Comprehensive theoretical study towards the accurate proton affinity values of naturally occurring amino acids

ARTICLE *in* INTERNATIONAL JOURNAL OF QUANTUM CHEMISTRY · NOVEMBER 2006

Impact Factor: 1.43 · DOI: 10.1002/qua.21117

3 AUTHORS, INCLUDING:

Tandabany C. Dinadayalane
Clark Atlanta University
66 PUBLICATIONS **1,156** CITATIONS
SEE PROFILE

G Narahari Sastry
Indian Institute of Chemical Technology
262 PUBLICATIONS **5,297** CITATIONS
SEE PROFILE

Comprehensive Theoretical Study Towards the Accurate Proton Affinity Values of Naturally Occurring Amino Acids

T. C. DINADAYALANE,[1] G. NARAHARI SASTRY[2] JERZY LESZCZYNSKI[1]

[1]*Computational Center for Molecular Structure and Interactions, Department of Chemistry, Jackson State University, 1400 JR Lynch Street, P. O. Box 17910, Jackson, Mississippi 39217, USA*

[2]*Molecular Modelling Group, Organic Chemical Sciences, Indian Institute of Chemical Technology, Hyderabad 500 007, India*

Received 20 July 2005; accepted 13 February 2006
Published online 12 July 2006 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/qua.21117

ABSTRACT: Systematic quantum chemical studies of Hartree–Fock (HF) and second-order Møller–Plesset (MP2) methods, and B3LYP functional, with a range of basis sets were employed to evaluate proton affinity values of all naturally occurring amino acids. The B3LYP and MP2 in conjunction with 6-311+G(*d*,*p*) basis set provide the proton affinity values that are in very good agreement with the experimental results, with an average deviation of ~1 kcal/mol. The number and the relative strength of intramolecular hydrogen bonding play a key role in the proton affinities of amino acids. The computational exploration of the conformers reveals that the global minima conformations of the neutral and protonated amino acids are different in eight cases. The present study reveals that B3LYP/6-311+G(*d*,*p*) is a very good choice of technique to evaluate the proton affinities of amino acids and the compounds derived from them reliably and economically. Int J Quantum Chem 106: 2920–2933, 2006

Key words: proton affinity; quantum chemical calculations; amino acids; global minima; potential energy surface

Correspondence to: J. Leszczynski; e-mail: jerzy@ccmsi.us or G. N. Sastry; e-mail: gnsastry@yahoo.com

This article contains supplementary material available via the Internet at http://www.interscience.wiley.com/jpages/0020-7608/supppmat/.

Contract grant sponsor: NIH-SCORE.
Contract grant number: S06-GM008047.

Introduction

Naturally occurring α-amino acids constitute the building blocks of proteins and peptides, which participate in a myriad of biological processes. The tendency of a molecule to accept a proton in the reaction has long been of interest to chemists. Gas phase basicity (GB) and proton affinity (PA) are among the most important properties playing a crucial role in a number of biological reactions. Knowledge of these properties provides important information about the participation of these molecules in proton transfer reactions, the intramolecular forces and in the gas phase acid–base equilibria [1, 2]. Measurement of these fundamental chemical properties of amino acids has become possible with a wide range of modern mass spectrometric techniques such as fast atom bombardment (FAB) [3, 4], secondary ion mass spectrometry (SIMS) [5], and the soft ion techniques, for example electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) [1, 6–8]. High-pressure mass spectrometry [9], Fourier transform ion cyclotron resonance (FTICR) mass spectrometry [7], and kinetic methods [3, 8, 10–22] have all been used in evaluating the gas phase basicity (GB) and proton affinity (PA).

Although accurate experimental determination is the ultimate choice for obtaining these properties, the high-level ab initio and DFT calculations have become an attractive alternative when the experimental determination is difficult or ambiguous. The computations play central role to validate the experimental results since a wide range of techniques has been employed in measuring the GB and PA values [18–27]. Interplay between theory and experiment is crucial for deeper understanding and appreciating the relevance of the observed proton affinities of amino acids [21–34] and DNA bases [35]. Harrison [1] pointed out the limitations of the methods that have been used to evaluate the gas phase basicities and proton affinities of amino acids and small peptides. Bojesen [3], Wu and Fenselau [12], and Isa et al. [14] independently reported the order of proton affinities of all the naturally occurring α-amino acids based on the experimental results, unfortunately, the exact proton affinity values were not reported. A comparison of the trends obtained in these studies shows similar orders. Table I gives the proton affinity values of amino acids reported by Hunter and Lias [36, 37] published in the National Institute of Standards and Technology (NIST) along with the other reported values. A quick comparison reveals a fairly good agreement among the reported values, with minor discrepancies for certain amino acids.

TABLE I

Experimental proton affinity values (in kcal/mol) of natural amino acids reported in the literature.

Amino acid	Hunter and Lias[a]	Bojesen and Breindahl[a]	Gorman et al.[c]	Li and Harrison[d]
Gly	211.9	—	211.7 ± 3.2	—
Ala	215.5	—	222.1 ± 2.9	—
Cys	215.9	—	214.6 ± 2.7	—
Ser	218.6	217.2	220.2 ± 2.1	—
Val	217.6	218.1	222.1 ± 2.9	215.0
Leu	218.6	218.7	225.0 ± 3.1	216.1
Asp	217.2	218.1	220.2 ± 2.1	—
Ile	219.3	219.2	225.0 ± 3.1	216.8
Thr	220.5	219.2	228.5 ± 2.2	216.6
Phe	220.6	219.9	227.2 ± 2.2	217.4
Tyr	221.3	220.7	227.2 ± 2.2	217.9
Asn	222.0	222.1	227.2 ± 2.2	217.8
Met	223.6	221.0	228.5 ± 2.2	218.0
Pro	220.0	222.4	231.0 ± 3.3	218.5
Trp	226.8	223.5	231.0 ± 3.3	220.8
Glu	218.2	222.3	240.6 ± 1.9	—
Gln	224.1	226.9	228.5 ± 2.2	221.4
His	236.1	230.5	228.5 ± 2.2	—
Lys	238.1	228.7	242.6 ± 3.4	222.9
Arg	251.2	>242.8	>243.2	—

[a] Values taken from Refs. [36, 37].
[b] Values taken from Ref. [10].
[c] Values taken from Ref. [7].
[d] Values taken from Ref. [11].

Recently, Poutsma and coworkers [22] reported that the use of extended kinetic method is required to obtain reliable proton affinities for the compounds where the intramolecular hydrogen bonding plays an important role. The combined mass spectroscopic and theoretical investigations carried out by Marino et al. [26] of proline revealed the dichotomy of global minima conformations between neutral and protonated species. The role of conformations in determining proton affinities of proline and modified proline was also highlighted in the work of Tabet and coworkers [24]. This study further validates Marino et al. [26] that the most stable neutral and protonated conformers are different for proline. Thus, the conformation corresponding to the global minima on the neutral sur-

face need not correspond to the global minima on the protonated amino acid surface and vice-versa.

So far, the available theoretical studies, except the investigation carried out by Maksic and Kovacevic [23] on the proton affinities of amino acids were restricted to the selected amino acids [21, 22, 24–34]. Although Maksic and Kovacevic [23] have reported Hartree–Fock (HF) and second-order Møller–Plesset (MP2) calculations on the proton affinities of all the naturally occurring amino acids, their study did not consider the conformational changes upon protonation. Previously, the HF method and B3LYP functional were employed to obtain fairly satisfactory proton affinity values for amino acids [21–34]. However, to our knowledge, the performance of these methods and the quality of the basis set employed were not critically examined. In this study, the conformational potential energy surface (PES) was systematically explored to locate and analyze the most stable conformation of neutral as well as protonated amino acids. Different levels of approximation were applied in calculating the proton affinity values. The study is also focused to probe the relationship between the structural changes upon protonation and the proton affinity values. The most commonly used three-letter nomenclature was followed to specify the amino acids throughout this study [38]. The values reported by Hunter and Lias [36, 37] were used as a reference for comparison to assess the performance of theoretical methods and to acquire the general conclusions.

Computational Details

The global minima for neutral and protonated forms of all the naturally occurring amino acids were obtained using HF/6-31G(*d*) by thoroughly searching the PES considering several chemically reasonable structures in each case. The nitrogen atoms were protonated for all the amino acids. The lowest-energy conformers obtained using HF/6-31G(*d*) were then optimized with B3LYP/6-31G(*d*) and characterized as true minima based on the vibrational frequency calculations. The single-point calculations were performed using HF and B3LYP with 6-311+G(*d*,*p*) and cc-pVTZ basis sets; we have also carried out the single-point calculations using MP2 method with 6-31G(*d*) and 6-311+G(*d*,*p*) basis sets. B3LYP/6-31G(*d*) optimized geometries were taken for all the single point calculations. Dunning correlation consistent polarized split-valence triple-ζ basis set (cc-pVTZ) has also been employed in the present study in addition to the Pople style basis sets to examine the Pople versus Dunning basis set on proton affinity values of amino acids. Moreover, studies on calculating proton affinities used Dunning correlation consistent basis set [39, 40]; hence we would like to address whether cc-pVTZ basis set can be used to obtain the accurate proton affinity values of the naturally occurring amino acids. Previous computational studies demonstrated that for several compounds, B3LYP functional provides proton affinities that are in good agreement with those reported at the CCSD(T) and G2 levels [31, 34]. All the calculations were carried out using the Gaussian 98 suite of program package [41].

The proton affinities were computed using the following equation:

$$\mathrm{PA(AA)} = -\Delta H_{298} = -\{[E(\mathrm{AAH}^+) - E(\mathrm{AA})] + [H^{\mathrm{corr}}(\mathrm{AAH}^+) - H^{\mathrm{corr}}(\mathrm{AA})] - (5/2)\mathrm{RT}\}, \quad (1)$$

where E and H^{corr} are the total energy and enthalpy correction at 298 K; AA and AAH^+ denote the amino acid and its protonated form, respectively; (5/2)RT is the classical estimation of the effect of losing three translational degrees of freedom [(3/2)RT], plus the PV term (RT). The enthalpy correction values obtained using B3LYP/6-31G(*d*) were employed to calculate the proton affinities with different methods considered except for the HF/6-31G(*d*). Direct protonation and deprotonation were performed on the neutral global minima and protonated global minima structures of amino acids, respectively, and those structures were optimized and characterized using HF/6-31G(*d*) level.

Results and Discussion

EQUILIBRIUM GEOMETRIES

The important intramolecular hydrogen bonding distances obtained using HF method and B3LYP function with 6-31G(*d*) basis set for all the neutral and protonated amino acids are shown in Figure 1. Considering all the structures, there are different types of hydrogen bonds such as N—H···N, N—H···O, O—H···N, O—H···O, S—H···N, and N—H···S. Detailed analysis of the geometries will give the insight into the relative strength of the above-mentioned hydrogen bonds and also how

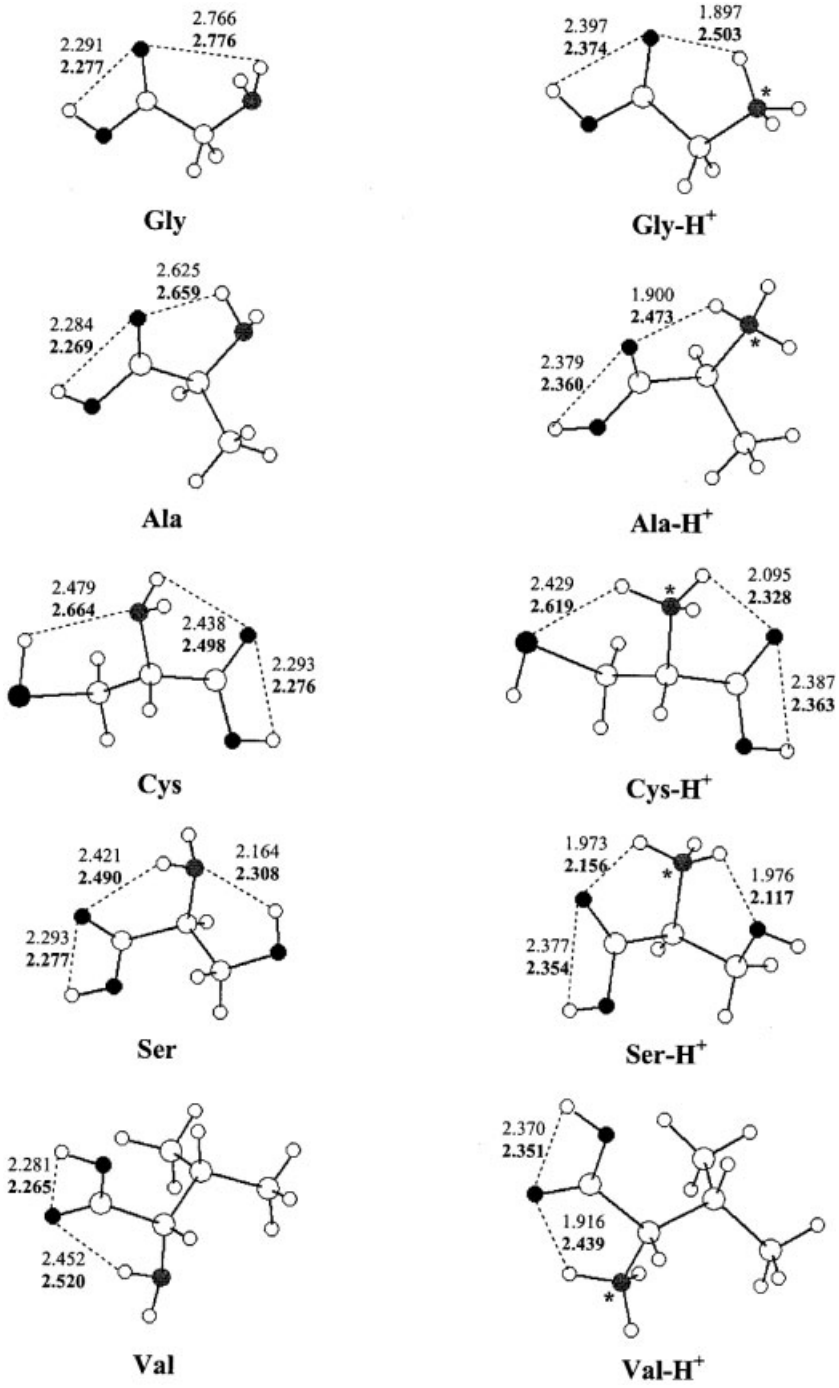

FIGURE 1. B3LYP/6-31G(*d*) optimized structures of all the neutral and protonated α-amino acids. H-bond distances [in (Å)] obtained using B3LYP (plain) and HF (bold) are given. Filled small and big circles correspond to oxygen and sulfur atoms. The dotted filled circles represent nitrogen atoms. Asterisk (*) marks indicate the protonated nitrogen atom.

the hydrogen bond strength varies between the neutral and protonated forms. A quick look at Figure 1 indicates that HF method consistently underestimates the hydrogen bonding distances of O—H···O═C by ~0.01–0.02 Å compared with B3LYP for all the neutral and protonated amino

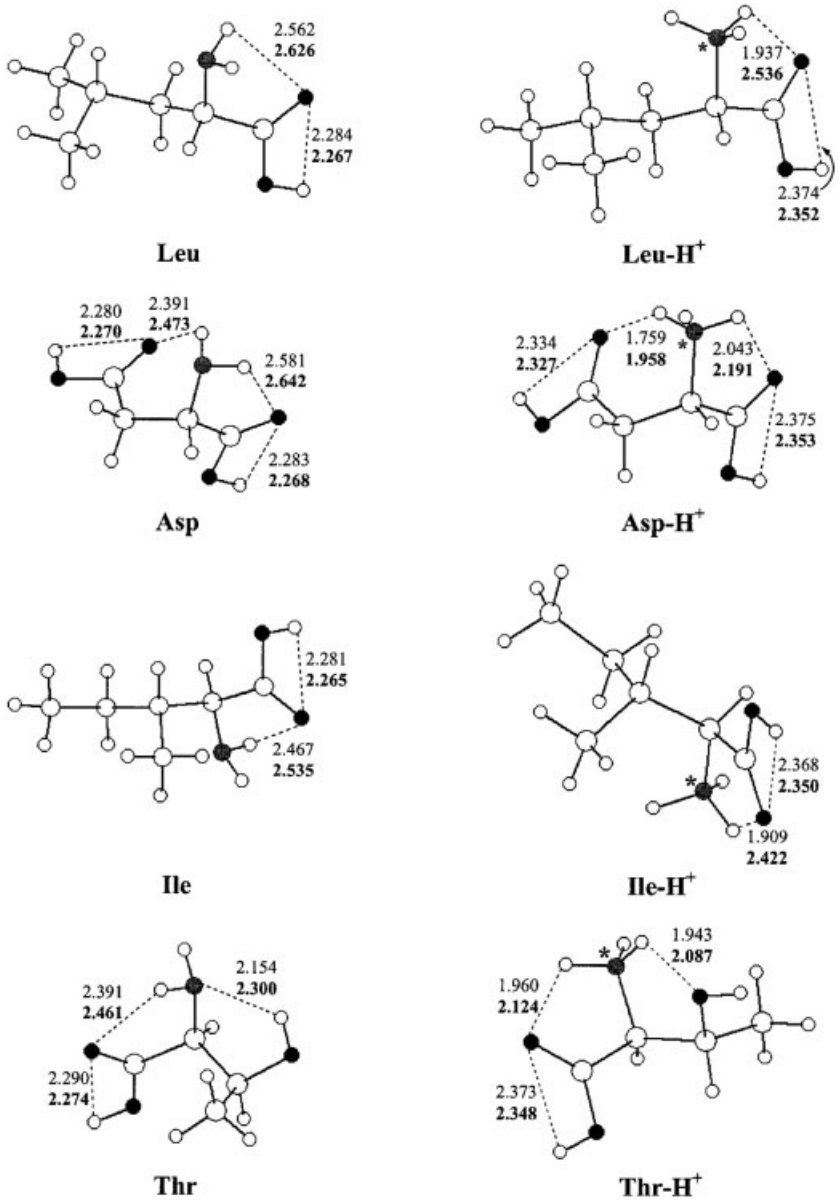

FIGURE 1. Continued.

acids. However, the other hydrogen bonding distances, in most cases, are predicted to be longer at the HF level and the extent of overestimation is substantial when the hydrogen bond occurs in protonated nitrogen, i.e. N*—H···O and N*—H···N. Earlier computational studies indicate that B3LYP is appropriate for the geometries of systems with significant hydrogen bonding [21, 22, 28]. Hence, the discussion of the structural features will be based on the B3LYP geometries. Intramolecular hydrogen bonds furnish increased stability of the species. In both neutral and protonated species, the N—H bond(s) create hydrogen bonding with carbonylic oxygen rather than hydroxylic oxygen due to higher basicity of the former.

As evidenced from Figure 1, the number of hydrogen bonds is identical in the neutral and its protonated structures for 16 amino acids; the additional intramolecular hydrogen bonds, which usually enhance the stability, arise in case of **Met**—$\mathbf{H^+}$, **Glu**—$\mathbf{H^+}$, **Lys**—$\mathbf{H^+}$, and **Arg**—$\mathbf{H^+}$. Protonation leads to significant shortening of the hydrogen bond distances, particularly in N*—H···X (X═O, N, or S). As reported in previous studies [22, 32–34],

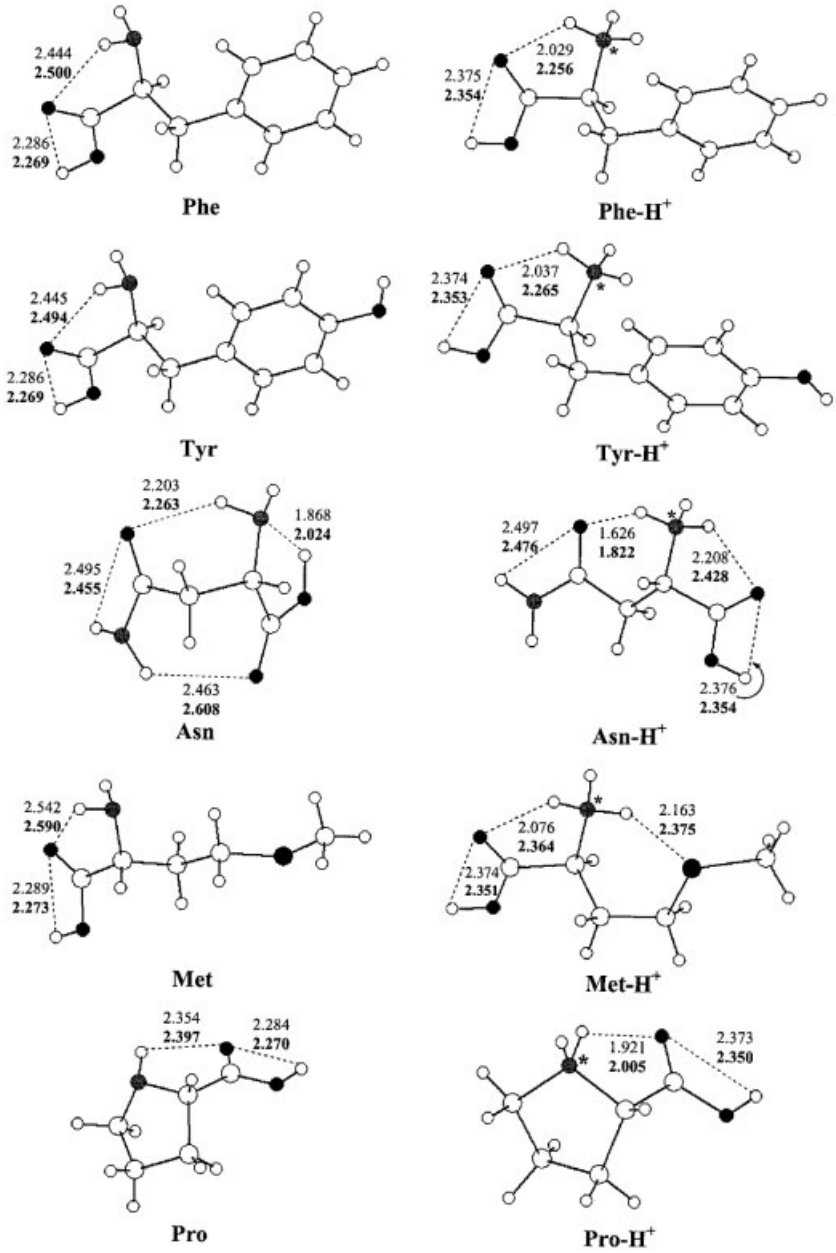

FIGURE 1. Continued.

the bifurcated hydrogen bonds are observed between the NH_2 group and the carbonylic oxygen for the most stable structure of some of the neutral amino acids, e.g., **Gly**, **Ala**, **Cys**, and **Leu**. The lowest-energy structures of protonated and the neutral amino acids exhibit cyclic structures; at least one five-membered ring is present in each structure. **Cys**, **Ser**, and **Thr** and their protonated forms comprise cyclic structures of two five-membered rings due to the intramolecular hydrogen bonds. Both five- and six-membered ring architecture are produced due to the intramolecular hydrogen bonds in case of **Asp**, **His** and their protonated species of the lowest-energy structures. It is worth mentioning that the protonation to **Met** leads to one additional hydrogen bond of N*—H···S, which fashioned a six-membered ring, compared with the neutral structure.

Two of the amino acids, **Glu** and **Gln**, as well as their protonated species were characterized with

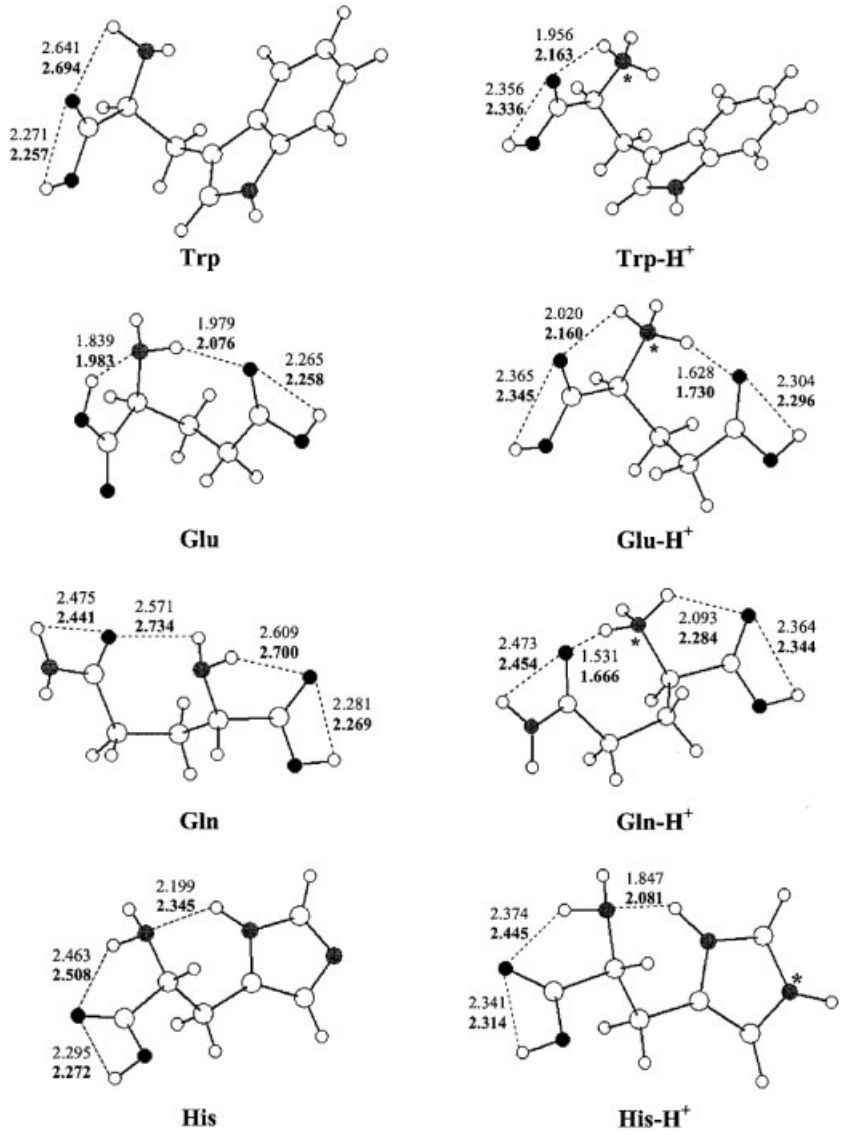

FIGURE 1 Continued.

cyclic arrangements of a seven- and five-membered ring in each of the structures due to the intramolecular hydrogen bonds. Interestingly, the hydrogen bonds generate the cyclic structure with five-, six-, and seven-membered rings in the **Asn**, whereas the seven-membered ring architecture has been lost in its corresponding protonated species. One of the N*—H···O hydrogen bond distances is very short in case of **Asn**—$\mathbf{H^+}$, **Glu**—$\mathbf{H^+}$, and **Gln**—$\mathbf{H^+}$, indicating the high bond strength of this hydrogen bond. The last structure possesses the shortest hydrogen bond distance of 1.531 Å among all the protonated amino acids studied here. The global minimum geometries obtained using B3LYP/6-31G(*d*) level for the neutral and protonated glutamic acid (**Glu**) are virtually identical to the B3LYP/6-31+G(*d*) geometries obtained by Sun et al. [28]. Interestingly, while there is no intramolecular hydrogen bond between the two amino groups in the neutral lysine (**Lys**), the protonation leads to the hydrogen bonding between the amino groups in the lowest-energy conformer. The intramolecular hydrogen bonds generate an eight- and seven-membered ring arrangements in the protonated **Lys** with the loss of a five-membered ring in neutral species. The lowest-energy structures of **Lys** and its protonated form obtained in the current study are very similar to those reported in the recent study of

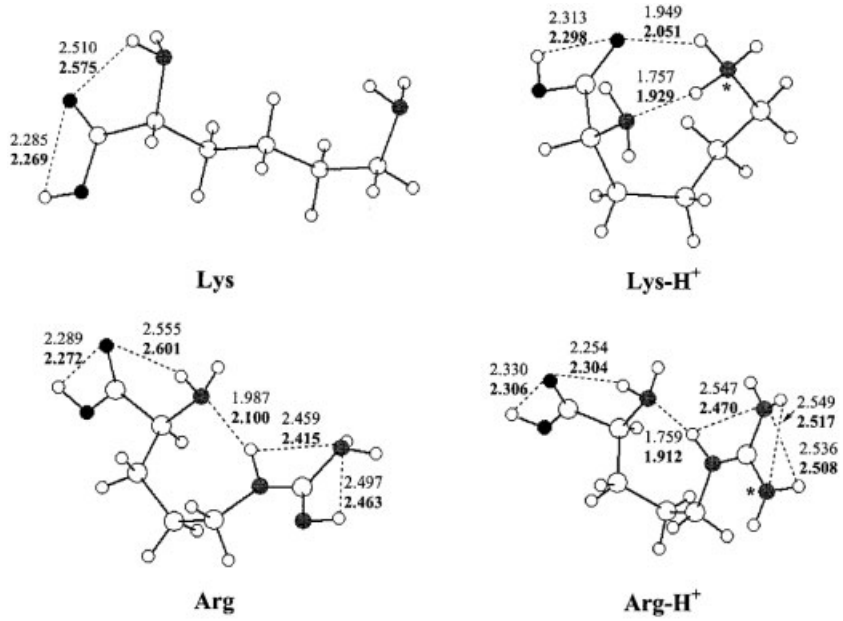

FIGURE 1 Continued.

Poutsma and coworkers [22] using B3LYP/6-31+G(*d*) level. The maximum number of intramolecular hydrogen bonds (5 and 6 H-bonds) is observed in the case of **Arg** and **Arg**—$\mathbf{H^+}$.

ENERGETICS

There has been considerable interest in the proton affinities and the site of protonation of the amino acids [22–34]. The preferred protonation site of the amino acids is known and in many cases, there is a unique α-amino nitrogen center. However, the amino acids such as **Trp**, **Asn**, **Gln**, **His**, **Lys**, and **Arg** contain more than one nitrogen atom available for protonation. To identify the most favorable site of proton attack for these six amino acids, the protons were added to all the available nitrogen sites individually to evaluate the proton affinities of designated sites. The proton affinity values only for the most favorable site and that protonated amino acid structures for the six abovementioned amino acids are given in the present study. In case of **Trp**, Rogalewicz et al. [42] have suggested indole nitrogen as the most preferred site of protonation. In contrast, Maksic and Kovacevic [23] and Lioe et al. [43] reported independently that the α-amino nitrogen is the most favorable site of protonation. Our theoretical study is in agreement with their result; the reason may be traced to the involvement of lone pair electron of indole nitrogen in the delocalization and the α-amino nitrogen lone pair is readily available for protonation. Although **Asn**, **Gln**, and **Lys** contain more than one nitrogen atoms, the most favored site of protonation is the α-amino nitrogen. However, the most preferred site of protonation for **His** and **Arg** is the unsaturated nitrogen in the ring and the side chain, respectively. **Pro** is a unique structure since it has only one nitrogen atom as secondary amine for protonation unlike other naturally occurring amino acids. Several controversies were found in the literature reports on the proton affinities of **Pro** [8, 24–26, 36]. In general, the most stable conformers in both neutral and protonated forms of amino acids possess the maximum number of intramolecular hydrogen bonds. Hydrogen atom attached to the preferred protonation site (nitrogen atom) involves hydrogen bonding with carbonylic oxygen atom in all the protonated amino acids except **His**—$\mathbf{H^+}$ and **Arg**—$\mathbf{H^+}$.

All efforts to identify the global minimum conformations of the neutral and protonated forms were taken manually. We have taken the HF/6-31G(*d*) study carried out by Maksic and Kovacevic [23] as a reference to substantiate whether the minimum energy structures are the lowest-energy ones and the values for comparison are given in Table II. Our exhaustive efforts revealed that 13 species in

TABLE II

Total energies (in hartrees) obtained at the HF/6-31G(*d*) level in this study and previously reported values[a] along with the difference (in kcal/mol) between them for neutral and protonated amino acids.

Amino acids	Neutral			Protonated		
	Our work	Reported[a]	Our work–reported	Our work	Reported[a]	Our work–reported
Gly	−282.83110	−282.83110	0.00	−283.18699	−283.18699	0.00
Ala	−321.86794	−321.86794	0.00	−322.23018	−322.23018	0.00
Cys	−719.37443	−719.37443	0.00	−719.73491	−719.73304	−1.17
Ser	−396.71829	−396.71829	0.00	−397.08540	−397.08456	−0.53
Val	−399.93591	−399.93591	0.00	−400.30209	−400.29902	−1.93
Leu	−438.97067	−438.97067	0.00	−439.33709	−439.33709	0.00
Asp	−509.48552	−509.48553	0.00	−509.85235	−509.85235	0.00
Ile	−438.96851	−438.96669	−1.14	−439.33635	−439.33246	−2.44
Thr	−435.75663	−435.75663	0.00	−436.12635	−436.12523	−0.71
Phe	−551.41190	−551.41190	0.00	−551.78524	−551.78524	0.00
Tyr	−626.26673	−626.26673	0.00	−626.64055	−626.64055	0.00
Asn	−489.65268	−489.65249	−0.12	−490.02949	−490.02871	−0.49
Met	−797.44367	−797.44367	0.00	−797.81472	−797.81472	0.00
Pro	−398.76603	−398.76268	−2.10	−399.14529	−399.13697	−5.22
Trp	−682.18001	−682.17547	−2.85	−682.55720	−682.54653	−6.70
Glu	−548.51672	−548.51601	−0.45	−548.89766	−548.89766	0.00
Gln	−528.68688	−528.68138	−3.45	−529.07517	−529.07412	−0.66
His	−545.53315	−545.53315	0.00	−545.92831	−545.92623	−1.31
Lys	−493.98985	−493.98847	−0.87	−494.38846	−494.38846	0.00
Arg	−602.89238	−602.89238	0.00	−603.31644	−603.31644	0.00

[a] Values were taken from Ref. [23].

neutral surface and ten in the protonated form have identical total energies. However, 7 neutral and 10 protonated lowest-energy conformers that we report here are lower in energy compared with those of Maksic and Kovacevic [23]. A careful comparison of the proline conformers is made with those reported by Kuntz et al. [21]. We have located eight local minima on proline PES including both minima reported in Ref. [21]; the lowest-energy conformer is marginally (0.4 kcal/mol) lower in energy than that reported earlier.

The conformational changes upon protonation and deprotonation of global minima neutral and protonated structures, respectively, using HF/6-31G(*d*) level for all the amino acids were evaluated; the values are given in Table III. Upon protonation of the conformation corresponding to the global minima of neutral amino acids, 12 (**Gly**, **Ala**, **Val**, **Leu**, **Asp**, **Ile, Phe**, **Tyr**, **Trp**, **Gln**, **His**, **Arg**) out of 20 yield the lowest-energy protonated counterparts. Thus, eight amino acids undergo a conformational change upon protonation, especially in **Thr**, **Glu**, and **Lys**. These conformational changes followed by an energetic gain in excess of 15 kcal/mol. Therefore, the protonation of amino acids or polypeptides of **Thr**, **Met**, **Glu**, and **Lys** residues are expected to trigger significant changes in the conformation. Similarly, the deprotonation was done from the global minima of protonated amino acids. While 10 structures upon deprotonation lead to structures higher in energy compared with the lowest-energy neutral conformer, the difference in energies is rather small. Most differences are within 4 kcal/mol, indicating no significant energetic gain upon deprotonation. The systematic computational study highlights the crucial aspects of conformational changes upon protonation and deprotonation of amino acids, which certainly have significant influence on protein and peptide structures. Especially, caution should be taken when the amino acid residues **Thr**, **Met**, **Glu**, and **Lys** are the protonated species. Similarly, deprotonation from protonated **Met**, **Asn**, **Gln**, and **Lys** residues may induce subtle changes in the 3D-structures of polypeptides and proteins. The present computational study also reestablishes that hydrogen bonding plays a key role in deciding the relative energy differences.

TABLE III

Total energies (in hartrees) obtained at the HF/6-31G(*d*) level for the conformers by protonation on neutral global minima and deprotonation from protonated global minima structures.*

	Neutral			Protonated		
Amino acid	Global minima (I) (hartrees)	Removal of H^+ from prot global minima (II) (hartrees)	I–II (kcal/mol)	Global minima (III) (hartrees)	Addition of H^+ on neut global minima (IV) (hartrees)	III–IV (kcal/mol)
Gly	−282.83110	−282.83110	0.00	−283.18699	−283.18699	0.00
Ala	−321.86794	−321.86794	0.00	−322.23018	−322.23018	0.00
Cys	−719.37443	−719.37154	−1.81	−719.73491	−719.73304	−1.17
Ser	−396.71829	−396.71587	−1.52	−397.08540	−397.08456	−0.53
Val	−399.93591	−399.93591	0.00	−400.30209	−400.30209	0.00
Leu	−438.97067	−438.97067	0.00	−439.33709	−439.33709	0.00
Asp	−509.48552	−509.48552	0.00	−509.85235	−509.85235	0.00
Ile	−438.96851	−438.96851	0.00	−439.33635	−439.33635	0.00
Thr	−435.75663	−435.75470	−1.21	−436.12635	−436.09161	−21.80
Phe	−551.41190	−551.41190	0.00	−551.78524	−551.78524	0.00
Tyr	−626.26673	−626.26567	−0.67	−626.64055	−626.64055	0.00
Asn	−489.65268	−489.64871	−2.49	−490.02949	−490.02591	−2.25
Met	−797.44367	−797.43780	−3.68	−797.81472	−797.80360	−6.98
Pro	−398.76603	−398.76541	−0.39	−399.14529	−399.14487	−0.26
Trp	−682.18001	−682.18001	0.00	−682.55720	−682.55720	0.00
Glu	−548.51672	−548.51475	−0.64	−548.89766	−548.86700	−19.24
Gln	−528.68688	−528.68050	−4.00	−529.07517	−529.07517	0.00
His	−545.53315	−545.53315	0.00	−545.92831	−545.92831	0.00
Lys	−493.98852	−493.98412	−2.76	−494.38846	−494.36452	−15.02
Arg	−602.89238	−602.89238	0.00	−603.31644	−603.31644	0.00

* The difference in energies (in kcal/mol) between the global minima and the conformer for protonation/deprotonation is also given.

Table IV presents the computed proton affinity values with the HF and MP2 methods, and B3LYP functional along with the reported experimental values of Hunter and Lias [36, 37]. The proton affinity values calculated using B3LYP/6-311+G(*d,p*) and MP2/6-311+G(*d,p*) are in excellent agreement with the experimental results, with the deviations within 1 kcal/mol for 10 of the amino acids. It should be noted that the deviation is large with all the methods employed for three amino acids **Pro**, **Glu**, and **Gln**. Recent experimental studies on the proton affinities of **Pro** reported the value of 224.9 kcal/mol [21], which is ~5 kcal/mol higher than the value suggested by Hunter and Lias, and this value is in good agreement with our computed results. There are no recent experimental reports on the proton affinities of **Glu** and **Gln** after Hunter and Lias. Both experimental and computational studies have limitations, and the previous results exhibited more deviations particularly for the amino acids where the intramolecular hydrogen bonding plays crucial role [22, 23, 28]. Thus, the present computational study suggests the experimental reevaluation of the proton affinities for **Glu** and **Gln** to confirm and report the most accurate proton affinity values for these two cases.

The average deviation of the proton affinities calculated with different methods is given in Table V. A quick look at Tables IV and V clearly indicates that, as expected, the Hartree–Fock method overestimates the proton affinity values of amino acids. The B3LYP and MP2 with 6-311+G(*d,p*) basis set appears to give very good agreement with the experimental results except for **Glu** and **Gln**. The average deviation of proton affinities, excluding two amino acids (**Glu** and **Gln**) in the set of 20, with the B3LYP/6-311+G(*d,p*) and MP2/6-311+G(*d,p*) with respect to the experimental values is ~1 kcal/mol. This analysis shows that the B3LYP/6-31G(*d*) has turned worst among the methods employed with ~3–7 kcal/mol deviation compared with the experi-

TABLE IV

Proton affinities of natural amino acids at various levels of theory calculated using Eq. (1).*

Amino acids	HF			MP2		B3LYP			
	6-31G(*d*)	6-311+G(*d*,*p*)[a]	cc-pVTZ[a]	6-31G(*d*)[a]	6-311+G(*d*,*p*)[a]	6-31G(*d*)	6-311+G(*d*,*p*)[a]	cc-pVTZ[a]	Expt[b]
Gly	214.9	213.4	215.6	215.6	211.9	216.8	211.5	214.2	211.9
Ala	218.9	217.4	219.6	218.9	215.1	220.8	215.6	218.1	215.5
Cys	217.9	217.5	220.0	218.6	215.8	220.6	216.3	218.9	215.9
Ser	222.2	221.1	222.9	222.4	217.7	223.2	217.8	220.3	218.6
Val	221.5	220.8	222.7	221.3	218.0	223.6	218.9	221.1	217.6
Leu	221.6	220.7	222.7	221.7	218.2	223.8	218.9	221.2	218.6
Asp	221.9	219.9	222.7	223.1	218.8	226.3	219.9	223.2	217.2
Ile	222.5	221.8	223.7	222.3	219.0	224.5	220.0	222.1	219.3
Thr	223.9	223.2	225.1	223.5	219.2	225.2	220.2	222.6	220.5
Phe	225.9	224.2	226.0	227.1	223.1	228.1	222.7	224.7	220.6
Tyr	226.2	224.6	226.7	227.8	223.8	229.2	223.5	226.0	221.3
Asn	228.1	227.7	230.2	226.6	222.7	229.4	224.5	227.3	222.0
Met	224.5	224.0	225.9	227.0	224.5	229.6	225.1	227.1	223.6
Pro	229.5	228.9	230.5	229.2	225.9	230.7	226.1	228.0	220.0
Trp	228.3	226.6	228.7	230.2	226.1	231.5	226.2	228.6	226.8
Glu	231.1	230.9	233.2	230.2	226.3	232.0	226.6	229.5	218.2
Gln	235.4	234.1	236.7	237.1	232.6	241.3	235.2	238.4	224.1
His	240.2	238.4	240.5	237.6	234.0	241.1	236.4	239.0	236.1
Lys	240.9	237.9	238.6	245.6	240.0	246.1	238.4	240.0	238.1
Arg	258.8	256.0	258.7	254.8	249.6	258.8	252.8	256.0	251.2

* Experimental values are also given for comparison. All values are in kcal/mol.

[a] Single-point calculations were done on B3LYP/6-31G(*d*) geometries; thermal correction to enthalpy values was taken from B3LYP/6-31G(*d*).

[b] Values were taken from Refs. [36, 37].

mental values. It should be mentioned that the 6-31G(*d*) basis set is not adequate to compute the proton affinities either with B3LYP or MP2. Thus, it appears to be important to utilize basis sets that have polarization functions for hydrogen atoms and also diffuse function, required to properly account for H-bonding. The present study further indicates that a basis set of triple-ζ quality with diffuse and polarization function appears to be necessary for obtaining reliable results. Figure 2 shows good correlation between the calculated and experimental proton affinity values.

Consistent with previous studies [1, 10, 23, 34], **Gly** is the least basic amino acid among the whole series of naturally occurring amino acids and is followed by **Ala**, which has 3–4 kcal/mol higher proton affinity value than **Gly**. The slight enhancement of proton affinity for **Cys** compared with **Ala** and further increase for **Ser** indicate the effect of the introduction of β-SH or β-OH group on the proton affinities. Comparing the proton affinity values of **Leu** and **Ile** indicates that the position of attaching methyl group has small effect on the proton affinity. The former where the methyl group attached to γ-position has slightly less proton affinity value than that of the latter. A small increase in the proton affinity value of **Tyr** than that of **Phe** may be attributed to the introduction of —OH group at the *para*-position in the phenyl ring. The amino acids **Asn** and **Gln** possess side chain amide group (—$CONH_2$), whereas the carboxylic acid group (—COOH) is present in **Asp** and **Glu**. Proton affinity values of **Asn** and **Gln** are ~4.5 and 8.5 kcal/mol higher, respectively, compared with the corresponding acids. The high proton affinity values for **Glu** and **Gln** may be attributed to the strong H-bond formation upon protonation by two of the hydrogens attached to protonated nitrogen atom with carbonylic oxygen. Similarly, the high proton affinity value for **Lys** could be traced to the formation of strong intramolecular H-bond between one of the hydrogens of protonated nitrogen and N atom of α-NH_2 group. In addition, one more hydrogen attached to protonated nitrogen forms H-

TABLE V

Average deviation of the proton affinity values (in kcal/mol) at different levels of theory with respect to the experimental values.

Level of theory	Considering all the amino acids	Excluding **Pro**, **Glu** and **Gln**	Excluding **Glu** and **Gln**[a]
HF/6-31G(*d*)	4.60	3.42	3.49
HF/6-311+G(*d*,*p*)[a]	3.55	2.32	2.41
HF/cc-pVTZ[a]	5.41	4.12	4.21
MP2/6-31G(*d*)[a]	4.91	3.76	3.79
MP2/6-311+G(*d*,*p*)[a]	2.06	1.10	1.09
B3LYP/6-31G(*d*)	7.01	5.79	5.79
B3LYP/6-311+G(*d*,*p*)[a]	2.25	1.14	1.14
B3LYP/cc-pVTZ[a]	4.32	3.11	3.11

[a] The experimental proton affinity value for proline (**Pro**) was taken from Ref. [21]; for remaining 17 amino acids are taken from Refs. [36, 37].
[b] Single-point calculations were done on the B3LYP/6-31G(*d*) geometries.

bond with carbonyl oxygen. These hydrogen bonds are formed due to the conformational flexibility by the presence of four methylene groups. Previous computational studies have reported such conformers upon protonation of **Lys** and its model compound [22, 23, 27].

Lys and **His** are the most basic amino acids among the naturally occurring amino acids next to **Arg**. In most of the previous studies, **His** is reported to be more basic than **Lys** [1, 10, 12]. However, Wu and Fenselau [13] used the extended kinetic method for the PA measurements of **Lys** and **His** and reported higher entropy changes in **Lys** due to intramolecular hydrogen bonding. They reported the values of 235.3 and 234.0 kcal/mol for **Lys** and **His**, respectively. Hunter and Lias [36] have also reported **Lys** (238.1 kcal/mol) to be more basic than **His** (236.1 kcal/mol). Our results authenticate with the values reported by Hunter and Lias. Both experimentally and theoretically, the proton affinities increase in the series **Gly**, **Ala**, **Val**, and **Leu**. Besides, proton affinities of **Met**, **Glu**, and **Gln** are ~7–11 kcal/mol higher than that of **Cys**, **Asp**, and **Asn**, respectively, indicating the effect of the number of methyl or methylene groups in the tail of α-amino acids.

Conclusions

The systematic quantum chemical calculations were performed with the HF and MP2 methods, and B3LYP functional to obtain the proton affinity values for natural amino acids. The proton affinities calculated using B3LYP/6-311+G(*d*,*p*) and MP2/6-

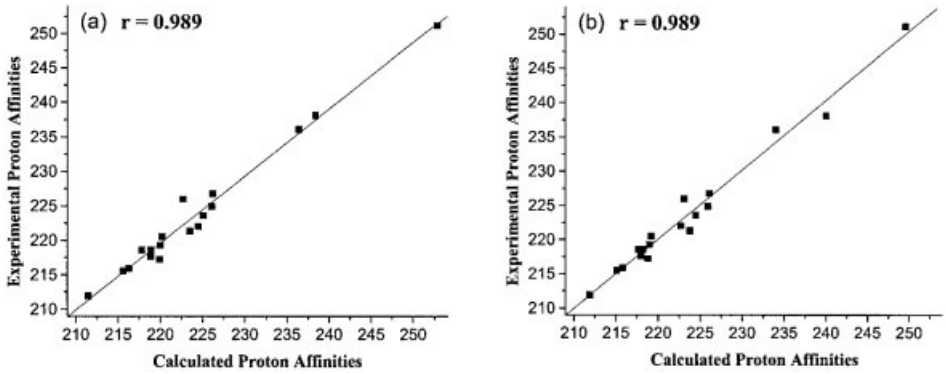

FIGURE 2. Correlation between the calculated and the experimental proton affinities. Calculated values with (a) B3LYP/6-311+G(*d*,*p*) and (b) MP2/6-311+G(*d*,*p*), using B3LYP/6-31G(*d*) geometries.

311+G(*d*,*p*) are in excellent agreement with the experimental reports. Electron correlation is important and HF method appears to be highly inadequate. The basis set of 6-31G(*d*) should be avoided and choosing the right basis set with B3LYP and MP2 seems to be crucial to obtain the highly reliable proton affinity values. Good correlation is obtained between calculated and experimental proton affinities. The number of hydrogen bonds largely controls the conformational preferences. The number and the relative strength of the hydrogen bonds play central role in obtaining reliable proton affinity values. The current study suggests for the experimental reevaluation of the proton affinity of **Glu** and **Gln** to resolve the discrepancy between theory and experiment. The present study further indicates that the global minima of neutral amino acid need not give the global minima of protonated conformer and vice-versa. Thus, this study recommends for thorough conformational search to find out the global minima of neutral and protonated species of amino acids to obtain reliable information about the proton affinities.

ACKNOWLEDGMENTS

Several useful discussions with Dr. M. Vairamani on the various experimental techniques of gas phase proton affinity measurements are greatly acknowledged.

Supplementary Material Available

The optimized geometries and the proton affinity values obtained for the different sites of the amino acids that contain more than one nitrogen atom, the total energies and the B3LYP/6-31G(*d*) optimized Cartesian coordinates for neutral and protonated species of all the amino acids.

References

1. Harrison, A.G. Mass Spectrom Rev 1997, 16, 201.
2. Taft, R. W. In Proton Transfer Reactions; Caldin, E.; Gold, V., Eds.; Chapman & Hall: London, 1975; pp 31–77.
3. Bojesen, G. J Am Chem Soc 1987, 109, 5557.
4. Caprioli, R. M., Ed. Continuous-Flow Fast Atom Bombardment Mass Spectrometry; Wiley: New York, 1990.
5. Benninghoven, A.; Roudenauer, F. G.; Werner, H. W., Eds. Secondary Ion Mass Spectrometry: Basic Concepts, Instrumental Aspects, Applications, and Trends; Wiley: New York, 1987.
6. Cole, R. B., Ed. Electrospray Inonization Mass Spectrometry: Fundamentals, Instrumentation, and Applications; Wiley: New York, 1997.
7. Gorman, G. S.; Speir, J. P.; Turner, C. A.; Amster, I. J. J Am Chem Soc 1992, 114, 3986.
8. Mirza, S. P.; Prabhakar, S.; Vairamani, M. Rapid Commun Mass Spectrom 2001, 15, 957.
9. Meot-Ner, M.; Hunter, E. P.; Field, F. H. J Am Chem Soc 1979, 101, 686.
10. Bojesen, G.; Breindahl, T. J Chem Soc Perkin Trans 2, 1994, 1029.
11. Li, X.; Harrison, A. G. Org Mass Spectrom 1993, 28, 366.
12. Wu, Z.; Fenselau, C. Rapid Commun Mass Spectrom 1992, 6, 403.
13. Wu, Z.; Fenselau, C. Rapid Commun Mass Spectrom 1994, 8, 777.
14. Isa, K.; Omote, T.; Amaya, M. Org Mass Spectrom 1990, 25, 620.
15. Cooks, R. G.; Kruger, T. L. J Am Chem Soc 1977, 99, 1279.
16. McLuckey, S.A.; Cameron, D.; Cooks, R. G. J Am Chem Soc 1981, 103, 1313.
17. Cooks, R. G.; Koskinen, J. T.; Thomas, P. D. J Mass Spectrom 1999, 34, 85.
18. Cooks, R. G.; Patrick, J. S.; Kotiaho, T.; McLuckey, S. A. Mass Spec Rev 1994, 13, 287.
19. Armentrout, P. B. J Am Soc Mass Spectrom 2000, 11, 371.
20. Afonso, C.; Modeste, F.; Breton, P.; Fournier, F.; Tabet, J. C. Eur J Mass Spectrom 2000, 6, 443.
21. Kuntz, A. F.; Boynton, A. W.; David, G. A.; Colyer, K. E.; Poutsma, J. C. J Am Soc Mass Spectrom 2002, 13, 72.
22. Schroeder, O. E.; Andriole, E. J.; Carver, K. L.; Colyer, K. E.; Poutsma, J. C. J Phys Chem A 2004, 108, 326.
23. Maksic, Z. B.; Kovacevic, B. Chem Phys Lett 1999, 307, 497.
24. Mezzache, S.; Afonso, C.; Pepe, C.; Karoyan, P.; Fournier, F.; Tabet, J.-C. Rapid Commun Mass Spectrom 2003, 17, 1626.
25. Pepe, C.; Rochut, S.; Paumard, J.-P.; Tabet, J.-C. Rapid Commun Mass Spectrom 2004, 18, 307.
26. Marino, T.; Russo, N.; Tocci, E.; Toscano, M. J Mass Spectrom 2001, 36, 301.
27. Bliznyuk, A. A.; Schaefer, H. F., III; Amster, I. J. J Am Chem Soc 1993, 115, 5149.
28. Sun, W.; Kinsel, G. R.; Marynick, D. S. J Phys Chem A 1999, 103, 4113.
29. Perez, P.; Contreras, R. A Chem Phys Lett 1998, 293, 239.
30. Baeten, A.; De Proft, F.; Geerlings, P. Int J Quantum Chem 1996, 60, 931.
31. Smith, B. J.; Radom, L. J Am Chem Soc 1993, 115, 4885.
32. Topol, I. A.; Burt, S. K.; Russo, N.; Toscano, M. J Am Soc Mass Spectrom 1999, 10, 318.
33. Topol, I. A.; Burt, S. K.; Toscano, M.; Russo, N. J Mol Struct (Theochem) 1998, 430, 41.
34. Noguera, M.; Rodriguez-Santiago, L.; Sodupe, M.; Bertran, J. J Mol Struct (Theochem) 2001, 537, 307.
35. Podolyan, Y.; Gorb, L.; Leszczynski, J. J Phys Chem A 2000, 104, 7346.

36. Hunter, E. P. L.; Lias, S. G. J Phys Chem Ref Data 1998, 27, 413.
37. NIST database of PA values (http://webbook.nist.gov/chemistry).
38. Solomons, T. W. G. In Fundamentals of Organic Chemistry, Wiley: Wiley, edn. 4, 1994; p 974.
39. Del Bene, J. E. J Phys Chem 1993, 97, 107.
40. Merril, G. N.; Kass, S. R. J Phys Chem 1996, 100, 17465.
41. 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.; Rega, N.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; 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.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. Gaussian 98; Revision A.11.2; Gaussian: Pittsburgh, PA, 2001.
42. Rogalewicz, F.; Hoppilliard, Y.; Ohanessian, G. Int J Mass Spectrom 2000, 195/196, 565.
43. Lioe, H.; O'Hair, R. A. J.; Reid, G. E. J Am Soc Mass Spectrom 2004, 15, 65.