Zwitterions are the most stable form for neutral arginylglycine in gas phase: Clear theoretical evidence

Cuihong Wang [a,b,c], Zijing Lin [a,*], Ruiqin Zhang [b,*]

[a] Department of Physics, University of Science and Technology of China, Hefei, China
[b] Department of Physics and Materials Science, City University of Hong Kong, Hong Kong, China
[c] School of Science, TianJin Institute of Urban Construction, Tianjin, China

ARTICLE INFO

ABSTRACT

The canonical and zwitterionic conformers of gaseous dipeptide arginylglycine, ArgGly, were thoroughly researched. A large number of trial structures were generated by considering all combinations of internal single bond rotamers of three ArgGly tautomers. The structures were successively optimized with the methods of PM3, HF/3-21G*, BHandHLYP/6-31G* and BHandHLYP/6-31+G**. The conformational energies were finally calculated by the approaches of DFT/BHandHLYP, DFT/B3LYP, DFT/B97D and MP2 using the basis set of 6-311++G (2df, 2p). The results of the four methods are quite different on various aspects. However, all the computational approaches find that the most stable structure of ArgGly is zwitterionic. The properties of the low energy conformers such as the rotational constants, dipole moments, vertical ionization energies, temperature dependent conformational distributions and IR spectra are analyzed. These data should be helpful for understanding the experiments as well as for the differentiation of the four theoretical approaches by the experiments.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Zwitterions are neutral compounds that contain oppositely charged centers. Zwitterions are known to occur in amino acids and peptides in solution and in solids as their dipolar character experiences strong electrostatic stabilization in such environments. Such zwitterions are thought to play important roles in the structures and functions of peptides and proteins [1]. In the gas phase, however, the zwitterionic charge separation is not stabilized by the environment. For most amino acids, the zwitterionic structure does not even correspond to a local minimum on the gas-phase potential energy surface (PES) [2–4]. Arginine (Arg) and histidine (His) are the only amino acids known for the zwitterionic structures to correspond to local minima on their PESs, but the most stable zwitterions of Arg and His are respectively 4 and 11 kcal/mol less stable than their canonical counterparts [5,6]. However, zwitterions may be stable for amino acids and small peptides if they are protonated, metalated or solvated by one or more water molecules or accompanied with an excess electron [6–12]. Indeed, all existing experimental and theoretical studies may only conclude the stability of zwitterions for non-isolated amino acids and small peptides [6–17]. Nevertheless, oligopeptides with sufficient intramolecular bonding flexibilities should be able to

stabilize the zwitterionic structure as the global minimum by the interactions of the charged groups with polar groups in the molecule [18]. It would be interesting to know what may be the smallest neutral peptide with the zwitterionic structure as the most stable conformation.

A dipeptide is the natural starting point to find the smallest peptide zwitterions. There have been numerous experimental and theoretical studies of the neutral dipeptide conformations. With the exception of arginyl glycine, ArgGly, all available experimental and theoretical results show that only the canonical forms are the stable ones [19–38]. However, a recent computational study by Prell et al. has identified a zwitterionic structure of ArgGly that is energetically very close to the global minimum of the canonical form [39]. ArgGly appears to be a promising candidate to be the smallest peptide with stable zwitterions.

The computational results by Prell et al. are based on conformational screening by the molecular mechanics force field that is prone to miss important structures [39]. Therefore, advanced computational searches are desired in order to provide more reliable results. In this work, thorough computational searches of the conformations of the gaseous ArgGly are performed. The purpose of this investigation is to locate all low energy gaseous ArgGly conformers and to obtain knowledge about their relative stabilities on the energy surfaces. For verification by future experiments, essential theoretical data are given that include the rotational constants, dipole moments, vertical ionization energies (VIEs),

* Corresponding authors. Tel.: +86 55163606345; fax: +86 55163606348.
  E-mail addresses: zjlin@ustc.edu.cn (Z. Lin), aprqz@cityu.edu.hk (R. Zhang).

equilibrium distributions and IR spectra of the low energy ArgGly conformers.

2. Computational methods

Similar to Arg, [5] neutral ArgGly has two canonical and one zwitterionic tautomers. The three tautomers of Arg-Gly, RG1, RG2, Z, are illustrated in Fig. 1. The conformational spaces of the three ArgGly tautomers are searched separately and their low energy conformers are then compared to determine the stable conformations of ArgGly. To ensure a reliable description of the PESs and locate all the low energy conformers, trial structures were generated by allowing for all combinations of internal single-bond rotamers [5]. A total of 209,952 and 314,928 trial structures were generated for the trans peptide bond isomers of RG1 and RG2, respectively. As the trans form isomers were found to be overwhelmingly preferred by various tests, the search for the low energy conformers of cis peptide bond isomers of RG1 and RG2 were carried out with a moderate set of trial structures, 46,656 structures for cis RG1 and 93,312 structures for cis RG2. Both the trans and cis peptide bond isomers were equally considered and a total of 349920 trial structures were generated for the tautomer Z.

All the trial structures were optimized by the semi-empirical method of PM3. The unique structures thus obtained were sorted by their HF/3-21G* energies. There are 4361 conformers for trans RG1, 1 conformer for cis RG1, 5508 conformers for trans RG2, 9 conformers for cis RG2 and 3760 conformers for Z that are within the range of 16 kcal/mol from their respective global minima. These low energy conformers were then optimized using the HF/3-21G* method. The low energy conformers in the 16 kcal/mol range of the global minimum at the level of HF/3-21G* were further optimized using the BHandHLYP/6-31G* method. The conformers thus obtained to be within 5 kcal/mol of the global minimum were finally optimized at the BHandHLYP/6-31+G** level of theory. The single point energies of the conformers were computed by the approaches of DFT/BHandHLYP [40–44], DFT/B3LYP [40,45,46], MP2 [47,48] and DFT/B97D [49,50] with the basis set of 6-311++G(2df,2p). All the low energy conformers of RG1 and RG2 were found to possess the trans peptide bonds.
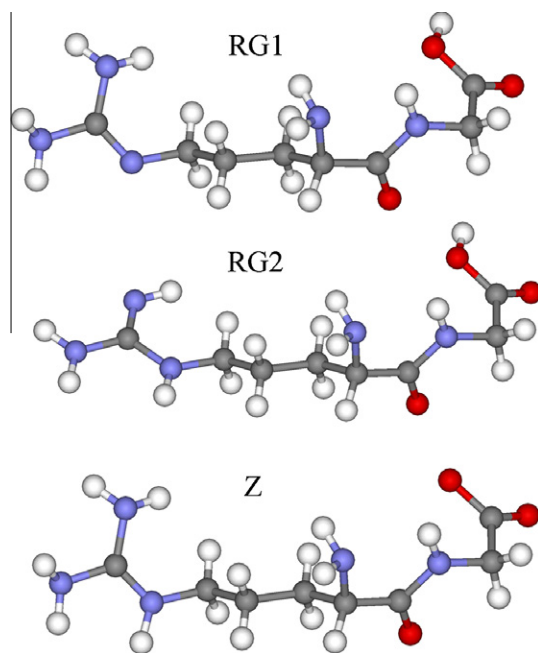

Fig. 1. Three tautomers of arginyl glycine dipeptide.

The vibrational frequencies were calculated at the BHandHLYP/6-31+G** level of theory and modified with due scaling factors. The scaling factors were 0.9581 and 0.9288 for wavenumbers below and above 2000 cm^{-1}, respectively [51,52]. The scaling factors for the zero-point vibrational energies (ZPVEs) and the thermal corrections to the Gibbs free energy were 0.9498 and 0.9288, respectively [51,52]. The dipole moments and VIEs of the conformers were determined at the BHandHLYP/6-311++G(2df,2p) level of theory.

All calculations were carried out on our PC Cluster with the GAUSSIAN03 suite of programs [53].

3. Results and discussion

3.1. Conformations and energies

A large number of local minima in the PES of ArgGly are found. The conformations are denoted as cn or zn in this work. Here the letters c and z refer to the canonical and zwitterionic structures, respectively. n is a numeral denoting the conformational stability ordered by the ascending BHandHLYP/6-311++G(2df,2p) energy. The symbol is underlined if the conformer possesses a cis peptide bond. For example, z1, z4 and c6 are a trans peptide bond zwitterion, a cis peptide bond zwitterion and a (trans peptide bond) canonical conformer with the lowest, the forth lowest and the sixth lowest BHandHLYP/6-311++G(2df,2p) energies, respectively. During the process of geometry optimization, it is noticed that the trial structures for zwitterions may be transformed into low energy canonical conformers, and vice versa. This seems to imply that it suffices to optimize either the canonical or the zwitterionic trial structure set if the goal is to find a few representative low energy conformers instead of a complete set of low energy conformers.

Table 1 shows relative electronic energies, relative ZPVEs, relative thermal corrections to the free energy at the standard state, dipole moments, rotational constants, and VIEs for conformers with the BHandHLYP/6-311++G(2df,2p) energies that are in the 3 kcal/mol range of the global minimum. The relative electronic energies calculated by the approaches of BHandHLYP, B3LYP, MP2 and B97D are denoted in the table as E^{BH}, E^{B3LYP}, E^{MP2} and E^{B97D}, respectively. Among the conformers shown in Table 1, only z2, c6, c11 and z30 have been located by Prell et al. [39]. That is, most of the low energy conformers were missed by their molecular mechanics force field based Monte Carlo searches, demonstrating the necessity of the thorough search performed in this work. Besides, Table 1 shows clearly that most of the low energy conformers are zwitterions.

Fig. 2 shows the structures of some representative low energy conformers of neutral ArgGly. The low energy conformers of ArgGly usually have very compact structures that are helpful for forming multiple hydrogen bonds (H-bonds) with the lowered electronic energies. There are three H-bond donors (—COOH, —NH$_2$ and —CN (H) C—) and five H-bond acceptors (HOC=O, C=NH, NHC=O, C=NC and –NH$_2$) in ArgGly. The existence of H-bond is determined here by a geometric criterion of taking a cutoff distance of 2.8 Å between the H-bond donor and acceptor [5]. By this definition, there are at least three H-bonds in every conformer shown in Table 1. Similar to the cases for valine [54] and arginine [5], all three H-bonds, the Y...H—X red shifting bond, the C—H...H—X dihydrogen bond and the X...H—C blue shifting bond (X, Y=N, O), are found in the low energy conformers of ArgGly. The three H-bond types are found to coexist in z2, z9, z14, z26 and z29.

The low energy conformers may be roughly classified into two types of configurations. In type A configuration, the guanidine group and the two carbonyl oxygens are roughly in the same plane with one or two strong H-bonds of CO...HN. In type B configuration, the guanidine plane tends to be roughly parallel with the COO plane to form a face to face structure. In general, type B structures are more