* **1997**, *119*, 11988−11989. (b) Wyttenbach, T.; Bushnell, J. E.; Bowers, M. T. *J. Am. Chem. Soc.* **1998**, *120*, 5098−5103. (c) Wyttenbach, T.; Matthias, W.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 3458−3464. (d) Wyttenbach, T.; Matthias, W.; Bowers, M. T. *Int. J. Mass Spectrom.* **1999**, *183*, 243−252.

(27) (a) Chapo, C. J.; Paul, J. B.; Provencal, R. A.; Roth, K.; Saykally, R. J. *J. Am. Chem. Soc.* **1998**, *120*, 12956−12957. (b) Jockusch, R. A.; Price, W. D.; Williams, E. R. *J. Phys. Chem. A* **1999**, *103*, 9266−9274. (c) Strittmatter, E. F.; Lemoff, A. S.; Williams, E. R. *J. Phys. Chem. A* **2000**, *104*, 9793−9796. (d) Gutowski, M.; Skurski, P.; Simons, J. *J. Am. Chem. Soc.* **2000**, *122*, 10159−10162. (e) Julian, R. R.; Beauchamp, J. L.; Goddard, W. A., III. *J. Phys. Chem. A* **2002**, *106*, 32−34. (f) Lemoff, A. S.; Bush, M. F.; Williams, E. R. *J. Am. Chem. Soc.* **2003**, *125*, 13576−13584. (g) Wyttenbach, T.; Paizs, B.; Barran, P.; Breci, L.; Liu, D.; Suhai, S.; Wysocki, V. H.; Bowers, M. T. *J. Am. Chem. Soc.* **2003**, *125*, 13768−13775.

(28) Kato, H.; Suzuki, T. *Biochemistry* **1971**, *10*, 972−980.

(29) (a) Volpe, A. R.; Fontecchio, G.; Carmignani, M. *Immunopharmacology* **1999**, *44*, 87−92. (b) Volpe, A. R.; Giardina, B.; Preziosi, P.; Carmignani, M. *Immunopharmacology* **1996**, *33*, 287−290.

(30) Abbas, S. A.; Sharma, J. N.; Yusof, A. P. M. *Immunopharmacology* **1999**, *44*, 93−98.

(31) Stewart, J. M.; Gera, L.; Chan, D. C.; Bunn, P. A., Jr.; York, E. J.; Simkeviciene, V.; Helfrich, B. *Can J. Physiol. Pharmacol.* **2002**, *80*, 275−280.

(32) (a) Young, J. K.; Hicks, R. P. *Biopolymers* **1994**, *34*, 611. (b) Cann, J. R.; Vatter, A.; Vavrek, R. J.; Stewart, J. M. *Peptides* **1986**, *7*, 1121−1130. (c) Cann, J. R.; London, E. R. Unkefer, C. J.; Vovrek, R. J.; Stewart, J. M. *Int. J. Pept. Protein Res.* **1987,** *29*, 486−496. (d) Lee, S. C.; Russel, A. F.; Laidig, W. D. *Int. J. Pept. Protein Res.* **1990,** *35*, 367−377.

(33) Guo, Y.; Wang, J.; Javahery, G.; Thomson, B. A.; Siu, K. W. M. *Anal. Chem.* **2005**, *77*, 266−275.

(34) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908−5920.

(35) (a) Wu, Y.-D.; Zhao, Y.-L. *J. Am. Chem. Soc.* **2001**, *123*, 5313−5319. (b) Masunov, A.; Dannenberg, J. J. *J. Phys. Chem. B* **2000**, *104*, 806−810. (c) Zhao, Y.-L.; Wu, Y.-D. *J. Am. Chem. Soc.* **2002**, *124*, 1570−1571.

(36) *SPARTAN*, version 5.0; Wavefunction, Inc.: Irvine, CA.

(37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.11; Gaussian, Inc.: Pittsburgh, PA, 1998.

(38) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939−947.

(39) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257−2261.
(40) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(41) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(42) Rak, J.; Skurski, P.; Simons, J.; Gutowski, M. *J. Am. Chem. Soc.* **2001**, *123*, 11695−11707.
(43) Gdanitz, R. J.; Cardoen, W.; Windus, T. L.; Simons, J. *J. Phys. Chem. A* **2004**, *108*, 515−518.
(44) Lii, J. H.; Ma, B.; Allinger, N. L. *J. Comput. Chem.* **1999**, *20*, 1593−1603.
(45) Espinosa, E.; Molins, E.; Lecomte, C. *Chem. Phys. Lett.* **1998**, *285*, 170−173.
(46) Humbel, S. *J. Phys. Chem. A* **2002**, *106*, 5517−5520.
(47) Lukin, O.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 6775−6782.
(48) (a) Venkatachalam, C. *Biopolymers* **1968**, *6*, 1425. (b) Perzel, A.; McAllister, M. A.; Császár, P.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1993**, *115*, 4849−4858.
(49) Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. *J. Am. Chem. Soc.* **2000**, *122*, 4750−4755.
(50) Von Helden, G.; Hsu, M. T.; Gotts, N. G.; Bowers, M. T. *J. Phys. Chem.* **1993**, *97*, 8182−8192.
(51) Shvartsburg, A. A.; Jarrold, M. F. *Chem. Phys. Lett.* **1996**, *261*, 86−91.
(52) Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *J. Phys. Chem.* **1996**, *100*, 16082−16086.
(53) Addario, V.; Guo, Y.; Chu, I. K.; Ling, Y.; Ruggerio, G.; Rodriquez, C. F.; Hopkinson, A. C.; Siu, K. W. M. *Int. J. Mass Spectrom.* **2002**, *219*, 101−114.