Computational and Theoretical Chemistry 1020 (2013) 14–21

Contents lists available at ScienceDirect

Computational and Theoretical Chemistry

journal homepage: www.elsevier.com/locate/comptc

Thorough theoretical search of conformations of neutral, protonated and deprotonated glutamine in gas phase

Rui Pang [a], Minghao Guo [a], Sanliang Ling [b], Zijing Lin [a,*]

[a] Department of Physics, University of Science and Technology of China, Hefei, Anhui 230026, China
[b] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

ARTICLE INFO

Article history:
Received 21 March 2013
Received in revised form 12 July 2013
Accepted 13 July 2013
Available online 23 July 2013

Keywords:
First principle calculation
Structure
Hydrogen bond
Proton affinity
Proton dissociation energy
Entropy effect

ABSTRACT

There are large discrepancies among the theoretical and experimental results of the proton affinity (PA) of glutamine (Gln). To provide a reliable basis for the theoretical investigation, extensive conformational searches have been performed for neutral, protonated and deprotonated Gln in gas phase by optimizing the trial structures generated by allowing for all combinations of internal single-bond rotamers. The structures and hydrogen bonding features, relative electronic energies, zero point vibrational energies, rotational constants, dipole moments, vertical ionization energies of the low energy conformers and equilibrium conformational distributions are presented. PA, GB (gas phase basicity), PDE (proton dissociation enthalpy) and GA (gas phase acidity) of Gln were computed by the theoretical approaches of BHandHLYP, B3LYP, B97D, MP2, G3MP2B3, M062X and CCSD. The computed relative conformational energies and PA, GB, PDE and GA are dependent on the theoretical approach and the basis set. Analysis of the computational results shows that the extended kinetic method provides an accurate estimate of PA and overestimate of the entropic effect, while all other experiments underestimate PA of Gln. The best theoretical estimates of PA, GB, PDE, GA and the protonation and deprotanation entropies for Gln are 987.2 ± 4.0, 945.1 ± 5.8, 1385.3 ± 9.0, 1362.9 ± 9.1 kJ/mol and −32.4 ± 6.2 and 33.9 ± 5.1 J/mol/K, respectively.

1. Introduction

Studying biomolecules in gas phase is important for revealing their intrinsic properties free of the influence of the interacting environment [1,2] and the gas phase properties are also indicative of that in solution [3]. Many experimental approaches have been employed to determine the structures and properties of gaseous amino acids, such as proton affinity (PA) and gas phase basicity (GB) [4–9], proton dissociation enthalpy (PDE) and gas phase acidity (GA) [10,11], dipole moments, rotational constants [12], IR and UV spectra [13,14], ionization potentials [15], two-photon circular dichroism [16], etc. With the rapid development in quantum chemistry methods and computer hardware, increasingly more computational investigations are reported [17–21]. The advantage of computation is that it offers results in the exact ideal situation without experimental uncertainties. It also provides information such as geometries and hydrogen bondings that are difficult to measure directly by the experiment.

PA, GB, PDE and GA are important thermochemical properties of molecule. Early literatures on the thermo-chemical properties of amino acids have been summarized by Harrison [22] and Hunter and Lias [5]. However, due to the intrinsic difficulties associated with the handling of these involatile and thermally labile molecules and with the methods of measuring thermochemical properties, the obtained data should be used with care [6–8]. For example, the results obtained by equilibrium constant measurements [4] and thermo kinetic methods [5] are often somewhat different and require corrections using the Hunter and Lias scale [23].

Because of their fundamental importance, the thermochemical quantities for most amino acids have been repeatedly measured over the last decades and the results have been compiled in a recent review by Bouchoux [23]. Overall speaking, the experimental data are in acceptable agreement with each other and with the most reliable computational results. However, glutamine (Gln) is a severe exception. The GB of Gln measured by the equilibrium method is smaller than that obtained with the extended kinetic method (EKM) by about 50 kJ/mol [23]. The PA of Gln determined by EKM is larger than the average measured by the simple kinetic method (SKM) by over 30 kJ/mol [23]. However, the theoretical results show a relatively small spread and are closer to the results of EKM [23,24]. Nevertheless, the results by EKM are believed to be clearly too high [23], an assertion that is not well justified.

Gln is an elemental amino acid that composes proteins in biological systems [25,26]. Gln is the main source of nitrogen in human bodies and comprises approximately 50% of the

* Corresponding author.
E-mail address: zjlin@ustc.edu.cn (Z. Lin).

2210-271X/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.comptc.2013.07.016

whole-body pool of free amino acid. It is the most important amino acids in muscle growth, and is involved in the synthesis of a variety of enzymes. It is considered as important fuel for many kinds of cells. The polar groups in Gln tend to form hydrogen bonds (H-bonds). However, the tendency to form H-bonds and the strengths of the H-bonds may be significantly enhanced in protonated Gln. The large entropic difference in neutral and protonated Gln is ignored in SKM, but easily captured in EKM. Moreover, the reported EKM data show that the isothermal point may be determined with a limited uncertainty for Gln [27]. Therefore, the EKM result should be more reliable than that of SKM and the assertion that there is a large error in the EKM result is unjustified. Fortunately, the computational method is ideally suited to resolve the dispute as it may easily take the entropic effect into consideration. Nevertheless, a high quality computational study is required to draw a convincing conclusion.

Numerous theoretical studies on the PA of Gln have been reported. Maksic and Kovacevic calculated the PA of Gln at the MP2/6-31+G(d,p) level, but the conformations of neutral and protonated Gln structures were not reliably determined [28]. Dinadayalane et al. searched the conformations of neutral and protonated Gln based on chemical intuition [29]. Bleiholder et al. [30] improved the conformational search by using a simulated annealing technique combined with an empirical Hamiltonian as a pretreatment to deal with a large number of trail structures, and the final result of PA was determined at the level of B3LYP/6-31G(p,d) and G2MP2. As the stochastic nature of the simulated annealing technique has a considerable possibility of missing the global minimum [18] and the basis set of 6-31G(p,d) is often insufficient for obtaining accurate results, the cause for the difference between the theoretical and experimental results is uncertain. Bertran and coworkers [31] used a Monte Carlo multiple minimum technique combined with the MMFF94s force field [32] and the non-local meta-hybrid MPWB1K density functional in the search of Gln conformations. They found three low energy structures based on the assumption that there is one intramolecular H-bond in the most stable conformer. Bouchoux determined the PA value based on the G3MP2B3 theory [23]. Guo performed detailed conformational searches and calculated the PA of Gln at the MP2/6-311++G** level [24]. The theoretical basis for Guo's result seems solid. Nevertheless, further validation may be needed as substantial differences in different DFT approaches and other traditional first principle calculations have been observed [33].

Though to a less extent, the difference between the theoretical and experimental PDE results of Gln is also notable. The experimental values are 1388 and 1385 kJ/mol as determined by O'Hair et al. [11] and by Jones et al. [10], respectively. Jones et al. also reported a theoretical value of 1378 kJ/mol obtained at the B3LYP/6-311++G** level. However, a value of 1368 kJ/mol is obtained by Guo at the MP2/6-311++G** level based on an improved conformational search [24]. Though the latter is not very different from the former theoretical result, its difference from the experimental ones is about 20 kJ/mol and uncomfortably large. It is meaningful to present a more thorough theoretical examination.

In this study, systematic searches of the conformational spaces of neutral, protonated and deprotonated Gln by varying all reasonable rotational degrees of freedom were performed. A series of local minima on the potential energy surfaces of these Gln species were obtained by systematic search of all the reasonable rotamers [18]. A new set of PA, GB, PDE and GA data are obtained and compared with previous experimental and theoretical results. Discussion on the theoretical and experimental difference is given to support the current theoretical results. Analysis also shows that the extended kinetic method is a reliable way of determining PA and PDE of amino acid.

2. Computational method

The representative structures of neutral, protonated and deprotonated Gln are shown in Fig. 1. The conformational spaces of the three Gln species are thoroughly searched by optimizing trial structures generated by combinations of all reasonable internal single-bond rotamers [18]. For Gln, as the C–N bond rotation in the acyl group is prevented by the electron conjugation on the C and N atoms, there are 6, 5 and 5 bond rotational degrees of freedom for the neutral canonical, protonated and deprotonated Gln, respectively. The bond rotational degrees of freedom are illustrated in Fig. 2 for neutral canonical Gln. As a result, a total of 7776 trial structures were generated for canonical Gln. All these trial structures were optimized at the PM3 level [34,35], resulting in 1200 unique structures. These structures were re-optimized at the HF/3-21G* level, and the unique structures thus obtained were further refined at the BHandHLYP/6-31G* level [21]. A total of 143 conformers were found for neutral canonical Gln. The 21 lowest energy conformers that spanned an energy range of 3 kcal/mol were further optimized at the level of BHandHLYP/6-311++G**. The vertical ionization energies (VIEs) were determined at the BHandHLYP/6-311++G(2df,2pd) level. For comparison, VIEs were also computed with the outer valence Green's function (OVGF) [36] method. The 6-311++G** basis set is used for the OVGF calculations in order to save the computational cost. As a side note, the zwitterionic forms of Gln were also similarly searched and no zwitterions were found to correspond to the local minima in the potential energy surface (PES) of neutral Gln.

The numbers of trial structures for protonated and deprotonated Gln were 2592 and 3888, respectively. The trial structures were first optimized at the HF/3-21G* level and the unique structures thus obtained were re-optimized at the level of BHandHLYP/6-31G*. The lowest energy conformers in the range of 3 kcal/mol were further optimized at the BHandHLYP/6-311++G** level.

Single point energies of the low energy conformers were calculated using the computational approaches of BHandHLYP, B3LYP [37–40], MP2 [41], B97D [42], M062X [43,44] and CCSD [45], combined with one or more of the following basis sets: 6-311++G**, 6-311++G(2df,2pd), cc-PVTZ and cc-PVQZ. The vibrational frequencies were determined at the BHandHLYP/6-311++G** level and scaled with a factor of 0.93 [46]. The zero point vibrational energies (ZPVEs), the thermal corrections for enthalpy and free energy are scaled with a factor of 0.95 and 0.94, respectively [46]. These results were combined with the electronic energies at various levels to determine the conformational distributions of canonical, protonated and deprotonated Gln species.

The low energy conformers were also optimized at the BHandHLYP/cc-PVTZ level to check their basis set dependencies. They are also optimized with the B97D and M062X methods to see the influence of the DFT functionals on the geometries.

The standard PA (GB) is calculated as the negative of the enthalpy (Gibbs free energy) change for the gas-phase protonation reaction at room temperature, $T = 298$ K. The enthalpy of H^+, $H(H^+)$, is the sum of the translational energy of H^+ and the PV work from the reaction and is calculated as $H(H^+) = E + PV = 5/2RT$. The proton free energy is calculated to be −26.2 kJ/mol [21]. The free energy and enthalpy for a given species were obtained through weighted averaging over its conformations. Moreover, the free energy calculations also take into account the entropy of mixing, $-R\sum_i x_i \ln x_i$, where x_i is the population of conformer i [23]. The PDE and GA of Gln were similarly calculated based on the gas-phase deprotonation reaction.

The CCSD calculations were performed with Molpro [47]. All other calculations were carried out with the GAUSSIAN09 suite of