ChemInform Abstract: Gas Phase Basicities of Polyfunctional Molecules. Part 3. Amino Acids

ARTICLE *in* CHEMINFORM · DECEMBER 2012

Impact Factor: 0.74 · DOI: 10.1002/chin.201249267

1 AUTHOR:

Guy Bouchoux

Université Paris-Sud 11

231 PUBLICATIONS **2,804** CITATIONS

SEE PROFILE

GAS PHASE BASICITIES OF POLYFUNCTIONAL MOLECULES. PART 3: AMINO ACIDS

Guy Bouchoux*
Département de Chimie, Laboratoire des Mécanismes Réactionnels, Ecole Polytechnique, 91120 Palaiseau, France

Received 23 June 2011; accepted 22 August 2011

Published online 5 October 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/mas.20349

The present article is the third part of a general overview of the gas-phase protonation thermochemistry of polyfunctional molecules (first part: Mass Spectrom. Rev., 2007, 26:775-835, second part: Mass Spectrom. Rev., 2011, in press). This review is devoted to the 20 proteinogenic amino acids and is divided in two parts. In the first one, the experimental data obtained during the last 30 years using the equilibrium, thermokinetic and kinetic methods are presented. A general re-assignment of the values originating from these various experiments has been done on the basis of the commonly accepted Hunter&Lias (1998) gas-phase basicity scale in order to provide an homogeneous set of data. In the second part, theoretical investigations on gaseous neutral and protonated amino acids are reviewed. Conformational landscapes of both types of species were examined in order to provide theoretical protonation thermochemistry based on the truly identified most stable conformers. Proton affinities computed at the presently highest levels of theory (i.e. composite methods such as Gn procedures) are presented. Estimates of thermochemical parameters calculated using a Boltzmann distribution of conformers at 298K are also included. Finally, comparison between experiment and theory is discussed and a set of evaluated proton affinities, gas-phase basicities and protonation entropies is proposed. © 2011 Wiley Periodicals, Inc. Mass Spec Rev 31: 391–435, 2012

Keywords: *protonation thermochemistry; proton affinity; protonation entropy; protonation site; intramolecular hydrogen bond; conformational landscapes*

I. INTRODUCTION

Detailed knowledge on the structural and energetic aspects of the protonation of amino acids is of obvious interest in many areas of chemistry and biochemistry. These molecules, which constitute the basic structural units of the proteins of living species, are indeed polyfunctional compounds particularly rich from the point of view of their conformational structures and of their protonation sites. Moreover, the structure determination of amino acids and their polymers by mass spectrometry involves, most of the time, elucidation of the fragmentation pathways of their protonated forms. Thus the knowledge of the sites of proton attachment as well as the corresponding energetic aspect is essential for the understanding of the fragmentation reactions and the structural information obtained. For these reasons the intrinsic acid-bases properties of the 20 naturally occurring α-amino acids (Schemes 1–5) have attracted the interest of a number of researchers since several decades. We will present in this review a survey and, when possible, a critical evaluation, of the experimental and theoretical results gained during the last 30 years on the gas phase protonation of these fundamental molecules.

Two reviews related to the gas phase basicities and proton affinities of amino acids and peptides appeared almost 15 years ago in the present journal (Green & Lebrilla, 1997; Harrison, 1997). The methods of establishing GB and PA were summarized and a set of recommended values has been proposed for amino acids and some peptides. At that time however, the gas phase basicity scale was in the state of revision. In particular, reassignment of the upper part of the basicity scale was under debate (for a detailed description of this question see Mautner, 2003). One year later, an extended compilation of basicity data was appearing (Hunter & Lias, 1998) which proposes evaluated data sometimes at variance from the Green & Lebrilla and Harrison's estimates. In addition, during the latest decade, new experiments on amino acids were designed and, as detailed in the introductory part of this review (Bouchoux, 2007), means to obtain thermochemical data from mass spectrometry experiments were critically evaluated and improved. One goal of the present review was to include these latest results and to compare them with the previously obtained experimental data (Section II).

The extraordinary development of quantum chemistry methods during the last years has been at the origin of a large number of theoretical investigations on the structures and energies of neutral and protonated amino acids. These studies contributed deeply to the knowledge of the protonation thermochemistry of these molecules in providing both thermochemical data of increasing accuracy and invaluable information's on the protonation sites and conformational preferences. Proton affinities and protonation entropies, together with structural aspects of the protonation processes, recently obtained by theoretical methods will be presented in Section III.

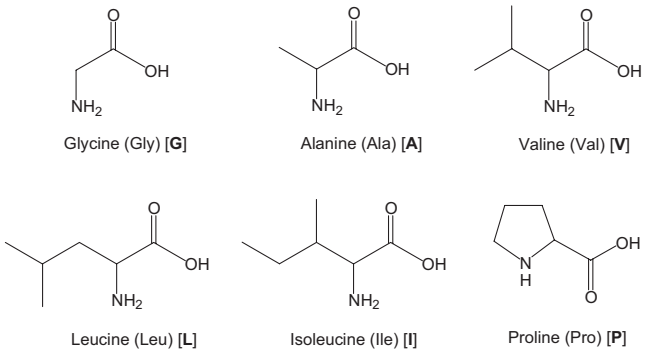
Before entering into the examination of the experimental and theoretical results concerning the gas phase proton thermochemistry of amino acids, several relevant definitions may be briefly recalled. Proton affinity, $PA_T(M)$, and gas phase basicity, $GB_T(M)$, of a molecule M at a temperature T are defined as the standard enthalpy, $\Delta_1 H_T^\circ$, and standard Gibbs free energy, $\Delta_1 G_T^\circ$, of reaction (1), respectively:

$$\underset{(gas)}{MH^+} \rightarrow \underset{(gas)}{M} + \underset{(gas)}{H^+} \qquad (1)$$

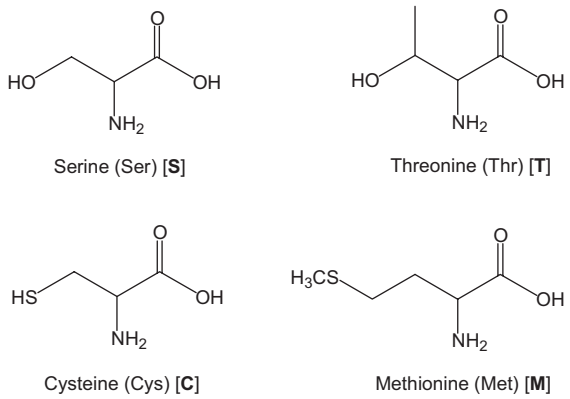
Introducing the entropy of reaction (1), $\Delta_1 S_T^\circ$, the thermodynamic relationship $\Delta G_T^\circ = \Delta H_T^\circ - T\Delta S_T^\circ$ obviously

Dedicated to Pr. A.G. Harrison at the occasion of his 80th birthday and in recognition to his outstanding contributions to gas phase ion chemistry and thermochemistry.

**Correspondence to:* Guy Bouchoux, Département de Chimie, Laboratoire des Mécanismes Réactionnels, Ecole Polytechnique, 91120 Palaiseau, France. E-mail: bouchoux@dcmr.polytechnique.fr



SCHEME 1. Alkyl substituted amino acids.



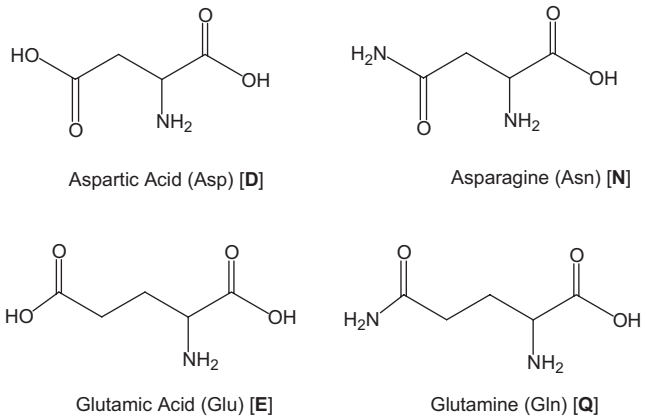
SCHEME 2. Oxygen and sulfur substituted amino acids.

leads to Equation (2):

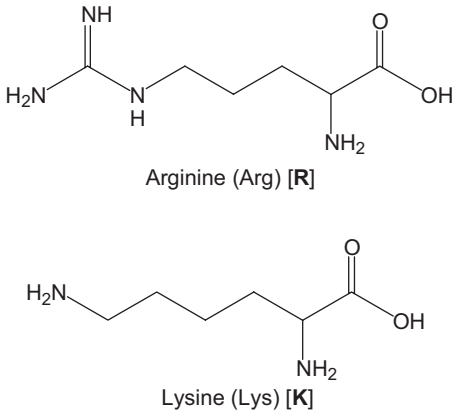
$$GB_T(M) = PA_T(M) - T\Delta_1 S^\circ_T \quad (2)$$

If we call the difference $\Delta_p\ S_T^\circ(M) = S_T^\circ(MH^+) - S_T^\circ(M)$ the ''protonation entropy'' of the species M (Hunter & Lias, 1998), then, the entropy of reaction (1), $\Delta_1\ S_T^\circ$, reduces to $\Delta_1\ S_T^\circ = S_T^\circ(H^+) - \Delta_p\ S_T^\circ(M)$, and consequently, the gas phase basicity and proton affinities are interrelated following Equation (3):

$$GB_T(M) = PA_T(M) - T[S^\circ_T(H^+) - \Delta_p S^\circ_T(M)] \quad (3)$$



SCHEME 3. Carbonylated amino acids.

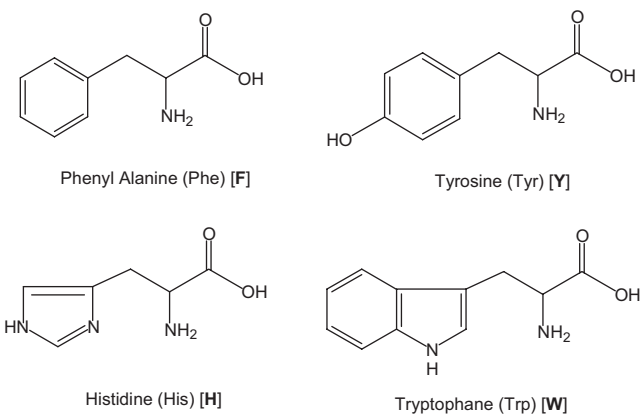


SCHEME 4. Nitrogen substituted amino acids.

All the quantities $PA_T(M)$, $GB_T(M)$, and $\Delta_p\ S_T^\circ(M)$ are generally given, and tabulated, at the temperature of 298 K. In the following lines, the simplified notations PA(M), GB(M), and $\Delta_p\ S^\circ(M)$ will be used to denote proton affinity, gas phase basicity, and protonation entropy, respectively, at 298 K.

II. EXPERIMENTAL PROTONATION THERMOCHEMISTRY OF AMINO ACIDS

This part is organized in two sections dealing with the general means to access thermochemical parameters: proton transfer equilibrium method, thermokinetic method and kinetic methods. Results obtained by the two formers methods are gathered in Section A (since both provide directly Gibbs free energies of protonation) while the applications of the ''simple'' and ''extended'' kinetic methods are presented in Section B. The first precaution to be taken when considering literature experimental data is to identify the bases which have been taken as reference and to adjust the original results to a common proton affinity scale. In the present review we have uniformly reevaluated the original data to the presently widely used Hunter & Lias scale (1998). All the results, that is, gas phase basicities, proton affinities and protonation entropies, coming from experiment and anchored to this gas phase basicity scale are gathered in Table 1.



SCHEME 5. Aromatic amino acids.

TABLE 1. Experimental gas-phase protonation thermochemistry of α-amino acids*

M	Method	GB(M) kJ.mol^{-1}	PA(M) kJ.mol^{-1}	$\Delta_p S°(M)$ J.K^{-1}.mol^{-1}
Glycine	equilibrium	850.2±0.8[a]	882.5±0.8[a]	2±8 [a]
		856.6[b]		
	thermokinetic	856.6±1.9[c]		
	simple kinetic		889.2[d]	
			887.1[l]	
			891.8[f]	
	extended kinetic	855.4±3.6[l]	886.3±3.1[l]	2±6[l]
	average	**854.7±3.1**	**887.4±3.5**	
Alanine	equilibrium	874.7[a]/ 864.5[a]	900.1[a]	11[a]
		869.3[b]		
	thermokinetic	867.6±0.4[c]		
	simple kinetic		[900.3] 902.7[d]	
			899.9[f]	
	extended kinetic	869.6[j]	902.0[j]	0[j]
	average	**869.1±3.7**	**901.2±1.4**	
Valine	equilibrium	877.6[a]/ 876.3[a]/874.7[a]		
		881.2[b]		
	simple kinetic		[909.7] 914.1[d]	
			917.3[e]	
			917[f]	
	average	**877.5±2.8**	**916.1±1.8**	
Leucine	equilibrium	881.2[a]/ 877.6[a]/877.1[a]		
		886.2[b]		
	simple kinetic		[911.2] 916.0[d]	
			920.3[e]	
			918.1[f]	
	average	**882.4±5.4**	**918.1±2.2**	
Isoleucine	equilibrium	890.3[b]		
	simple kinetic		[914.1] 919.6[d]	
			922.8[e]	
			921.0[f]	
	average	**890.3**	**921.1±1.6**	
Proline	equilibrium	898.7[a]/ 896.8[a]	932.0[a]	-3
		911.0[b]		
	thermokinetic	898.3±1.6[c]		
	simple kinetic		[930.2] 939.3[d]	
			938.5[e]	
			946.9[f]	
	extended kinetic	902.1[h]	(926.2)[h]	(-28)[h]
		909.2[i]	941.0±6.7[i]	-8 [i]
		905.8[m]	(947.5±4.7)[m]	(-31±8)[m]
	average	**903.1±5.6**	**941.4±3.8**	

TABLE 1. (*Continued*)

Serine	equilibrium	880.3[b]		
	thermokinetic	876.2±4.3[c]		
	simple kinetic		[906.6] 910.4[d]	
			912.9[e]	
			911.4[f]	
	average	**878.3±2.9**	**911.6±1.3**	
Threonine	equilibrium	888.5[b]		
	simple kinetic		[913.4] 918.7[d]	
			922.8[e]	
			921.4[f]	
	average		**921.0±2.1**	
Cysteine	equilibrium	868.8[b]		
	simple kinetic		[901.4] 904.0[d]	
			901.9[f]	
	average		**903.0±1.5**	
Methionine	equilibrium	901.3[b]		
	bracketing	899.0		
	simple kinetic		[921.3] 928.4[d]	
			931.6[e]	
			927.2[n]	
			936.5[f]	
			928.7[o]	
	average		**{930.5±3.7}**	
	extended kinetic	898.5±3.2[o]	937.5±2.9[o]	-22±5[o]
			(924.7-931.4)[n]	
	average	**899.6±1.5**		
Aspartic Acid	equilibrium	879.8[b]		
	simple kinetic		917.3[e]	
			915.9[f]	
			912.0[l]	
	average		**{915.1±2.7}**	
	extended kinetic	883.6±2.3[l]	926.8±1.8[l]	-36±5[l]
	average	**881.7±2.7**		
Asparagine	equilibrium	(894.6)[b]		
	simple kinetic		[918.5] 934.6[d]	
			937.1[e]	
			943.2[f]	
			937.0[l]	
	average		**{938.0±3.7}**	
	extended kinetic	(920.0±5.6)[l]	(965.2±5.2)[l]	(-43±7)[l]
Glutamic acid	equilibrium	(878.9)[b]		
	simple kinetic		938.1[e]	
			944.4[f]	
			935.0[l]	
			932.1[v]	
	average		**{937.4±5.3}**	
	extended kinetic	904.5±3.0[s]	945.3±2.8[s]	-28±4[s]

TABLE 1. (*Continued*)

Glutamine	equilibrium	(888.2)[b]		
	simple kinetic		[937.1] 947.7[d]	
			960.7[e]	
			958.0[f]	
			961.0[l]	
	average		**{956.9±6.2}**	
	extended kinetic	(939.3±4.2)[l]	(988.1±3.6)[l]	(-55±7)[l]
Arginine	simple kinetic		(>1039)[e]	
			1041[g]	
			1037.6[p]	
	average		**{1040.1±1.3}**	
	extended kinetic	1004.3±2.2[p]	1043.9±1.9[p]	-24±4[p]
		1005.9±3.1[p+g]	1047.7±2.6[p+g]	-31±6[p+g]
	average	**1005.1±1.1**	**1045.8±2.7**	
Lysine	equilibrium	943.9[b]		
	thermokinetic	942.6±7.5[c]		
	simple kinetic		[940.9] 952.2[d]	
			969.6[e]	
			969.5[f]	
			972.0[l]	
	average		**{965.8±9.2}**	
	extended kinetic	948.9[r]	993.8[r]	-42[r]
		(951.1)[k]	(1007±41)[k]	(-80±110)[k]
		952.8±3.4[l]	993.9±3.1[l]	-29±5[l]
	average	**947.0±4.7**	**993.9±0.1**	
Phenylalanine	equilibrium	882.1[a]/877.9[a]		
		895.4[b]		
	simple kinetic		[916.7] 922.7[d]	
			926.2[e]	
			927.7[f]	
			921.0[q]	
	average		**{924.4±3.1}**	
	extended kinetic	892.0±1.3[q]	931.3±1.1[q]	-23±2[q]
	average	**886.9±8.2**		
Tyrosine	equilibrium	905.4[b]		
	simple kinetic		[919.2] 925.7[d]	
			930.2[e]	
			929.8[f]	
			922.8[q]	
	average		**{927.1±3.5}**	
	extended kinetic	894.6±2.6[q]	934.2±2.3[q]	-24±4[q]
	average	**900.0±7.6**		

TABLE 1. (*Continued*)

Histidine	equilibrium	950.0[b]		
	thermokinetic	942.6±5.4[c]		
	simple kinetic		978.5[e]	
			972.3[f]	
			972.0[l]	
	average		**{974.3±3.7}**	
	extended kinetic	950.3[r]	987.9[r]	-17[r]
		952.6±3.2[l]	996.0±2.8[l]	-37±5[l]
	average	**948.9±4.3**	**992.0±5.7**	
Tryptophan	equilibrium	919.6[b]		
	simple kinetic		[931.5] 940.8[d]	
			944.0[e]	
			946.6[f]	
	average		**{943.8±2.9}**	
	extended kinetic	908.4±2.1[q]	945.6±2.0[q]	-16±2[q]
		(898.0)[h]	(934.6)[h]	(-14)[h]
	average	**915.3±6.1**		

*Data anchored to GB(NH_3) = 819.0 kJ/mol and PA(NH_3) = 853.6 kJ/mol (Hunter & Lias, 1998) and corrected to the Hunter & Lias scale using linear correlation between the original and present basicity scale. Values in parentheses are considered as suspect and are not included in the average calculation.

a: Mautner, Hunter, & Field (1979), equilibrium constant measurements at variable temperature in a high pressure mass spectrometer, as given in Hunter & Lias (1998); b: Locke, Hunter, & McIver (1979), Locke (1981), Locke et al. (1983), equilibrium constant measurements in a ion cyclotron resonance mass spectrometer (values adapted by Hunter & Lias, 1998 by assuming a temperature of 350 K); c: Bouchoux & Salpin (2003), thermokinetic method (without ΔG_a correction; Bouchoux, 2006b); d: Li & Harrison (1993), apparent proton affinities determined by the simple kinetic method, corrected to the Hunter & Lias (1998) scale using the relationships (i) PA(Hunter & Lias) =PA(Li & Harrison)*1.0321–21.892 kJ/mol obtained by linear regression of the PA values of the 21 reference base (amines) (in brackets), and (ii) (PA(Hunter & Lias) =PA(Li & Harrison)*1.2618–224.76 kJ/mol) obtained from a correlation between the **amino acids** proton affinities; e: Bojesen & Breindahl (1994), apparent proton affinities determined by the simple kinetic method, corrected to the Hunter & Lias scale using the relationship PA(Hunter & Lias) =PA(Bojesen & Breindahl)*1.179–158.5 kJ/mol obtained by linear regression of the PA values of the 17 reference bases (amines) used by Bojesen and Breindahl; f: Afonso et al. (2000), apparent proton affinities determined by the simple kinetic method, corrected to the Hunter & Lias scale using relationship obtained by linear regression of the PA values of the five reference bases (Ser, Leu, Thr, Met, Trp); g: Wu & Fenselau (1992), simple kinetic method with MIKE experiments (reference base: amines); h: Mirza, Prabhar, & Vairamani (2001), extended kinetic method originally using the Bojesen & Breindahl PA scale. The values indicated here are corrected to the Hunter & Lias scale using the same relationship as in e (reference base: amino acids); i: Kuntz et al. (2002) (reference base: amines); j: Hahn & Wesdemiotis (2003) (reference base: amines); k: Schroeder et al. (2004) (reference base: amines); l: (Bouchoux et al., 2004a) (reference base: amines); m: Mezzache et al. 2005b (reference base: amines); n: Lioe et al. (2007) (reference base: amino acids); o: Desaphy, Malosse, & Bouchoux (2008) (reference base: amines); p: Bouchoux et al. (2008) (reference base: amines); q: Bouchoux et al. (2009a) (reference base: amines); r: Wu & Fenselau (1994) (reference base: amines); s: Bouchoux, Bourcier, & Nacer (2009b) (reference base: amines).

A. Gas Phase Basicities From Equilibrium and Thermokinetic Measurements

The first information concerning the gas phase basicity of amino acids came from the determination of proton transfer equilibrium constants, either by high pressure mass spectrometry (Mautner, Hunter, & Field, 1979) or ion cyclotron resonance mass spectrometry (Locke, Hunter, & McIver, 1979; Locke, 1981; Locke & McIver, 1983). The former experiments were done for a limited number of amino acids and allowed, for glycine, alanine, and proline, the determination of the proton affinity and protonation entropy, $\Delta_p S^\circ$, from equilibrium constant measurements at variable temperature. Thermochemical parameters deduced from these HPMS experiments and corrected to the Hunter & Lias scale (1998) are reported in Table 1.

The set of data coming from ICR experiments encompasses almost all the naturally occurring amino acids, the only

exception being arginine. Note that for most of the Locke results, the compilations of Lias et al. (1984) and Hunter and Lias (1998) are the only easily accessible sources of data. Measurements were done at only one temperature originally assumed to be equal to 300 K but subsequently estimated to 320 K (Lias, Liebman, & Levin, 1984). In their 1998 compendium, Hunter and Lias applied a second correction to the Locke results by adjusting the Gibbs free energy values to a temperature of 350 K (Hunter & Lias, 1998). These latter values are those reported in Table 1. When a comparison is possible between the gas phase basicities obtained by these two equilibrium techniques, it appears (Table 1 and Fig. 1) that the ICR (Locke, 1981) estimates are always above the HPMS (Mautner, Hunter, & Field, 1979) values. The mean absolute deviation (MAD) is equal to 7.9 kJ/mol. There is no evident explanation for this systematic deviation since the major sources of errors in the equilibrium method [i.e., (i) error in the estimate of the temperature of the experiment and (ii) error in the measurement of the pressure of the considered molecule M, particularly due to the presence of more volatile impurities or decomposition products] may affect both ICR or HPMS experiments.

Proton transfer reaction bracketing has been also used by several authors to propose a gas phase basicity order of amino acids. As discussed in the Part 1 of this review (Bouchoux, 2007), the uncertainty attached with GB values obtained by this method may be unpredictably high when the reaction efficiency is only qualitatively estimated. The results obtained by this procedure (Gorman et al., 1992) are thus not presented in Table 1 (the reader may find the corresponding data, re-evaluated to $GB(NH_3) = 819.0$ kJ/mol in Harrison, 1997 and Hunter & Lias, 1998). However, a quantitative use of bracketing experiments is possible when reaction efficiencies, determined for a sufficiently large set of reference bases, are available. This procedure has been applied to glycine (Wu & Lebrilla, 1993; Zhang et al., 1993), alanine (Cassady et al., 1995), serine (Carr & Cassady, 1996), lysine (Carr & Cassady, 1996), histidine (Carr & Cassady, 1996), and proline (Ewing, Zhang, & Cassady, 1996). The values presented in Table 1 and in Figure 1 were obtained by a thermokinetic treatment (Bouchoux, Salpin, & Leblanc, 1996; Bouchoux, 2007) of the original data (Bouchoux & Salpin, 2003). We note that, for these six amino acids, there is a correct agreement between the equilibrium methods and thermokinetic results (MAD = 4.5 kJ/mol), even for histidine and lysine where appreciable protonation entropies are operating.

B. Protonation Energetics of Amino Acids Determined by the Kinetic Methods

A number of studies concerning amino acid proton affinities have been carried out using the kinetic method where the competitive dissociations of a series of proton bound dimers $[MHB_i]^+$ (where M is the molecule of interest and B_i reference bases) are considered (Scheme 6).

Proton affinity orders have been qualitatively established by simply considering the sign of $y_i = \ln[MH^+]/[B_iH^+]$ (Bojesen, 1987; Isa, Omote, & Amaya, 1990). In fact, as detailed in the first part of this review (Bouchoux, 2007), y_i is related to the difference in gas phase basicity between M and B_i, at an "effective" temperature T (assumed to be the same for all the protonated dimers in a set of experiments) and its relationship with the proton affinities is given by the fundamental equation (4):

$$y_i = \ln\left(\frac{[MH]^+}{[B_iH]^+}\right) = \frac{1}{RT[PA_{298}(M) - PA_{298}(B_i) - T\Delta_i S^\circ_{298}]} \quad (4)$$

It immediately appears from examination of Equation (4) that the intercept of the $\ln(MH^+)/(B_iH^+)$ versus $PA(B_i)$ line with the x-axis corresponds to an apparent proton affinity given by:

$$PA_{app}(M) = PA_{298}(M) + T\Delta_i S^\circ_{298} \quad (5)$$

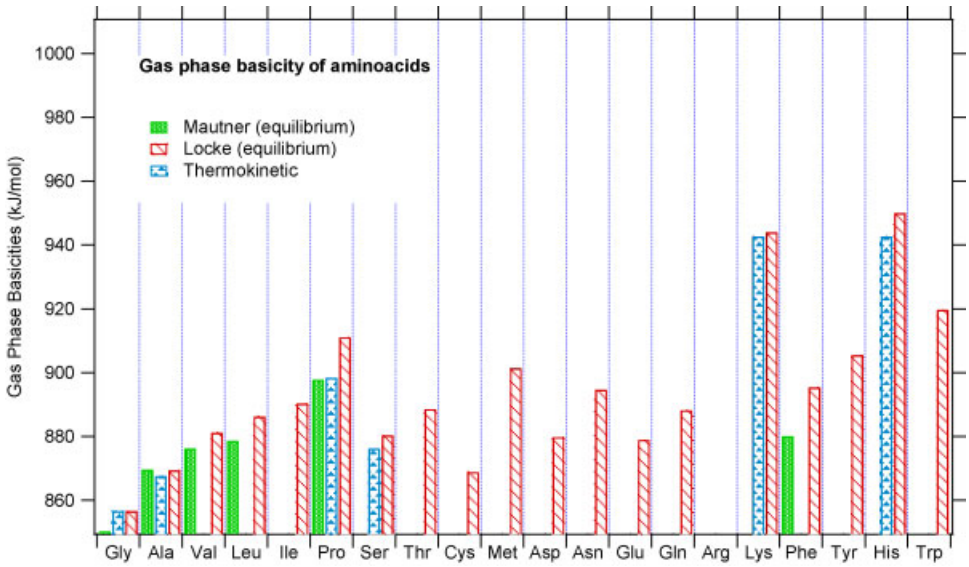
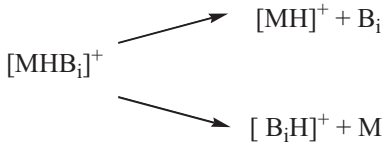


FIGURE 1. Summary of gas-phase basicities determined by the equilibrium and thermokinetic methods.



SCHEME 6

Consequently the proton affinity of the unknown, $PA_{298}(M)$ is associated with $y_i = 0$ if the difference in protonation entropy:

$$\Delta_i S^\circ_{298} = \Delta_p S^\circ(M) - \Delta_p S^\circ(B_i) = [S^\circ(MH^+) - S^\circ(M)] - [S^\circ(B_iH^+) - S^\circ(B_i)] \quad (6)$$

is equal to zero. In the first age of the kinetic method, this limitation was not fully evidenced and $\Delta_i S^\circ_{298}$ was assumed to be negligible and thus $PA_{298}(M)$ assigned to the apparent proton affinity value $PA_{app}(M)$. This simple version of the kinetic method has been applied by several authors to amino acids. Most of the time monoamines have been used as reference bases B_i thus limiting $\Delta_p S^\circ(B_i)$ to small values, typically in the range −5 to +5 J/mol K (Hunter & Lias, 1998). Sometimes, several amino acids were taken as reference bases with the aim to compensate both the $\Delta_p S^\circ(M)$ and $\Delta_p S^\circ(B_i)$ terms. Unfortunately, the protonation entropies of amino acids are strongly dependent on the structure of the lateral chain (the range of values lies between −80 and 10 J/mol K, Table 1). Thus, the neglect of the $\Delta_i S^\circ_{298} = \Delta_p S^\circ(M) - \Delta_p S^\circ(B_i)$ difference is not always possible. Moreover, for those amino acids possessing significant $\Delta_p S^\circ$, the true proton affinity is often known with a poor accuracy (see below). Consequently, the data obtained by the kinetic methods (either in its simple or extended form) using amino acids as reference bases should be considered with care.

If, to be practiced, the simple kinetic method needs $\Delta_i S^\circ_{298} \sim 0$, only one series of experiments using n_i reference bases B_i at only one effective temperature is necessary to derive $PA_{298}(M)$. By contrast, when $\Delta_i S^\circ_{298}$ is different from zero it is necessary to use n_j experiments realized under different conditions of activation of the adduct ions (and thus corresponding to different effective temperatures T_j in Eq. 7) for each of the n_i reference bases B_i. The statistical treatment of the $[n_i, n_j]$ set of experiments is at the basis of the "extended" kinetic method. A straightforward method for extracting thermochemical information from the extended kinetic method consists in the use of the entire set of experimental observables y_{ij}:

$$y_{ij} = \frac{\Delta S^\circ_i}{R} + \frac{[PA_{298}(M) - PA_{298}(B_i)]}{RT_j} \quad (7)$$

The y_{ij} versus $PA_{298}(B_i)$ points may be fitted by a set of regression lines intersecting in a point of coordinate $x_0 = PA_{298}(M)$ and $y_0 = <\Delta S^\circ_i>/R$, called the "isothermal" or "isoequilibrium" point. A statistical treatment of Equation (7), based on the orthogonal distance regression (ODR) method (Boggs et al., 1992) leading to $PA_{298}(M)$ and $<\Delta S^\circ_i>/R$ has been proposed by Ervin & Armentrout (2004). Figure 2 illustrates kinetic method plots for an ideal case

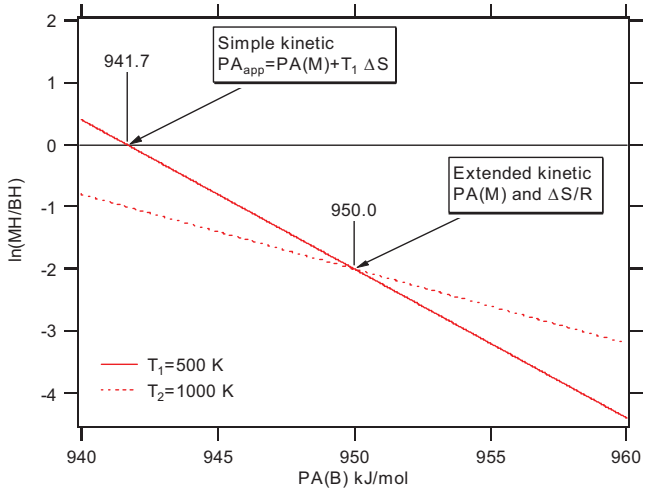


FIGURE 2. Extended kinetic method plot illustrating the difference between PA_{app} and the true PA of a molecule M having a non-negligible $\Delta_p S^\circ_{298}(M)$.

where the molecule of interest M possesses a negative protonation entropy $\Delta_p S^\circ_{298}(M)$ (arbitrarily equal to −2R) while the reference bases B_i exhibit no entropy change upon protonation ($\Delta_p S^\circ_{298}(B_i) = 0$). This means that $<\Delta S^\circ_i>$, defined by the difference $\Delta_p S^\circ_{298}(M) - <\Delta S^\circ_{298}(B_i)>$, is simply equal to $\Delta_p S^\circ_{298}(M)$. Two sets of experiments, corresponding to two different effective temperatures T_1 and T_2 equal to 500 and 1,000 K, respectively, are presented in Figure 2. Following Equation (7), the two fitting lines intercept at $x_0 = PA_{298}(M)$ and $y_0 = \Delta_p S^\circ(M)/R$ (the isothermal point) thus allowing the direct determination of $PA_{298}(M)$ and $\Delta_p S^\circ(M)$. By contrast, the simple kinetic method which assigns the $PA_{298}(M)$ value to the x intercept $PA_{app}(M) = PA_{298}(M) + T.\ \Delta S^\circ_i$ (Eq. 5) lead to serious underestimate. The difference $PA_{298}(M) - PA_{app}(M)$ is equal to 8.3 kJ mol^{-1} for $T_1 = 500$ K and attains 16.6 kJ mol^{-1} for $T_2 = 1{,}000$ K. It is clear that the simple kinetic method should not be used for the determination of the proton affinity of molecule possessing noticeable protonation entropy, even limited to 15 J mol^{-1} K^{-1} as it is the case in this simplified example.

1. Apparent Proton Affinities From the Simple Kinetic Method

Li and Harrison generate MHB_i^+ adducts by chemical ionization of a mixture of M = amino acid and B_i = reference base and studied their collision induced decompositions in the R.f.-only quadrupole of a BEqQ mass spectrometer (Li & Harrison, 1993). The authors used a set of 21 monoamines as reference bases and collision energies (laboratory frame) ranging from 10 to 50 eV. At that time, the proton affinity scale was in the course of a re-evaluation. This led the authors to use two proton affinity scales called the "Lias scale" (Lias, Liebman, & Levin, 1984) and the "Mautner scale" (Mautner & Sieck, 1991). The original data have been reconsidered here to anchor the PA values to the Hunter & Lias proton affinity scale (1998). Accordingly, when considering the 21 reference amines or the 16 examined amino acids, linear correlations are

observed between PA(Hunter & Lias) and PA(Lias, Liebman, & Levin). Applying these relationships to the original data of Li and Harrison (1993), two sets of corrected PA values are obtained and were reported in Table 1.

Using a comparable set of monoamine reference bases, but studying MHB_i^+ adduct ions generated by FAB and monitoring their metastable dissociations occurring in the first field free region of a E-B mass spectrometer by a linked B/E scan, Bojesen and Breindahl obtained also apparent PA values (Bojesen & Breindahl, 1994). In their original article, the authors anchored their PA to the Aue & Bowers scale (1979). We thus correct the PA values by considering the linear correlation observed between the two (Aue & Bowers and the Hunter & Lias) proton affinity scales for the 17 monoamines used as reference bases by Bojensen and Breindahl. The results are also quoted in Table 1.

In a more recent study, Afonso et al. (2000) explored the decompositions of adducts produced in an external electrospray source of an ion trap mass spectrometer. By contrast with the two preceding groups, the authors do not use monoamines as reference bases but five amino acids: serine, leucine, threonine, methionine, and tryptophan. Moreover, in their original article, Afonso et al. used PA(amino acids) values from the ''Lias scale'' (Lias, Liebman, & Levin, 1984) rather than the Hunter & Lias scale (1998). The results presented in Table 1 correct the original Afonso et al. data by using a linear correlation between the PA values they originally used (906.2; 911.7; 914.6; 925.5; 938.5 kJ/mol,) and the PA values evaluated by Hunter & Lias (1998) (914.6; 914.6; 922.5; 935.4; 948.9 kJ/mol) for the five reference amino acids (serine, leucine, threonine, methionine, and tryptophan, respectively). No attempt was made to consider entropy effects or readjustment of the proton affinity of the reference amino acids to more recent determinations.

Figure 3 allows a visual comparison between various estimates of the apparent proton affinities of amino acids coming from the use of the simple kinetic method. Comparison of the three sets of corrected data shows a rough general agreement. Data of Afonso et al. (2000) closely follow that from Bojesen & Breindahl (1994). The mean average deviation between the two scales is less than 1 kJ/mol with a standard deviation of 3.9 kJ/mol. The largest deviations correspond to proline (+8.4 kJ/mol) and histidine (−6.2 kJ/mol). The results of Li & Harrison (1993) are generally lower than those of other experiments, particularly in the high proton affinity region. The situation is particularly acute for Asn, Gln, and Lys since the deviations may attain ∼20 kJ/mol. This point has been underlined by Harrison himself who suggest that, for an unknown reason, a compression of the basicity scale has occurred (Harrison, 1997). We note that the best agreement is obtained by using PA(Li & Harrison) values anchored to the Hunter & Lias scale based on amino acids proton affinities (MAD = 3.8 ± 5.0 kJ/mol) rather than on the set of simple amines (MAD = 10.4 ± 7.7 kJ/mol). Data reported in Figure 3 are PA(Li & Harrison) corrected to the amino acids based Hunter & Lias scale.

A spread of apparent proton affinity values is not unexpected when considering different experimental conditions for systems subjected to large entropy effect. Accordingly, Equation (5) predicts that the apparent proton affinity, $PA_{app}(M)$ should be lower than $PA_{298}(M)$ if $\Delta_i S^\circ$ is negative. Moreover, this deviation is dependent on the experimental conditions, reflected by the ''effective temperature'' *T*. It is worthy to note that the Li & Harrison experiments involve collision induced dissociations associated with a high effective temperature T (∼500–600 K) while Bojesen & Breindahl and Afonso et al. explored low internal energy adduct ions (either metastable ions or species thermalized in a trapping device) characterized by a low effective temperature ($T \sim 300$ K). This may explain the differences observed for Asn, Gln, and Lys where entropy effects are possible.

Finally, the three PA values obtained for glycine, alanine, and proline by the equilibrium method at variable temperature (Mautner, Hunter, & Field, 1979) are also presented in Figure 3. It appears a clear agreement between these values and the apparent proton affinities determined by the simple

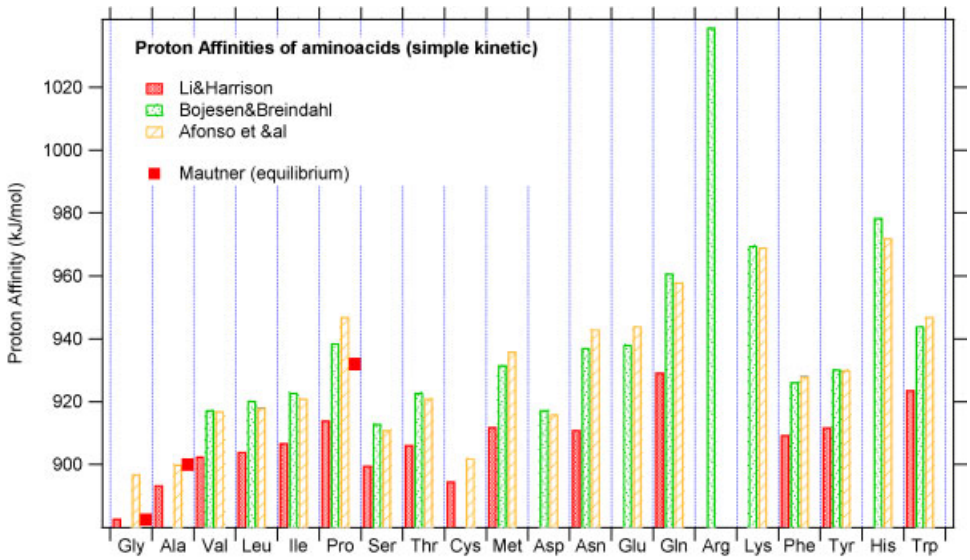


FIGURE 3. Apparent proton affinities of amino acids determined by the simple kinetic method.

kinetic method. This observation is not unexpected since a small protonation entropy value is attached to these three amino acids thus leading to a negligible $T\Delta_i S^\circ$ term and consequently to $PA_{app}(M) \sim PA_{298}(M)$.

2. *Proton Affinities and Protonation Entropies From the Extended Kinetic Method*

As recalled above, the extended kinetic method, which makes use of Equation (7) at variable temperature for several reference bases B_i, allows in principle the determination of the true proton affinity PA_{298} and the protonation entropy $\Delta_p S^\circ(M)$. Systematic errors are however associated with the method when important entropy variation is associated with the protonation process. The extent of these systematic deviations has been estimated by microcanonical rate constant calculations considering model systems (Drahos & Vekey, 2003; Ervin & Armentrout, 2004; Drahos, Peltz, & Vekey, 2004) and was delineated by considering the experimental data obtained for a large set of bi- or tri-dentate bases (Bouchoux, 2006a). The main conclusions of these studies are that the extended kinetic method provide generally an underestimate of the proton affinity PA(M) and of the absolute value of the protonation entropy $\Delta_p S^\circ(M)$ but that, importantly enough, the corresponding gas phase basicity GB(M) is estimated within a few kJ/mol. On the other hand, considerable random errors may be associated with the extended kinetic method if the number of reference bases and the range of effective temperatures are too low.

Despite these limitations, the extended kinetic method is probably the easiest mean to approach the correct protonation thermochemistry of non-volatile polyfunctional molecules such as amino acids and peptides. Indeed, the determination of proton affinities of amino acids by the extended kinetic method has been considered in several recent studies (Wu & Fenselau, 1994; Mirza, Prabhar, & Vairamani, 2001; Kuntz et al., 2002; Hahn & Wesdemiotis, 2003; Schroeder et al., 2004; Bouchoux et al., 2004; Mezzache et al., 2005b; Lioe et al., 2007; Desaphy, Malosse, & Bouchoux, 2008; Bouchoux et al., 2008, 2009a; Bouchoux, Bourcier, & Nacer, 2009b; Bouchoux, Bourcier, & Riffet, 2011). In all, except two of, these studies (Mirza, Prabhar, & Vairamani, 2001; Lioe et al., 2007), the reference bases were monofunctional amines.

The earlier study by Wu & Fenselau (1994) was one of the first attempts to apply the extended kinetic method to polyfunctional molecule. Adduct ions MHB_i^+ produced in a FAB ion source (with M = Lys or His and B_i = dipropylamine, dibutylamine, diethylmethylamine, disec-butylamine, triethylamine) were selected by the E_1B_1 stage of a four sector mass spectrometer; their metastable or collision induced dissociations were analyzed by scanning the E_2B_2 part of the instrument. The original data were anchored to the ''Lias'' proton affinity scale (Lias, Liebman, & Levin, 1984) and are consequently corrected in Table 1 to the Hunter & Lias scale (1998). For this purpose, linear correlations between the two scales have been established for the limited set of five monoamines used by Wu and Fenselau. The resulting corrected values, PA(Lys) = 993.8 kJ/mol and PA(His) = 987.9 kJ/mol, are quoted in Table 1 (note that the precision on the proton affinity determination indicated in the original article was 4–8 kJ/mol).

Mirza, Prabhar, & Vairamani (2001) used the extended kinetic method to derive the proton affinities of tryptophan and proline. Their experiments were done in a tandem triple quadrupole mass spectrometer equipped with an ESI source. The reference bases used by the authors (Glu, Gln, Asp, and Met) were all expected to have noticeable protonation entropy but no specific value has been introduced in the extended kinetic treatment. Moreover, the proton affinities values of these reference bases were anchored to the data obtained by Bojesen and Breindahl (1994), that is, the apparent proton affinities discussed above. The values, PA(Trp) = 834.6 and PA(Pro) = 926.2 kJ/mol indicated in Table 1, result from a simple correction based on the linear correlation existing between the Bojesen and Breindahl and Hunter & Lias scales.

Proton bound dimers of Ala and either monoamines (propargylamine, methylamine, allylamine, benzylamine) or amino acids (gly, cys, asp, val) were formed by FAB ionization and studied by Hahn & Wesdemiotis (2003) in an EBhQ tandem mass spectrometer. No significant protonation entropy (less than 1 J/mol K) and identical proton affinities were determined by the extended kinetic method in both of the two series of experiments.

In addition to the study of Mirza, Prabhar, & Vairamani (2001) discussed above, proline has been the subject of two other investigations using the extended kinetic method (Kuntz et al., 2002; Mezzache et al., 2005b). Both used an ion-trap device equipped with an external ESI source which is known to provide a limited range of collision energies thus rendering delicate the practice of the extended kinetic method. Mezzache et al. thus complete their data by performing experiments in a triple quadrupole mass spectrometer. The proton affinities determined by both groups were in correct agreement if we take into consideration the relevant experimental uncertainties, they are however well above the Mirza et al. result (Mirza, Prabhar, & Vairamani, 2001).

Another example of the extended kinetic results obtained by performing experiments in an ion-trap is provided by the determination of the proton affinity of Lys by Schroeder et al. (2004). Proton bound dimers of lysine and one of a series of monoamine bases were formed by electrospray ionization. The isolated ions were allowed to undergo collision induced dissociations with the background helium atom at activation amplitude of 15%, 50%, and 85%. These experiments correspond to a very limited range of excitation energies as revealed by the proximity of the corresponding three effective temperatures of 356, 370, and 384 K. Moreover, due to the occurrence of significant protonation entropy, the isothermal point cannot be located in the explored proton affinity domain (Fig. 4). In such circumstances, the extended kinetic method shows generally large uncertainties on the deduced proton affinity and protonation entropy (see below a comparable situation for Gln). This is clearly illustrated by the standard deviations reported in Table 1 for the Schroeder et al. data revaluated here using the ODR method.

Protonation thermochemistry of Gly, Asp, Asn, His, Lys, Glu, and Gln has been explored using a triple quadrupole mass spectrometer equipped with an ESI source (Bouchoux et al., 2004). The center of mass collision energies of the MHB^+ adducts were situated between zero and 4 eV and a large set of monofunctional molecules was used as reference bases. The statistical treatment of the data used the ODR

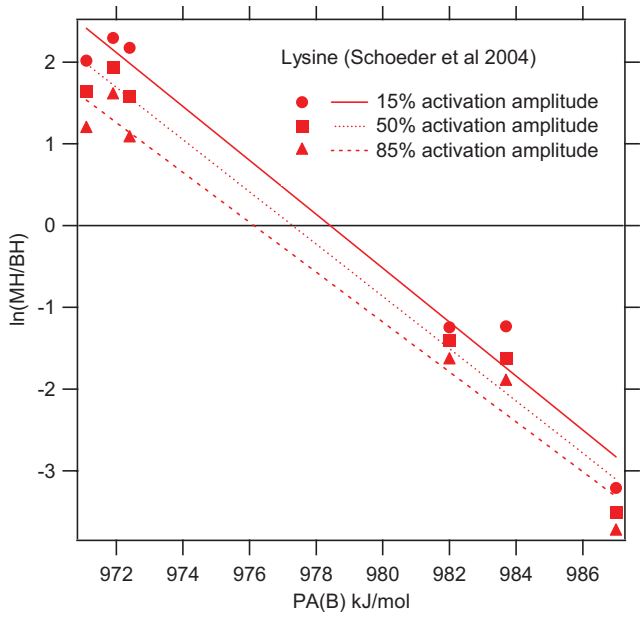


FIGURE 4. Extended kinetic plot for protonation of lysine.

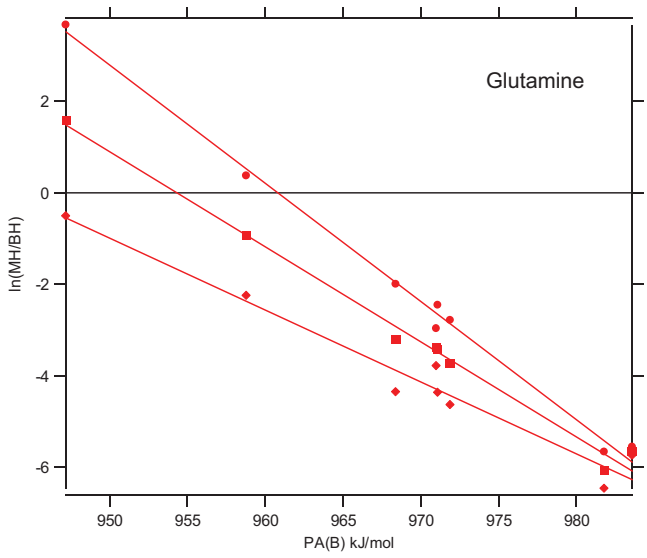


FIGURE 6. Extended kinetic plot for protonation of glutamine.

method and shows good accuracy for Gly, Asp, His, and Lys but larger uncertainties for Asn, Glu, and Gln. Two examples of the ln(MH/BH) versus PA(B) data are presented in Figures 5 and 6. Practically, difficulties to obtain accurate MH/BH ratio above $\sim 10^{2}$ or below $\sim 10^{-2}$ limit the explored proton affinity range in the use of the extended kinetic method. In the case of Gly, which clearly demonstrates a negligible protonation entropy, the isothermal point falls close to the *x*-axis and is situated inside the PA(B) range (Fig. 5). It can be consequently accurately located by interpolation using a statistical method such as the ODR method. By contrast, when a negative protonation entropy is associated with the molecule M, the isothermal point may be situated outside the PA(B) range, as it is for example observed for Gln (Fig. 6). In such situation, extrapolation of the isothermal point coordinates may be associated with important errors. This limitation was particularly dramatic for Glu in the original study (Bouchoux et al., 2004). Recently, a re-examination of this amino acid using a more sensitive ionization source (nano-ESI Waters) and tandem mass spectrometry device (Q-TOF Waters Premier) allows a more confident determination of PA(Glu) and $\Delta_p S^{\circ}$(Glu) (Bouchoux, Bourcier, & Nacer, 2009b) (Table 1).

Comparable experiments involving formation of protonated amino acid/amine adducts by electrospray ionization were undertaken on Arg (Bouchoux et al., 2008), Phe and Tyr (Bouchoux et al., 2009a), and Trp (Bouchoux & Bourcier & Riffet, 2011). With Arginine, the most basic of the 20 naturally occurring amino acids, the extended kinetic method was successful with imino reference bases: 1,1,3,3-tetramethylguanidine (TMG), 1,5-diazabicyclo[4.3.0]non-5-ene (DBN), and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU). Collision induced dissociations of the MHB^{+} adducts were studied in a Q-TOF instrument by floating the potential of a linear ion trap located between the two mass analyzers. Results obtained at three different center of mass collision energies are illustrated by Figure 7. It may be noted that in a previous study, Wu & Fenselau (1992) were able to measure the $[MH]^{+}$ and $[BH]^{+}$

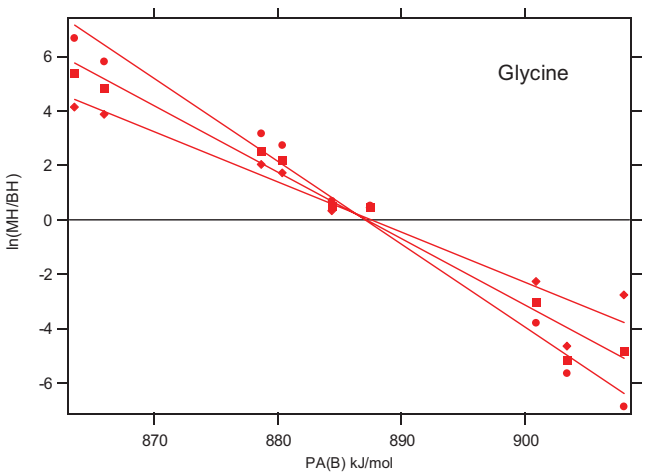


FIGURE 5. Extended kinetic plot for protonation of glycine.

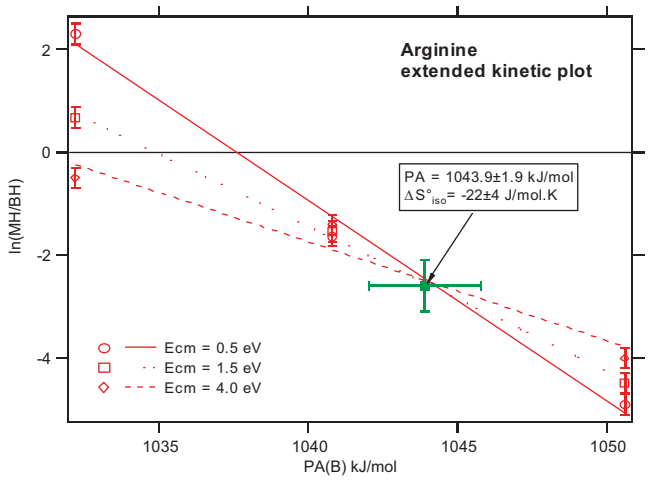


FIGURE 7. Extended kinetic plot for protonation of Arginine (E_{cm}: center of mass collision energy; ΔS°_{iso}: isothermal protonation entropy).

peak intensities (M = Arg and B = TMG, DBN, and DBU) from the dissociations of metastable $[MHB]^+$ ions produced in a FAB ion source and traveling a EB-EB sector mass spectrometer. Their data were obtained without activation of the precursor ions and are thus characterized by only one effective temperature T but are nicely comparable to that obtained in the ESI-Q-TOF experiments. When considering altogether these data, the ODR method leads to results comparable to what is obtained from the ESI-Q-TOF experiments alone (Table 1).

Investigation of the basicities of aromatic amino acids: Phe, Tyr, and Trp were done in a triple quadrupole mass spectrometer equipped with an ESI source (Bouchoux et al., 2009a; Bouchoux, Bourcier, & Riffet, 2011). In the three cases, noticeable protonation entropies have been detected by the extended kinetic method. In parallel, and as expected, the proton affinities derived from the isothermal point is slightly higher (by 5–10 kJ/mol) than that obtained previously by the simple kinetic method (Table 1).

Lioe et al. (2007) presented recently a tentative of determination of the proton affinity of Met by the simple and the extended kinetic methods. Since they used amino acids which may present significant protonation entropies (Asn, Glu, Tyr, Phe) as reference base, the authors find difficulties to correctly anchor their data to a suitable basicity scale. This leads to the estimated bracketing of 925–938 kJ/mol indicated in Table 1 for the corresponding proton affinity. At the same time, Desaphy, Malosse, and Bouchoux (2008) derive a PA(Met) value of 937.5 kJ/mol with a standard deviation of 2.9 kJ/mol by using the extended kinetic method with monoamine as reference bases (Table 1).

Averaged proton affinities values determined by the simple and the extended kinetic methods are compared in Figure 8. Data obtained by Mautner, Hunter, and Field (1979) from equilibrium constant determinations at variable temperature are also reported. It is evident from examination of Figure 8 that, for a number of cases, the proton affinity provided by the extended kinetic method is significantly larger than the value given by the simple kinetic method. Significant deviations are evidenced for Asn (27.2 kJ/mol), Gln (31.2 kJ/mol), Lys (28.1 kJ/mol), His (17.7 kJ/mol), Asp (11.7 kJ/mol), Glu (7.9 kJ/mol), and for Tyr, Phe, and Met (∼7 kJ/mol). Such differences in PA are expected if the protonation entropies of the considered amino acids are negative. According to Equation (5), it results indeed that the apparent value given by the simple kinetic method (PA_{app}) is equal to the true $PA_{298}(M)$ value plus the product $T\Delta_i\ S^\circ_{298}$ (where $\Delta_i\ S^\circ_{298} = \Delta_p\ S^\circ(M) - \Delta_p\ S^\circ(B_i)$, Eq. 6). If $\Delta_p S^\circ(M)$ is negative and $\Delta_p\ S^\circ(B_i)$ close to zero, the inequality $PA_{app} < PA_{298}(M)$ holds.

Protonation entropy $\Delta_p\ S^\circ$ plays a crucial part in the protonation thermochemistry of amino acids, experimental values are indicated in Table 1 when available. In their compilation, Hunter & Lias (1998) generally assumed that $\Delta_p\ S^\circ = -5$ J mol^{-1} K^{-1}, for α-amino acids not bearing a basic site in their side chain. In fact this value has been assigned by comparison with methylamine for which $\Delta_p\ S^\circ = -7$ J mol^{-1} K^{-1}. The lowering in entropy observed for the protonation of methylamine is indeed reproduced by the calculation (East, Smith, & Radom, 1997). The expectation of an entropy loss during protonation of simple α-amino acids should not be surprising if one consider that the protonated forms may enjoy an intramolecular hydrogen bond between the amino group and the oxygen of the carbonyl (see Section III). However, the three $\Delta_p\ S^\circ$ values experimentally obtained for glycine, alanine, and proline, from a van't Hoff plot using the protonation equilibrium constants at variable temperature (Mautner, Hunter, & Field, 1979), are positive, though relatively small (Table 1). By contrast, for α-amino acids bearing a basic site in their side chain, a significant entropy loss is expected upon protonation. This may particularly concern Asp, Asn, Glu, Gln, Arg, Lys and, to a lesser extend, Ser, Thr, Cys, Met and the aromatic amino acids: Phe, Tyr, His, and Trp. Accordingly, using the extended kinetic method, $\Delta_p\ S^\circ$ values ranging from −30 to −80 J mol^{-1} K^{-1} have been determined for Arg, Asp, Asn, Glu, Gln, and Lys.

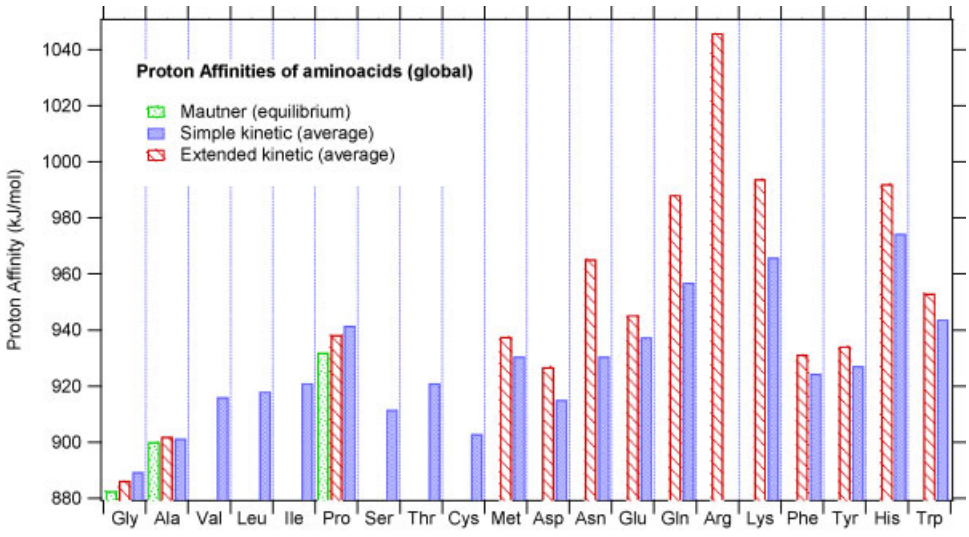


FIGURE 8. Comparison of proton affinities of amino acids determined by the simple and extended kinetic methods.

These values compare reasonably well with that discussed for 1,3-, 1,4-, and 1,5-diamines (−30/−50; −45/−60; −50/−80 J mol^{-1} K^{-1}) (Part II of this review, under press). Less important, but noticeable, negative protonation entropies (in the range −10/−20 J mol^{-1} K^{-1}) have been detected for Met, Phe, Tyr, His, and Trp.

The applicability of the extended kinetic method to the protonation thermochemistry of molecules having large protonation entropies is more evident when gas phase basicities are considered. This point has been theoretically predicted by microcanonical rate constant modeling (Drahos & Vekey, 2003; Drahos, Peltz, & Vekey, 2004; Ervin & Armentrout, 2004) and demonstrated from consideration of a large set of experimental data (Bouchoux, 2006a). Accordingly, re-examination of a series of 36 experiments shows that the gas phase basicity values $GB_{298}(M)$ obtained by combining the proton affinity and the protonation entropy given by the extended kinetic method are generally in agreement, within less than 3 kJ/mol, with gas-phase basicities obtained by the equilibrium method. The gas phase basicity values obtained by the equilibrium, thermokinetic and extended kinetic methods (Table 1) are graphically compared in Figure 9. The agreement between the various set of data is clearly satisfactory for most of the amino acids considered. First of all, for the amino acids where a comparison is possible, that is, Pro, Lys, and His, the extended kinetic method results obtained from different laboratories agree with each other to less than 2.5 kJ/mol if we consider only experiments obtained with monoamines as reference bases. Further, the gas phase basicities obtained by the extended kinetic method compare reasonably well with those obtained by the equilibrium method. An average difference $GB_{equ} - GB_{ext.kin.}$ close to zero, with a standard deviation of 6.1 kJ/mol, is observed when considering glycine, alanine, proline, methionine, aspartic acid, lysine, phenylalanine, tyrosine, tryptophan, and histidine. Severe discrepancies are however evident for Gln (−51.1 kJ/mol), Asn (−25.4 kJ/mol), Glu (−25.6 kJ/mol), and for Phe, Tyr, and Trp. As discussed above, the extended kinetic method may provide erroneous results if the isothermal point falls well apart from the explored range of proton affinities. This difficulty may be suspected for Asn and Gln (see Fig. 6) but probably not for the new experiments on Glu (Bouchoux, Bourcier & Nacer, 2009b), Phe, Tyr (Bouchoux et al., 2009a), and Trp (Bouchoux, Bourcier, & Riffet, 2011). We also note that the Locke gas phase basicity values obtained for Glu and Gln were considered too low by Harrison (1997) who suggests that thermal decomposition of the amino acids may have altered the estimate of its neutral concentration. Similar difficulties may be expected for Phe, Tyr, and Trp.

C. General Evaluation of the Experimental Data

All the available experimental data, re-evaluated as described in the preceding paragraphs, are gathered in Table 1. When available the error associated with each measurement is indicated. Averaged values of the proton affinities and gas-phase basicities are calculated and indicated in bold in Table 1. Beside these values are presented their associated standard deviation (which should not be confused with the experimental error but provide an idea of the confidence limits). Values suspected to be erroneous are indicated in parentheses and are not included in the averaging procedure.

Gas-phase basicities of Ile, Thr, and Cys have been the subject of only one experimental determination (by the equilibrium method), the uncertainties on these values are unknown. As discussed above, for a number of amino acids, the range of experimentally determined GB values is important: 57 kJ/mol (Gln), 25 kJ/mol (Asn, Glu), ∼15 kJ/mol (Phe, Tyr, Trp), and ∼5 kJ/mol (Leu, Pro). For one reason or one other, some equilibrium or extended kinetic experiments should be suspected to be erroneous (values indicated in parentheses in Table 1).

Concerning proton affinities, a larger set of experiments is available, thus leading to a more meaningful statistical treatment. Averaged PA have been calculated by considering all the data, whatever the methods, for amino acids showing negligible protonation entropy, that is: Gly, Ala, Val, Leu, Ile, Pro, Ser, Thr, and Cys. For amino acids suspected to present

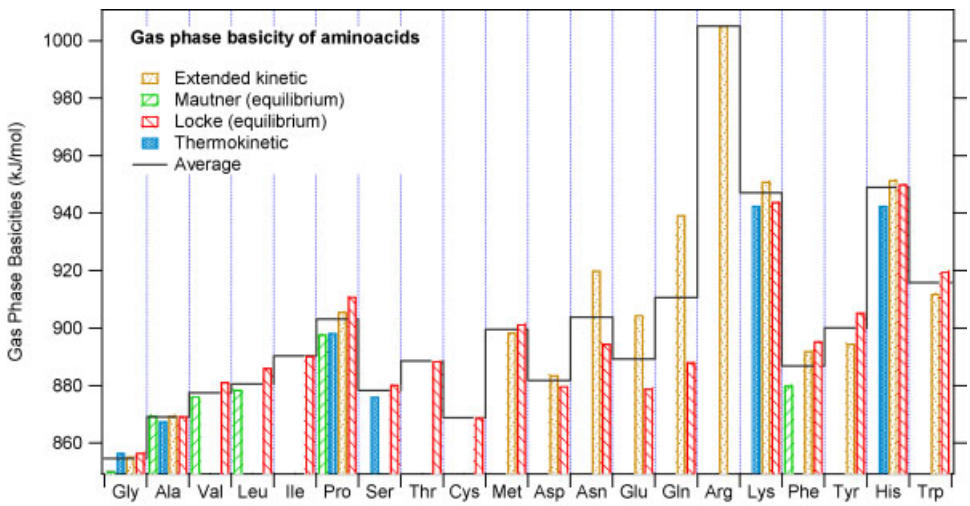


FIGURE 9. Summary of gas-phase basicities of amino acids determined by various methods.

negative protonation entropies, averaged PA_{app} values obtained by the simple kinetic method are given in Table 1, these figures should provide the lower limit for the true proton affinities. Proton affinities obtained by the extended kinetic method would approach at best the true PA, however, for Asn and Gln these values are clearly too high and are thus indicated in parentheses in Table 1.

III. THEORETICAL PROTONATION THERMOCHEMISTRY OF AMINO ACIDS

Beside the experimental determinations presented above, there have been numerous information of the proton affinities and gas phase basicities of amino acids obtained using molecular orbital calculations. The energetic and structural findings revealed by these studies will be presented in the present section. The peculiarities attached to the theoretical estimate of thermochemical quantities associated with reaction (1) will be first briefly recalled (for a more extended discussion see Bouchoux, 2007).

Since proton affinity is defined as the enthalpy of reaction (1), this quantity, at a given temperature T, is given by:

$$\begin{aligned} PA_T(M) &= \Delta_1 H_T^\circ = H_T^\circ(M) - H_T^\circ(MH^+) + H_T^\circ(H^+) \\ &= H_T^\circ(M) - H_T^\circ(MH^+) + 5/2RT \end{aligned} \quad (8)$$

This molar quantity may be evaluated if the compositions of each mol of neutral, M, and protonated forms, MH^+, are well defined. Computational approaches usually assume that each mol of species M and MH^+ contains only one type of conformer, generally the most stable. Under these circumstances, Equation (8) may be directly used to derive a ''most stable conformers'' proton affinity, at 298 K:

$$\begin{aligned} PA_{msc}(M) &= H_{298}{}^\circ(M)_{moststable} - H_{298}{}^\circ(MH^+)_{moststable} \\ &\quad + 6.2\,\text{kJ mol}^{-1} \end{aligned} \quad (9)$$

However, if a comparison between theory and experiment is to be done, this ''most stable conformers'' method may be questioned. Indeed, for polyfunctional and flexible molecules, a large number of conformers of neutral and protonated species may be populated in the experimental condition. A distribution of conformers should be consequently considered for M and MH^+. One possibility is to consider that M and MH^+ are in thermal equilibrium at a (possibly ''effective'') temperature T. Assuming a Boltzmann distribution, the molar fractions x_i of each conformer are then calculated using Equation (10):

$$x_i = \exp\left(\frac{-G_i}{RT}\right) \sum_1^N \exp\left(\frac{-G_i}{RT}\right) \quad (10)$$

where G_i are the individual Gibbs free energies. A molar (average) proton affinity may be defined by considering the population of neutral (x_i) and protonated (x'_i) conformers at the temperature T:

$$PA_{average} = \sum_1^N x_i H_T^\circ(M)_i - \sum_1^{N'} x'_i H_T^\circ(MH^+)_i + 5/2RT \quad (11)$$

When looking at the protonation entropy, $\Delta_p S^\circ(M)$, and consequently at the gas phase basicity, GB(M), another important point to consider is the entropy of mixing. Rigorously, the total entropy of one mole of a mixture of N distinguishable conformers with molar fractions x_i is indeed given by:

$$S^\circ_{T\,average} = \sum_1^N x_i (S_T^\circ)_i - R \sum_1^N x_i \ln x_i \quad (12)$$

where the second term is called the ''entropy of mixing.'' Thus, computation of a molar protonation entropy, $\Delta_p S_T^\circ(M)_{average}$, should make use of Equation (12). Further, the corresponding gas phase basicity $GB_{298}(M)_{average}$, may be obtained using Equation (3).

The extraordinary development of quantum chemistry calculations during the last 10 years have brought computational methods able to predict enthalpy variations at the ''benchmark'' and ''chemical'' accuracy levels (i.e., with deviation less than 2–4 kJ mol^{-1}). Presently, techniques which may approach this level of accuracy are essentially composite methods such as Gn (Curtiss & Raghavachari, 2002; Curtiss, Redfern, & Raghavachari, 2007a, b), CBSn (Montgomery et al., 1999), and Wn (Martin & de Oliveira, 1999). Complete sets of theoretical proton affinity values of α-amino acids have been however obtained at lower correlated levels such as MP2/6-311+G(d,p) (Maksic & Kovacevic, 1999a) or B3LYP/6-31+G(d,p) (Bleiholder, Suhai, & Paizs, 2006; Dinadayalane, Sastry, & Leszczynski, 2006). In their recent studies, Bleiholder et al. (2006) and Gronert, Simpson, and Conner (2009) present results obtained by means of the G2MP2 and G3MP2 composite methods. In addition, a number of more specific computational studies devoted to selected amino acids appeared during the last decade. A critical examination of the theoretical data however reveals, in several cases, considerable spread of values. The origins of the discrepancies are multiple as revealed by a careful reconsideration of the panel of published data. Theoretical thermochemistry is obviously dependent on the level of theory used (see above). Moreover, care should be taken to compare theoretical data where the zero point energies and thermal correction for enthalpies at finite temperature (298 K) have been conveniently considered. Most importantly, the significance of the calculations depends on the effectiveness to identify the most stable conformers of neutral and ionized structures. This important and delicate point will be detailed in the following part before examining the published theoretical protonation thermochemistry. Finally, a comparison between theory and experiment will be done to propose evaluated protonation thermochemistry of α-amino acids.

A. Structure of the Neutral and Protonated Forms of Amino Acids

It is known for a long time that, in the crystalline state and in aqueous solution, amino acids exist mainly in their zwitterionic forms: $H_3N^+CHRCO_2^-$. By contrast, in the gas phase, isolated amino acids appear to be more stable in their neutral forms: $H_2NCHRCO_2H$. For example isolated glycine, which has received much attention from this point of view, is more stable than the isolated glycine zwitterion by ca. 70 kJ/mol.

Since glycine is also the less basic amino acid it is also the worst candidate for its existence as a zwitterion in the gas phase. At the other side of the gas phase basicity scale, arginine is expected to offer better chance to form a stable zwitterion. However, present knowledge based on quantum chemical computation seem to indicate that arginine is also more stable in its neutral form, although its most stable zwitterionic structure is situated only few kJ/mol above. It may be emphasized that zwitterionic amino acid structures may be stabilized by addition of a few number of molecules of solvent (e.g., water) or charged species (metal and organic ions). This aspect of the gas phase amino acids chemistry is out of the scope of the present review but the reader may have an idea of the corresponding literature in the references given in recent articles (Jensen & Gordon, 1995; Wang et al., 2002 and 2003a; Aikens et al., 2006; Blom et al., 2007; McLain et al., 2007; Wu & McMahon, 2008a; Dunbar et al., 2009; Mullin & Gordon, 2009; Armentrout et al., 2010; O'Brien et al., 2010).

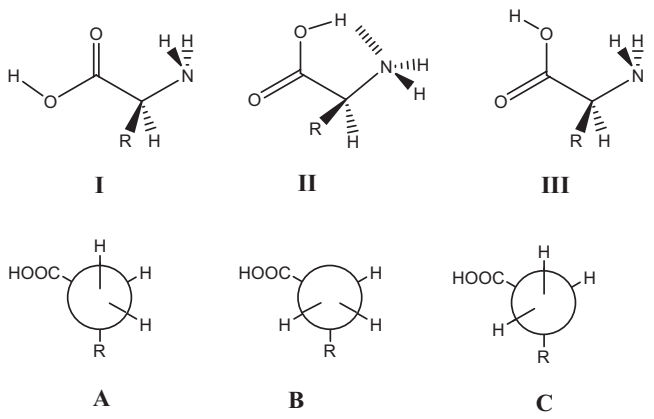
An important structural characteristic of gas phase, neutral and protonated, α-amino acids is the possibility of stabilization by intramolecular hydrogen bonding. Favorable interaction may first occur inside the acidic group, between the acidic hydrogen and the carbonyl oxygen, if the COOH arrangement is *syn*. The second source of intramolecular stabilization is offered by interactions between the amino group and the acidic moiety. A general nomenclature will be used thereafter to classify the studied α-amino acids according to the atoms arrangement of the $CH(NH_2)COOH$ part (conformations **I–III**, Scheme 7). Conformers **I** and **III** present a *syn* HOCO arrangement and possibilities of hydrogen bonds between the hydrogen atoms of the amino group and either the carbonyl oxygen atom (conformers **I**) or the hydroxylic oxygen (conformers **III**). Although *anti* HOCO is intrinsiquely less stable than *syn* HOCO by ca. 20 kJ mol^{-1}, the former arrangement may be present in low lying conformers of amino acids if its allow the formation of a hydrogen bond between the acidic hydrogen and a neighboring basic site. It is the case for conformers of type **II** which are characterized by the formation of a strong OH ... NH_2 hydrogen bond inside the amino acid framework. A second possibility is the formation of a hydrogen bond between the acidic hydrogen and a basic function of the side chain, as observed for example for lysine or arginine. This situation will be denoted **I′** thereafter (Scheme 7). Each type of conformers, **I**, **I′**, **II**, or **III**, may have C–N rotamers and, depending upon the C(1)C(2)NH(H) dihedral angle, conformations **A**, **B**, and **C** may be distinguished as depicted in Scheme 7. Finally, a systematic nomenclature using *gauche* (g+ or g−) and *antiperiplanar* (a) arrangements has been adopted through this review to characterize the side chains R conformations.

As will be seen in detail in this section, the thermochemically favored site of protonation of α-amino acids is generally the amino group. The exceptions are arginine, lysine and histidine where the side chain bears a more basic site. Among the possible conformers resulting from the N-protonation of amino acids, the most stable structure corresponds generally to conformation **HI** where a strong hydrogen bond is established between one H of the protonated amino group and the carbonyl oxygen (Scheme 8). The second possible conformation, **HII** (Scheme 8), involves the OH group as hydrogen bond acceptor. Due to the lower basicity of the OH moiety of a carboxylic acid with respect to the carbonyl oxygen, conformers **HII** are less stable and situated ca. 10–20 kJ mol^{-1} above conformers **HI**.

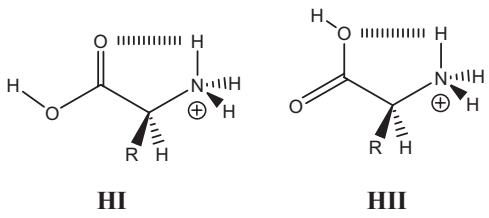
1. Alkyl Substituted Amino Acids

a. Glycine

As the simplest amino acids, neutral glycine has been extensively studied, both experimentally and theoretically, from the point of view of its structure and its conformations. Conformational analysis of neutral glycine by quantum chemistry calculations reveals the existence of four minima in the potential energy surface clearly more stable than the remaining conformers by more than 10 kJ/mol (Csaszar, 1992, 1995; Barone, Adamo, & Lelj, 1995; Zhang & Chung-Phillips, 1998b, 1999; Balta et al., 2000; Noguera et al., 2001; Pacios, Galvez, & Gomez, 2001; Wang et al., 2003b; Miller & Clary, 2004; Falzon & Wang, 2005; Bouchoux & Chia, 2009; Yang et al., 2009). Conformers **GlyI$_C$**, **GlyIII**, and **GlyI$_A$** (Fig. 10) are characterized by a *syn* HOCO arrangement, bearing its own favorable electrostatic interaction, and by additional NH_2 ... OC or NH_2 ... OH hydrogen bonding type interaction. By contrast, conformer, **GlyII** (Fig. 10) presents an *anti* HOCO arrangement allowing the formation of a strong OH ... NH_2 hydrogen bond. The particular strength of this latter hydrogen bond may be underlined since it has been even claimed that only conformer of type **II** exhibits a true internal hydrogen bond. This view is indeed supported theoretically by an important electron density, characteristic of a bonding interaction, between the H and N atoms of the OH ... NH_2 system



SCHEME 7



SCHEME 8

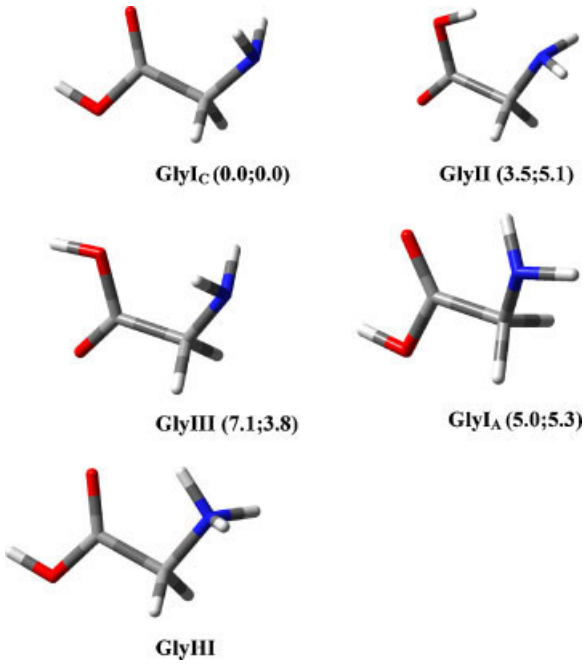


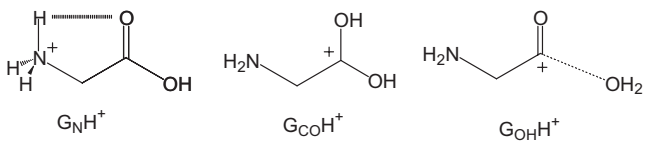
FIGURE 10. Most stable conformers of neutral and protonated glycine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G3 results).

(Pacios, Galvez, & Gomez, 2001; Wang et al., 2003b). As a first consequence, the destabilizing effect of the *anti* HOCO arrangement, of more than 20 kJ mol^{-1} with respect to *syn* HOCO, is almost fully counterbalanced by the OH ... NH_2 hydrogen bond energy. Accordingly, **GlyII** is predicted at the G3 level to be situated only 3.5 kJ mol^{-1} above **GlyI$_C$** in the enthalpy scale (Fig. 10). It may be also indicated that the barrier for the *syn* → *anti* isomerization is close to 60 kJ mol^{-1}. It is clearly higher than the rotational barriers along the C–C and C–N bonds (∼10–20 kJ mol^{-1}). The second consequence of the very efficient internal hydrogen bonding occurring in **GlyII**, is a noticeable increase of the C(1)–C(2) and C(2)–N rotational barriers and therefore a decrease of the corresponding entropy contributions. It has been already noted that a lower vibrational entropy is attached to conformer **GlyII** with respect to conformer **GlyI$_C$**, **GlyIII**, and **GlyI$_A$** (Miller & Clary, 2004). The obvious result of the lower entropy of conformer **GlyII** is the parallel increase of its free energy and its reduced participation to the conformer population, particularly at high temperature. Consideration of the conformational equilibrium of neutral glycine has been done from Gibbs free energies calculated at the MP4/6-31+G(2d,2p)//MP2/6-31+G(d) and G3MP2B3 levels (Zhang & Chung-Phillips, 1998b; Bouchoux, Huang, & Inda, 2011). In the hypothesis of a Boltzmann distribution at 298 K, the resulting population is dominated by **GlyI$_C$** which represent 70–80% of the mixture of conformers.

A number of experiments were designed to identify the conformers of neutral glycine present in the gas phase using microwave spectroscopy (Suenram & Lovas, 1978, 1980; Godfrey & Brown, 1995; Godfrey, Brown, & Rodgers, 1996), electron diffraction (Iijima, Tanaka, & Onuma, 1991), jet-cooled Raman spectroscopy (Balabin, 2010), gas phase IR spectroscopy (Linder et al., 2008), IR spectroscopy in low-temperature argon matrices (Stepanian et al., 1998a; Ivanov, Sheina, & Blagoi, 1999; Espinoza et al., 2010), core binding energies determined by photoelectron spectroscopy (Slaughter & Banna, 1988; Plekan et al., 2007) and gas-phase basicity measurements (Locke, Hunter, & McIver, 1979; Locke & McIver, 1983). The central conclusion of these studies is that gaseous glycine mainly exists in conformation **GlyI$_C$** (to ca. 75%) in the 200–300 K temperature range and that the remaining contribution corresponds to a mixture of conformers **GlyII** and **GlyIII**. The fourth structure **GlyI$_A$** has not been yet firmly identified, its occurrence has been suggested only very recently (Balabin, 2010). These finding are consequently in correct agreement with theory in predicting **GlyI$_C$** as the major conformer. The difficulties to detect conformer **GlyI$_A$**, has been suggested earlier (Godfrey, Brown, & Rodgers, 1996; Miller & Clary, 2004) to result from its rapid collapse to conformer **GlyI$_C$** by C–N rotation.

Glycine presents three evident sites for protonation, the nitrogen and the two oxygen atoms, with possible intramolecular hydrogen bond in the various protonated forms. A number of theoretical calculations has been devoted to the isomeric structures and conformations of protonated glycine (Bouchonnet & Hoppilliard, 1992; Wu & Lebrilla, 1993; Zhang et al., 1993; Uggerud, 1997; Topol et al., 1998; Zhang & Chung-Phillips, 1998a,b, 1999; Balta et al., 2000; O'Hair et al., 2000; Rogalewicz & Hoppilliard, 2000; Noguera et al., 2001; Bouchoux et al., 2004; Bouchoux & Chia, 2009). All the authors conclude that the most basic site is the amino nitrogen, followed by the carbonyl oxygen and, further, by the hydroxyl oxygen. Protonation of the oxygen atoms of the carbonyl and the hydroxyl functions lead to structures, **Gly$_{CO}$H$^+$** and **Gly$_{OH}$H$^+$** which are situated ∼110–130 and ∼160 kJ/mol above **Gly$_N$H$^+$**, respectively (Scheme 9).

Three conformers of the N-protonated glycine were located as minima on the potential energy surface. The most stable form of the N-protonated glycine is conformer **GlyHI** where a single intramolecular hydrogen bond is created between one N–H hydrogen and the oxygen of the carbonyl group (Fig. 10). The second conformer, **GlyHII** (see Scheme 8 for the nomenclature used) is less stable by ∼20 kJ mol^{-1} since the amino hydrogens are under the effect of the less basic hydroxyl oxygen. As observed with neutral glycine, the *syn* HOCO arrangement brings its part to the stability of the corresponding protonated forms thus favoring conformations **GlyHI** and **GlyHII**. It is noteworthy that the energy difference between the *syn* and *anti* conformers appears to be amplified in the protonated species as indicated by the high relative enthalpy of the *anti* form of conformer **GlyHI** (∼35 kJ mol^{-1}). The free energy difference between **GlyHI** and **GlyHII** (17.1 kJ mol^{-1}, G3 result) lead to the presence of 99.9% of **GlyHI** in the conformers population at 298 K, the



SCHEME 9

participation of **GlyHII** may be consequently discarded in most reported experiments.

Protonation of glycine at the carbonyl oxygen gives rise to $\mathbf{Gly_{CO}H^+}$ structures situated 110–130 kJ/mol above $\mathbf{Gly_NH^+}$. The most stable $\mathbf{Gly_{CO}H^+}$ conformer is characterized by a *syn–anti* HOCOH arrangement, its *syn–syn* homologue is situated 25 kJ/mol above while the *anti–anti* species collapses to a N-protonated conformer by a 1,4-hydrogen migration (Balta et al., 2000). Of interest is the fact that the hydroxyl protonated structure $\mathbf{Gly_{OH}H^+}$ exhibits a quite elongated C ... OH_2 bond (2.15 Å at the B3LYP/6-31G(d) level) thus leading to the conclusion that it may be more properly seen as an ion-neutral complex rather than a covalent structure (Bouchonnet & Hoppilliard, 1992; Uggerud, 1997; O'Hair et al., 2000; Rogalewicz & Hoppilliard, 2000). The calculated energy of this structure is in the range 120–180 kJ/mol thus well above **GlyHI**.

The situation detailed here for the glycine molecule may be applied to other α-amino acids provided the α-side chain R does not significantly interact with the $CH(NH_2)COOH$ moiety during the protonation process. In other word, the most stable conformers of the neutral molecules in a ∼10 kJ/mol range should pertain to the series **I–III** (Scheme 7) while the protonated form should be only in the **HI** conformation (Scheme 8). This is obviously expected for the alkyl-substituted α-amino acids Ala, Val, Leu, Ile (Scheme 1) and indeed observed.

b. Alanine

In 1993, Godfrey et al. examined the rotational spectrum of gaseous alanine and interpreted their observation by the existence of two conformers, one of type **I** and one of type **II**. This assignment was supported by HF/6-31G(d,p) calculations on six conformers (Godfrey et al., 1993; Godfrey, Brown, & Rodgers, 1996). During the last decade, new experiments confirmed this original conclusion. These new evidences came from matrix infrared spectroscopy (Stepanian et al., 1998b), electron diffraction (Iijima & Nakano, 1999), molecular beam Fourier transform microwave spectroscopy (Blanco et al., 2004) and X-ray photoemission and photoabsorption (Feyer et al., 2008). At the same time, a number of theoretical studies have been devoted to isolated neutral alanine (Godfrey et al., 1993; Csaszar, 1995, 1996; Godfrey, Brown, & Rodgers, 1996; Iijima & Nakano, 1998; Stepanian et al., 1998b; Topol et al., 1998; Pecul et al., 2004; Abirami et al., 2005; Kapitan et al., 2006; Degtyarenko et al., 2007; Yang et al., 2009; Jaeger et al., 2010; Bouchoux, Huang, & Inda, 2011). The finding of these theoretical studies is in line with the experimental observations. Six conformers are situated in a 6 kJ mol^{-1} enthalpy range (Fig. 11). By order of increasing enthalpy we find $\mathbf{AlaI_C}$, $\mathbf{AlaII_A}$, $\mathbf{AlaII_B}$, $\mathbf{AlaI_B}$, and $\mathbf{AlaI_A}$, followed by conformer **AlaIII**. This order is slightly modified when looking at the free energy differences placing **AlaIII** below $\mathbf{AlaI_B}$, $\mathbf{AlaI_A}$, and $\mathbf{AlaII_B}$. The population of conformers deduced from the G3MP2B3 free energies is 63/10/10/8/6/6% for $\mathbf{AlaI/AlaII_A/AlaIII/AlaI_B/AlaI_A/AlaII_B}$ thus explaining the experimental observation of a major contribution of a conformer of type I, $\mathbf{AlaI_C}$, beside a minor participation attributed to a conformer of type II, $\mathbf{AlaII_A}$ (Bouchoux, Huang, & Inda, 2011). The lack of experimental detection of the other conformers, despite non-negligible molar fraction at 298 K,

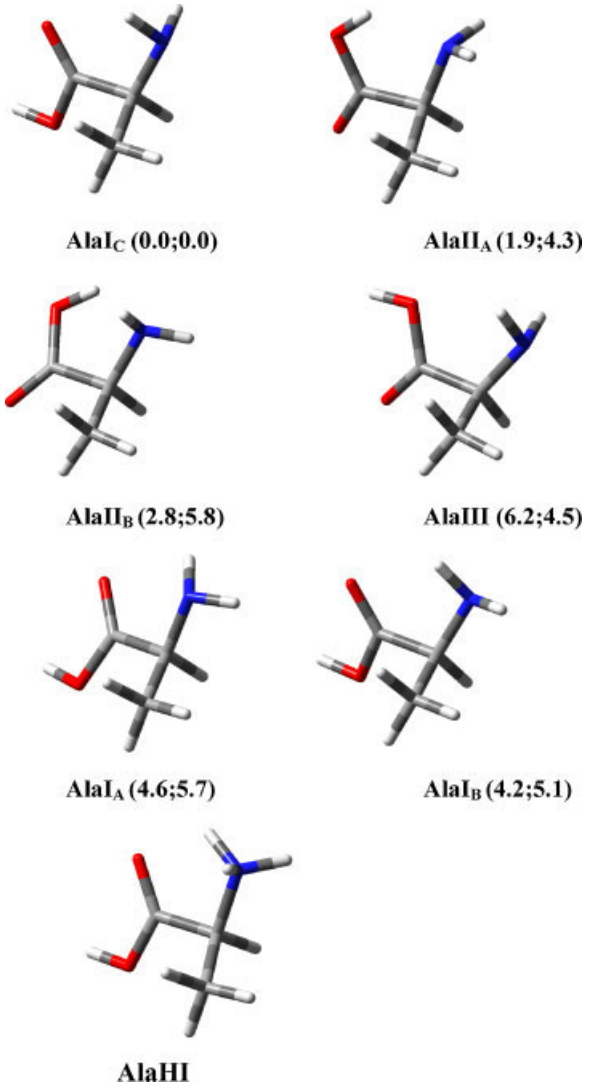


FIGURE 11. Most stable conformers of neutral and protonated alanine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G3MP2B3 results).

may be explained by their facile interconversion to the most stable conformers $\mathbf{AlaI_C}$ and $\mathbf{AlaII_A}$ by simple rotations around C(1)C(2) and C(2)N bonds, as already suggested for glycine (Godfrey, Brown, & Rodgers, 1996; Miller & Clary, 2004). The interconversion between $\mathbf{AlaI_C}$ and $\mathbf{AlaII_A}$ is not allowed because it involves the *anti* HOCO → *syn* HOCO isomerization. Indeed, this process is associated with an energy barrier of ca. 60 kJ mol^{-1}, clearly higher than the rotational barriers along the CC and CN bonds which are in the 10 kJ mol^{-1} range (Blanco et al., 2004).

Computational studies on the protonated forms of alanine are much more limited (Topol et al., 1998; Abirami et al., 2005; Yang et al., 2009; Bouchoux, Huang, & Inda, 2011). The two most stable conformers **AlaHI** and **AlaHII** are separated by differences in H°_{298} and G°_{298} of 16 and 15 kJ/mol, respectively, at the G3MP2 level. These enthalpy and free energy differences compare closely to that observed for glycine. In particular it may be considered that protonated alanine is present at 298 K to more than 99% under its conformation **AlaH** (Fig. 11).

c. Valine

Computational studies of neutral valine are much more limited (Stepanian et al., 1999; Lesarri et al., 2004; Yang et al., 2009; Dokmaisrijan, Lee, & Nimmanpipug, 2010; Bouchoux, Huang, & Inda, 2011). Investigation of the conformational landscape of neutral valine at the G3MP2B3 level places seven conformers in the first 5 kJ mol^{-1} range in the order of decreasing enthalpy: **ValI$_C$g−** > **ValII$_B$g−** > **ValI$_C$g+** > **ValII$_A$g−** > **ValII$_A$a** > **ValIII$_C$g+** ∼ **ValI$_C$a**. When looking at the free energy ordering, this number reduces to five (Fig. 12) and the ordering becomes **ValI$_C$g−**, **ValI$_C$g+**, **ValIII$_C$g+**, **ValII$_A$g−**, and **ValII$_B$g−**. The predicted population at 298 K of these conformers is 47/16/16/14/7, respectively (Bouchoux, Huang, & Inda, 2011).

Experimental matrix isolation IR spectroscopy of gaseous valine at 14 K was interpreted, with the help of B3LYP/6-311++G(d,p) calculation, by the occurrence of **ValI$_C$g−** and **ValII$_{(B?)}$g−**, with possible traces of **ValIII$_C$g+** (Stepanian et al., 1999). Similarly, the microwave spectrum of valine, vaporized by laser ablation, after analysis by MP2/6-311++G(d,p) calculations, seems to be explainable by the existence of conformers **ValI$_C$g−** and **ValII$_{(B?)}$g−** (Lesarri et al., 2004). Again, the experiments are apparently able to identify only the most stable conformers (in the enthalpy scale) of types **I** and **II**. Facile relaxation by rotation around the CC bonds in **ValI$_C$g+** and **ValIII$_C$g+** is indeed expected to lead to the most stable conformer **ValI$_C$g−** which may hardly interconvert with **ValII$_B$g−**.

Concerning protonated valine, the three conformers resulting from rotation around C(2)–C(3) on the **ValHI** structure (denoted g+, g− and a) were investigated (Kamariotis et al., 2006; Bouchoux, Huang, & Inda, 2011). The most stable conformers, **ValHIg−** is presented in Figure 12. The conformer's population **ValHIg−/ValHIa/ValHIg+**, computed from the 298 K free energies computed at the G3MP2B3 level is equal to 86/9/5.

d. Leucine

Only two systematic searches for the most stable conformers of leucine seem to have been undertaken to date (Dokmaisrijan, Lee, & Nimmanpipug, 2010; Bouchoux, Huang, & Inda, 2011). Results of the H°_{298} calculations at the G3MP2B3 level of theory show that the most stable conformers of Leucine are, by order of decreasing stability in the first 5 kJ mol^{-1}, **LeuI$_C$g+g+** > **LeuI$_C$aa** ∼ **LeuIIaa** > **LeuII$_B$g+g+** > **LeuIIg−a** (Fig. 13). This situation is changed when considering the free energy order since conformers **II** are entropically disfavored and the three most stable conformers become: **LeuI$_C$g+g+** > **LeuI$_C$aa**> **LeuIIIg+g+**, with relative population of 69/19/12. The jet

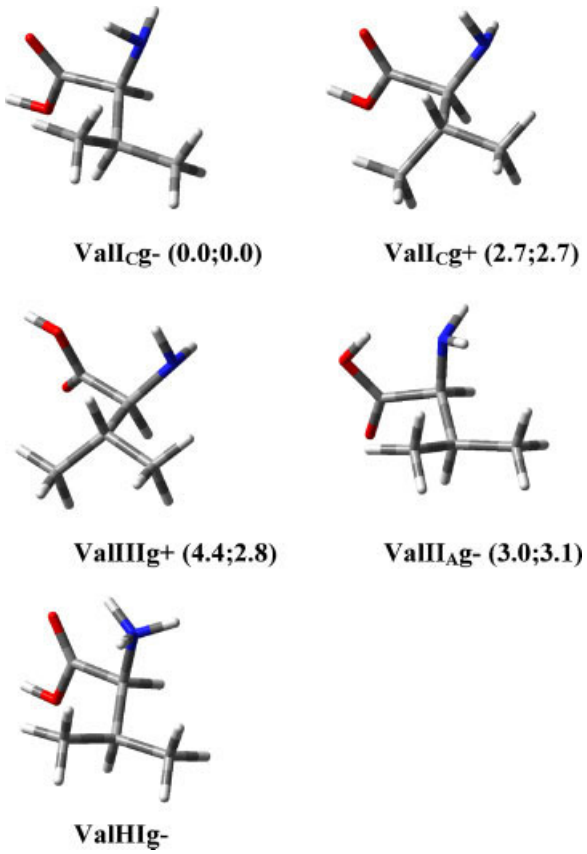


FIGURE 12. Most stable conformers of neutral and protonated valine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G3MP2B3 results).

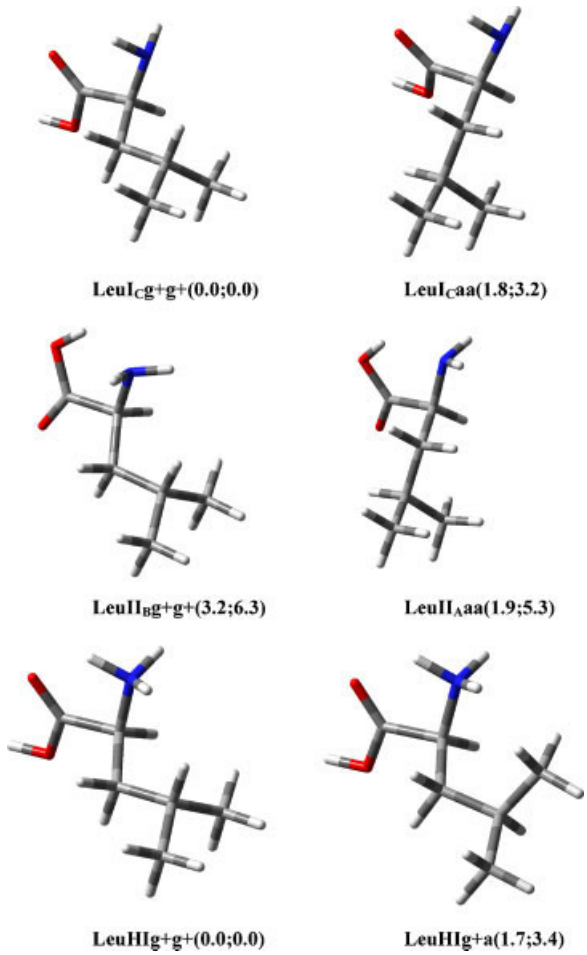


FIGURE 13. Most stable conformers of neutral and protonated leucine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G3MP2B3 results).

cooled rotational spectrum of neutral leucine has been interpreted by the existence of conformer **LeuI$_C$g+g+** and, surprisingly, in admixture with the highest energy conformer **LeuII$_B$g+g+** (Cocinero et al., 2007).

Protonated leucine may a priori adopt nine conformations if we consider only the N-protonated structure **LeuHI**. G3MP2B3 computations show that six conformers are predicted to lie in the first 5 kJ mol^{-1} of 298 K enthalpy but that this number reduces to only two when free energy is considered. Structures **LeuHIg+g+** and **LeuHIg+a** (see Fig. 13) are expected to represent ∼75% of the population of protonated leucine conformers at 298 K with a ratio 4/1.

e. Isoleucine

The various conformations of isoleucine have been explored at the MP2/6-311++G(d,p) and G3MP2B3 levels (Lesarri et al., 2005; Bouchoux, Huang, & Inda, 2011). It was found that 11 conformers are situated in the first 5 kJ mol^{-1} 298 K enthalpy. The two conformers predicted to be the most stable at the G3MP2B3 level are **IleI$_C$g−a** and **IleII$_B$g−a**, they are separated by only 1.2 kJ mol^{-1} on the 298 K enthalpy scale. They are followed by **IleII$_B$g−g−**, **IleI$_C$g+a**, **IleI$_C$ag−**, and **IleI$_C$g−g−**. As already observed for glycine, alanine, valine and leucine, it is noteworthy that conformations of type **III** are, as whole, energetically disfavored. However, one of them, namely **IleIII$_C$g+a** is predicted to be as stable as **IleI$_C$g+g+** and **IleI$_C$aa**. With regards to free energy, the most stable conformer is still **IleI$_C$g−a**, but **IleII$_B$g−a** passes at the fourth position, conformers **IleI$_C$g+a** and **IleIII$_C$g+a** being more favored. These four conformers (Fig. 14) represent 65% of the total population of neutral gaseous isoleucine and are predicted to be present in a ratio 57/18/14/11 at 298 K. Experimental characterization of isoleucine conformers has been done using Fourier transform microwave spectrometry after vaporization of the sample by laser ablation (Lesarri et al., 2005). Two conformers have been detected and identified using computations at the MP2/6-311++G(d,p) level. The proposed conformers, namely Ia1 and IIa1 in ref 24, correspond to structures **IleI$_C$g−a** and **IleII$_B$g−a** identified in the present study as the two more stable conformers on the enthalpy scale at the G3MP2B3 level. Again the lack of observation of conformers such as **IleI$_C$g+a** and **IleIII$_C$g+a** may be due to facile internal rotations leading to **IleI$_C$g−a**.

A search of the most stable conformers has been done for the protonated forms of isoleucine of the **HI** type (Bouchoux, Huang, & Inda, 2011). Seven of them have been examined at the G3MP2B3 level. Four conformers are situated in the 5 kJ mol^{-1} 298 K enthalpy range, the most stable being **IleHIg−a**. When considering 298 K free energies, the number of conformers reduces to only two (**IleHIg−a/IleHIag−**, Fig. 14), with a population ratio of 83/17.

f. Proline

Five membered rings are expected to present ten envelope conformations, E, characterized by the position of the out of plane atom *i* which may be either *"exo-i"* (noted E$_i$) or *"endo-i"* (noted ^{i}E) with respect to a given reference substituent. The passage from one ^{i}E conformer to its nearest homologues E$_{j=i+1}$ or E$_{j=i-1}$ is assumed by a pseudorotation of the ring allowing formation of twisted forms ^iT_j through low energy barriers. In the case of pyrrolidine ring, up to 20 E and T conformations are consequently possible. Considering now proline, account should be made of *syn* or *anti* conformations of the COOH moiety and of hydrogen bonding interactions with the NH group. Obviously, only one hydrogen bond between the HN moiety and the carbonyl oxygen is possible in proline, thus ruling out possibilities of conformers of type **I$_C$** where two amino hydrogens are involved in bifurcated hydrogen bonds. It remains however to investigate the occurrence of conformers of type **II**, **III** and the analogous of **I$_{B,C}$** denoted **IV**. A number of them has been explored by several research groups (Ramek, KeltererR, & Nikoli, 1997; Stepanian et al., 2001; Marino et al., 2001; Benzi et al., 2002; Lesarri et al., 2002; Mezzache et al., 2003; Czinki & Csaszar, 2003; Pecul et al., 2004; Mata et al., 2009; Bouchoux, Huang, & Inda, 2011).

The general conclusion of these studies is that four conformers of proline are situated in the first ∼5 kJ/mol of enthalpy or Gibbs free energy, the two most stable are of type **II** while the two other are of type **IV** (Fig. 15). It is interesting to note that conformers **III** of neutral proline are greatly

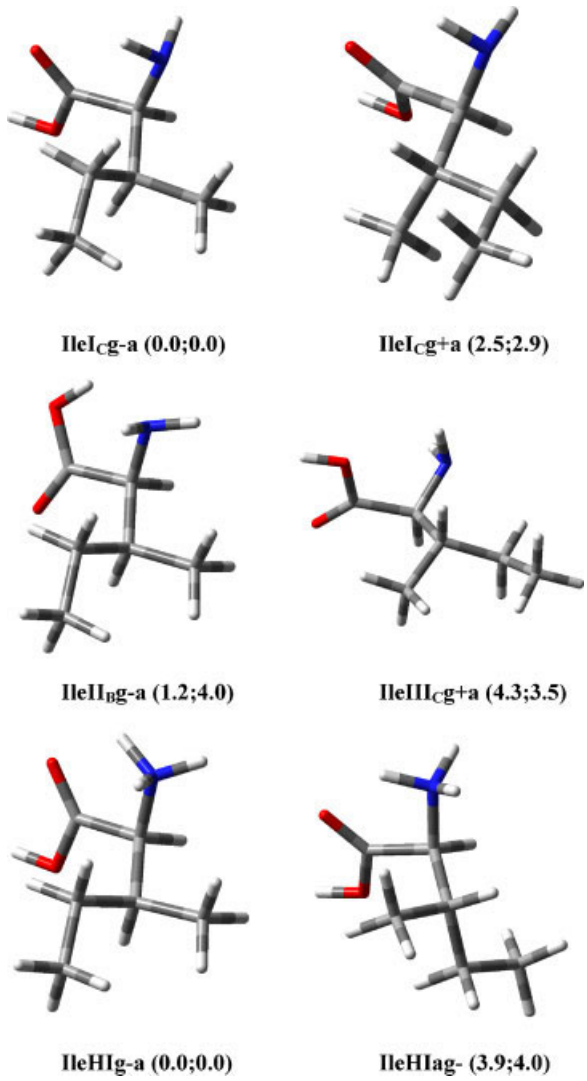


FIGURE 14. Most stable conformers of neutral and protonated isoleucine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G3MP2B3 results).

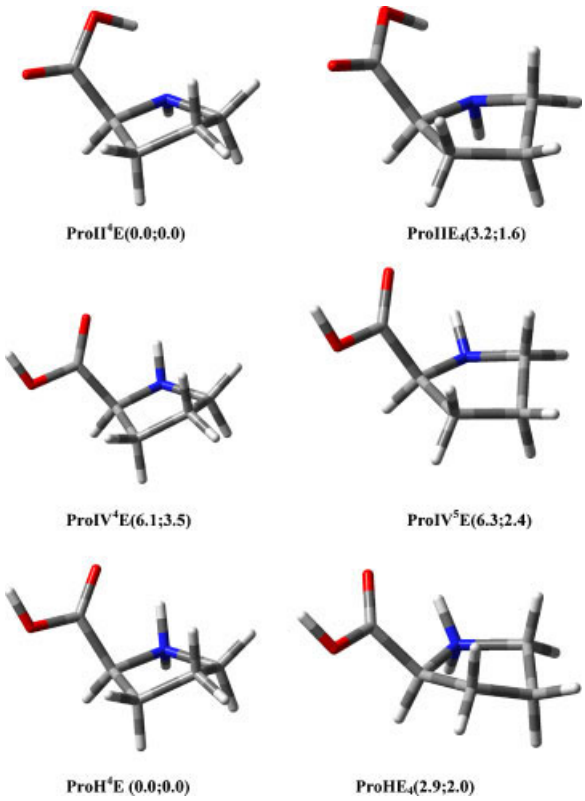


FIGURE 15. Most stable conformations of neutral and protonated proline (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G3MP2B3 results).

destabilized with respect to **II** since they are situated 14 kJ mol^{-1} above (Czinki & Csaszar, 2003). G3MP2B3 calculations show that **ProII^{4}E**(*endo-C4*) and **ProIIE$_4$**(*exo-C4*) are enthalpically favored over **ProIV^{4}E**(*endo-C4*) and **ProIV^{5}E**(*endo-C5*) (Bouchoux, Huang, & Inda, 2011). However, the situation is more balanced when entropy is considered as illustrated by the Gibbs free energy differences reported in Figure 15. It seems nevertheless that conformer **ProII^{4}E** (*endo-C4*) mostly participate to the description of proline at room temperature (the overall population ratio is 46/24/12/18).

Experimentally, IR spectrum of matrix isolated proline (Stepanian et al., 2001), rotational (microwave) spectrum of laser ablated proline (Lesarri et al., 2002; Mata et al., 2009) and valence and core-level photoelectron spectroscopy (Pecul et al., 2004) provided information on the lowest energy conformers. The presence of two bands in the C=O stretching region lead Stepanian et al. (2001) to the conclusion that two conformers, namely **ProII^{4}E**(*endo-C4*) and **ProIV^{5}E**(*endo-C5*), were present in argon matrix. One year later, Lesarri et al. (2002) concluded that **ProII^{4}E**(*endo-C4*) is the major conformer sampled under their experimental conditions (gaseous proline is desorbed by laser ablation, transported by a carrier gas and expanded supersonically into a microwave resonator) and that a second, minor, conformer, probably **ProIIE$_4$**(*exo-C4*), is also present. Finally, in 2009, the same group (Mata et al., 2009) detected two more conformers to which they assigned the structures **ProIV^{4}E**(*endo-C4*) and **ProIV^{5}E**(*endo-C5*).

Low lying conformers of protonated proline have been theoretically examined by DFT (Marino et al., 2001; Mezzache et al., 2003; Pecul et al., 2004) and G3MP2B3 (Bouchoux, Huang, & Inda, 2011) methods. Two conformers **ProHI^{4}E**(*endo-C4*) and **ProHIE$_4$**(*exo-C4*) of comparable energies were identified (Fig. 15). Using the G3MP2B3 298 K Gibbs free energies to calculate the conformers fractional abundances, a **ProHI^{4}E**(*endo-C4*)/**ProHIE$_4$**(*exo-C4*) ratio equal to 69/31 is obtained. IRMPD spectrum of protonated proline (Wu & McMahon, 2008b) is fully interpretable by the existence of **ProHI^{4}E**(*endo-C4*) conformer as a major component, however participation of **ProHIE$_4$**(*exo-C4*)cannot be excluded.

2. *Oxygen and Sulfur Substituted Amino Acids*

a. Serine

Explorations of the conformational space of neutral serine were done using Hartree-Fock (Gronert & O'Hair, 1995; Lakard, 2004), B3LYP (Noguera et al., 2001; Lambie, Ramaekers, & Maes, 2004; Miao et al., 2005; Pecul, 2006; Dinadayalane, Sastry, & Leszczynski, 2006; Jones et al., 2007), MP2 (Blanco et al., 2007), G2(MP2) (Bleiholder et al., 2006), G3(MP2) (Gronert, Simpson, & Conner, 2009) and a large panoplie, up to the G4 (Riffet, Frison, & Bouchoux, 2011), methods (note that enantiomer 2R rather than 2S was considered in Gronert & O'Hair, 1995; Noguera et al., 2001). At the G4 level, four conformers were predicted to be situated in a 5 kJ/mol enthalpy range, namely **SerI$_C$g**+, **SerII$_B$g**−, **SerI$_A$g**−, and **SerII$_B$a** (Fig. 16).

In the two conformers of type I (**SerI$_C$g**+, **SerI$_A$g**−) which, by definition, present the $NH_2 \ldots O = C–OH$(*syn*), arrangement, a second stabilizing interaction occurs between the hydrogen of the alcohol function and the nitrogen atom: $CH_2OH \ldots NH_2$. Conformers **SerII$_B$g**− and **SerII$_B$a** which allow a $H_2N \ldots HO–C{=}O$(*anti*), (type II) internal hydrogen bonding are also characterized by a hydrogen bond involving the hydrogen of the alcohol moiety and the oxygen of the carbonyl group. According to the G4 relative free energies the three most stable conformers of neutral serine, represent ∼80% of the total population at 298 K, in the ratio **SerI$_C$g**+/**SerII$_B$g**−/**SerI$_A$g**− of 47/35/18. This expectation is in correct agreement with experimental observations. Accordingly, the infrared spectrum of neutral serine, isolated in low temperature argon matrices, has been recorded and interpreted by the occurrence of a mixture of conformers **SerI$_C$g**+, **SerII$_B$g**−, **SerII$_B$a**, and **SerI$_A$g**− (Lambie, Ramaekers, & Maes, 2004). A more recent microwave study of gaseous serine produced by laser ablation has been interpreted by the existence of a population of conformers **SerI$_C$g**+, **SerII$_B$g**−, **SerI$_A$g**−, **SerII$_B$a** but also of two other conformers (**SerIII$_C$g**− and **SerII$_A$g**+) (Blanco et al., 2007).

Protonated serine conformational space has been extensively explored by Miao et al. (2005) (B3LYP/6-311+G(d,p) level) and by Riffet, Frison, & Bouchoux (2011) (up to the G4 level). The four most stable structures **SerHIg**−, **SerHIg**+, **SerHIIg**−, and **SerHIIg**+ present a *syn* HOCO arrangement of the acidic moiety allowing the establishment of a strong hydrogen bond with one of the hydrogen of the NH_3^+ group. The most favorable interaction obviously occurs with the oxygen of the carbonyl group thus leading to the two most

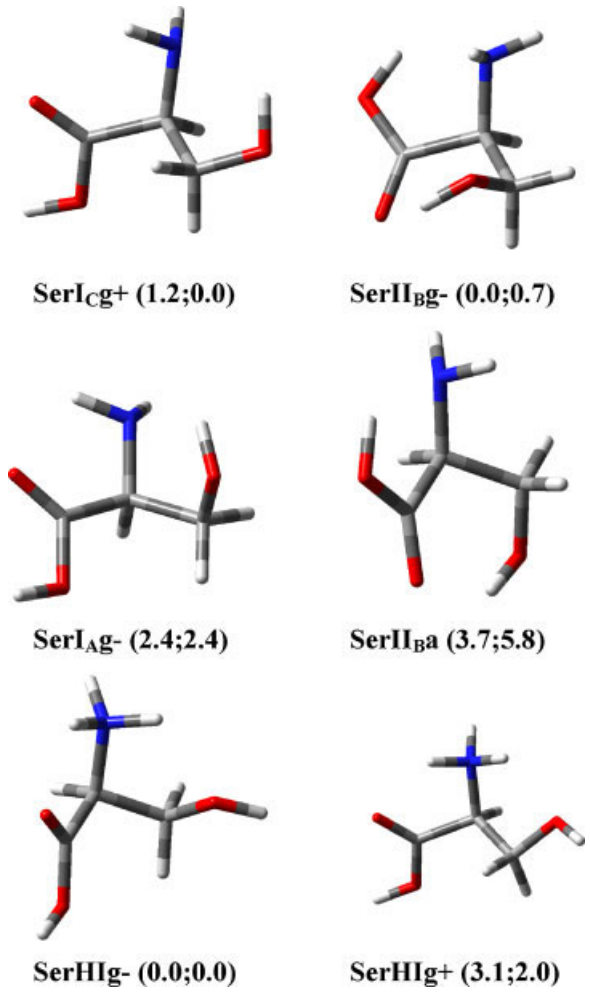


FIGURE 16. Most stable conformations of neutral and protonated serine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G4 results).

stable conformers **SerHIg−** and **SerHIg+** (Fig. 15). The other possibility, that is, interaction between NH_3^+ and the oxygen of the hydroxyl group is less favorable and leads to conformers **SerHIIg−** and **SerHIIg+** situated ∼15 kJ/mol above **SerHIg−** and **SerHIg+**. The structural difference between **SerHIg−** and **SerHIg+** (or **SerHIIg−** and **SerHIIg+**) originates from the NCCO(H) dihedral angle which induces different distances between NH_3^+ and the oxygen of the alcohol function. This distance is shorter for conformers **g−** thus explaining why they are more stable than their counterparts **g+**. The population of protonated serine at 298 K is obviously dominated by conformers **SerHIg−** and **SerHIg+** which amount for 69% and 31%, respectively, at the G4 level. Experimentally, IRMPD spectrum of protonated serine in the 1,000–2,000 cm^{-1} range has been recorded (Wu & McMahon, 2008b). Three characteristics absorption bands were observed at 1,158 cm^{-1} (HOCO bend), 1,460 cm^{-1} (NH_3 umbrella) and 1,794 cm^{-1} (CO stretching) and attributed to the presence of conformer **SerHIg−** (SH01 in the original article). It is noteworthy that conformer **SerHIg+** (which has not been considered by Wu and McMahon, 2008b) presents closely similar vibrational frequencies values.

b. Threonine

Conformational landscape of neutral L-threonine has been the subject of recent theoretical investigations (Lakard, 2004; Zhang & Lin, 2006; Szidarovsky, Czako, & Csaszar, 2009; Xu & Lin, 2010; Riffet, Frison, & Bouchoux, 2011). As expected conformational landscape of threonine strongly resemble that of serine. Thus, the first most stable conformers (Fig. 17) are similar to that reported above for serine. Slight differences are however observed due to the presence of the methyl group. Four conformers, **ThrII$_B$g−**, **ThrI$_C$g+**, **ThrI$_A$g−**, and **ThrIII$_C$g+** (Fig. 17), essentially describe gaseous neutral threonine at 298 K.

From the G4 Gibbs free energies it is deduced a population of conformers **ThrII$_B$g−/ThrI$_C$g+/ThrI$_A$g−/ThrIII$_C$g+** at 298 K in the ratio 34/24/17/10%, the remaining 15% being shared between five minor conformers (Riffet, Frison, & Bouchoux, 2011). Comparable results based on MP2/6-311+G(2df,p) calculations have been reported by Lin & Xu (2010). From an experimental point of view, Fourier transform microwave spectroscopy was applied to L-threonine vaporized by laser ablation (Alonso et al., 2009). Analysis of the resulting rotational spectra leads the authors to conclude that seven conformers were present in the experimental conditions, the most abundant of which being **ThrII$_B$g−**, **ThrI$_C$g+**, and **ThrI$_A$g−** in agreement with the G4 expectations.

Two conformers of close energies were identified for protonated threonine (Xu & Lin, 2010; Riffet, Frison, & Bouchoux, 2011). These two structures, **ThrHIg−** and **ThrHIg+**, are predicted to represent the equilibrium population at 298 K in the ratio of 62/38% (Riffet, Frison, &

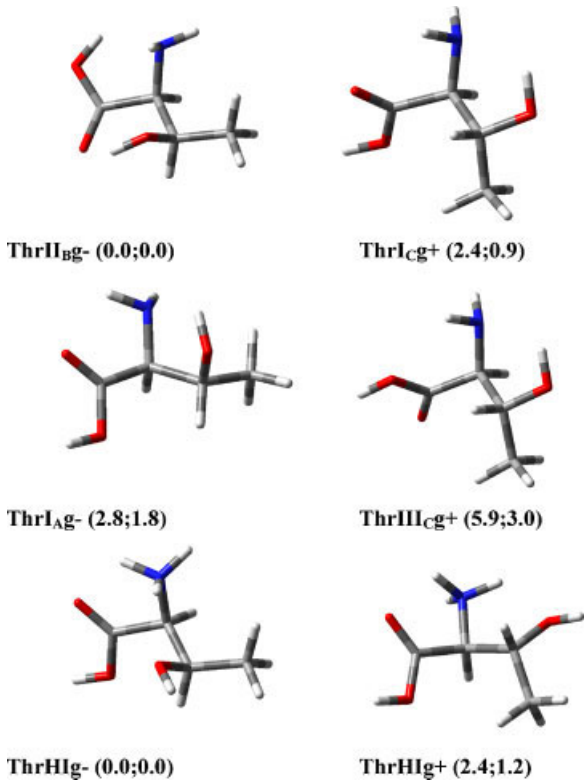


FIGURE 17. Most stable conformations of neutral and protonated threonine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G4 results).

Bouchoux, 2011). These two conformations present particularly strong intramolecular hydrogen bonds between the NH_3^+ moiety and (i) the oxygen of the carbonyl group, and, (ii) the hydroxylic oxygen (Fig. 17). Two other structures, **THIIg+** and **THIIg−**, characterized by a less efficient $NH_3^+ \ldots$ O(H)–C=O bonding are situated more than 10 kJ/mol above **THIg−** and **THIg+**. This situation is obviously reminiscent to that encountered with serine.

c. Cysteine

Theoretical works on neutral cysteine reported results obtained at the HF (Gronert & O'Hair, 1995), B3LYP (Fernandez-Ramos et al., 2000; Noguera et al., 2001; Pecul, 2006; Dobrowolski, Rode, & Sadlej, 2007; Tian et al., 2007), MP2 (Sanz et al., 2008; Wilke et al., 2009) and composite (Riffet, Frison, & Bouchoux, 2011) levels of theory. Some of these studies however are concerned by the non-natural 2S configuration of cysteine (Gronert & O'Hair, 1995; Fernandez-Ramos et al., 2000; Noguera et al., 2001; Wilke et al., 2009). There is a general consensus in placing conformer **$CysII_Bg-$** as the most stable species. At the G4 level, the order of stability of the cysteine conformers in terms of H°_{298} is: **$CysII_Bg-$** > **$CysI_Cg-$** > **$CysI'_Cg-$** > **$CysI_Ag-$**, **$CysI_Cg+$** > **$CysII_Ag+$** > **$CysIII_Cg-$**. It is noteworthy that the enthalpy difference between **$CysII_Bg-$** and the conformers situated immediately above is as large as ~6 kJ/mol while this difference in stability is significantly reduced in the free energy scale G°_{298} at the G4 level.

The origin of this difference is the shift of the conformers of type I (**$CysI'_Cg-$**, **$CysI_Cg-$**, **$CysI_Cg+$**, **$CysI_Ag-$**) toward low G°_{298}. This was expected from their high third law entropies (near 380 J mol^{-1} K^{-1}, as compared to 367 J mol^{-1} K^{-1} for **$CysII_Bg-$**) (Riffet, Frison, & Bouchoux, 2011). This entropy difference is obviously related to the strong character of the two major internal hydrogen bonds occurring in **$CysII_Bg-$** ($H_2N \ldots$ HOC = O(*anti*) and NH ... SHC(3)). The population of conformers at 298 K is predicted to be **$CysII_Bg-$**/**$CysI'_Cg-$**/**$CysI_Cg-$**/**$CysI_Cg+$**/**$CysI_Ag-$**/**$CysIII_Cg-$**: 35/24/15/12/8/6% at the G4 level.

Experimentally, no less than six conformers have been claimed to be identified (Sanz et al., 2008). Accordingly, microwave spectra of vapor phase cysteine have been interpreted by the major presence of **$CysII_Bg-$**, **$CysI_Cg+$**, and **$CysI_Cg-$** conformers beside minor contribution of **$CysIII_Cg-$** and two other conformers (**$CysII_Ag+$**, **$CysIII_Ca$**) (Sanz et al., 2008). Surprisingly enough, conformers **$CysI'_Cg-$** and **$CysI_Ag-$** are not mentioned despite the fact that they are predicted to be more stable than **$CysII_Ag+$**, **$CysIII_Ca$**, or **$CysIII_Cg-$**. It may be underlined however that rotational constants of **$CysI'_Cg-$** and **$CysI_Ag-$** are very close to that of **$CysI_Cg-$** probably rendering a clear structural assignment uneasy.

Among the 21 conformers identified for protonated cysteine, four were predicted to be situated in a ~6 kJ/mol of(298 K) enthalpy or free energy range (Riffet, Frison, & Bouchoux, 2011). These four structures, **CysHIg−**, **CysHIg+**, **CysHI′g−**, and **CysHI′g+**, are stabilized by $NH_4^+ \ldots$ O = COH(*syn*) and $NH_4^+ \ldots$ S hydrogen bonding interactions. The difference between conformers **CysHIg−** and **CysHI′g−**, or **CysHIg+** and **CysHI′g+**, lies on a rotation around the C(3)–S bond (Fig. 18). The distribution of conformers **CysHIg−**/**CysHI′g+**/**CysHI′g−**/**CysHIg+**, computed using the G4 free energies, is equal to 54/28/11/7% (Riffet, Frison, & Bouchoux, 2011).

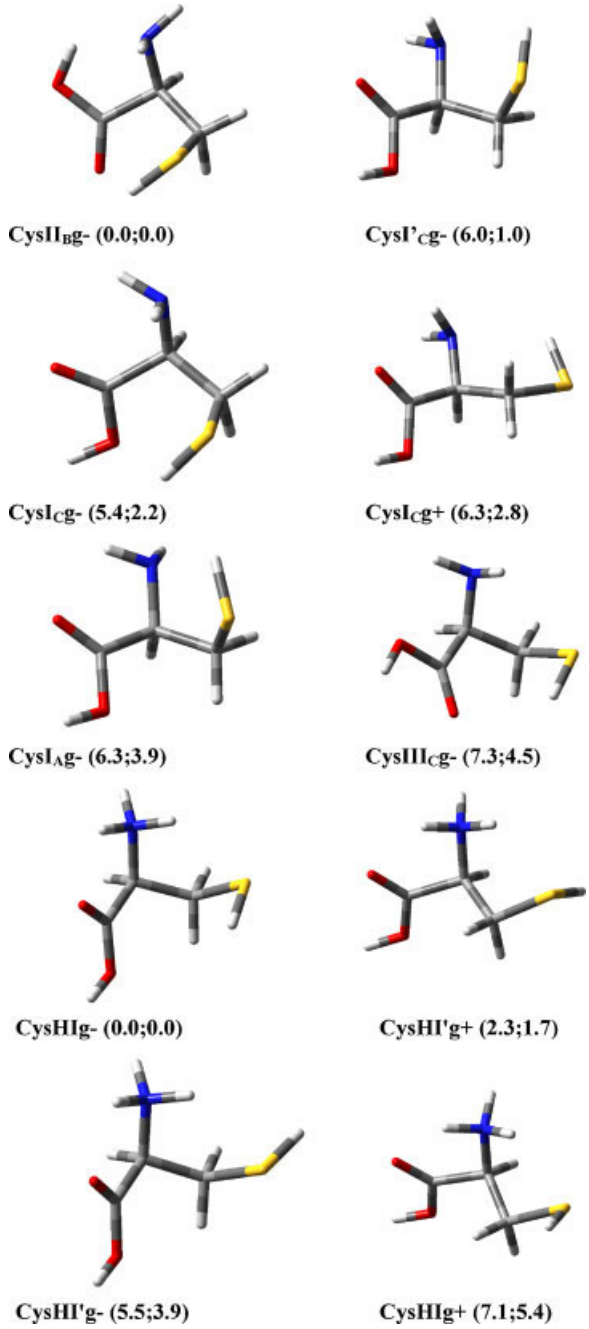


FIGURE 18. Most stable conformations of neutral and protonated cysteine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G4 results).

d. Methionine

Neutral methionine potential energy surface has been examined by several groups at various levels of theory (Bleiholder et al., 2006; Dinadayalane, Sastry, & Leszczynski, 2006; Jones et al., 2007; Lioe et al., 2007; Plekan et al., 2007; Linder et al., 2008; Desaphy, Malosse, & Bouchoux, 2008; Armentrout, Gabriel, & Moision, 2009; Gronert, Simpson, & Conner, 2009; Riffet, Frison, & Bouchoux, 2011). Most of the time, a

limited number of conformers has been considered, however the most recent investigation reports results obtained using several DFT (B97-D, B3LYP, M06-2X) and composite (CBS-QB3, G3, G4) methods on a panel of 29 conformer (Riffet, Frison, & Bouchoux, 2011). It emerges from this study that five structures of type I (**MetI$_C$g+g+g+**, **MetI$_C$g+ag−**, **MetI$_C$g+ag+**, **MetI$_C$g+aa**, **MetI$_C$g+g+g−**, Fig. 19) and three structures of type II (**MetII$_B$g−g+a**, **MetII$_A$g+g−g−**, **MetII$_B$g−g+g+**, Fig. 19) are situated in a 6 kJ/mol range of the computed G4 H°_{298} scale. When passing in the G4 G°_{298} scale, only the former conformers are located in the first 6 kJ/mol. This is so because the third law entropies of conformers of type II (~450 J mol^{-1} K^{-1}) are significantly lower than that of conformers of type I (~435 J mol^{-1} K^{-1}). The theoretical population of neutral methionine conformers at 298 K calculated at the G4 level is **MetI$_C$g+g+g+/MetI$_C$g+ag−/MetI$_C$g+ag+/MetI$_C$g+aa/MetI$_C$g+g+g−**: 58/15/11/9/7%.

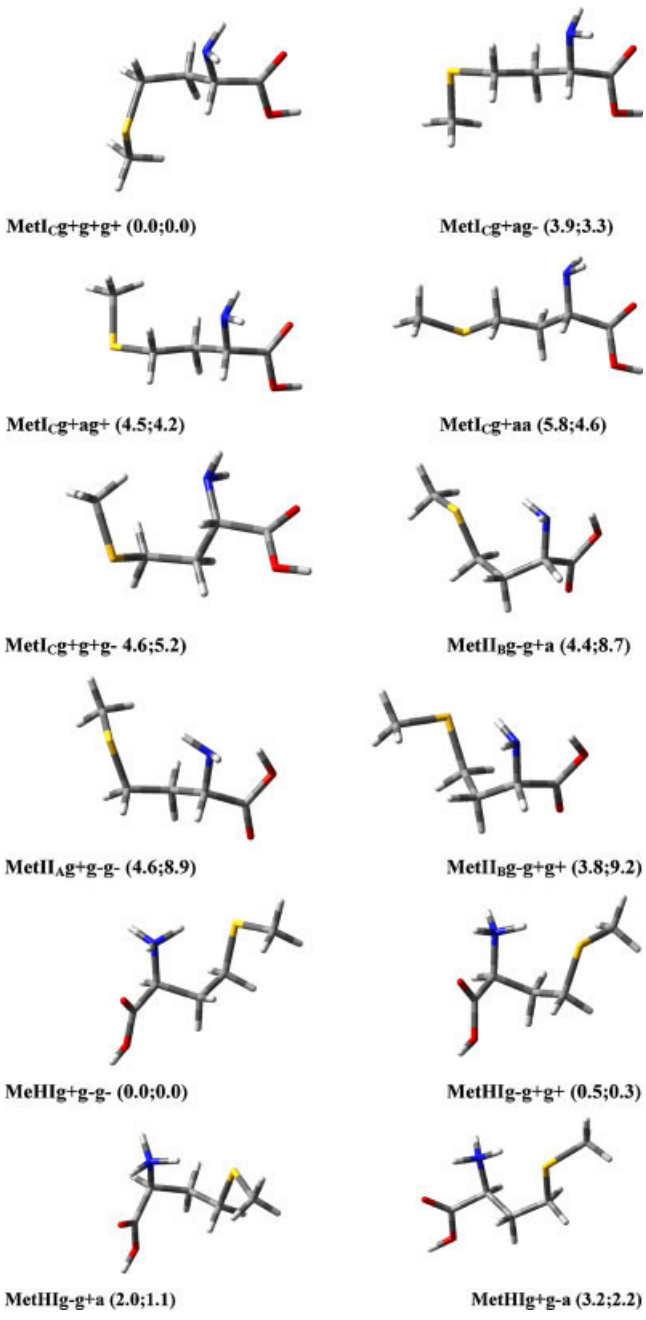


FIGURE 19. Most stable conformations of neutral and protonated methionine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G4 results).

Neutral methionine has been studied experimentally by valence core photoelectron spectroscopy in the VUV and soft X-ray regions (Plekan et al., 2007) and by gas phase Fourier transform IR spectroscopy (Linder et al., 2008). However, observations were interpreted by computations on conformers situated more than 10 kJ/mol above the most stable.

Only two conformers of protonated methionine were previously considered (Bleiholder et al., 2006; Dinadayalane, Sastry, & Leszczynski, 2006; Lioe et al., 2007; Desaphy, Malosse, & Bouchoux, 2008; Gronert, Simpson, & Conner, 2009). A more recent theoretical study (Riffet, Frison, & Bouchoux, 2011) reveals that four unique structures are situated in a ~3.0 kJ/mol energy (both H°_{298} and G°_{298}) window (**MetHIg+g−g−**, **MetHIg−g+g+**, **MetHIg−g+a** and **MetHIg+g−a**, Fig. 19). Beside the $NH_3^+ \ldots$ OCOH(syn) interaction, the four structures are stabilized by a second internal hydrogen bond involving the sulfur atom as a proton acceptor: $NH_3^+ \ldots$ S. It is noteworthy that the NH ... S distance is remarkably constant for the four conformers (~2.134 ± 0.007 Å) thus suggesting comparable stabilization effect. Clearly, **MetHIg+g−g−** and **MetHIg−g+g+**, are of identical stabilities and should be hardly distinguished experimentally. At the G4 level, the calculated relative population of conformers, **MetHIg+g−g−/MetHIg−g+g+/MetHIg−g+a/MetHIg+g−a** is equal to 34/30/22/14% at 298 K.

3. *Carbonylated Amino Acids*

a. Aspartic acid

Neutral aspartic acid conformations have been examined during the course of extensive calculations of proton affinities of the 20 natural amino acids (Bouchoux et al., 2004; Bleiholder et al., 2006; Dinadayalane, Sastry, & Leszczynski, 2006; Gronert, Simpson, & Conner, 2009) and in a recent microwave spectroscopy study (Sanz, Lopez, & Alonso, 2010). Re-examination of the proposed structures at the G4MP2 level reveals that at least eight conformers present relative G°_{298} values less than 5 kJ/mol (Bouchoux & Riffet, 2011a). The six most abundant structures, as predicted by G4MP2 computations, are presented in Figure 20. Structures **AspII$_B$g−**, **AspI$_C$g−**, and **AspII$_B$g+** have been also examined at the B3LYP/6-311+G(d,p) level (Heaton, Moision, & Armentrout, 2008). The calculated population ratio of the six conformers presented in Figure 20 is, **AspII$_B$g−/AspI$_C$g−/AspII$_B$g+/AspIII$_B$g−/AspI$_A$g+/AspIII$_A$**: 13/28/10/23/13/13% at 298 K. Sanz et al. (2010) claimed that they have identified six conformers of aspartic acid from analysis of Fourier transform microwave spectra of vaporized sample. These structures correspond to the first five conformers presented in Figure 20 plus one conformer of type I of higher energy. It is noteworthy that the population of conformers deduced from experiment differs significantly from the theoretical prediction. This has been interpreted by a relaxation of some conformers to more stable species by internal rotations during the jet expansion

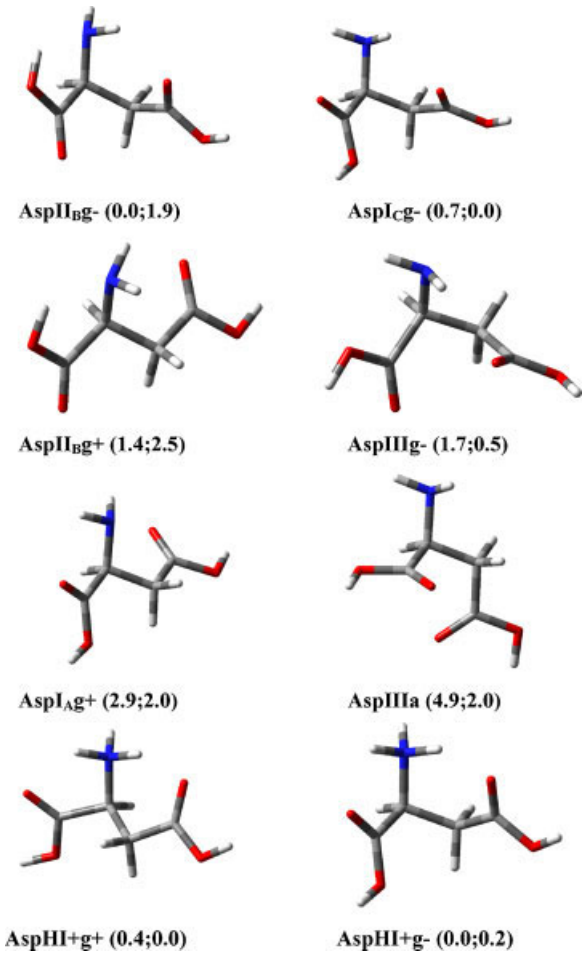


FIGURE 20. Most stable conformations of neutral and protonated aspartic acid (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G4MP2 results).

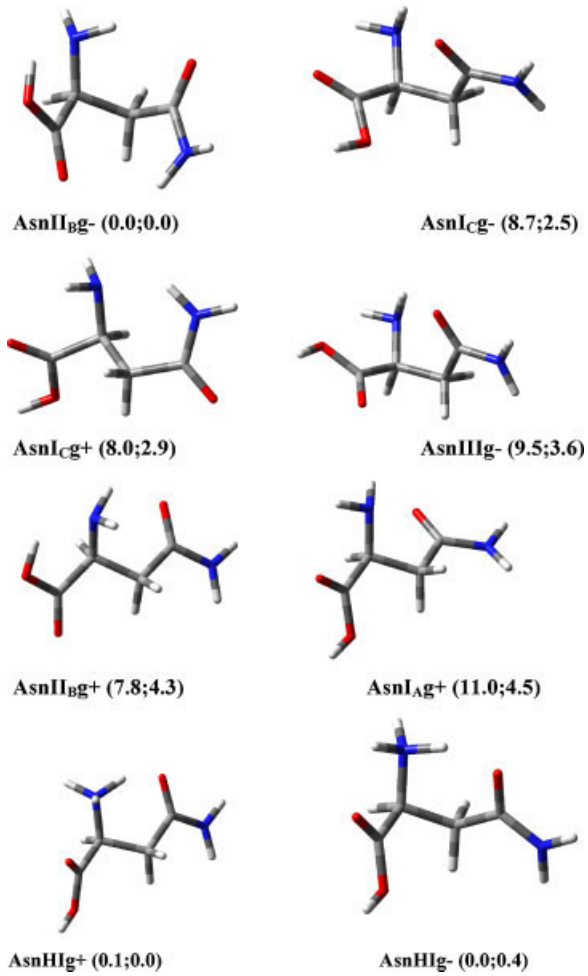


FIGURE 21. Most stable conformations of neutral and protonated asparagine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G4MP2 results).

allowing the vaporization of the sample (Sanz, Lopez, & Alonso, 2010).

Protonated aspartic acid presents two conformations of comparable energies (**AspHI**+**g**+ and **AspHI**+**g**−, Fig. 20). One or the other structures have been used to calculate the proton affinity of aspartic acid (Bouchoux et al., 2004; Dinadayalane, Sastry, & Leszczynski, 2006; Bleiholder et al., 2006; Gronert, Simpson, & Conner, 2009; Bouchoux & Riffet, 2011a).

b. Asparagine

An extensive search of stable conformations of neutral asparagines has been performed by Chen, Huang, & Lin (2005). A total of 62 conformers was identified and their energies were calculated at the MP2/6-311++G(d,p)//B3LYP/6-311++G(d,p) + ZPE level. Four structures were found in the 0–5 kJ/mol H°_0 range. In a recent re-examination of the most stable asparagine conformers at the G4MP2 level (Bouchoux & Riffet, 2011a) a total of six structures were located in the 5 kJ/mol of relative G°_{298}, namely: **AsnII$_{Bg}$**−, **AsnII$_{Bg}$**+, **AsnI$_{Cg}$**−, **AsnI$_{Cg}$**+**(anti)**, **AsnI$_{Ag}$**+, **AsnIIIg**− (Fig. 21). There is a clear agreement in most of the studies concerned with neutral asparagines to identity **AsnII$_{Bg}$**− as the most stable conformer (Bouchoux et al., 2004; Chen, Huang, & Lin, 2005; Dinadayalane, Sastry, & Leszczynski, 2006; Heaton, Moision, & Armentrout, 2008; Gronert, Simpson, & Conner, 2009). This structure is stabilized by the usual (*anti*)-O=COH ... NH_2 internal hydrogen bond of type II conformers plus two additional interactions: NH_2 ... O=CNH_2 and O=CNH_2 ... O=COH. According to the G4MP2 calculation this conformer represents ∼45% of the population of the six structures presented in Figure 21.

Protonated asparagine may exist in two conformations of comparable energies, and consequently of ∼50/50 population (at 298 K): **AsnIHg**+ and **AsnIHg**− (Fig. 21) (Heaton & Armentrout, 2009; Heaton et al., 2009; Bouchoux & Riffet, 2011a). Accordingly, the IRMPD spectrum of protonated asparagines is fully explainable by a mixture of both conformers **AsnIHg**+ and **AsnIHg**− (Heaton et al., 2009).

c. Glutamic acid

A first extensive examination of the conformational landscape of neutral and protonated glutamic acid has been done 12 years ago at the B3LYP/6-31G(d,p) level (Sun, Kinsel, & Marynick,

1999). Two more restricted works (Bouchoux et al., 2004; Heaton, Moision, & Armentrout, 2008) and a recent G3MP2B3 study completed the information on this system (Bouchoux et al., 2009b). Four conformations are situated in a 5 kJ/mol enthalpy range (Fig. 22). Only the two conformers of type I remain in the 5 kJ/mol of 298 K Gibbs free energy (**GluI$_C$g+g+** and **GluI$_C$g+a**, Fig. 22). Note that, for these two conformers, the side chain acidic group is in its *syn* configuration. The corresponding *anti* conformers, which allow a stabilizing O=COH ... NH_2 interaction, are situated ~7 kJ/mol above (Bouchoux et al., 2009b; Bouchoux & Riffet, 2011a).

The two conformers **GluIH+g−g+** and **GluIH+g+g−** (Fig. 22) are of comparable H°_{298} and G°_{298} (Bouchoux et al., 2009b; Bouchoux & Riffet, 2011a). As expected, they find their stability from a double internal hydrogen bonding network HOC=O ... HNH_2^+ ... O=COH(side chain). Conformers **GluI$_C$g+g+** and **GluIH+g−g+** were used during computations of proton affinity of glutamic acid (Bouchoux et al., 2004; Dinadayalane, Sastry, & Leszczynski, 2006; Gronert, Simpson, & Conner, 2009).

d. Glutamine

Several conformers of neutral glutamine were identified (Bouchoux et al., 2004; Heaton, Moision, & Armentrout, 2008) but, to our knowledge, no systematic examination of the conformational landscape of glutamine has been published. A G3MP2B3 investigation of Gln conformers places 3 and 8 structures in a 5 kJ/mol H°_{298} and G°_{298} range, respectively (Bouchoux & Riffet, 2011a). Conformer **GlnI$_C$g+g+(s)** turns out to be the most stable in the Gibbs free energy scale while **GlnII$_A$ag−(s)** was the conformer of lowest enthalpy (Fig. 23).

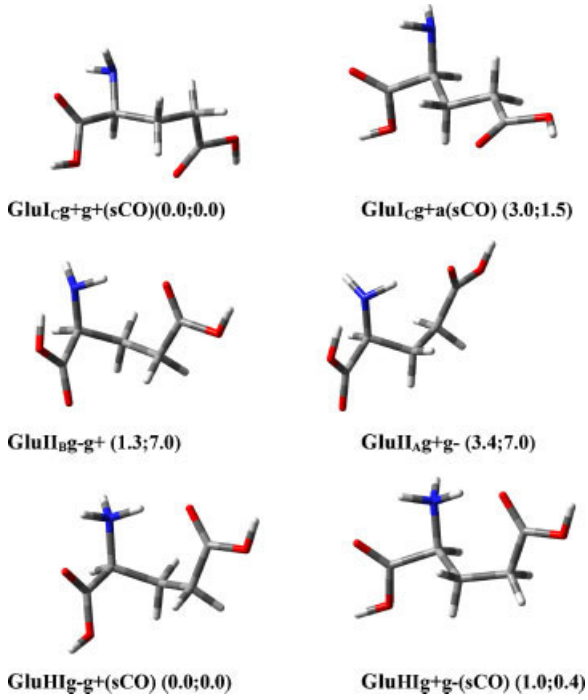


FIGURE 22. Most stable conformations of neutral and protonated glutamic acid (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G3MP2B3 results).

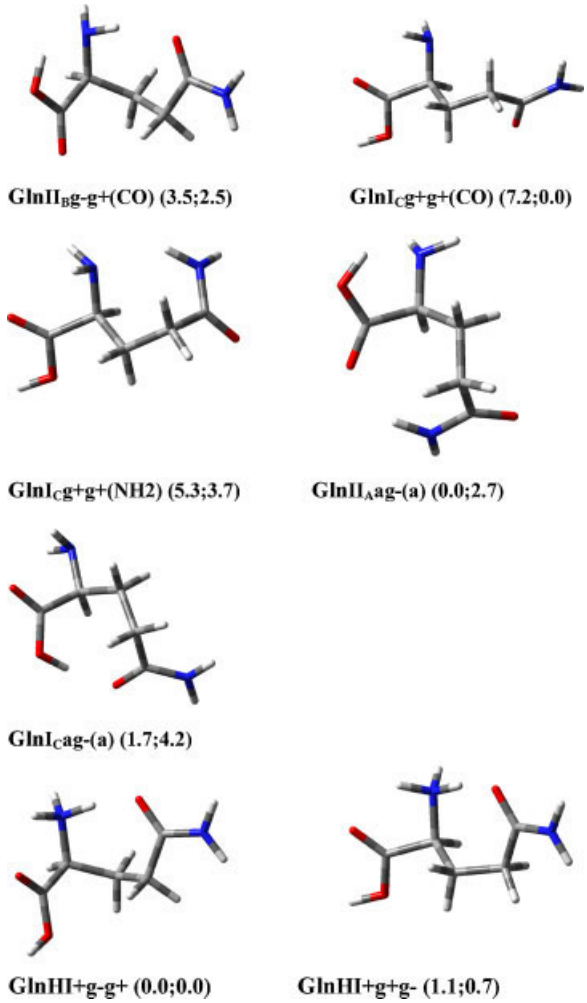


FIGURE 23. Most stable conformations of neutral and protonated glutamine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G3MP2B3 results).

As observed for glutamic acid, two protonated forms of glutamine are predicted to be of comparable stabilities (**GlnIHg−g+(s)** and **GlnIHg+g−(s)**, Fig. 23).

4. Nitrogen Substituted Amino Acids

Arginine and lysine have been frequently examined by mass spectrometry techniques for several reasons. The first point is the fact that these amino acids are the most basic in the series of natural α-amino acids and that the most probable protonation site is a nitrogen atom situated on the side chain, not in α of the acidic group. Moreover their long chain bidentate aliphatic structure has been early suspected to give rise to strong intramolecular hydrogen bonding in the protonated form (Campbell et al., 1992). The strong proton affinities of arginine and lysine could outweigh the energy necessary for the charge separation occurring in zwitterionic species. The present theoretical and experimental evidence show that arginine and lysine do not exist as stable, isolated, zwitterionic forms in the gas phase. By contrast, an addition of single water molecules or counterions or clusters formation can strongly stabilize zwitterionic structures (Julian, Beauchamp,

& Goddard, 2002; Lemoff et al., 2006; Bush et al., 2007a,b; Forbes et al., 2007; Wu & McMahon, 2008a; O'Brien et al., 2010).

a. Lysine

Extensive search for the conformers of neutral and protonated lysine has been reported recently (Leng et al., 2008). Previous studies were generally limited to a few structures (Bouchoux et al., 2004; Bleiholder et al., 2006; Dinadayalane, Sastry, & Leszczynski, 2006; Lemoff et al., 2006; Gronert, Simpson, & Conner, 2009). Six conformers of neutral lysine were reported to be situated in a 5 kJ/mol H°_{298} range, this number amounts to eight when considering the G°_{298} scale (MP2/6-311G(2df,p)//B3LYP/6-311G(d,p) calculations) (Leng et al., 2008). These structures were re-examined here at the G3MP2B3 level, relevant data are reported in Figure 24. The three most stable conformers on the H°_{298} scale are of type I with, however, a conformation *anti* of the acidic group (denoted I′), **LysI′ag−g+g+**, **LysI′g+g+g+g−**, and **LysI′ag−g+g−**. This arrangement allows the formation of a strong OH ... NH_2(side chain) internal hydrogen bond. The three following structures (**LysII$_B$g+g+g−g+**, **LysII$_A$g+ag+g+**, and **LysII$_B$g−ag−g−**, Fig. 24) are conformers of type II for which the flexibility of the $(CH_2)_4$ aliphatic chain is hampered by the formation of a (amino acid)NH_2 ... NH_2(side chain) hydrogen bond. Finally, the fully linear conformer of type I, **LysIaaaa**, is also reported in Figure 24 since it represent the most stable structure on the G°_{298} scale. Accordingly, its floppy character leads to a very large third law entropy S°_{298} of 467 J mol^{-1} K^{-1} (G3MP2B3 calculation) while S°_{298} is close to 430 J mol^{-1} K^{-1} for conformers of type I′. This difference in entropy induces a $T\Delta S$ term of ∼10 kJ/mol, at the origin of the observed inversion in stability ordering between the H°_{298} and G°_{298} scales.

Leng et al. (2008) identified three conformers of protonated lysine in a ∼3 kJ/mol H°_{298} or G°_{298} energy range, results which are confirmed at the G3MP2B3 level (**LysHIg−g−g+g−**, **LysHIg−g+g−g+A**, and **LysHIg−g+g−g+B**, Fig. 24). These structures are characterized by a double internal hydrogen bonding: NH_3^+ ... NH_2 and NH_3^+ ... O=COH although this latter interaction is weaker in conformer **LysHIg−g+g−g+B**. Experimentally, protonated lysine was studied by resonant IRMPD spectroscopy in the 2,800–3,600 cm^{-1} region (Vaden et al., 2007) and in the 1,100–2,000 cm^{-1} region (Bush et al., 2007a; Wu & McMahon, 2008b). The experimental IRMPD spectra were indeed interpreted by protonation at the terminal (side chain) nitrogen and by the existence of at least a NH_3^+ ... NH_2 hydrogen bond, as well as the absence of NH_3^+ ... OH interaction as indeed provided by any of the three structures **LysHIg−g−g+g−**, **LysHIg−g+g−g+A**, or **LysHIg−g+g−g+B**.

b. Arginine

Extensive searches of stable conformations of neutral arginine have been tentatively done during the last 10 years (Rak et al., 2001; Gdanitz et al., 2004; Ling et al., 2006; Schlund et al., 2008). Due to the very large number of possible conformer and to the uncertainties on the method used in the first selection of the low energy conformers, several discrepancies appears between these different studies. The most recent theoretical study (Ling et al., 2006) clearly illustrate this aspect

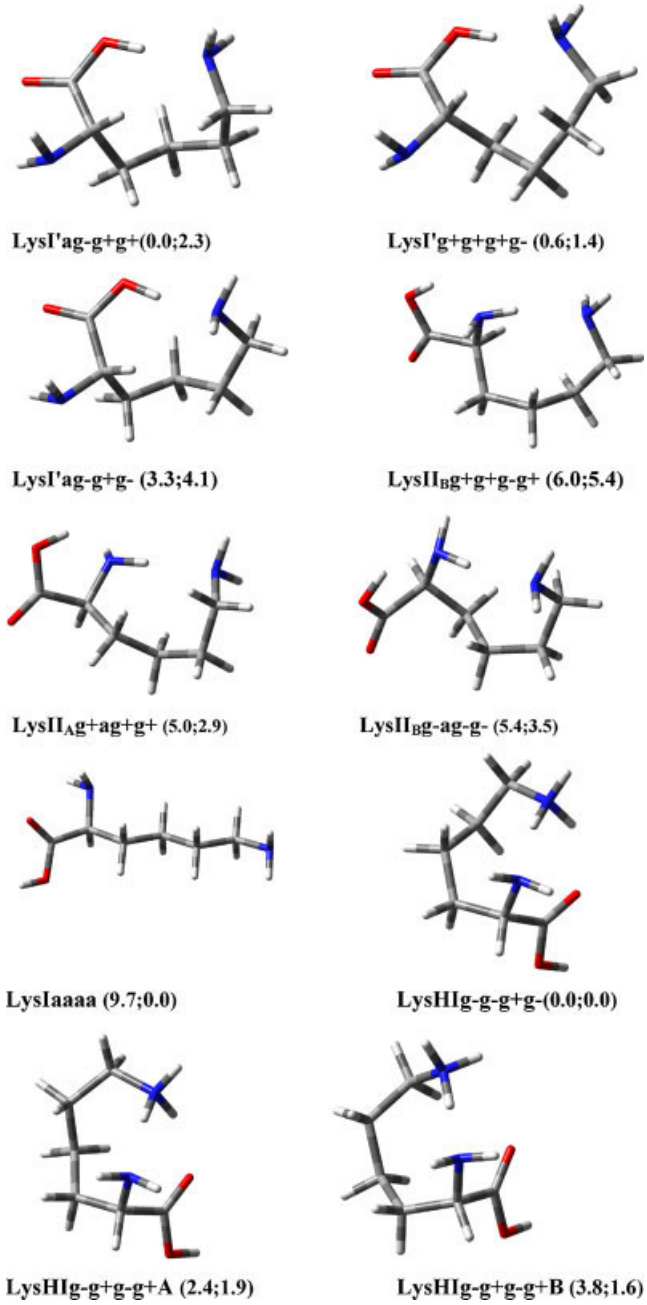
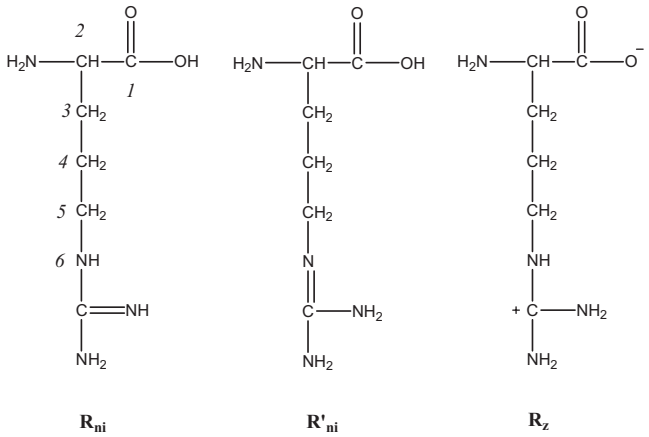


FIGURE 24. Most stable conformations of neutral and protonated lysine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G3MP2B3 calculations).

since it presents no less than 11 new canonical and zwitterionic forms which are lower in energy by 8–10 kJ mol^{-1} than the previously identified structures (Rak et al., 2001; Gdanitz et al., 2004). The first question concerns the most stable tautomeric state of arginine in the gas phase (Scheme 10). There is a general agreement in placing the zwitterionic forms (R_z; Scheme 10) ca. 15 kJ mol^{-1} above the non-ionic forms R_{ni} (–NH–C(NH_2)=NH) or R'_{ni} (–N=C(NH_2)$_2$) (Julian, Beauchamp, & Goddard, 2002; Gdanitz et al., 2004; Lemoff et al., 2006; Ling et al., 2006; Bush et al., 2007a,b; Forbes et al., 2007; Wang, Ohanessian, & Wesdemiotis, 2008; O'Brien et al., 2010). This is consistent with the fact that arginine



SCHEME 10

exists in a non-zwitterionic form when vaporized in a supermolecular beam (Chapo et al., 1998).

Conformers of non-ionized arginine corresponding to the two possible tautomeric forms of the guanidine moiety, R_{ni} and R'_{ni} (Scheme 10), seem to present similar stabilities. In their recent theoretical study, Ling et al. (2006) find that the most stable conformers of neutral arginine pertain generally to the R'_{ni} tautomeric form thus having one imino nitrogen and two NH_2 groups in the guanidine moiety. By contrast, Schlund et al. (2008) considered only R_{ni} structures in their conformational analysis of arginine in the gas phase. We re-examined here, at the G4MP2 level, the most stable conformers presented in both studies (i.e., structures C1–C6 from Ling et al. (2006) and N1–N3 from Schlund et al. (2008), Note that C3 and N1 correspond to the same conformer). Results are reported in Figure 25.

Six structures are predicted to be situated in the first 5 kJ/mol of the G4MP2 H°_{298} scale: structures C1, C2, C4, and C5 identified by Ling et al. (2006) which correspond to R'_{ni} tautomers and structures N1(C3) and N2, pertaining to the R_{ni} tautomeric form, identified by Schlund et al. (2008). Conformers C4 (**ArgIag−g+A**) and C5 (**ArgIag−g+B**) present an *anti* HOCO arrangement and a bifurcated $NH_2 \ldots O{=}C$ hydrogen bonding in the amino acid moiety (type I conformers). Additional stabilization is brought by another hydrogen bond involving the imino nitrogen of the guanidine group and the hydroxyl hydrogen. Conformers C1 (**ArgII$_C$g−g−g+A**) and C2 (**ArgII$_C$g−g−g+B**) are also characterized by a *anti* HOCO arrangement, but, by contrast with C4 and C5, the stabilization in the amino acid part is offered by one hydrogen bond between the hydrogen of the hydroxyl group and the nitrogen atom of α-amino group as typical in type II conformers. Another hydrogen bond involving the imino nitrogen of the guanidine group and one H of the α-amino group participates to the overall stabilization of these two conformers. Conformers C4 and C5 in one hand, and C1 and C2 in the other, differ only in the relative arrangement of the two amino groups of the guanidine moiety. The two R_{ni} tautomers N1(C3) (**ArgII$_C$g−g−g+C**) and N2 (**ArgI′g+g−g+**) are conformers of type II and I, respectively. Their extra stabilization is provided by favorable interactions between the iminogroup of the guanidine moiety and the hydroxyl. When

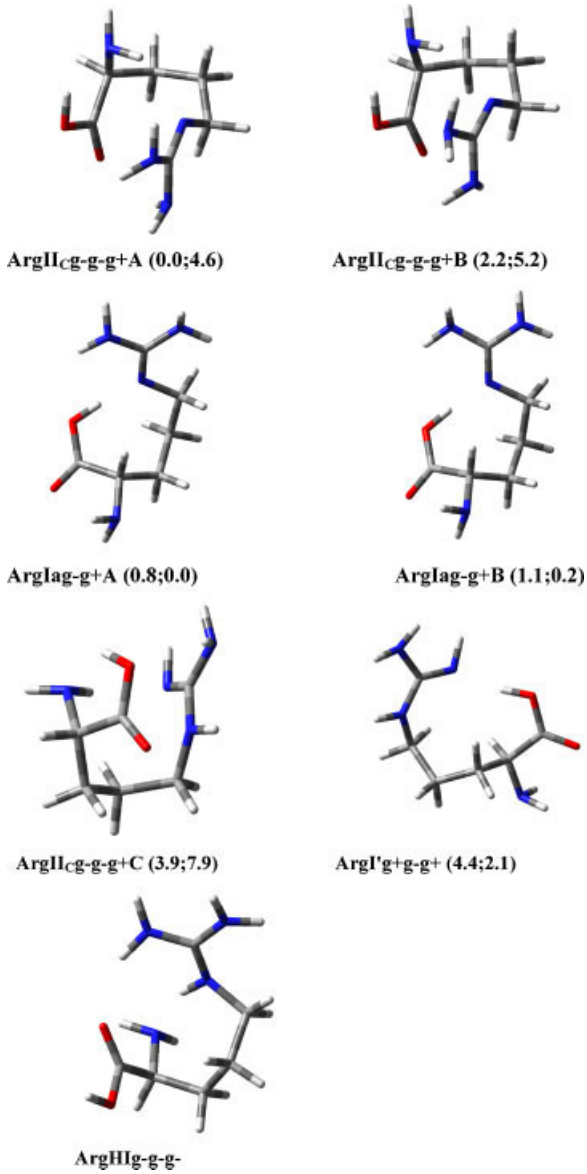


FIGURE 25. Most stable conformations of neutral and protonated arginine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G4MP2 calculations).

looking at the G4MP2 G°_{298} scale, it turns out that C4 and C5, followed by N2 becomes the most stable conformers. The relative population of these three major conformers is 43/39/18.

Infrared spectroscopy of protonated arginine unambiguously demonstrated that protonation occurs on the side chain (Bush et al., 2007b; Forbes et al., 2007). This is in line with a larger basicity of the guanidine moiety (PA(guanidine) = 986 kJ/mol; Hunter & Lias, 1998) with respect to the amino group (PA(hexylamine) = 928 kJ/mol; Hunter & Lias, 1998). The structure of the most stable conformation of protonated arginine reached a large consensus in the community (Maksik & Kovacevic, 1999b; Rak et al., 2001; Bleiholder et al., 2006; Ling et al., 2006; Bush et al., 2007b; Forbes et al., 2007; Bouchoux et al., 2008). This structure (**ArgHIg−gg−**, Fig. 25) is stabilized by two internal

hydrogen bonds, one between the H(N) of the protonated guanidine and the nitrogen atom of the α-amino group, and the second between one H of the NH_2 group of protonated guanidine and the oxygen of the carbonyl function. The second protonated form of arginine in order of increasing 298 K enthalpy or Gibbs free energy is situated ∼8 kJ/mol above **ArgHIg−g+g−** and does not participate significantly to the population of conformers at that temperature (Rak et al., 2001; Ling et al., 2006).

5. *Aromatic Amino Acids*

The presence of an aromatic ring in the side chain of phenylalanine, tyrosine, tryptophan, and histidine is at the origin of stabilizing interactions between polarized hydrogen and the π-electron system both in the neutral and in the protonated forms. Theoretical investigation of the neutral system benzene/NH_3 reveals that the most stable structure is a monodentate complex characterized by a weak binding energy of approximately 7–9 kJ/mol (Tsusuki et al., 2000; Tarakeshwar, Choi, & Kim, 2001). By contrast, the interaction energy calculated between the NH_4^+ cation and aromatic molecules such as benzene, toluene, and indole is in the range 70–100 kJ/mol (Aschi, Mazza, & Di Nola, 2002; Reddy & Sastry, 2005). Experimentally a 298 K enthalpy of complexation between benzene and NH_4^+ equal to 71 kJ/mol has been determined from variable temperature equilibrium constant determinations (Deakyne & Mautner, 1985). The present theoretical picture describes the ammonium–cation π-electrons interaction as essentially governed by electrostatic effect, with significant contribution from the polarization (Aschi, Mazza, & Di Nola, 2002). From this point of view it is worthy to note that the polarizability passes from ∼10 $Å^3$ for the phenyl moiety to ∼15 $Å^3$ for the indole ring, in keeping with the increase of interaction energy given by the calculations. These elements lead to the expectation that $NH_3^+ \ldots \pi$-electrons interaction may be an important clue for the conformational preference and the protonation energetics of aromatic amino acids. From this point of view, it is noteworthy that the structure of protonated serotonin (a homologue of tryptophan) maximizes the intramolecular $NH_3^+ \ldots \pi$ interaction, as demonstrated recently from IRMPD spectroscopy (Lagutschenkov et al., 2010).

a. *Phenylalanine*

Conformational analysis of phenylalanine has been conducted at various correlated levels (Snoek et al., 2000; Lee et al., 2003; Ebata et al., 2006; Huang et al., 2006a; Kaczor et al., 2006; Bouchoux et al., 2009a; Olsztynska-Janus et al., 2009; Baek et al., 2010; Riffet & Bouchoux, 2011). Five conformers were identified in a 5 kJ/mol enthalpy range at the G4MP2 level (Riffet & Bouchoux, 2011). Among them, two pertain to the type I (**$PheI_Cg-$** and **$PheI_Ca$**, Fig. 26), and three to the type II (**$PheII_Ag+$**, **$PheII_Bg-$**, and **$PheII_Bg+$**, Fig. 26) subgroups. These conformers involve the classical $NH_2 \ldots O{=}C$ (type I) or $OH \ldots NH_2$ (type II) single intramolecular hydrogen bonding in the α-amino acid moiety but also additional stabilizing interactions between one H of the amino group and the π-electrons of the aromatic ring (**$PheII_Ag+$**, **$PheII_Bg-$**, and **$PheII_Bg+$**) or between one ortho hydrogen of the aromatic ring and the COOH moiety (**$PheI_Cg+$**, **$PheI_Ca$**, **$PheII_Bg-$**).

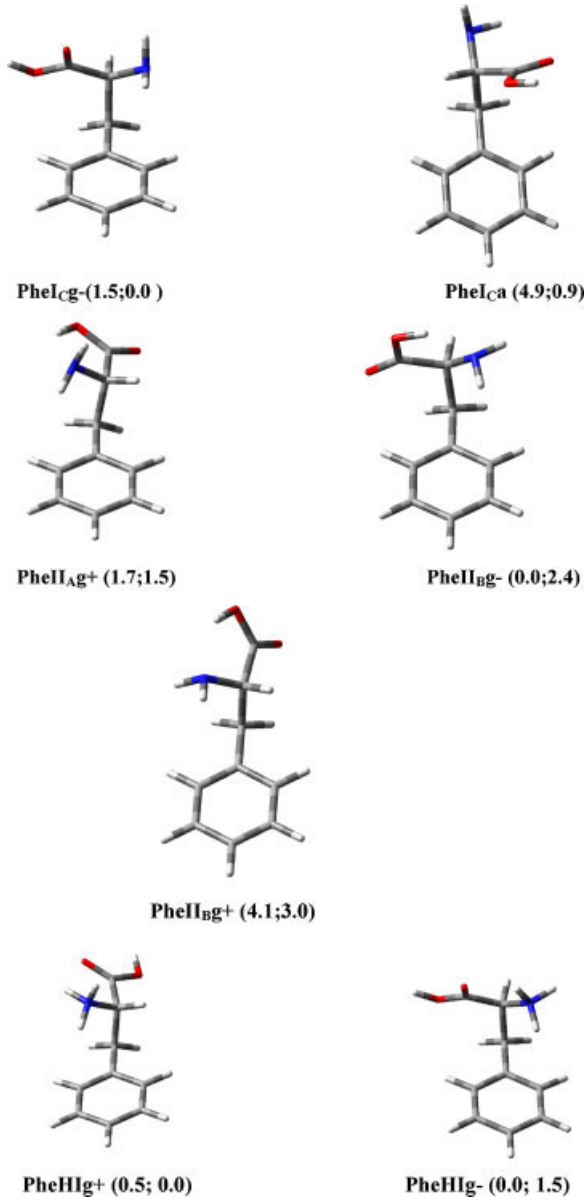


FIGURE 26. Most stable conformations of neutral and protonated phenylalanine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G4MP2 calculations).

When passing to the free energy scale, the situation is more complex since no less than ten conformers are found in the first 5 kJ/mol at the G4MP2 level (Riffet & Bouchoux, 2011). There is however five major conformers which represent ∼75% of the overall population at 298 K, namely **$PheI_Cg-$**, **$PheI_Ca$**, **$PheII_Ag+$**, **$PheII_Bg-$**, and **$PheII_Bg+$**. Their relative populations are predicted to be 25/18/14/10/8%, respectively.

Mass selected resonant two photons ionization on the jet-cooled phenylalanine and UV and IR ion depletion spectra of phenylalanine were interpreted by the existence of population of six conformers (Snoek et al., 2000; Lee et al., 2002b, 2004; Ebata et al., 2006). Comparable conclusions have been deduced from experiments involving vertical ionization energies determinations (Lee et al., 2002a, 2003), fluorescence

lifetime measurement (Hashimoto et al., 2006) and matrix infrared spectroscopy (Kaczor et al., 2006). This proposal provided also a basis for the interpretation of the resonance enhanced multiphoton ionization (REMPI) spectrum of phenylalanine (Cohen et al., 2000). The most recent experimental studies (Lee et al., 2004; Ebata et al., 2006) conclude to the best stability of conformer **PheI$_C$g−** followed by **PheII$_B$g−**, **PheII$_A$g+**, and **PheI$_C$a** as demonstrated by the quantum chemical calculations presented above.

The most stable conformers of protonated phenylalanine, **PheHIg+** and **PheHIg−** (Fig. 26), are of type **HI**, that is, they are stabilized by a (*syn*)HOCO ... H_3N^+ internal hydrogen bond (Stearns et al., 2007; Wu & McMahon, 2008b; Bouchoux et al., 2009a; Riffet & Bouchoux, 2011). What is also evident from examination of Figure 26 is that structures **PheHIg+** and **PheHIg−** are also characterized by a NH_3^+ ... π interaction. It is noteworthy that the structure **PheHIa** is situated ca. 20 kJ/mol above **PheHIg+**, in keeping with the fact that no NH ... π interaction is possible in the *antiperiplanar* arrangement of the NC(2)C(3)C(4) backbone in **PheHIa**. The experimental IRMPD spectrum of protonated phenylalanine has been interpreted recently by the coexistence of both conformers **PheHIg+** and **PheHIg−** (Wu & McMahon, 2008b). Similarly, UV photofragmentation spectra of protonated phenylalanine and their vibronically resolved bands were attributed to the existence of two different stable conformers which present NH_3^+ ... O=C and NH_3^+ ... π interactions as demonstrated for **PheHIg+** and **PheHIg−** (Stearns et al., 2007).

b. Tyrosine

In many aspects, the behavior of tyrosine is comparable to that of phenylalanine. Obviously, since the orientation of the phenol hydroxyl group may take two distinguishable positions, the number of conformers is multiplied by two with respect to phenylalanine. This phenomenon has been claimed to be at the origin of the doublets observed in the REMPI (Cohen et al., 2000; Grace et al., 2002) and laser-induced fluorescence (Inokuchi et al., 2007) spectra of tyrosine. Conformational space of neutral tyrosine has been examined at the B3LYP/6-31++G(d,p) (Ramaekers et al., 2005) (note that a R configuration of the asymmetric carbon has been assumed in this study), B3LYP/6-31++G(d,p) (Bouchoux et al., 2009a), MP2/6-311G(2df,p)//B3LYP/6-311++G(d,p) (Zhang, Huang, & Lin, 2005) and G4MP2 (Riffet & Bouchoux, 2011) levels of theory. Some of the recent data concerning the most stable conformers of tyrosine are presented in Figure 27.

As observed with phenylalanine, conformers of type **I$_C$g−** and **I$_C$a** in both their "right" **(r)** and "left" **(l)** arrangements of the OH group are predicted to be the most stable in the Gibbs free energy scale. They are followed by **TyrII$_B$g−(l)**, **TyrII$_A$g+(l)**, **TyrII$_A$g+(l)**, and **TyrII$_B$g−(r)**. This height structures represent more than 70% of the population of conformers at 298 K (G4MP2 data). From an enthalpic point of view, conformers **TyrII$_B$g−(l)** and **TyrII$_B$g−(r)** become slightly favored.

The most stable protonated forms of tyrosine, **TyrHIg+** and **TyrHIg−** (both **(l)** and **(r)**), presents the usual *syn* HOCO arrangement, and two single NH ... CO bond and a NH ... π interactions. They fall in a very small energy (H°_{298} or G°_{298}) range of ∼1 kJ/mol while structures of type **HII** are well above by ca 15 kJ/mol. The existence of four stable conformers of protonated tyrosine has been demonstrated experimentally by Rizzo and co-workers (Boyarkin et al., 2006a; Stearns et al., 2007) from vibronically resolved bands of UV photofragmentation spectra. Similarly, photoexcitation spectrum of protonated tyrosine has been interpreted by taking into account these four conformers (Kwon et al., 2009).

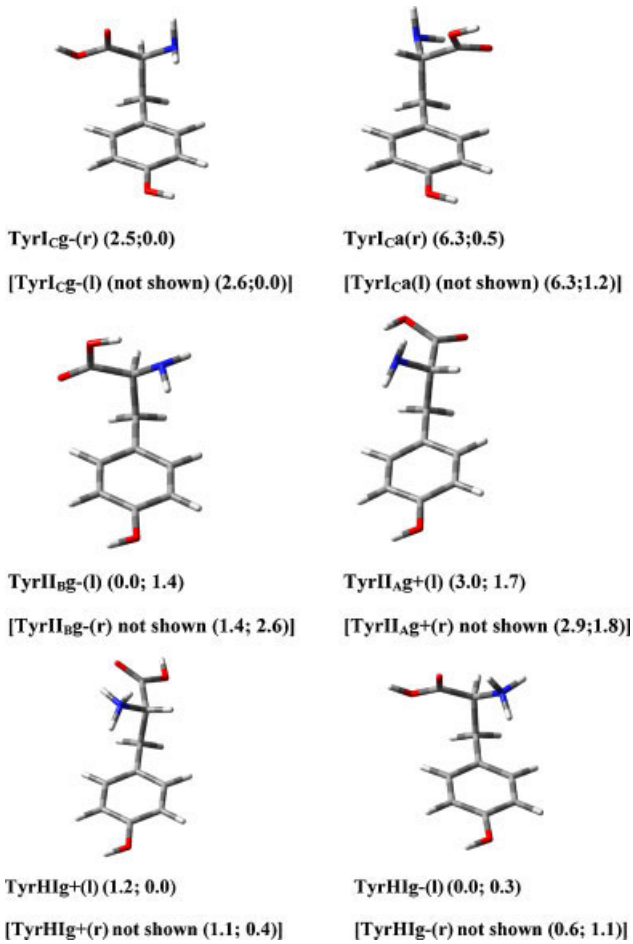


FIGURE 27. Most stable conformations of neutral and protonated tyrosine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G4MP2 calculations).

c. Tryptophan

Thorough investigations of the conformational landscape of neutral tryptophan has been done by Simons and co-workers (Snoek et al., 2001) (B3LYP/6-31+G(d) and MP2/6-311+G(d,p)//B3LYP/6-31+G(d) calculations), Jarrold and co-workers (Compagnon et al., 2001) (B3LYP/6-31G(d) and MP2/6-31G(d) calculations), Huang & Lin (2005) (B3LYP/6-311G(d) and MP2/6-311++G(d,p) calculations), Kaczor et al. (2007) (B3LYP/6-311++G(d,p) level) and Bouchoux, Bourcier, and Riffet (2011) (CBS-QB3). Up to 17 conformers have been identified in the 10 kJ/mol energy range relative to the most stable (Kaczor et al., 2007). According to the used theoretical level, different orders of stabilities have been reported for the tryptophan conformers. However, all the above-mentioned studies agree to assign the same structure to the most stable conformer: **TrpII$_C$g−(r)** (Fig. 28). This type

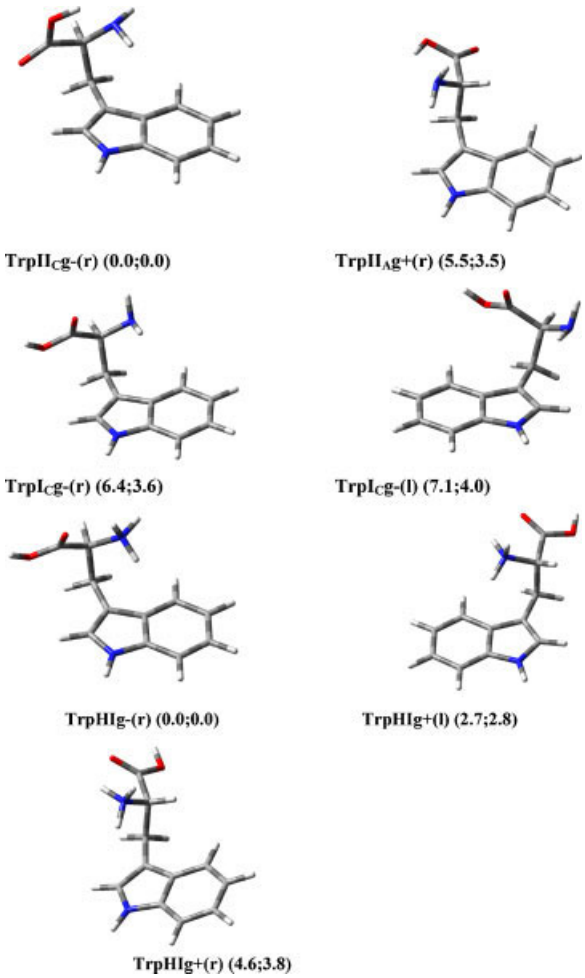


FIGURE 28. Most stable conformations of neutral and protonated tryptophan (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, CBS-QB3 calculations).

II conformation is stabilized not only by the archetypal single intramolecular OH ... NH_2 hydrogen bond but also by further favorable interactions between one of the amino hydrogen atoms and the π-electron system of the aromatic ring and between the carbonyl oxygen and a neighboring H atom of the indole ring (Bouchoux, Bourcier, & Riffet, 2011). Calculation of the conformational distribution based on the MP2/6-311++G(d,p) (Huang & Lin, 2005) or CBS-QB3 (Bouchoux, Bourcier, & Riffet, 2011) free energies shows that the most stable conformer, **TrpII$_{Cg}$−(r)**, constitutes ca. 50% of the distribution at 298 K while consideration of the six most stable conformers covers the conformational distribution to approximately 90%. It is also concluded by Huang & Lin (2005) that only one conformer **TrpII$_{Cg}$−(r)** is present (to more than 99%), at 85 K. This confirm the conclusion drawn by Compagnon et al. (2001) from measurements of permanent electric dipole moment, that only one conformation of tryptophan is dominant at this temperature. A larger number of conformers have been detected by Piuzzi et al. (2000) after laser vaporization of tryptophan thus pointing for slightly more energetic conditions. The resonantly enhanced two-photon ionization and fluorescence spectra of tryptophan have been similarly interpreted by the occurrence of several conformers in cold supersonic molecular beam (Rizzo et al., 1986a; Rizzo, Park, & Levy, 1986b).

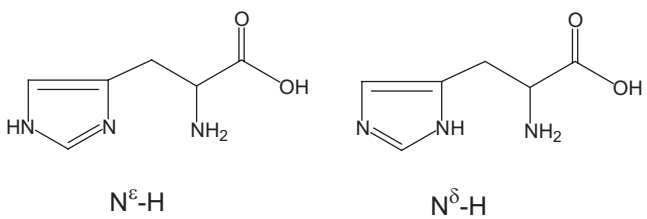
Protonation of tryptophan occurs preferentially on the amino group. However three conformers fall in the 8 kJ/mol Gibbs free energy range. The most stable, **TrpNHIg−(r)** (Fig. 28) represents more than 60% of the equilibrium mixture at 298 K (CBS-QB3 calculations; Bouchoux, Bourcier, & Riffet, 2011). This structure is characterized by a single internal CO ... HNH_2^+ hydrogen bond accompanied by a second favorable interaction between another H of the NH_3^+ group and the π-electrons of the indole moiety (Fig. 28). This interaction is stronger in Trp than in phenylalanine and tyrosine. Protonation of tryptophan has been sometime suspected to occur on the aromatic ring (Harrison, 1997; Rogalewicz, Hoppilliard, & Ohanessian, 2000). The preferred protonation on the carbon C(3) of the indole molecule (Somers et al., 2004) led us to examine the possibility of an *ipso* protonation of the tryptophan molecule. At the B3LYP/6-31+G(d,p) level, the resulting structures are situated ∼40 kJ/mol above **TrpNHIg−(r)** and are thus probably not involved during low energy protonation processes (Riffet & Bourcier & Bouchoux, 2011). The unfavorable protonation ability of the indole nitrogen has been also confirmed in the tryptophan molecule since an energy difference of 95 kJ/mol has been calculated (B3LYP/6-31G(d) calculations) between **Trp$_I$H$^+$** and the most stable protonated form of tryptophan, **TrpNHIg−(r)** (Lioe et al., 2004; Riffet & Bouchoux, 2011).

Electronic spectra of gaseous protonated tryptophan have been experimentally studied in the recent years (Kang et al., 2005a,b; Talbot et al., 2005; Boyarkin et al., 2006b; Mercier et al., 2006; Fujihara et al., 2008). Seemingly, interpretation of the experimental data was based on high energy conformers and should be consequently re-examined.

d. Histidine

The imidazole group of the histidine side chain may present two tautomeric forms, N$^{\varepsilon}$–H and N$^{\delta}$–H (Scheme 11). It is well known that the two tautomers are in equilibrium in solution, but, at the same time, the amino acid moiety is its zwitterionic form (Farr-Jones et al., 1993). According to the most recent calculations (Kovacevic et al., 2005; Huang, Yu, & Lin, 2006b; Huang, Lin, & Song, 2007; Tehrani, Tavasoli, & Fattahi, 2010; Bouchoux & Riffet, 2011b) the N$^{\varepsilon}$–H and N$^{\delta}$–H tautomers seems of comparable stability in the gas phase although a major contribution of conformers of type N$^{\varepsilon}$–H is evidenced.

By analogy with the other aromatic amino acids, it may be expected that the preferred conformations of histidine will



SCHEME 11

contain the *anti* HOCO arrangement associated with an intramolecular simple hydrogen bond OH ... NH_2 allowing the development of a system of cooperative hydrogen bonds with the aromatic ring. According to the tautomeric form considered, this hydrogen bond may be anticipated between one aromatic H and the oxygen of the carbonyl (N^{δ}–H form) or between one H of the amino acid group and the imidazole (N^{ε}–H form). Four important studies have been devoted to the structures and conformations of neutral histidine (Kovacevic et al., 2005; Huang, Yu, & Lin, 2006b; Huang, Lin, & Song, 2007; Tehrani, Tavasoli, & Fattahi, 2010), a general comparison at the G4 level is under investigation and the most salient results are presented here (Bouchoux & Riffet, 2011b).

At the G4MP2 level, the most stable conformer of neutral histidine is the structure **His$^{\varepsilon}$II$_B$g−(r)** (Fig. 29). This N^{ε}–H tautomer is stabilized by a OH ... NH_2 hydrogen bond

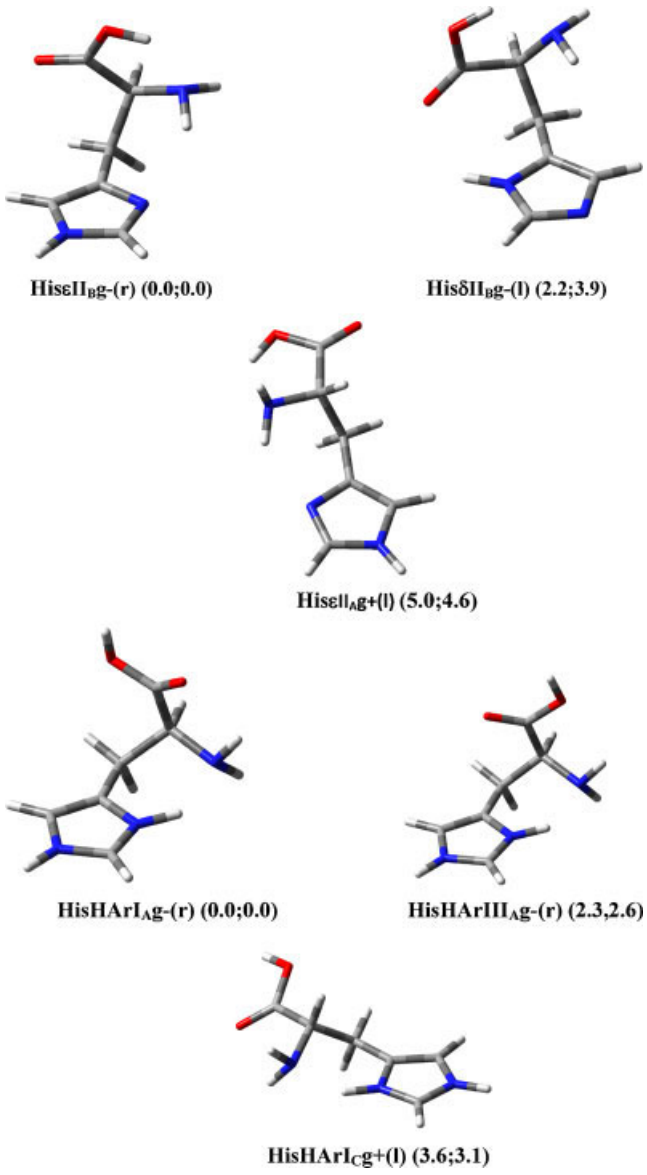


FIGURE 29. Most stable conformations of neutral and protonated histidine (in parentheses: ΔH°_{298}; ΔG°_{298} in kJ mol^{-1}, G4MP2 calculations).

(type II conformer) plus a NH_2 ... N^{δ} interaction and, but to a lesser extend, a C=O ... HC(imidazole) interaction. The structure **His$^{\delta}$IIg−(l)** (Fig. 29) presents noticeable interactions between the acidic NH of the heterocycle and the carbonyl oxygen. At the G4MP2 level **His$^{\delta}$IIg−(l)** is slightly less stable than its tautomer **His$^{\varepsilon}$II$_B$g−(r)** (by 2.2 and 3.9 kJ/mol in H°_{298} and G°_{298}, respectively, Fig. 29). Note that **His$^{\varepsilon}$II$_B$g−(r)** has been identified as the most stable in the N^{ε}–H series (Huang, Yu, & Lin, 2006b; Wang, Ohanessian, & Wesdemiotis, 2008; Tehrani, Tavasoli, & Fattahi, 2010) but that **His$^{\delta}$IIg−(l)** was retained as the most stable in most studies (Kovacevic et al., 2005; Bleiholder et al., 2006; Huang, Lin, & Song, 2007; Dunbar et al., 2009). The two following structures by order of increasing energy are **His$^{\varepsilon}$II$_A$g+(l)** and **His$^{\varepsilon}$II$_C$a(r)** (Fig. 29). These four conformers represent more than 85% of the overall population at 298 K, their relative proportions **His$^{\delta}$IIg−(l)/His$^{\varepsilon}$IIg−(r)/His$^{\varepsilon}$II$_A$g+(l)** and **His$^{\varepsilon}$II$_C$a(r)** are 69/14/11/6% (G4MP2 calculations; Riffet & Bouchoux, 2011b).

The most favorable protonation site of histidine appears to be the imino nitrogen of the aromatic ring. Obviously, protonation of both N^{ε}–H and N^{δ}–H tautomers in those positions would produce the same ionized form. The most stable conformer of these aromatic protonated forms of histidine is structure **HisHArI$_A$g−(r)** (Fig. 29), as revealed by MP2/6-311+G(d) (Huang, Lin, & Song, 2007) or G4MP2 calculations (Riffet & Bouchoux, 2011b). Two other conformers, **HisHArIII$_A$g−(r)** and **HisHArI$_C$g+(l)** (Fig. 29), has been predicted by theory to be ∼3 kJ/mol above **HisHArI$_A$g−(r)** (Kovacevic et al., 2005; Huang, Lin, & Song, 2007; Riffet & Bouchoux, 2011b). Protonation of N^{ε}–H neutral histidine on the amino nitrogen leads to species situated ∼8 kJ/mol above **HisHArI$_A$g−(r)** in the G°_{298} scale (Huang, Lin, & Song, 2007; Riffet & Bouchoux, 2011b). Protonation of N^{δ}–H tautomers would produce protonated forms of too high energies to be considered as accessible in most experiments. The population of conformers **HisHArI$_A$g−(r)/HisHArIII$_A$g−(r)/HisHArI$_C$g+(l)** is calculated to be equal to 61/21/18% (G4MP2 calculations).

6. Computed Protonation Thermochemistry of Amino Acids

Table 2 gathers the computed proton affinity values of the 20 naturally occurring α-amino acids, based on the neutral and protonated conformers of lowest H°_{298} (''most stable conformers'' proton affinity, PA_{msc}, see Eq. 9). The analogous quantities GB_{msc}, computed using the most stable conformers in the G°_{298} scale, are also reported in Table 2.

Examination of Table 2 reveals that comparable PA_{msc}(M) are obtained when using composite methods Gn or CBSn. Small differences arise between computed proton affinities for several amino acids mainly because the considered neutral and (or) protonated structures were not really the most stable conformers. The more secure PA_{msc}(M) values, that is, computed at the highest composite level G3 or G4 and considering the ascertained most stable conformers, are noted in bold.

Consideration of the population of conformers at 298 K for both neutral and protonated species lead to averaged proton affinities, $PA_{average}$, protonation entropies, $\Delta_p S^{\circ}_{average}$, and gas phase basicities, $GB_{average}$ (see Eqs. 10–12). These data are reported in Table 3. The quantities $\Delta_p S^{\circ}_{average}$, and

TABLE 2. Computed gas-phase protonation thermochemistry of α-amino acids: monoconformer GB_{msc}, PA_{msc}, and $\Delta_p S^\circ_{msc}$ values*

M	GB_{msc}(M) kJ.mol^{-1}	PA_{msc}(M) kJ.mol^{-1}	$\Delta_pS°_{msc}$(M) J.mol^{-1}.K^{-1}
Glycine	854.9[r]	**887.6**[r]	-0.9[r]
	855.8[o]	**886.6**[o]	
		890.0[d]	
		892.4[k]	
		895.0[b]	
		886.5[a]	
		886.6(884.9)[l]	
Alanine	869.7[r]	**902.2**[r]	-0.5[r]
	870.4[o]	**901.6**[o]	
		904.7[k]	
		901.2[a]	
		904[g]	
		900.0(902.1)[l]	
Valine	883.6[r]	**915.0**[r]	3.5[r]
	884.0[o]	**914.1**[o]	
		916.5[k]	
		906.2[a]	
		912.1(915.9)[l]	
Leucine	883.2[r]	**915.8**[r]	-0.7[r]
	882.9[o]	**914.1**[o]	
		919.1[k]	
		912.9[a]	
		912.9(915.9)[l]	
Isoleucine	886.3[r]	**919.3**[r]	-1.9[r]
	887.8[o]	**918.3**[o]	
		920.5[k]	
		911.6[a]	
		916.3(920.5)[l]	
Proline	909.9[r]	**941.4**[r]	3.0[r]
	909.6[o]	941.6[o]	
		942.6[k]	
		934.2[a]	
		941.8[f]	
		945.2(946.0)[l]	
Serine	879.9[s]	**912.7**[s]	-5.1[s]
	880.0[o]	**912.5**[o]	
		917.1[k]	
		909.5[a]	
		911.3[h]	
		910.9(911.3)[l]	
Threonine	886.9[s]	**918.6**[s]	2.7[s]
	887.4[o]	918.0[o]	
		920.4[k]	
		915.8[a]	
		917.1(921.3)[l]	

TABLE 2. (*Continued*)

Cysteine	872.1[s]	**903.6**[s]	3.2[s]
	873.7[o]	902.1[o]	
		902.0[k]	
		897.8[a]	
		902.9(905.0)[l]	
Methionine	900.3[s]	**938.8**[s]	-19.5[s]
	901.7[o]	**936.4**[o]	
		942.2[k]	
		933.4[a]	
		937.2[n]	
		939.3(941.8)[l]	
Aspartic Acid	883.7[t]	**918.4**[t]	-8.4[t]
	884.3[o]	916.3[o]	
		920.4[k]	
		915.0[a]	
		915.5(920.1)[l]	
Asparagine	906.4[t]	**938.3**[t]	-9.2[t]
	904.4[o]	**936.0**[o]	
		936.7[k]	
		939.7[a]	
		931.8(939.3)[l]	
Glutamic acid	909.1[u]	**948.0**[u]	-21.5[u]
	908.5[o]	**947.7**[o]	
		946.3[k]	
		951.0[a]	
		946.8(948.1)[l]	
Glutamine	937.5[u]	**971.7**[u]	-9.6[u]
	933.6[o]	971.5[o]	
		970.6[k]	
		980.7[a]	
		973.2(984.1)[l]	
Arginine	1012.1[t]	**1045.0**[t]	-4.2[t]
	1013.6[o]	1046.4[o]	
		1053.0[k]	
		1045.9[a]	
		1042[e,m]	
		1072.4[j]	
		1053[n]	
		1044.3(1057.7)[l]	
Lysine	955.6[u]	**991.6**[u]	-44.3[u]
	958.5[o]	994.5[o]	
		998.2[k]	
		1000.8[a]	
		1003[c]	
		1004.2(997.5)[l]	
		993.3[p]	

TABLE 2. (*Continued*)

Phenylalanine	893.3[t]	**926.3**[t]	-5.2[t]
	894.9[o]	923.4[o]	
		925.8[k]	
		940.1[a]	
		933.5(931.8)[l]	
Tyrosine	896.6[t]	**928.9**[t]	-4.0[t]
	894.2[o]	926.8[o]	
		927.5[k]	
		934.2[a]	
		936.4(935.1)[l]	
Histidine	946.8[t]	**979.2**[t]	0.5[t]
	946.2[o]	**978.6**[o]	
		976.1[k]	
		975.7[i]	
		967.3[a]	
		979.1(989.1)[l]	
		967.3[q]	
Tryptophan	910.2[t]	**942.5**[t]	0.8[t]
	909.0[o]	**940.1**[o]	
		942.6[k]	
		929.2[a]	
		946.0(946.4)[l]	

*Monoconformers gas phase basicities and proton affinities are calculated using the most stable G°_{298} and H°_{298} conformers, respectively. $\Delta_p S^{\circ}_{msc}$ is estimated using the most stable conformer in the G°_{298} scale.

a: Maksic & Kovacevic (1999a), MP2/6-311+G(d,p)//HF/6-31G(d)+ZPE calculations, anchored here to PA(Gly) = 886.5 kJ/mol. b: Uggerud (1997), G2(MP2) calculation. c: Kinser, Nicol, & Ridge (2002) calculations. d: Bouchoux & Chia (2009). e: Maksik & Kovacevic (1999b), MP2/6-311+G(d,p)//MP2/6-31G(d) calculations. f: Lemoff, Bush, & Williams (2005) B3LYP/6-31++G(d,p) corrected to 298 K by adding 6.2 kJ/mol. g: Abirami et al. (2005), B3LYP/6-311+G(3df,2p)/B3LYP/6-31+G(d)+ZPE corrected to 298 K by adding 6.2 kJ/mol. h: Miao et al. (2005), B3LYP/6-311+G(d,p) corrected to 298 K by adding 6.2 kJ/mol. i: Kovacevic et al. (2005). j: Rak et al. (2001), CCSD/6-31++G(d,p)//MP2/6-31++G(d,p) calculations. k: From Bleiholder et al. (2006). PA(0 K) corrected to 298 K values by (i) the addition of 6.2 kJ/mol and (ii) by correction to the most stable conformer enthalpy at 298 K calculated at the G3MP2B3 level. l: Dinadayalane, Sastry, & Leszczynski (2006), MP2/6-311+G(d,p)//B3LYP/6-31G(d) and, in parentheses, B3LYP/6-311+G(d,p)//B3LYP/6-31G(d) calculations. m: Bouchoux et al. (2008), B3LYP/6-311+G(3df,2p)//B3LYP/6-31+G(d,p) calculations, anchored to PA(guanidine) = 987.4 kJ/mol. n: Ling et al. (2006), CCSD(T)/6-31++G(d,p)//MP2/6-31++G(d,p) calculations. o: Gronert, Simpson, & Conner (2009), G3MP2 calculations. p: Leng et al. (2008), MP2/6-311G(2df,p)//B3LYP/6-311++G(d,p) calculations. q: Huang, Lin, & Song (2007), MP2/6-311+G(2df,p)//B3LYP/6-311G(d) calculations. r: Bouchoux et al. (2011), G3MP2B3 calculations. s: Riffet, Frison, & Bouchoux (2011), G4 calculations. t: Bouchoux & Riffet (unpublished results), G4MP2 calculations. u: Bouchoux & Riffet (unpublished results), G3MP2B3 calculations.

$GB_{average}$ include the contribution of the entropy of mixing (Eq. 12). However, for information, the values which do not include the mixing correction are indicated in parentheses in Table 3.

Limitations on the G°_{298} and the x_i calculations may be mentioned since it has consequence on the calculation of averaged thermochemical quantities. To fully appreciate these difficulties, it may be emphasized that estimate of x_i with a precision of 10% need the knowledge of relative G_i to ca. 0.5 kJ mol^{-1} at 298 K. In the quantum chemistry suites of programs, standard statistical thermodynamic formulae are used to obtain the electronic, translational, rotational, and vibrational contributions to entropy. The latter terms are estimated using the harmonic oscillator approximation. However, it is known that the lowest frequencies, particularly internal rotations, are generally highly anharmonic and are thus poorly described by the harmonic oscillator approximation. Unfortunately, the lowest frequencies are also those which give the largest contributions to the vibrational entropy. In such situation, a means to more correctly estimate the

TABLE 3. Computed gas-phase protonation thermochemistry of α-amino acids: average thermochemical parameters GB_{av}, PA_{av}, and $\Delta_p S°_{av}$*

M	GB_{av}(M) kJ.mol^{-1}	$GB_{av\ mix}$(M) kJ.mol^{-1}	PA_{av}(M) kJ.mol^{-1}	$\Delta_p S°_{av}$(M) J.mol^{-1}.K^{-1}	$\Delta_p S°_{av\ mix}$(M) J.mol^{-1}.K^{-1}
Glycine[a]	856.2	854.4	888.7	-0.3	-6.3
Alanine[a]	871.6	868.4	903.7	0.9	-9.8
Valine[a]	885.1	881.6	916.3	4.1	-7.6
Leucine[a]	883.6	882.9	916.6	-2.1	-4.3
Isoleucine[a]	887.6	884.8	920.3	-1.0	-10.3
Proline[a]	910.5	908.9	943.1	-0.7	-6.0
Serine[c]	881.4	878.3	914.4	-1.5	-12.1
Threonine[c]	888.1	885.5	920.4	0.7	-8.3
Cysteine[c]	873.0	870.7	906.6	-3.7	-11.4
Methionine[c]	903.1	901.0	941.2	-18.8	-25.7
Aspartic Acid[b]	885.3	881.7	920.8	-9.9	-22.2
Asparagine[b]	908.8	905.5	944.2	-9.9	-21.0
Glutamic acid[a]	910.2	909.3	948.9	-20.8	-23.9
Glutamine[a]	939.4	936.4	976.9	-16.8	-26.9
Arginine[b]	1013.3	1009.6	1046.7	-3.0	-15.2
Lysine[a]	956.4	954.7	995.0	-20.4	-26.2
Phenylalanine[b]	894.7	891.0	929.5	-7.8	-20.3
Tyrosine[b]	898.1	896.1	932.7	-6.8	-18.7
Histidine[b]	947.3	946.3	980.0	-0.5	-3.8
Tryptophan[b]	908.8	912.3	946.3	-5.1	-17.0

*Values computed using a 298 K Boltzmann distribution of conformers (see eq. 10–12). The indice (mix) means that the entropy of mixing has been included in the calculation of $\Delta_p S°_{av}$ and GB_{av}.

a: Bouchoux & Riffet (unpublished results), G3MP2B3 calculations. b: Bouchoux & Riffet (unpublished results), G4MP2 calculations. c: Riffet, Frison, & Bouchoux (2011), G4 calculations.

vibrational entropy is to treat separately each internal rotation in the frame of a hindered rotor model. This correction has been done in a few cases (East & Radom, 1997; Bouchoux et al., 2008; Desaphy, Malosse, & Bouchoux, 2008; Bouchoux et al., 2009a; Bouchoux, Bourcier, & Nacer, 2009b) and is generally not straightforward.

7. Comparison With Experiment and Evaluated Protonation Thermochemistry of Amino Acids

Comparison between experimental and theoretical proton affinities or gas phase basicities is now possible. As developed in the Section II of the present review, severe discrepancies are observed in some of the available experimental data (Table 1). This has been noted particularly for Asn and Gln but large ranges of experimental GB values were also observed for Leu, Pro, Glu, Phe, Tyr, and Trp. The experimental proton affinities and gas phase basicities presented in Table 4 are the averaged PA and GB values presented in Table 1 for Gly, Ala, Val, Leu, Ile, Pro, Ser, Thr, and Cys. For the other amino acids, where generally a protonation entropy is suspected, the retained PA value is that obtained by the extended kinetic method, except for Asn, Gln, and His where these values are clearly too high (indicated in parentheses in Table 4). For these three particular cases, a bracketing of the real PA value is possible using the PA_{app} given by the simple kinetic method. As a result, PA_{exp} values indicated in Table 4 for Asn, Gln, and His, have been taken as the middle of this experimental range, it thus constitutes only rough estimates.

Concerning the computed thermochemical data, we select the values expected to be the most accurate, that is, those obtained by using composite methods (Gn's or CBS's). However, even with these procedures, small deviations may be expected. Accordingly, we observe that $PA(NH_3)$ and $GB(NH_3)$ are systematically overestimated with respect to the presently recognized reference values (Table 5).

A mean to account for these deviations is to correct the computed PA or GB of amino acids by the difference in experimental and computed $PA(NH_3)$ and $GB(NH_3)$. These corrections are included in the "computed" data reported in Table 4.

A second important question concerning the computed thermochemical parameters is the choice between the monoconformer "msc" or the average "ave" PA and GB values (see Section B). As a rule, PA_{msc} are lower than PA_{ave} (compare data presented in Tables 2 and 3). This is due to the larger number of conformers for the neutral species in a given energy range which consequently leads to a larger value for its averaged molar enthalpy. When considering the 20 amino acids, the averaged difference ($PA_{msc} - PA_{ave}$) is equal to -2.3 ± 1.4 kJ/mol. This systematic shift in enthalpy is counterbalanced by an opposite shift in the $-TS$ term due to the entropy of mixing which is also more important for the mixture of neutral conformers. As a result, the ($GB_{msc} - GB_{ave}$) differences become positive, their averaged value being equal to $+1.0 \pm 1.0$ kJ/mol. Clearly, the difference between GB_{msc} and GB_{ave} are limited and, in the average, close to (or lower than) experimental uncertainties. This explains why the "msc" PA and GB values offer generally correct estimates of the gas-phase thermochemical data.

The selected experimental and computed "msc" or "ave" proton affinities presented in Table 4 are graphically compared in Figure 30. It may be seen from examination of this Figure 30 (and consideration of Table 4) that PA_{exp} values are larger than PA_{ave} for Asn, Asp and His by 8.1, 6.7, and 3.9 kJ/mol, respectively, while PA_{exp} is underestimated by 3.5 kJ/mol for Gln (it may be recalled that the PA_{exp} of Table 4 are only rough estimates for Asn, Gln, and His). It is comforting to observe that, excluding Asn, Asp, Gln, and His from the data, excellent linear correlations with a slope close to 1.0 and a negligible ordinate at the origin are obtained when plotting PA_{msc} or PA_{ave} with PA_{exp} (Fig. 30). The best fit is obtained with the averaged proton affinities: $PA_{ave} = -0.4 + 1.0003\ PA_{exp}$ ($Pr = 0.9989$). Using this latter correlation more reliable PA values can be proposed for Asn, Asp, His, and Gln: PA(Asn) = 943.6 kJ/mol, PA(Asp) = 920.2 kJ/mol, PA(His) = 978.9 kJ/mol, PA(Gln) = 976.1 kJ/mol.

Comparison between GB_{ave} (or GB_{msc}) with GB_{exp} (Table 4) leads also to linear correlation as depicted in Figure 31. The largest deviations between GB_{exp} and GB_{ave} occur for Gln (+21.5 kJ/mol), Lys (+6.6 kJ/mol), Ile (−6.6 kJ/mol), and Trp (−6.5 kJ/mol). Excluding Gln from the correlation, a relationship $GB_{ave} = -28.4 + 1.0309\ GB_{exp}$ ($Pr = 0.9950$) is found. Again, approximate GB_{exp} values are used for Asn and Gln in constructing Figure 31, a better choice is probably to adjust their gas-phase basicity values to the linear fit GB_{ave} versus GB_{exp}. This leads to GB(Asn) = 905.0 kJ/mol and GB(Gln) = 934.8 kJ/mol. Improvement could be also proposed for GB_{exp} values associated with large uncertainties such as GB(Lys) = 952.6 kJ/mol, GB(Pro) = 908.1 kJ/mol, GB(Trp) = 909.1 kJ/mol, GB(Phe) = 891.0 kJ/mol, GB(Tyr) = 895.9 kJ/mol. It is noteworthy that these GB values agree closely to the experimental values determined from the extended kinetic method (see Tables 1 and 4).

Significant negative $\Delta_p S^\circ$ are computed (Tables 3 and 4) for Met, Asp, Asn, Glu, Gln, Arg, Lys, Phe, Tyr, and Trp. It is satisfying to observe that negative $\Delta_p S^\circ$ are also experimentally detected for these amino acids (Table 1). Uncertainties on these latter quantities are however too large to allow more detailed comment than this qualitative agreement.

Finally, combining the information gathered in Table 4 and the relevant discussions, a set of evaluated thermochemical values may be proposed (Table 4, last line). Comparisons between these proposed thermochemical parameters and (i) calculated values, and, (ii) data presented in other available compilations is done graphically in Figures 32 and 33. As expected computed "msc" and "average" PA and GB values correlate nicely with the set of presently recommended values (Pr is always better than 0.999). Concerning comparison with previous compilations, it is evident from examination of Figures 30 and 31 that a close correspondence is found with the Harrison's estimates (1997). The only exception is Gln, the basicity of which has been also underestimated in the Hunter & Lias scale (1998) whereas several experimental studies pointed to higher values (Bojesen & Breindahl, 1994; Afonso et al., 2000; Bouchoux et al., 2004). Presently recommended PA and GB values of Asn, Pro, Glu, and His are noticeably higher than that suggested in the Hunter & Lias tabulation.

TABLE 4. Summary of evaluated protonation thermochemistry of α-amino acids.

M	Method	GB(M) kJ.mol^{-1}	PA(M) kJ.mol^{-1}	$\Delta_pS°(M)^d$ J.K^{-1}.mol^{-1}
Glycine	experiment	854.7±3.1	887.4±3.5	2
	computed[e]	853.8/853.3	886.7/887.8	-1/-6
	recommended	852.2[a]	886.5[a]	-6[a]
		853.3[b]	885.8[b]	0[b]
		854[c]	**887**[c]	**0**[c]
Alanine	experiment	869.1±3.7	901.2±1.4	5
	computed[e]	868.6/867.3	901.3/902.8	0/-10
	recommended	867.7[a]	901.6[a]	-5[a]
		869.8[b]	902.3[b]	0[b]
		868[c]	**902**[c]	**0**[c]
Valine	experiment	877.5±2.8	916.1±1.8	-
	computed[e]	882.5/880.5	914.1/915.4	3/-8
	recommended	876.7[a]	910.6[a]	-5[a]
		880.1[b]	912.5[b]	0[b]
		881[c]	**915**[c]	**0**[c]
Leucine	experiment	882.4±5.4	918.1±2.2	-
	computed[e]	882.1/881.8	914.7/915.7	-1/-4
	recommended	880.6[a]	914.6[a]	-5[s]
		884.1[b]	916.5[b]	0[b]
		883[c]	**916**[c]	**0**[c]
Isoleucine	experiment	890.3(±?)	921.1±1.6	-
	computed[e]	885.2/883.7	918.2/919.4	-2/-10
	recommended	883.6[a]	917.4[a]	-5[a]
		889.5[b]	921.9[b]	0[b]
		885[c]	**919**[c]	**0**[c]
Proline	experiment	903.1±5.6	941.4±3.8	-5
	computed[e]	908.8/907.8	940.3/942.2	3/-6
	recommended	886.0[a]	920.5[a]	-7[a]
		905.1[b]	937.5[b]	0[b]
		908[c]	**942**[c]	**0**[c]
Serine	experiment	878.3±2.9	911.6±1.3	-
	computed[e]	878.3/876.7	911.2/912.9	-5/-12
	recommended	880.7[a]	914.6[a]	-5[a]
		874.3[b]	906.7[b]	0[b]
		878[c]	**912**[c]	**0**[c]
Threonine	experiment	888.5(±?)	921.0±2.1	-
	computed[e]	885.7/883.9	917.1/918.9	3/-9
	recommended	888.5[a]	922.5[a]	-5[a]
		893.5[b]	925.9[b]	0[b]
		886[c]	**919**[c]	**0**[c]

TABLE 4. (*Continued*)

Cysteine	experiment	868.8(±?)	903.0±1.5	-
	computed[e]	871.3/869.1	902.1/905.1	3/-12
	recommended	869.3[a]	903.2[a]	-5[a]
		868.9[b]	901.4[b]	0[b]
		870[c]	**903[c]**	**0[c]**
Methionine	experiment	899.6±1.5	(>938.0) 937.5±2.9	-22
	computed[e]	898.8/899.4	937.3/939.7	-19/-26
	recommended	901.5[a]	935.4[a]	-5[a]
		900.6[b]	933.0[b]	0[b]
		899[c]	**938[c]**	**-20[c]**
Aspartic Acid	experiment	881.7±2.7	(>915.1) 926.8±1.8	-36
	computed[e]	882.8/880.8	917.7/920.1	-8/-22
	recommended	875.0[s]	908.9[s]	-5[s]
		879.6[w]	912.1[w]	0 [w]
		882[t]	**920[t]**	**-15[t]**
Asparagine	experiment	(894.6-920.0)	(>930.5-965.2)	(-43)
	computed[e]	905.5/904.6	937.6/943.5	-9/-21
	recommended	891.5[s]	929.0[s]	-17[s]
		898.4[w]	930.8(944.2)[w]	(-42)[w]
		905[t]	**942[t]**	**-15[t]**
Glutamic acid	experiment	(878.9) 904.5±3.0	(>937.4) 945.3±2.8	-28
	computed[e]	908.0/908.2	947.1/948.0	-21/-24
	recommended	879.1[s]	913.0[s]	-5[s]
		910.9[w]	943.3(956.6)[w]	0 (-42)[w]
		908[t]	**947[t]**	**-20[t]**
Glutamine	experiment	(888.2-939.3)	(>956.9-988.1±3.6)	(-55)
	computed[e]	936.4/935.3	970.8/976.0	-10/-27
	recommended	900.0[s]	937.8[s]	-18[s]
		904.6[w]	937.0(950.4)[w]	0 (-42)[w]
		935[t]	**975[t]**	**-25[t]**
Arginine	experiment	1005.1±1.1	(>1039.2) 1045.8±2.7	-28
	computed[e]	1011.2/1008.7	1044.3/1046.0	-4/-15
	recommended	1006.6[s]	1051.0[s]	-40[s]
		1006.3[w]	1038.7(1052.0)[w]	0 (-42)[w]
		1007[t]	**1046[t]**	**-20[t]**
Lysine	experiment	947.0±4.7	(>965.8) 993.9±3.1	-35
	computed[e]	954.5/953.6	990.7/994.1	-44/-26
	recommended	951.0[s]	996.0[s]	-42[s]
		965.2[w]	997.7[w]	0 [w]
		952[t]	**994[t]**	**-30[t]**
Phenylalanine	experiment	893.7±1.7	(>924.4) 931.3±1.1	-23
	computed[e]	892.4/890.1	925.6/928.8	-5/-20
	recommended	888.9[s]	922.9[s]	-5[s]
		893.5[w]	925.9[w]	0 [w]
		892[t]	**930[t]**	**-15**

TABLE 4. (*Continued*)

Tyrosine	experiment	894.6±2.6	(>927.1) 934.2±2.3	-24
	computed[e]	895.7/895.2	928.2/932.0	-4/-20
	recommended	892.1[s]	926.0[s]	-5[s]
		899.7[w]	932.1[w]	0[w]
		895[t]	**933**[t]	**-15**
Histidine	experiment	948.9±4.3	(>974.3-992.0)	(-27)
	computed[e]	945.7/945.4	978.5/979.3	0/-4
	recommended	950.2[s]	988.0[s]	-17[s]
		947.0[w]	979.4(992.8)[w]	0 (-42)[w]
		947[t]	**979**[t]	**0**[t]
Tryptophan	experiment	908.4±2.1	(>943.8) 945.6±2.0	-16
	computed[e]	909.3/907.9	941.8/945.6	1/-17
	recommended	915.0[s]	948.9[s]	-5[s]
		913.1[w]	945.5(958.9)[w]	0 (-42)[w]
		909[t]	**945**[t]	**-10**[t]

a: Evaluated by Hunter and Lias (1998). b: Evaluated by Harrison (1997). c: Proposed evaluation. d: The two $\Delta_p S^\circ$ msc and averaged with mixing are given as msc/average. e: Computed GB and PA are anchored to GB(NH_3) = 819.0 kJ/mol [and PA(NH_3) = 853.6 kJ/mol] (Hunter & Lias, 1998) since G4, G4MP2, G3MP2, and G3MP2B3 methods provide values slightly overestimated by 1.6 [1.5], 0.9 [0.7], 1.5 [1.4], and 1.1 [0.9] kJ/mol, respectively, see Table 5.

IV. CONCLUSION

Protonation thermochemistry of isolated amino acids has been explored experimentally since more than 30 years. In the present article, we tried to critically review the corresponding literature, bearing in mind the difficulties associated with (i) the handling of such involatile and thermally labile molecules, and (ii) the methods of measuring proton affinities and gas-phase basicities (Section II). Computation on neutral and protonated amino acids using high level quantum chemistry methods is detailed in Section III of this review. Accordingly, invaluable structural information is given by the theoretical conformational analysis reported in this section (Figs. 10–29). Moreover, the formidable development of theoretical methods during the last 10 years allows computing thermochemical parameters with a precision close to experiment. The complementarity of these experimental and theoretical approaches has been exploited in this review to suggest a consistent set of recommended GB and PA values. In this regard, gas-phase basicities and proton affinities listed in Tables 1–4 provide the largest comprehensive set of thermochemical data concerning the gas-phase protonation thermochemistry of amino acids to date.

During the examination of this corpus of data, serious difficulties in experimentally evaluating proton affinities of several amino acids have been evidenced. Due to the low volatility and thermal stability of amino acids, several GB values obtained by the equilibrium method appeared to be erroneous by no less than ∼10 kJ/mol (see Ala, Pro, Asn, Glu, Gln, Lys, Phe, and Tyr). The extended kinetic method, expected to

TABLE 5. Proton affinity and gas-phase basicity of ammonia as evaluated by different composite computational methods

Method	PA(NH_3) kJ.mol^{-1}	ΔPA(NH_3) kJ.mol^{-1}	GB(NH_3) kJ.mol^{-1}	ΔGB(NH_3) kJ. mol^{-1}
Reference	853.6	0	819.0	0
G4	855.1	1.5	820.6	1.6
G4MP2	854.3	0.7	919.9	0.9
G3B3	855.7	2.1	921.3	2.3
G3MP2	855.0	1.4	820.5	1.5
G3MP2B3	854.5	0.9	820.1	1.1
CBS-QB3	854.3	0.7	819.8	0.8

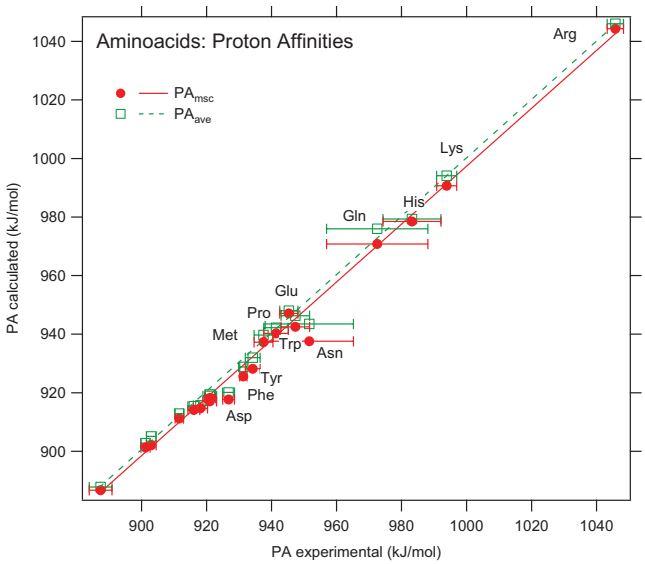


FIGURE 30. Calculated proton affinities (PA_{msc} and PA_{ave}) as a function of experimental averaged PA_{exp} values.

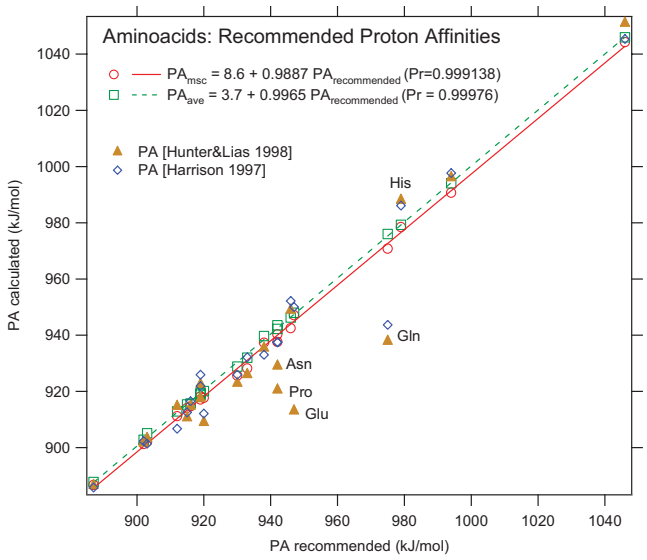


FIGURE 32. Calculated proton affinities (PA_{msc} and PA_{ave}) as a function of the presently recommended PA values (previously recommended data from Hunter & Lias, 1998 and Harrison, 1997 are indicated for comparison).

provide both PA and $\Delta_p S^\circ$, shows also severe discrepancies in several cases. For example, PA determined by this method for Asn, Gln, and His were significantly overestimated and, simultaneously, the $\Delta_p S^\circ$ too negative. However, the gas phase basicities deduced from extended kinetic method measurements are generally correct, there is only three exceptions with Pro, Asn, and Gln. Despite these difficulties a reasonable consensus is appearing for most of the 20 α-amino acids. One reason of this agreement, and a crucial point when considering experimental data, is the re-anchoring of the published data to a common basicity scale. As detailed in Section II, we chose the Hunter & Lias (1998) scale because of its large popularity among the physicochemists community, even though several flaws were identified.

Proton affinities computed at the Gn levels (see Tables 3 and 4) are presently the most accurate and probably the most reliable PA values by comparison with some experiments. Difficulties however remain in theoretically estimating entropies and consequently $\Delta_p S^\circ$ and molar fractions in a population of conformers. As a result, the theoretical GB values may be

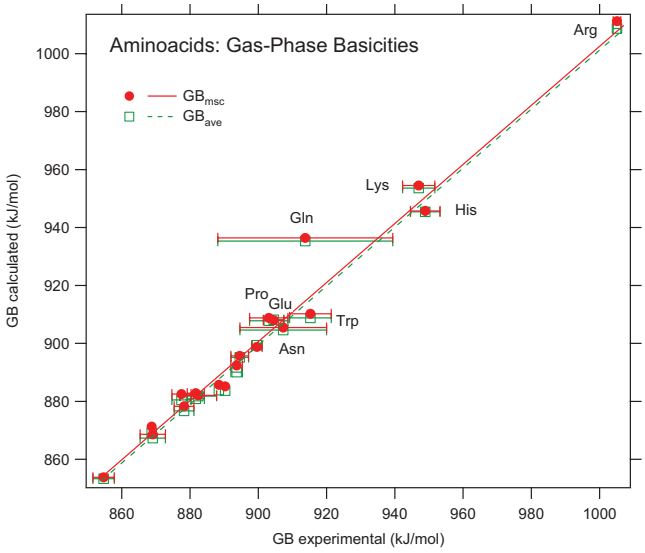


FIGURE 31. Calculated gas-phase basicities (GB_{msc} and GB_{ave}) as a function of experimental averaged GB_{exp} values.

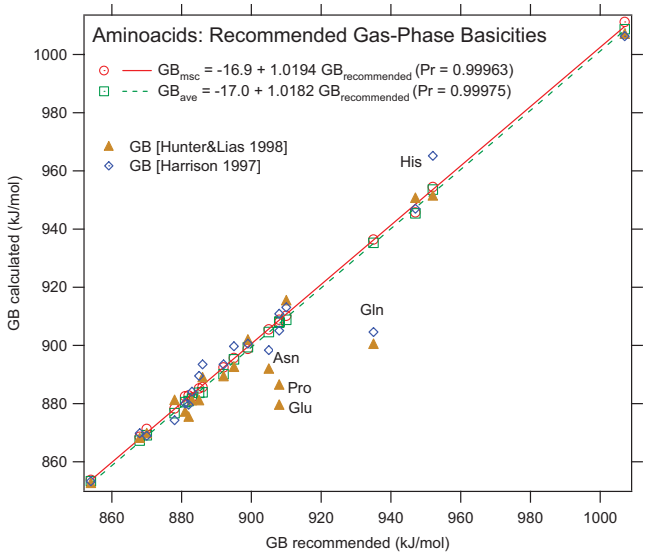


FIGURE 33. Calculated gas-phase basicities (GB_{msc} and GB_{ave}) as a function of the presently recommended GB values (previously recommended data from Hunter & Lias, 1998 and Harrison, 1997 are indicated for comparison).

attached with error up to a few kJ/mol for amino acids having large $\Delta_p S^\circ$, that is, where the side chain may significantly interact with the amino acid moiety.

REFERENCES

Abirami S, Xing YM, Tsang CW, Ma NL. 2005. Theoretical study of α/β-alanine and their protonated/alkali metal cationized complexes. J Phys Chem A 109:500–506.

Afonso C, Modeste F, Breton P, Fournier F, Tabet JC. 2000. Proton affinities of the commonly occurring L-amino acids by using electrospray ionization-ion trap mass spectrometry. Eur J Mass Spectrom 6:443–449.

Aikens CM, Gordon MS. 2006. Incremental solvation of nonioized and zwitterionic glycine. J Am Chem Soc 128:12835–12850.

Alonso JL, Perez C, Sanz ME, Lopez JC, Blanco S. 2009. Seven conformers of L-threonine in the gas phase: A LA-MB-FTMW study. PhysChemChemPhys 11:617–627.

Armentrout PB, Gabriel A, Moision RM. 2009. An experimental and theoretical study of alkali metal cation/methionine interactions. Int J Mass Spectrom 283:56–68.

Armentrout PB, Armentrout EI, Clark AA, Cooper TE, Stennett EMS, Carl DR. 2010. An experimental and theoretical study of alkali metal cation interactions with cysteine. J Phys Chem B 114:3927–3937.

Aschi M, Mazza F, Di Nola A. 2002. Cation-π interactions between ammonium ion and aromatic rings: An energy decomposition study. J Mol Struct 587:177–188.

Aue DH, Bowers MT. 1979. In: Bowers MT, editor. Gas phase ion chemistry, Vol. 2. New York: Academic Press. pp 2–51.

Baek KY, Hayashi M, Fujimura Y, Lin SH, Kim SK. 2010. Investigation of conformation dependent properties of L-phenylalanine in neutral and radical cations by using a density functional taking into account noncovalent interactions. J Phys Chem A 114:7583–7589.

Balabin RM. 2010. Conformational equilibrium in glycine: Experimental jet-cooled Raman spectrum. J Phys Chem Lett 1:20–23.

Balta B, Basma M, Aviyente V, Zhu C, Lifshitz C. 2000. Structures and reactivity of gaseous glycine and its derivatives. Int J Mass Spectrom 201:69–85.

Barone V, Adamo C, Lelj F. 1995. Conformational behavior of gaseous glycine by a density functional approach. J Chem Phys 102:364–370.

Benzi C, Improta R, Scalmani G, Barone V. 2002. Quantum mechanical study of the conformational behavior of proline and 4R-hydroxyproline dipeptide analogue in vacuum and in aqueous solution. J Comput Chem 23:341–350.

Blanco S, Lesarri A, Lopez JC, Alonso JL. 2004. The gas-phase structure of alanine. J Am Chem Soc 126:11675–11683.

Blanco S, Sanz ME, Lopez JC, Alonso JL. 2007. Revealing the multiple structures of serine. Proc Natl Acad Sci USA 51:20183–20188.

Bleiholder C, Suhai S, Paizs B. 2006. Revising the proton affinity scale of the naturally occurring α-aminoacids. J Am Soc Mass Spectrom 17:1275–1281.

Blom MN, Compagnon I, Polfer NC, van Helden G, Meijer G, Suhai S, Paizs B, Oomens J. 2007. Stepwise solvation of an aminoacid: The appearance of zwitterionic structures. J Phys Chem A 111:7309–7316.

Boggs PT, Byrd RH, Rogers JE, Schnabel RB. 1992. ODRPACK version 2.01 Software for weighted orthogonal distance regression. Report NISTIR 92-4834, National Institute of Standards and Technology, Gaithersburg, MD.

Bojesen G. 1987. The order of proton affinities of the 20 common L-α-amino acids. J Am Chem Soc 109:5557–5558.

Bojesen G, Breindahl T. 1994. On the proton affinity of some α-amino acids and the theory of the kinetic method. J Chem Soc Perkin Trans 2:1029–1037.

Bouchonnet S, Hoppilliard Y. 1992. Proton and sodium affinities of glycine and its sodium salt in the gas phase. Ab initio calculations. Org Mass Spectrom 27:71–76.

Bouchoux G. 2006a. Evaluation of the protonation thermochemistry obtained by the extended kinetic method. J Mass Spectrom 41:1006–1013.

Bouchoux G. 2006b. Microcanonical modeling of the thermokinetic method. J Phys Chem A 110:8259–8265.

Bouchoux G. 2007. Gas phase basicities of polyfunctional molecules. Part 1: Theory and methods. Mass Spectrom Rev 26:775–835.

Bouchoux G, Chia RX. 2009. Gas phase basicity of glycine. Croatica Chem Acta 82:47–61.

Bouchoux G, Riffet V. 2011a. Acid-base thermochemistry of gaseous carbonylated aminoacids (Asp, Asn, Glu, Gln). In preparation.

Bouchoux G, Riffet V. 2011b. Acid-base thermochemistry of gaseous histidine. In preparation.

Bouchoux G, Salpin JY. 2003. Gas-phase basicity of glycine, alanine, proline, serine, lysine, histidine and some of their peptides by the thermokinetic method. Eur J Mass Spectrom 9:391–402.

Bouchoux G, Salpin JY, Leblanc D. 1996. A relationship between the kinetics and thermochemistry of proton transfer reactions in the gas phase. Int J Mass Spectrom 153:37–48.

Bouchoux G, Buisson DA, Colas C, Sablier M. 2004. Protonation thermochemistry of α-aminoacids bearing a basic residue. Eur J Mass Spectrom 10:977–992.

Bouchoux G, Desaphy S, Bourcier S, Malosse C, Ngo Biboum Bimbong R. 2008. Gas-phase protonation thermochemistry of arginine. J Phys Chem B 112:3410–3419.

Bouchoux G, Bourcier S, Blanc V, Desaphy S. 2009a. Gas-phase protonation thermochemistry of phenylalanine and tyrosine. J Phys Chem B 113:5549–5562.

Bouchoux G, Bourcier S, Nacer F. 2009b. Gas-phase protonation thermochemistry of glutamic acid. J Phys Chem A 113:6666–6676.

Bouchoux G, Huang S, Inda BS. 2011. Acid-base thermochemistry of gaseous aliphatic α-aminoacids. PhysChemChemPhys 13:651–668.

Bouchoux G, Bourcier S, Riffet V. 2011. Gas-phase basicity and acidity of tryptophan. In preparation.

Boyarkin OV, Mercier SR, Kamariotis A, Rizzo TR. 2006a. Electronic spectroscopy of cold, protonated tryptophan and tyrosine. J Am Chem Soc 128:2816–2817.

Boyarkin OV, Mercier SR, Kamariotis A, Rizzo TR. 2006b. Microsolvation effects on the excited-state dynamics of protonated tryptophan. J Am Chem Soc 128:16938–16943.

Bush MF, Forbes MW, Jockusch RA, Oomens J, Polfer NC, Saykally RJ, Williams ER. 2007a. Infrared spectroscopy of cationized lysine and e-N-methyllysine in the gas phase: Effects of alkali-metal ion size and proton affinity on zwitterion stability. J Phys Chem A 111:7753–7760.

Bush MF, O'Brien JT, Prell JS, Saykally RJ, Williams ER. 2007b. Infrared spectroscopy of cationized arginine in the gas phase: Direct evidence for the transition from nonzwitterionic to zwitterionic structure. J Am Chem Soc 129:1612–1622.

Campbell S, Beauchamp JL, Rempe M, Lichtenberger DL. 1992. Correlations of lone pair ionization energies with proton affinities of amino acids and related compounds. Site specificity of protonation. Int J Mass Spectrom Ion Proc 117:83–99.

Carr SR, Cassady CJ. 1996. Gas phase basicities of histidine and lysine and their selected di- and tri-peptides. J Am Soc Mass Spectrom 7:1203–1210.

Cassady C, Carr SR, Zhang K, Chung-Philips A. 1995. Experimental and ab initio studies on protonations of alanine and small peptides of alanine and glycine. J Org Chem 60:1704–1712.

Chapo CJ, Paul JB, Provencal RA, Roth K, Saykally RJ. 1998. Is arginine zwitterionic or neutral in the gas phase? Results from IR cavity ringdown spectroscopy. J Am Chem Soc 120:12956–12957.

Chen M, Huang Z, Lin Z. 2005. Ab initio studies of gas phase asparagine conformers. J Mol Struct (Theochem) 719:153–158.

Cocinero EJ, Lesarri A, Grabow JU, Lopez JC, Alonso JL. 2007. The shape of leucine in the gas phase. ChemPhysChem 8:599–604.

Cohen R, Brauer B, Nir E, Grace L, de Vries MS. 2000. Resonance enhanced multiphoton ionization spectroscopy of dipeptides. J Phys Chem A 104:6351–6355.

Compagnon I, Hagemeister FC, Antoine R, Rayane D, Broyer M, Dugourd P, Hudgins RR, Jarrold MF. 2001. Permanent electric dipole and conformation of unsolvated tryptophan. J Am Chem Soc 123:8440–8441.

Csaszar AG. 1992. Conformers of gaseous glycine. J Am Chem Soc 114:9568–9575.

Csaszar AG. 1995. On the structure of free glycine and α-alanine. J Mol Struct 346:141–152.

Csaszar AG. 1996. Conformers of gaseous alpha-alanine. J Phys Chem 100:3541–3551.

Curtiss LA, Raghavachari K. 2002. Gaussian-3 and related methods for accurate thermochemistry. Theor Chem Acc 108:61–70.

Curtiss LA, Redfern PC, Raghavachari K. 2007a. Gaussian-4 theory. J Chem Phys 126:084108(1–12).

Curtiss LA, Redfern PC, Raghavachari K. 2007b. Gaussian-4 theory using reduced order perturbation theory. J Chem Phys 127:124105 (1–8).

Czinki E, Csaszar AG. 2003. Conformers of gaseous proline. Chem Eur J 9:1008–1019.

Deakyne CA, Mautner M. 1985. Unconventional ionic hydrogen bonds. 2. NH ... π complexes of onium ions with olefins and benzene derivatives. J Am Chem Soc 107:474–479.

Degtyarenko IM, Jalkanen KJ, Gurtovenko AA, Nieminen RM. 2007. L-Alanine in a droplet of water: A density functional molecular dynamic study. J Phys Chem B 111:4227–4234.

Desaphy S, Malosse C, Bouchoux G. 2008. Gas-phase basicity of methionine. J Mass Spectrom 43:116–125.

Dinadayalane TC, Sastry GN, Leszczynski J. 2006. Comprehensive theoretical study towards the accurate protonaffinity values of naturally occurring amino acids. Int J Quant Chem 106:2920–2933.

Dobrowolski JC, Rode JE, Sadlej J. 2007. Cysteine conformations revisited. J Mol Struct (Theochem) 810:129–134.

Dokmaisrijan S, Lee VS, Nimmanpipug P. 2010. The gas phase conformers and vibrational spectra of valine; leucine and isoleucine: An ab initio study. J Mol Struct (Theochem) 953:28–38.

Drahos L, Vekey K. 2003. Entropy evaluation using the kinetic method: Is it feasible? J Mass Spectrom 38:1025–1042.

Drahos L, Peltz C, Vekey K. 2004. Accuracy of enthalpy and entropy determination using the kinetic method: Are we approaching a consensus? J Mass Spectrom 38:1016–1024.

Dunbar RC, Hopkinson AC, Oomens J, Siu CK, Siu KWM, Steill JD, Verkerk UH, Zhao J. 2009. Conformation switching in gas-phase complexes of histidine with alkaline earth ions. J Phys Chem B 113:10403–10408.

East ALL, Radom ML. 1997. Ab initio statistical thermodynamic model for the computation of third law entropies. J Chem Phys 106:6655–6674.

East ALL, Smith BJ, Radom L. 1997. Entropies and free energies of protonation and proton transfer reactions. J Am Chem Soc 119:9014–9020.

Ebata T, Hashimoto T, Ito T, Inokuchi Y, Altunsu F, Brutschy B, Tarakeshwar P. 2006. Hydration profiles of aromatic amino acids: Conformations and vibrations of L-phenylalanine-(H_2O)n clusters. PhysChemChemPhys 8:4783–4791.

Ervin KM, Armentrout PB. 2004. Systematic and random errors in ion affinities and activation entropies from the extended kinetic method. J Mass Spectrom 39:1004–1015.

Espinoza C, Szczepanski J, Vala M, Polfer NC. 2010. Glycine and its hydrated complexes: A matrix isolation IR study. J Phys Chem A 114:5919–5927.

Ewing NP, Zhang X, Cassady CJ. 1996. Determination of the gas phase basicities of proline and its di- and tri-peptides with glycine: The enhanced basicity of prolylproline. J Mass Spectrom 31:1345–1350.

Falzon C, Wang F. 2005. Understanding glycine confirmation through molecular orbitals. J Chem Phys 123:214307.

Farr-Jones S, Wong WYL, Gutheil WG, Bachovchin WW. 1993. Direct observation of the tautomeric forms of histidine in 15N NMR spectra at low temperatures. Comments on intramolecular hydrogen bonding and on tautomeric equilibrium constants. J Am Chem Soc 115:6813–6819.

Fernandez-Ramos A, Cabaleiro-Lago E, Hermida-Ramon JM, Martinez-Nunez E, Pena-Gallego A. 2000. DFT conformational study of cysteine in gas phase and aqueous solution. J Mol Struct (Theochem) 498:191–200.

Feyer V, Plekan O, Richter R, Coreno M, Prince KC, Carravetta v. 2008. Core level study of alanine and threonine. J Phys Chem A 112:7806–7815.

Forbes MW, Bush MF, Polfer NC, Oomens J, Dunbar RC, Williams ER, Jockusch RA. 2007. Infrared spectroscopy of arginine cation complexes: Direct observation of gas phase zwitterions. J Phys Chem A 111:11759–11770.

Fujihara A, Matsumoto H, Shibata Y, Isnikawa H, Fuke K. 2008. Photodissociation and spectroscopic study of cold protonated dipeptides. J Phys Chem A 112:1457–1463.

Gdanitz RJ, Cardoen W, Windus TL, Simons J. 2004. Very large scale computations of the free energies of eight low lying structures of arginine in the gas phase. J Phys Chem A 108:515–518.

Godfrey PD, Brown RD. 1995. Shape of glycines. J Am Chem Soc 117:2019–2023.

Godfrey PD, Firth S, Hatherley LD, Brown RD, Pierlot AP. 1993. Millimeter wave spectroscopy of biomolecules: Alanine. J Am Chem Soc 115:9687–9691.

Godfrey PD, Brown RD, Rodgers FM. 1996. The missing conformers of glycine and alanine: Relaxation in seeded supersonic jets. J Mol Struct 376:65–81.

Gorman GS, Speir JP, Turner CA, Amster IJ. 1992. Proton affinities of the 20 common α-amino acids. J Am Chem Soc 114:3986–3988.

Grace LI, Cohen R, Dunn TM, Lubman DM, de Vries MS. 2002. The R2PI spectroscopy of tyrosine: A vibronic analysis. J Mol Spectrosc 215:204–219.

Green MK, Lebrilla CB. 1997. Ion molecule reactions as probes of gas phase structures of peptides and proteins. Mass Spectrom Rev 16:53–71.

Gronert S, O'Hair RAJ. 1995. Ab initio studies of amino acid conformations. 1. The conformers of alanine, serine and cysteine. J Am Chem Soc 117:2071–2081.

Gronert S, Simpson DC, Conner KM. 2009. A reevaluation of computed proton affinities for the common α-aminoacids. J Am Soc Mass Spectrom 20:2116–2123.

Hahn IS, Wesdemiotis C. 2003. Protonation thermochemistry of β-alanine. An evaluation of proton affinities and entropies determined by the extended kinetic method. Int J Mass Spectrom 222:465–479.

Harrison AG. 1997. The gas phase basicities and proton affinities of amino acids and peptides. Mass Spectrom Rev 16:201–217.

Hashimoto T, Tahasu Y, Yamada Y, Ebata T. 2006. Anomalous conformer dependent S1 lifetime of L-phenylalanine. Chem Phys Lett 421:227–231.

Heaton AL, Armentrout PB. 2009. Thermodynamic and mechanism of protonated asparagine decomposition. J Am Chem Soc Mass Spectrom 20:852–866.

Heaton AL, Moision RM, Armentrout PB. 2008. Experimental and theoretical studies of sodium cation interactions with the acidic amino acids and their amide derivatives. J Phys Chem A 112:3319–3327.

Heaton AL, Bowman VN, Oomens J, Steill JD, Armentrout PB. 2009. Infrared multiphoton dissociation spectroscopy of cationized

asparagine: Effects of metal cation size on gas-phase conformation. J Phys Chem A 113:5519–5530.

Huang Z, Lin Z. 2005. Detailed ab initio studies of the conformers and conformational distributions of gaseous tryptophan. J Phys Chem A 109:2656–2659.

Huang Z, Yu W, Lin Z. 2006a. Exploration of the full conformational landscape of gaseous aromatic aminoacid phenylalanine: An ab initio study. J Mol Struct (Theochem) 758:195–202.

Huang Z, Yu W, Lin Z. 2006b. First principle studies of gaseous aromatic amino acid histidine. J Mol Struct (Theochem) 801:7–20.

Huang Z, Lin Z, Song C. 2007. Protonation processes and electronic spectra of histidine and related ions. J Phys Chem A 111:4340–4352.

Hunter EP, Lias SG. 1998. Evaluated gas phase basicities and proton affinities of molecules: An update. J Phys Chem Ref Data 27:413–656.

Iijima K, Nakano M. 1999. Reinvestigation of molecular structure and conformation of gaseous l-alanine by joint analysis using electron diffraction data and rotational constants. J Mol Struct 485/486:255–260.

Iijima K, Tanaka K, Onuma S. 1991. Main conformers of gaseous glycine: Molecular structure and rotational barrier from electron diffraction data and rotational constants. J Mol Struct 246:257–266.

Inokuchi Y, Kobayashi Y, Ito T, Ebata T. 2007. Conformation of L-tyrosine studied by fluorescence detected UV-UV and IR-UV double resonance spectroscopy. J Phys Chem A 111:3209–3215.

Isa K, Omote T, Amaya M. 1990. New rules concerning the formation of protonated amino acids from protonated dipeptides using the proton affinity order determined from collisionally activated decomposition spectra. Org Mass Spectrom 25:620–628.

Ivanov AY, Sheina G, Blagoi YP. 1999. FTIR spectroscopic study of the UV-induced rotamerization of glycine in the low temperature matrices (Kr, Ar, Ne). Spcetrochim Acta A 55:219–228.

Jaeger HM, Schaefer HF III, Demaison J, Csaszar AG, Allen WD. 2010. Lowest lying conformers of alanine: Pushing theory to ascertain precise energetic and semiexperimental Re structures. J Chem Theory Comput 6:3066–3078.

Jensen JH, Gordon MS. 1995. On the number of water molecules necessary to stabilize the glycine zwitterion. J Am Chem Soc 117:8159–8170.

Jones CM, Bernier M, Carson E, Colyer KE, Metz R, Wischow E, Webb I, Andriole AJ, Poutsma JC. 2007. Gas phase acidities oft he 20 protein amino acids. Int J Mass Spectrom 267:54–62.

Julian RR, Beauchamp JL, Goddard WA III. 2002. Cooperative salt bridge stabilization of gas phase zwitterions in neutral arginine clusters. J Phys Chem A 106:32–34.

Kaczor A, Reva ID, Proniewicz LM, Fausto R. 2006. Importance of entropy in the conformational equilibrium of phenylalanine: A matrix-isolation infrared spectroscopy and density functional theory study. J Phys Chem A 110:2360–2370.

Kaczor A, Reva ID, Proniewicz LM, Fausto R. 2007. Matrix-isolated monomeric tryptophan: Electrostatic interactions as nontrivial factors stabilizing conformers. J Phys Chem A 111:2957–2965.

Kamariotis A, Boyarkin OV, Mercier S, Beck RD, Bush MF, Williams ER, Rizzo TR. 2006. Infrared spectroscopy of hydrated aminoacids in the gas phase: Protonated and lithiated valine. J Am Chem Soc 128:905–916.

Kang H, Dedonder-Lardeux C, Jouvet C, Gregoire G, Desfrançois C, Schermann JP, Barat M, Fayeton JA. 2005a. Control of bond-cleaving reactions of free protonated tryptophan ion by femtosecond laser pulses. J Phys Chem A 109:2417–2420.

Kang H, Jouvet C, Dedonder-Lardeux C, Martrenchard S, Gregoire G, Desfrançois C, Schermann JP, Barat M, Fayeton JA. 2005b. Ultrafast deactivation mechanisms of protonated aromatic aminoacids following UV excitation. PhysChemChemPhys 7:394–398.

Kapitan J, Baumruk V, Kopecky V Jr, Bour P. 2006. Conformational flexibility of L-alanine zwitterion determines shapes of Raman and Raman optical activity spectral bands. J Phys Chem A 110:4689–4696.

Kinser RD, Nicol G, Ridge DP. 2002. The proton affinity and entropy of protonation of lysinamide. The effects of intramolecular proton solvation. J Phys Chem A 106:9925–9929.

Kovacevic B, Rozman M, Klasinc L, Srzic D, Maksic ZB, Yanez M. 2005. Gas phase structure of protonated histidine and histidine methyl ester: Combined experimental mass spectrometry and theoretical ab initio study. J Phys Chem A 109:8329–8335.

Kuntz AF, Boynton AW, David GA, Colyer KE, Poutsma JC. 2002. The proton affinity of proline analogs using the kinetic method with full entropy analysis. J Am Soc Mass Spectrom 13:72–81.

Kwon JS, Choi CM, Kim HJ, Kim NJ, Jang J, Yang M. 2009. Combined theoretical modeling of photoexcitation spectrum of an isolated protonated tyrosine. J Phys Chem A 113:2715–2723.

Lagutschenkov A, Langer J, Berden G, Oomens J, Dopfer O. 2010. Infrared spectra of protonated neurotransmitters: Serotonin. J Phys Chem A 114:13268–13276.

Lakard B. 2004. Ab initio study of amino acids containing hydroxyl groups (serine, threonine and tyrosine). J Mol Struct (Theochem) 681:183–189.

Lambie B, Ramaekers R, Maes G. 2004. Conformational behavior of serine: An experimental matrix isolation FT-IR and theoretical DFT(B3LYP)/6-31++G** study. J Phys Chem A 108:10426–10433.

Lee KT, Sung J, Lee KJ, Park YD, Kim SK. 2002a. Conformation dependent ionization energies of L-phenylalanine. Angew Chem Int Ed 41:4114–4117.

Lee KT, Sung J, Lee KJ, Kim SK, Park YD. 2002b. Resonant two photon ionization study of jet cooled aminoacid: L-phenylalanine and its monohydrated complex. J Chem Phys 116:8251–8254.

Lee KT, Sung J, Lee KJ, Kim SK, Park YD. 2003. Conformation dependent ionisation of l-phenylalanine: Structures and energetics of cationic conformers. Chem Phys Lett 368:262–268.

Lee Y, Jung J, Kim B, Butz P, Snoek LC, Kroemer RT, Simons JP. 2004. Alanyl side chain folding in phenylalanine: Conformational assignment through ultraviolet rotational band contour analysis. J Phys Chem A 108:69–73.

Lemoff AS, Bush MF, Williams ER. 2005. Structures of cationized proline analogues: Evidence for the zwitterionic form. J Phys Chem A 109: 1903–1910.

Lemoff AS, Bush MF, O'Brien JT, Williams ER. 2006. Structures of lithiated lysine and structural analogues in the gas-phase: Effects of water and proton affinity on zwitterionic stability. J Phys Chem A 110:8433–8442.

Leng Y, Zhang M, Song C, Chen M, Lin Z. 2008. A semi-empirical and ab initio combined approach for the full conformational search of gaseous lysine and lysine-H_2O complex. J Mol Struct (Theochem) 858:52–65.

Lesarri A, Mata S, Cocinero EJ, Blanco S, Lopez JC, Alonso JL. 2002. The structure of neutral proline. Angew Chem Int Ed 43:605–610.

Lesarri A, Cocinero EJ, Lopez JC, Alonso JL. 2004. The shape of neutral valine. Angew Chem Int Ed 43:605–610.

Lesarri A, Sanchez R, Cocinero EJ, Lopez JC, Alonso JL. 2005. Coded amino acids in gas phase: The shape of isoleucine. J Am Chem Soc 127:12952–12956.

Leung K, Rempe SB. 2011. Ab initio dynamics study of glycine intramolecular proton transfer in water. J Chem Phys 122:184506.

Li X, Harrison AG. 1993. A kinetic approach to the proton affinities of amine bases. Org Mass Spectrom 28:366–371.

Lias SG, Liebman JF, Levin RD. 1984. Evaluated gas phase basicities and proton affinities of molecules; heats of formation of protonated molecules. J Phys Chem Ref Data 13:695–808.

Linder R, Seefeld K, Vavra A, Leinermanns K. 2008. Gas phase infrared spectra of nonaromatic amino acids. Chem Phys Lett 453:1–6.

Ling S, Yu W, Huang Z, Lin Z, Haranczyk M, Gutowski M. 2006. Gaseous arginine conformers and their unique intramolecular interactions. J Phys Chem A 110:12282–12291.

Lioe H, O'Hair RAJ, Reid GE. 2004. Gas-phase reactions of protonated tryptophan. J Am Soc Mass Spectrom 15:65–76.

Lioe H, O'Hair RA, Gronert S, Austin A, Reid GE. 2007. Experimental and theoretical proton affinities of methionine, methionine sulfoxide and their N- and C-terminal derivatives. Int J Mass Spectrom 267:220–232.

Locke MJ. 1981. Studies of the low volatility compounds by ion cyclotron resonance mass spectrometry. PhD Thesis. University of California.

Locke MJ, McIver RT Jr. 1983. Effect of solvation on the acid/base properties of glycine. J Am Chem Soc 105:4226–4232.

Locke MJ, Hunter RL, McIver RT Jr. 1979. Experimental determination of the acidity and basicity of glycine in the gas phase. J Am Chem Soc 101:272–273.

Maksic ZB, Kovacevic B. 1999a. Towards the absolute proton affinities of 20 α-amino acids. Chem Phys Lett 307:497–504.

Maksik Z, Kovacevic B. 1999b. Neutral vs zwitterionic form of arginine—An ab initio study. J Chem Soc Perkin 2:2623–2629.

Marino T, Russo N, Tocci E, Toscano M. 2001. Density functional computations of proton affinity and gas-phase basicity of proline. J Mass Spectrom 36:301–305.

Martin JML, de Oliveira G. 1999. Towards standard methods for benchmark quality *ab initio* thermochemistry—W1 and W2 theory. J Chem Phys 111:1843–1856.

Mata S, Vaquero V, Cabezas C, Pena I, Perez C, Lopez JC, Alonso JL. 2009. Observation of two new conformers of neutral proline. PhysChemChemPhys 11:4141–4144.

Mautner M. 2003. The proton affinity scale, and effects of ion structure and solvation. Int J Mass Spectrom 227:525–554.

Mautner M, Sieck LW. 1991. Proton affinity ladders from variable temperature equilibrium measurements 1. A reevaluation of the upper proton affinity range. J Am Chem Soc 113:4448–4460.

Mautner M, Hunter EP, Field FH. 1979. Ion thermochemistry of low volatility compounds in the gas phase.1. Intrinsic basicities of α-aminoacids. J Am Chem Soc 101:686–689.

McLain SE, Soper AK, Terry AE, Watts A. 2007. Structure and hydration of L-proline in aqueous solutions. J Phys Chem B 111:4568–4580.

Mezzache S, Afonso C, Pepe C, Karoyan P, Fournier F, Tabet JC. 2003. Proton affinity of proline and modified prolines using the kinetic method: Role of the conformation investigated by ab initio calculations. Rapid Commun Mass Spectrom 17:1626–1632.

Mezzache S, Pepe C, Karoyan P, Fournier F, Tabet JC. 2005a. Proton affinity of diastereoisomers of modified prolines using the kinetic method and density functional theory calculations: Role of the cis/trans substituent on the endo/exo ring conformation. Rapid Commun Mass Spectrom 19:2279–2283.

Mezzache S, Bruneleau N, Vekey K, Afonso C, Karoyan P, Fournier F, Tabet JC. 2005b. Improved proton affinity measurement for proline and modified prolines using triple quadrupole and ion trap mass spectrometers. J Mass Spectrom 40:1300–1308.

Miao R, Jin C, Yang G, Hong J, Zhao C, Zhu L. 2005. Comprehensive density functional theory study on serine and related ions in the gas phase: Conformations, gas phase basicities and acidities. J Phys Chem A 109:2340–2349.

Miller TF, Clary DC. 2004. Quantum free energies of the conformers of glycine on an ab initio potential energy surface. PhysChemChemPhys 6:2563–2571.

Mirza SP, Prabhar S, Vairamani M. 2001. Estimation of proton affinity of proline and tryptophan under electrospray ionization conditions using the extended kinetic method. Rapid Commun Mass Spectrom 15:957–962.

Montgomery JA Jr, Frisch MJ, Ochterski JW, Petersson GA. 1999. A complete basis set model chemistry. VI. Use of density functional geometries and frequencies. J Chem Phys 110:2822–2827.

Mullin JM, Gordon MS. 2009. Alanine: Then there was water. J Phys Chem B 113:8657–8669.

Noguera M, Rodriguez-Santiago L, Sodupe M, Bertran J. 2001. Protonation of glycine, serine and cysteine. Conformations, proton affinities and intrinsic basicities. J Mol Struct (Theochem) 537:307–318.

O'Brien JT, Prell JS, Berden G, Oomens J, Williams ER. 2010. Effects of anions on the zwitterions stability of Glu, His and Arg investigated by IRMPD spectroscopy and theory. Int J Mass Spectrom 297:116–123.

O'Hair RAJ, Broughton PS, Styles ML, Frink BT, Hadad CM. 2000. The fragmentation pathways of protonated glycine: A computational study. J Am Soc Mass Spectrom 11:687–696.

Olsztynska-Janus S, Szymborska K, Komorowska M, Lipinski J. 2009. Conformational changes of L-phenylalanine—Near infrared-induced mechanism of dimerization: B3LYP studies. J Mol Struct (Theochem) 911:1–7.

Pacios LF, Galvez O, Gomez PC. 2001. Intramolecular effects and relative stabilities of conformers of gaseous glycine. J Phys Chem A 105:5232–5241.

Pecul M. 2006. Conformational structures and optical rotation of serine and cysteine. Chem Phys Let 418:1–10.

Pecul M, Ruud K, Rizzo A, Helgaker T. 2004. Conformational effect on the optical rotation of alanine and proline. J Phys Chem A 108:4269–4276.

Piuzzi F, Dimicoli I, Mons M, Tardivel B, Zhao Q. 2000. A simple laser vaporization source for thermally fragile molecules coupled to a supersonic expansion: Application to the spectroscopy of tryptophan. Chem Phys Lett 320:282–288.

Plekan O, Feyer V, Richter R, Coreno M, de Simone M, Prince KC, Carravetta V. 2007. Investigation of the aminoacids glycine, proline and methionine by photoemission spectroscopy. J Phys Chem A 111:10998–11005.

Rak J, Skurski P, Simons J, Gutowski M. 2001. Low energy tautomers and conformers of neutral and protonated arginine. J Am Chem Soc 123:11695–11707.

Ramaekers R, Pajak J, Rospenk M, Maes G. 2005. Matrix isolation FT-IR spectroscopic study and theoretical DFT(B3LYP)/6-31++G** calculations of the vibrational and conformational properties of tyrosine. Spectrochim Acta A 61:1347–1356.

Ramek M, KeltererR A-M, Nikoli S. 1997. Ab initio and molecular mechanics conformational analysis of neutral L-proline. Int J Quant Chem 65:1033–1045.

Reddy AS, Sastry GN. 2005. Cation [M=H^+, Li^+, Na^+, K^+, Ca^{2+}, Mg^{2+}, NH_4^+, NMe_4^+] interactions with the aromatic motifs of naturally occurring aminoacids: A theoretical study. J Phys Chem A 109:8893–8903.

Riffet V, Bouchoux G. 2011. Acid-base thermochemistry of gaseous aromatic aminoacids (Phe, Tyr). In preparation.

Riffet V, Frison G, Bouchoux G. 2011. Acid-base thermochemistry of gaseous oxygen and sulfur substituted aminoacids (Ser, Thr, Cys, Met). PhysChemChemPhys, in press.

Rizzo TR, Park YD, Peteanu LA, Levy DH. 1986a. The electronic spectrum of the amino acid tryptophane in the gas phase. J Chem Phys 84:2534–2541.

Rizzo TR, Park YD, Levy DH. 1986b. Dispersed fluorescence of jet-cooled tryptophane: Excited state conformers and intramolecular exciplex formation. J Chem Phys 85:6945–6951.

Rogalewicz F, Hoppilliard Y. 2000. Low energy fragmentation of protonated glycine. An ab initio theoretical study. Int J Mass Spectrom 199:235–252.

Rogalewicz F, Hoppilliard Y, Ohanessian G. 2000. Fragmentation mechanisms of α-amino acids protonated under electrospray ionization: A collisional activation and ab initio study. Int J Mass Spectrom 195/196:565–590.

Sanz ME, Blanco S, Lopez JC, Alonso JL. 2008. Rotational probes of six conformers of neutral cysteine. Angew Chem Int Ed 47:6216–6220.

Sanz ME, Lopez JC, Alonso JL. 2010. Six conformers of neutral aspartic acid identified in the gas phase. PhysChemChemPhys 12:3573–3578.

Schlund S, Müller R, Grassmann C, Engels B. 2008. Conformational analysis of arginine in gas phase—A strategy for scanning the potential energy surface effectively. J Comput Chem 29:407–415.

Schroeder OE, Andriole EJ, Carver KL, Colyer KE, Poutsma JC. 2004. Proton affinity of lysine homologues from the extended kinetic method. J Phys Chem A 108:326–332.

Slaughter AR, Banna MS. 1988. Core-photoelectron binding energies of gaseous glycine: Correlation with its proton affinity and gas phase acidity. J Phys Chem 92:2165–2167.

Snoek LC, Robertson EG, Kroemer RT, Simons JP. 2000. Conformational landscapes in aminoacids: Infrared and ultraviolet ion-dip spectroscopy of phenylalanine in the gas phase. Chem Phys Lett 321: 49–56.

Snoek LC, Kroemer RT, Hockridge MR, Simons JP. 2001. Conformational landscapes of aromatics aminoacids in the gas phase: Infrared and ultraviolet ion-dip spectroscopy of tryptophan. PhysChemChemPhys 3:1819–1826.

Somers KRF, Kryachko ES, Ceulemans A. 2004. Theoretical study of indole: Protonation, indolyl radical, tautomers of indole and its interaction with water. Chem Phys 301:61–79.

Stearns JA, Mercier S, Seaiby C, Guidi M, Boyarkin OV, Rizzo TR. 2007. Conformation-specific spectroscopy and photodissociation of cold, protonated tyrosine and phenylalanine. J Am Chem Soc 129:11818–11820.

Stepanian SG, Reva ID, Radchenko ED, Rosado MTS, Duarte M, Fausto R, Adamowicz L. 1998a. Matrix isolation infrared and theoretical studies of the glycine conformers. J Phys Chem A 102:1041–1054.

Stepanian SG, Reva ID, Radchenko ED, Adamowicz L. 1998b. Conformational behavior of α-alanine. Matrix isolation infrared and theoretical DFT and ab initio study. J Phys Chem A 102:4632–4639.

Stepanian SG, Reva ID, Radchenko ED, Adamowicz L. 1999. Combined matrix isolation infrared and theoretical DFT and ab initio study of the nonioized valine conformers. J Phys Chem A 103:4404–4412.

Stepanian SG, Reva ID, Radchenko ED, Adamowicz L. 2001. Conformers of nonionized proline. Matrix isolation infrared and post-Hartree-Fock ab initio study. J Phys Chem A 105:10664–10672.

Suenram RD, Lovas FJ. 1978. Millimeter wave spectrum of glycine. J Mol Spectrosc 72:372–382.

Suenram RD, Lovas FJ. 1980. Millimeter wave spectrum of glycine: A new conformer. J Am Chem Soc 102:7180–7184.

Sun W, Kinsel GR, Marynick DS. 1999. Computational estimates of gas phase basicity and proton affinity of glutamic acid. J Phys Chem A 103:4112–4117.

Szidarovsky T, Czako G, Csaszar A. 2009. Conformers of gaseous threonine. Mol Phys 107:761–775.

Talbot FO, Tabarin T, Antoine R, Broyer M, Dugourd P. 2005. Photodissociation spectroscopy of trapped protonated tryptophane. J Chem Phys 122:074310.

Tarakeshwar P, Choi HS, Kim KS. 2001. Olefinic vs aromatic interaction: A theoretical investigation of the nature of interaction of first row hydrides with ethane and benzene. J Am Chem Soc 123:3323–3331.

Tehrani ZA, Tavasoli E, Fattahi A. 2010. Conformational behavior and potential energy profile of gaseous histidine. J Mol Struct (Theochem) 960:73–85.

Thar J, Zahn S, Kirchner B. 2008. When is a molecule properly solvated by a continuum model or in a cluster ansatz? A first principle simulation of alanine hydration. J Phys Chem B 112:1456–1464.

Tian Z, Pawlow A, Poutsma JC, Kass SR. 2007. Are carboxyl groups the most acidic sites in aminoacids? Gas phase acidity, H/D exchange experiments and computations on cysteine and its conjugate base. J Am Chem Soc 129:5403–5407.

Topol IA, Burt SK, Toscano M, Russo N. 1998. Protonation of glycine and alanine: Proton affinities, intrinsic basicities and proton transfer path. J Mol Struct (Theochem) 430:41–49.

Tsusuki S, Honda K, Uchimaru T, Mihami M, Tanabe K. 2000. Origin of the attraction and directionality of the NH/π interaction: Comparison with OH/π and CH/π interactions. J Am Chem Soc 122:11450–11458.

Uggerud E. 1997. The unimolecular chemistry of protonated glycine and the proton affinity of glycine: A computational model. Theor Chem Acta 97:313–316.

Vaden TD, de Boer TSJA, MacLeod NA, Marzluff EM, Simons JP, Snoek LC. 2007. Infrared spectroscopy and structure of photochemically protonated biomolecules in the gas phase: A noradrenaline analogue, lysine and alanyl alanine. PhysChemChemPhys 9:2549–2555.

Wang W, Zheng W, Pu X, Wong NB, Tian A. 2002. The 1:1 glycine-water complex: Some theoretical observations. J Mol Struct (Theochem) 618:235–244.

Wang W, Pu X, Zheng W, Wong NB, Tian A. 2003a. Some theoretical observations on the 1:1 glycine zwitterion-water complex. J Mol Struct (Theochem) 626:127–132.

Wang W, Pu X, Zheng W, Wong NB, Tian A. 2003b. Hyperconjugation versus intramolecular hydrogen bond: Origin of the conformational preference of gaseous glycine. Chem Phys Lett 370:147–153.

Wang P, Ohanessian G, Wesdemiotis C. 2008. The sodium ion affinities of asparagine, glutamine, histidine and arginine. Int J Mass Spectrom 269:34–45.

Wilke JJ, Lind MC, Schaefer HF III, Csaszar AG, Allen WD. 2009. Conformers of gaseous cysteine. J Chem Theory Comput 5:1511–1523.

Wu Z, Fenselau C. 1992. Proton affinity of arginine measured by the kinetic approach. Rapid Commun Mass Spectrom 6:403–405.

Wu Z, Fenselau C. 1994. Gas phase basicities and proton affinities of lysine and histidine measured from the dissociation of proton bound dimers. Rapid Commun Mass Spectrom 8:777–780.

Wu J, Lebrilla CB. 1993. Gas phase basicities and sites of protonation of glycine oligomers (GLY_n; n = 1–5). J Am Chem Soc 115:3270–3275.

Wu R, McMahon TB. 2008a. Stabilization of zwitterionic structures of aminoacids (Gly, Ala, Val, Leu, Ile, Ser and Pro) by ammonium ions in the gas phase. J Am Chem Soc 130:3065–3078.

Wu R, McMahon TB. 2008b. An investigation of protonation sites and conformations of protonated aminoacids by IRMPD spectroscopy. ChemPhysChem 9:2826–2835.

Xu X, Lin Z. 2010. Comprehensive ab initio study on the conformations of L-threonine and L-allo-threonine and related species in the gas phase. J Mol Struct Teochem 962:23–32.

Yang G, Zhou L, Zu Y, Fu Y, Zhu R, Liu C. 2009. Effects of side chains in gas-phase amino acids: Conformational analysis and relative stabilities. J Mol Struct (Theochem) 901:81–87.

Zhang K, Chung-Phillips A. 1998a. Conformers of gaseous protonated glycine. J Comput Chem 19:1862–1876.

Zhang K, Chung-Phillips A. 1998b. Gas phase basicity of glycine: A comprehensive ab initio study. J Phys Chem A 102:3625–3634.

Zhang K, Chung-Phillips A. 1999. A computational study of intramolecular proton transfer in gaseous protonated glycine. J Chem Inf Comput Sci 39:382–395.

Zhang M, Lin Z. 2006. Ab initio studies of the conformers and conformational distribution of the gaseous hydroxyamino acid threonine. J Mol Struct (Theochem) 760:159–166.

Zhang K, Zimmerman DM, Chung-Philips A, Cassady CJ. 1993. Experimental and ab initio studies of the gas phase basicities of polyglycines. J Am Chem Soc 115:10812–10822.

Zhang M, Huang Z, Lin Z. 2005. Systematic ab initio studies of the conformers and conformational distribution of gas-phase tyrosine. J Chem Phys 122:134313(1–7).