Acid-base thermochemistry of gaseous oxygen and sulfur substituted amino acids (Ser, Thr, Cys, Met)
Vanessa Riffet, Gilles Frison, Guy Bouchoux

To cite this version:
Vanessa Riffet, Gilles Frison, Guy Bouchoux. Acid-base thermochemistry of gaseous oxygen and sulfur substituted amino acids (Ser, Thr, Cys, Met). Physical Chemistry Chemical Physics, Royal Society of Chemistry, 2011, 13, pp.18561-18580. <10.1039/c1cp22206f>. <hal-00616361>

HAL Id: hal-00616361
https://hal.archives-ouvertes.fr/hal-00616361
Submitted on 12 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Acid–base thermochemistry of gaseous oxygen and sulfur substituted amino acids (Ser, Thr, Cys, Met)†

Vanessa Riffet, Gilles Frison and Guy Bouchoux*

Received 5th July 2011, Accepted 26th August 2011
DOI: 10.1039/c1cp22206f

Acid–base thermochemistry of isolated amino acids containing oxygen or sulfur in their side chain (serine, threonine, cysteine and methionine) have been examined by quantum chemical computations. Density functional theory (DFT) was used, with B3LYP, B97-D and M06-2X functionals using the 6-31+G(d,p) basis set for geometry optimizations and the larger 6-311++G(3df,2p) basis set for energy computations. Composite methods CBS-QB3, G3B3, G4MP2 and G4 were applied to large sets of neutral, protonated and deprotonated conformers. Conformational analysis of these species, based on chemical approach and AMOEBA force field calculations, has been used to identify the lowest energy conformers and to estimate the population of conformers expected to be present at thermal equilibrium at 298 K. It is observed that G4, G4MP2, G3B3, CBS-QB3 composite methods and M06-2X DFT lead to similar conformer energies. Thermochernical parameters have been computed using either the most stable conformers or equilibrium populations of conformers. Comparison of experimental and theoretical proton affinities and $\Delta_{\text{acid}}H$ shows that the G4 method provides the better agreement with deviations of less than 1.5 kJ mol$^{-1}$. From this point of view, a set of evaluated thermochemical quantities for serine, threonine, cysteine and methionine may be proposed:

PA $= 912, 919, 903, 938$; GB $= 878, 886, 870, 899$; $\Delta_{\text{acid}}H$ $= 1393, 1391, 1396, 1411$; $\Delta_{\text{acid}}G$ $= 1363, 1362, 1367, 1382$ kJ mol$^{-1}$. This study also confirms that a non-negligible $\Delta pS^{\circ}$ is associated with protonation of methionine and that the most acidic hydrogen of cysteine in the gas phase is that of the SH group. In several instances new conformers were identified thus suggesting a re-examination of several IRMPD spectra.

1. Introduction

Detailed knowledge on the structural and energetic properties of isolated neutral, deprotonated and protonated amino acids is of interest to numerous areas of chemistry and biochemistry. It is, for example, well established that proton transfer involving amino acids and their polymers plays an important role in biological processes. From another point of view, structure elucidation of amino acids and their polymers by mass spectrometry results, most of the time, from the observation and dissociation of deprotonated or protonated species. The intrinsic acidity and basicity of isolated amino acids provide essential clues for the understanding of these processes.

Thermochernical data associated with protonation (proton affinity, PA, and gas phase basicity, GB) or deprotonation ($\Delta_{\text{acid}}H$ and $\Delta_{\text{acid}}G$) of molecular species in the gas phase are defined by reactions (1) and (2).

$$\text{MH}^+ (\text{gas}) \rightarrow \text{M} + \text{H}^+ (\text{gas})$$

$$\text{PA}(\text{M}) = \Delta_{\text{acid}}H^o$$

$$\text{M} (\text{gas}) \rightarrow [\text{M} - \text{H}]^+ + \text{H}^+ (\text{gas})$$

$$\Delta_{\text{acid}}H(\text{M}) = \Delta_{\text{acid}}H^o$$

These quantities can be obtained experimentally from mass spectrometry measurements, and theoretically, from quantum

---

*Laboratoire des Mécanismes Réactionnels, UMR CNRS 7651,
Département de Chimie, Ecole Polytechnique, F-91128 Palaiseau cedex, France. E-mail: bouchoux@dcmr.polytechnique.fr;
Fax: 33 1 69 33 48 03; Tel: 33 1 69 33 48 42
† Electronic supplementary information (ESI) available: enthalpies, entropies and free energies of the most stable conformers of neutral, protonated and deprotonated serine, threonine, cysteine and methionine computed at 298 K using the B3LYP, B97-D and M06-2X functionals. The CBS-QB3, G3B3, G4MP2 and G4 composite methods are given in Tables T1 to T4. Relative $H_{298}^+$ and $G_{298}^o$ energies and optimized B3LYP/6-31+G(d,p) geometries of all neutral, protonated and deprotonated species are presented in Fig. S1 to S21. Signed deviations (method-G4 and method-experiment) on the computed deprotonation and protonation thermochemistry of the studied 2-amino acids are presented in Tables T5 to T8. Computed thermochemical data of reference compounds ammonia and benzoic acid are given in Table T9. The correspondence between conformers’ names in excel tables and charts are contained in Table T10. See DOI: 10.1039/c1cp22206f

---
chemistry computations. Moreover, the latter methods bring valuable molecular structure information on the neutral and ionized species hardly obtainable experimentally. A crucial point in the computation of thermochemical quantities concerning amino acids, or, more generally, molecular polyfunctional systems, is the correct consideration of non-covalent interactions. Accordingly, in these systems, the stability of a given structure results from equilibrium between different non-covalent interactions such as hydrogen bonds, salt bridges and van der Waals interactions that include London dispersion forces. Non-covalent interactions consequently influence the conformational population of these molecules and their protonated or deprotonated forms. Description of long-range interactions in terms of geometry and energy has been a challenge of theoretical chemistry during the latest years. It has long been recognized that the correlation energy of the system plays an essential part in its stabilization energy. Møller–Plesset theory and coupled cluster method are the usual post Hartee–Fock treatments allowing consideration of such an effect. Good illustrations are composite methods which have been developed for energetic quantities. The four selected methods, namely CBS-QB3, G3B3, G4MP2 and G4, have been assessed on a test set containing up to 454 energies and show average deviation from experiment generally below ~ 5 kJ mol⁻¹. A complete re-examination of the available experimental gas phase basicity and acidity data of serine, threonine, cysteine and methionine is also presented in order to allow a meaningful comparison with theory and to propose newly evaluated thermochemical data.

2. Methodology

2.1. Quantum chemistry

In the first step, starting geometries were generated through systematic manual conformational exploration and molecular dynamics and Monte-Carlo calculations with the AMOEBA force field. The resulting geometries of neutral (M) and ionized (MH⁺ and [M-H]⁻) amino acids have been subsequently optimized with quantum chemical tools in order to identify the most stable conformers. For this purpose, we select three types of density functional: firstly, the popular hybrid B3LYP functional since it is presently widely used as a computational tool by many researchers, secondly, the pure semi-local generalized gradient correlation (GGA) B97-D functionals where long-range dispersion effects are included through a R⁶ classical potential, and, thirdly, the meta-GGA M06-2X functional which includes a high percentage of Hartree–Fock exchanges and a parameterization based on various sets of reference data including inter alia interaction energies of non-covalent complexes. The reasons that prompt us to use this latter functional are that it is claimed to best describe hydrogen bonded interactions and to provide accurate thermodynamic data such as barrier heights and isomerization energies. Moreover, M06-2X has also been demonstrated to reproduce nicely conformer energies and populations in systems dominated by non-covalent and electrostatic interactions.

Conformer relative energies, protonation and deprotonation thermochemistry were then computed using a panel of composite methods which offer, at the present time, the best accuracy on energetic quantities. The four selected methods, namely CBS-QB3, G3B3, G4MP2 and G4, have been assessed on a test set containing up to 454 energies and show average deviation from experiment generally below ~ 5 kJ mol⁻¹. A complete re-examination of the available experimental gas phase basicity and acidity data of serine, threonine, cysteine and methionine is also presented in order to allow a meaningful comparison with theory and to propose newly evaluated thermochemical data.

![Scheme 1](image-url)
the methyamine, respectively, which are already available in the AMOEBA parameter set. After having defined the starting geometries of the various conformers by these prior procedures, geometries were optimized at a DFT level (i.e. B3LYP, B97-D and M06-2X) using the double-ζ quality basis set 6-31+G(d,p). This procedure furnishes then 74, 13 and 11 unique conformers for neutral, protonated and deprotonated serine, respectively (65, 12 and 11 for Thr; 87, 21 and 39 for Cys, and 32, 6, 24 for Met, but, in this latter case, in a limited 15 kJ mol⁻¹ G̃₂⁹₈ range). The most stable conformers located at these levels in a 15 kJ mol⁻¹ range were then subjected to more elaborate computations using the composite methods CBS-QB3, G3B3, G4MP2 and G4. Vibrational frequencies were calculated at the same level as the geometry optimization to confirm that the structures are true minima on the potential energy surfaces. Single point energy calculations were performed at the 6-31+G(d,p) optimized geometries using triple-ζ quality basis set 6-311+G(3df,2p) on several key conformers. Composite CBS-QB3,15 G3B3,16 G4MP217 and G418 methods are based on a series of quantum chemistry calculations combined assuming additivity of the energy terms. The complete basis set, CBS-QB3 method15 uses optimized geometry and frequencies calculated at the B3LYP/6-311G(2d,p) level. Then, MP2/6-31+G(2df,2p) energy and CBS extrapolation are computed. Further, MP4(SDQ)/6-31+G(d,p) and QCISD(T)/6-311G(d,p) single points are performed and finally, two empirical correction terms including effects of absolute overlap integral and spin contamination, are considered in the overall energy estimate. The G3B3 methods16 use geometry optimized at the B3LYP/6-31G(d) level, zero point vibrational energy is obtained from vibrational frequencies calculated at this level and scaled by a factor 0.96. In a second step, single point energy calculations are performed using (i) the frozen core QCISD(T)/6-311G(d) approximation, (ii) MP2(full) computation using the G3Large basis set (roughly a 6-31+G(2df,2p) basis for the first row elements) set and (iii) MP4(FC) computations using the 6-31G(d) and larger basis sets such as 6-31+G(d) and 6-31G(2df,p). Finally a “high level correction” (HLC) is introduced to account for remaining deficiencies in the energy computation. The G4 technique18 differs from G3B3 by (i) equilibrium structure and harmonic frequencies (scaled by 0.9854) calculated at the B3LYP/6-31G(2df,p) level, (ii) calculation of the Hartree–Fock energy limit, (iii) single point correlation energies are calculated at the CCSD(T)/6-31G(d) rather than the QCISD(T)/6-31G(d) level, (iv) MP2(full) computation uses the G3LargeXP basis set which consists of the G3Large basis improved by adding new d-polarization functions, and finally two additional parameters (A' and E) are added in the HLC correction. In the G4(MP2) method,19 MP3 and MP4 large basis set calculations are eliminated from the G4 procedure.

2.2. Thermochemistry

As recalled in the Introduction, the thermochemical quantities associated with the intrinsic basicity and acidity of a molecule M are defined as the standard enthalpy and standard Gibbs free energy of reactions (1) and (2), respectively.3 Starting from these definitions, PA(M) and ΔacidH(M) at temperature T, may be calculated using:

\[ \text{PA}_T(M) = \Delta H^o_T = \overline{H}_T^o(M) + \overline{H}_T^o(H^+) - \overline{H}_T^o(MH^+) \]  

(3)

\[ \Delta_{\text{acid}}H^o_T(M) = \Delta_2H^o_T = \overline{H}_T^o([M - H]^-) + \overline{H}_T^o(H^+) - \overline{H}_T^o(M) \]  

(4)

where \( \overline{H}_T \) represents the enthalpy of each species.

Since \( \text{PA}_T(M) \) and \( \Delta_{\text{acid}}H(M) \) are molar quantities, their estimate should in principle involve one mol of M and MH⁺ (or [M–H]⁻) in thermal equilibrium at temperature \( T \). The usual computational approach assumes that each mol of reactant and product of reactions (1) or (2) contains only the most stable conformers of each species. Under these circumstances, eqn (3) and (4) may be directly used to derive “most stable conformers” (msc) quantities, \( \text{PA}_T(M)_{\text{msc}} \) and \( \Delta_{\text{acid}}H(M)_{\text{msc}} \). At 298 K these equations reduce to eqn (5) and (6):

\[ \text{PA}_T(M)_{\text{msc}} = \overline{H}^o_{298}(M)_{\text{msc}} - \overline{H}^o_{298}(MH^+)_{\text{msc}} + 6.19 \text{ kJ mol}^{-1} \]  

(5)

\[ \Delta_{\text{acid}}H(M)_{\text{msc}} = \overline{H}^o_{298}([M - H]^-)_{\text{msc}} - \overline{H}^o_{298}(M)_{\text{msc}} + 6.19 \text{ kJ mol}^{-1} \]  

(6)

since the \( \overline{H}_T^o(H^+) \) term is simply equal to 5/2 \( RT \).

Rigorously, however, a distribution of conformers should be considered for both reactants and products. Assuming thermal equilibrium, a Boltzmann distribution may be used to derive the molar fractions \( x_i \) of each conformer at temperature \( T \) using eqn (7):

\[ x_i = \exp(-G_i/RT)/ \sum_i \exp(-G_i/RT) \]  

(7)

where \( N \) is the total number of conformers and \( G_i \) their individual Gibbs free energies.

Averaged enthalpy value of one mole of mixture may be determined from the enthalpies of the separate species according to:

\[ \overline{H}^o_T = \sum_i x_i(H^o_T) \]  

(8)

thus leading to:

\[ \text{PA}_T(M)_{\text{average}} = \overline{H}^o_{298}(M)_{\text{average}} - \overline{H}^o_{298}(MH^+)_{\text{average}} + 6.19 \text{ kJ mol}^{-1} \]  

(9)

\[ \Delta_{\text{acid}}H(M)_{\text{average}} = \overline{H}^o_{298}([M - H]^-)_{\text{average}} - \overline{H}^o_{298}(M)_{\text{average}} + 6.19 \text{ kJ mol}^{-1} \]  

(10)

Similarly, the “most stable conformer” and “averaged” definitions may be introduced for the gas phase basicities, GB(M) and for the corresponding acidity parameter \( \Delta_{\text{acid}}G(M) \). It should be noted however that the averaged entropy of a collection of conformers should include the entropy of mixing and be consequently estimated via eqn (11):

\[ [S^o_T]_{\text{average}} = \sum_i x_i(S^o_1) - R \sum_i x_i \ln x_i \]  

(11)

where the second component corresponds to the mixing contribution.
2.3. Nomenclature

Nomenclature used to describe the various conformers parallels that used previously in a preceding study of aliphatic amino acids.23 Briefly, three types of internal hydrogen bonds located in the C(H,R)NH₂COOH framework have been defined as I, II and III (Chart 1). The three possible orientations of the amino group inside the amino acid moiety are designated by letters A, B or C (Chart 1). Note that for type I conformers, a secondary hydrogen bond O(S)H···NH₂ is generally occurring. The denomination I will be used to describe the situation where NH₂···O(S)H. Conformers of types I and III discussed in the text of the present study are invariably associated with C–N rotamers C. For the purpose of clarity, the C letter is consequently not indicated. Concerning the conformers generated by rotation around the C(2)–C(3) bond, the nomenclature gauche (abbreviated g+ or g−) and antiperiplanar (abbreviated a) has been used as noted in Chart 2.

Most stable protonated amino acids conformers are characterized by two general types of arrangements where the internal hydrogen bonding is established either toward the carbonyl or the hydroxyl oxygen. These two situations are denoted HI and HII as defined in Chart 3. In the case of deprotonated amino acids the two possible orientations of the NH₂ group with respect to the carboxylate moiety are denoted -H_A and -H_B (Chart 3) by analogy with the nomenclature used to describe the neutral molecules (Chart 1).

3. Results and discussion

In the following section we will present first the essential results of our extensive exploration of the conformational space of neutral and ionized amino acids Ser, Thr, Cys and Met. Detailed data are given in the electronic supplementary information.† In the second and third parts, protonation and deprotonation thermochemistry are examined and compared with experimental data.

3.1 Conformational landscape of neutral, protonated and deprotonated amino acids

Serine. Previous explorations of the conformational space of neutral serine were performed using Hartree–Fock24,25 B3LYP,26–31 MP227b and G2(MP2)32 or G3(MP2)33 methods (note that enantiomer 2R rather than 2S was considered in ref. 24 and 26). Depending upon the theoretical level and the strategy used to explore the potential energy surface, different conformers were considered as the most stable. In the present study, our conformational search followed by geometry optimization at the B97-D, B3LYP and M06-2X levels using the 6-31+G(d,p) basis set leads to 74 stable forms among which the most stable have been further examined by the composite methods CBS-QB3, G3B3, G4MP2 and G4. Only the results concerning the eighth most stable conformers will be discussed here (Fig. 1 and 2), the full set of data is available in Table T1 concerning the eighth most stable conformers will be discussed here (Fig. 1 and 2), the full set of data is available in Table T1 and Fig. S1 and S2 of electronic supplementary information (ESI).† Figure 1 illustrates the influence of the theoretical level on the relative stability (H₂⁰ and G₂⁰) of the first eight conformers. It clearly appears that the four composite methods predict similar relative H₂⁰ and G₂⁰ energies. The situation is different for DFT methods. If M06-2X behaves similarly to the composite methods, the relative energy orders given by B97-D and B3LYP seriously disagree. A second observation is that the relative stability orders given by H₂⁰ and G₂⁰ are different, obviously because of different third law entropies between conformers. In order to understand this effect it is necessary to examine the geometrical characteristics of these conformers (Fig. 2).

Conformers Slₐg+ , Slₐg− , Sl’ₐg− present a NH₂···O= C–OH(syn), (type I), arrangement while ShIₐg+ and ShIₐg− correspond to the same kind of internal hydrogen bond but involve the less basic hydroxyl oxygen as acceptor: NH₂···OH–C=O(syn), (type III). Note that conformers Slₐc+ and Slₐc− present a supplementary source of stabilization with a CH₂OH–···NH₂ hydrogen bond while conformer Sl’cₐg− is characterized by a reverse NH₂···O(H)CH₂ interaction.
Finally, the three conformers SII B g/C0, SII B a and SII A g+ correspond to an anti conformation of the acidic moiety thus allowing H2N–HO–C=O(anti), (type II) internal hydrogen bonding. It is known that the H-bond interaction involved in type II conformers is particularly efficient since it involves a strong acidic hydrogen and a strong basic nitrogen atom.23,34 This is attested to by the short distances between the involved atoms which lie around 1.9 Å (see SII B g/C0, SII B a and SII A g+, Fig. 2). By contrast, the H-bond distance is situated between 2.2 and 2.5 Å for conformers of types I or III. Comparable observation can be made with the hydrogen bonds involving the hydrogen of the alcohol moiety and either the amino nitrogen (conformers SI C g+, SI A g/C0, SIII C g+) or the oxygen of the carbonyl group (conformers SII B a, SIII C g/C0). From this point of view, the case of conformer SII B a should be underlined since it presents the shortest CH2 OH/C1/C1/C1 NH2 distance (1.978 Å, Fig. 2). These differences in internal H-bonding induce differences in rotational barriers and, consequently, in entropies. It is thus not surprising to observe that conformer SI H a has the lowest $S_{298}^0$ (349 J mol$^{-1}$ K$^{-1}$), followed by conformer SI H g– ($S_{298}^0 = 353$ J mol$^{-1}$ K$^{-1}$), whereas conformers SI C g+ and SI C g– present $S_{298}^0$ close to 365 J mol$^{-1}$ K$^{-1}$.

This results in the observed change in stability order between $H_2^{298}$ (SI H g– < SI C g+ < SI A g– < SI H a < SI H c g–, SI A g+ < SI C g+, SI C g–) and $G_2^{298}$ (SI C g+ < SI H g– < SI A g– < SI H c g+). Using the G4 relative free energies of the eight more stable conformers of neutral serine, we can calculate the conformer population at 298 K, SI C g+ /SI H g–/SI A g–/SI H c g+/SI C g+/SI C g–/SI H a/SI A g+/SI H c g–: 38.1/28.4/14.4/4.9/4.6/3.8/3.1/2.7%.

It is satisfactory to observe that the infrared spectrum of neutral serine, isolated in a low temperature argon matrix,27 has been interpreted by the occurrence of a mixture of conformers SI C g+, SI H g–, SI A g and SI A g–. A more recent microwave study of gaseous serine produced by laser ablation has been interpreted by the existence of a population of conformers SI C g+, SI H g–, SI A g–, SI A a, SI I c g– and SI A g+.27h These conclusions are in correct agreement with our computational results.

13 unique structures have been identified for protonated serine after geometry optimizations at the B97-D, B3LYP and M06-2X levels using the 6-31+G(d,p) basis set. Enthalpies and free energies at 298 K were then computed using the selected

Fig. 1 Relative $H_2^{298}$ and $G_2^{298}$ (kJ mol$^{-1}$) of the eight most stable conformers of neutral L-serine.
composite methods (Table T1, Fig. S3 and S4 of ESI†). Results are illustrated by Fig. 3 for the four most stable conformers \( \text{SHIg}^−, \text{SHIg}^+, \text{SHIIg}^− \) and \( \text{SHIIg}^+ \). As already observed for neutral serine, DFT methods behave differently from the composite methods. A clear increase in the \( H^0_{298} \) differences between \( \text{SHIg}^− \) and the three other conformers is observed when passing from B97-D to CBS-QB3. By contrast, these differences remain constant between CBS-QB3 and the more sophisticated G4 method. Evolution of the free energies at 298 K follows exactly the same trends, with however a slight compression of the energy scale. As for neutral serine, this difference may again be explained by considering the entropies of these four conformers, consideration which may be enlightened by looking at their geometrical features (Fig. 4).

The four structures \( \text{SHIg}^−, \text{SHIg}^+, \text{SHIIg}^− \) and \( \text{SHIIg}^+ \) present a syn HOCO arrangement of the acidic moiety allowing the establishment of a strong hydrogen bond with one of the hydrogens of the \( \text{NH}_3^+ \) group. The most favorable interaction obviously occurs with the oxygen of the carbonyl group thus leading to the two most stable conformers \( \text{SHIg}^− \) and \( \text{SHIg}^+ \) (upper part of Fig. 4). The other possibility i.e. interaction between \( \text{NH}_3^+ \) and the oxygen of the hydroxyl group is less favorable and leads to conformers \( \text{SHIIg}^− \) and \( \text{SHIIg}^+ \) situated ~15 kJ mol\(^{-1}\) above \( \text{SHIg}^− \) and \( \text{SHIg}^+ \). The most significant structural difference between \( \text{SHIg}^− \) and \( \text{SHIg}^+ \) (or \( \text{SHIIg}^− \) and \( \text{SHIIg}^+ \)) lies in the \( \text{C} = \text{O} \cdots \text{H}_2\text{N}^+ \) distance. This distance is shorter for conformers \( \text{g}^− \) thus explaining why they are more stable than their counterparts \( \text{g}^+ \). The second consequence is that the third law entropies of \( \text{SHIg}^− \) and \( \text{SHIIg}^− \) are lower than those of \( \text{SHIg}^+ \) and \( \text{SHIIg}^+ \), respectively. As observed in Fig. 3, this entropy effect produces a decrease in the \( G^0_{298} \) difference, with respect to \( H^0_{298} \), between the couples \( \text{SHIg}^− \) and \( \text{SHIg}^+ \) on one hand and \( \text{SHIIg}^− \) and \( \text{SHIIg}^+ \) on the other.

It is evident from the present data that the two conformers \( \text{SHIg}^− \) and \( \text{SHIg}^+ \) will describe most of the population of protonated serine at 298 K. Indeed, using the \( G^0_{298} \) calculated at the G4 level, a \( \text{SHIg}^−/\text{SHIg}^+/\text{SHIIg}^−/\text{SHIIg}^+ \) ratio equal to 68.3/31.1/0.4/0.2% is estimated. The largest stability of \( \text{SHIg}^− \) observed both in enthalpy and in free energy in the present study up to the G4 level confirms earlier conclusions based on less accurate calculations.\(^{26,30,32,33}\) Experimentally, the IRMPD spectrum of protonated serine presents three absorption bands at 115 cm\(^{-1}\), 1460 cm\(^{-1}\) and 1794 cm\(^{-1}\).\(^{13,35}\) This was attributed to a HOCO bend, a \( \text{NH}_3 \) umbrella and CO stretching, respectively for conformer \( \text{SHIg}^− \) (SH01 in the original paper).\(^{35}\) Computation at the B3LYP/6-31G(2df,p) level leads to (unscaled) theoretical vibrational frequencies of 1194, 1467 ± 4 and 1837 cm\(^{-1}\) for both conformers \( \text{SHIg}^− \) and \( \text{SHIg}^+ \) thus leading to the conclusion that the presence of the two conformers may also account for the experimental observation.

Geometry optimization of the 162 trial structures for deprotonated serine \([\text{Ser}−H]^−\) at the DFT levels converged on 11 unique conformers (Table T1 and Fig. S5 and S6 in ESI†). The relative enthalpies and free energies at 298 K of the four most stable conformers \( \text{S-H}α, \text{S-H}α, \text{S-H}γ^− \) and \( \text{S-H}γ^+ \) are reproduced in Fig. 5 as a function of the theoretical method, their optimized geometries are presented in the lower part of Fig. 4. As shown in Fig. 5, conformers \( \text{S-H}α \) and \( \text{S-H}α \) are almost of identical stabilities at the composite CBS-QB3, G3B3, G4MP2 and G4 levels. As previously observed for neutral and protonated species, the DFT methods present variable results. Discrepancies with the composite methods are however limited to a few kJ mol\(^{-1}\). Differences in \( H^0_{298} \) and \( G^0_{298} \) relative values are less pronounced here, compared with the data obtained for neutral and protonated serine. Accordingly, entropies are close together for the four conformers: 345 J mol\(^{-1}\) K\(^{-1}\) for \( \text{S-H}α \) and \( \text{S-H}α \), 348 and 351 J mol\(^{-1}\) K\(^{-1}\) for \( \text{S-H}γ^− \) and \( \text{S-H}γ^+ \) in accordance with structural similarities. In fact, the three most stable structures \( \text{S-H}α, \text{S-H}α \) and \( \text{S-H}α \) are characterized by strong intramolecular \( \text{CH}_3\text{OH}−\cdots\text{OCO}− \) interactions.

![Fig. 2](image-url) Optimized geometries (B3LYP/6-31G(2df,p)) of the eight most stable conformers of neutral L-serine (in brackets, relative \( H^0_{298} \) and \( G^0_{298} \) in kJ mol\(^{-1}\) calculated at the G4 level).
**Fig. 3** Relative $H^{298}_o$ and $G^{298}_o$ (kJ mol$^{-1}$) of the four most stable conformers of protonated $l$-serine.

**Fig. 4** Optimized geometries (B3LYP/6-31G(2df,p)) of the four most stable conformers of protonated and deprotonated $l$-serine (in brackets, relative $H^{298}_o$ and $G^{298}_o$ in kJ mol$^{-1}$ calculated at the G4 level).
The essential structural differences between $\text{S-H}_\text{Aa}$ and $\text{S-H}_\text{Ba}$ arise from the orientation of the NH$_2$ group, the hydrogen of which being in favorable interaction with the second oxygen of the COO$^-$ moiety. Conformer $\text{S-HBg}$ is characterized by an internal H-bond (OH$\cdots$O$=$ distance: 1.695 Å, Fig. 4). In this arrangement the NH$_2$ group becomes pseudo axial with respect to the OC$_1$C$_2$C$_3$OH cycle and cannot approach the other carboxylate oxygen at a distance lower than 2.382 Å (Fig. 4). The net result is a conformer slightly destabilized by comparison with $\text{S-H}_\text{Aa}$ and $\text{S-H}_\text{Ba}$. At the G4 level the ratio of the conformer populations $\text{S-H}_\text{Ba}/\text{S-H}_\text{Aa}/\text{S-H}_\text{Bg}/\text{S-H}_\text{Ag}^+$ is equal to 52.0/43.2/4.5/0.3.

Threonine. The conformational landscape of neutral L-threonine has been the subject of recent theoretical investigations. A number of conformers situated between 56 and 74 have been found using the B3LYP functional and either 6-311++G(d,p) or 6-31G(d) basis sets. Our investigation of (2S,3R)-L-threonine leads to the characterization of a similar range of 65 stable structures. Full data concerning the twelve most stable conformers are given in the ESI (Table T2 and Fig. S7 and S8). As already observed for serine, composite and M06-2X methods predict comparable relative $H^\circ_{298}$ and $G^\circ_{298}$ energies. By contrast, B97-D and B3LYP functionals may lead to $H^\circ_{298}$ or $G^\circ_{298}$ with a variance of several kJ mol$^{-1}$ (Table T2, Fig. S7). The difference observed between the relative stability orders, when passing from $H^\circ_{298}$ to $G^\circ_{298}$, is also easily understood when considering the structural differences between conformers. Since at the G4 level five conformers ($\text{TI}_\text{Bg}^-$, $\text{TI}_\text{G}$, $\text{TI}_\text{Ag}^-$, $\text{TIH}_\text{cG}^+$, $\text{TIcG}^-$) are predicted to represent $\text{S-H}_\text{Aa}$ and $\text{S-H}_\text{Ba}$ are the most likely structures of deprotonated L-serine in the gas phase in perfect agreement with our data.

![Fig. 5](https://example.com/fig5.png) Relative $H^\circ_{298}$ and $G^\circ_{298}$ (kJ mol$^{-1}$) of the four most stable conformers of deprotonated L-serine.
\( \approx 90\% \) of the population of neutral threonine at 298 K, only these structures are presented in Fig. 6 and discussed below.

As observed for serine, conformers of type II present the strongest internal hydrogen bonding and consequently the lowest enthalpy but also the lowest absolute entropies. For serine and threonine, this situation arises for conformers \( S_{II}g^- \) and \( T_{II}g^- \). It is noteworthy that the same order in enthalpy is observed for the first conformers of both serine and threonine (i.e. \( H_{II}g^- < I_{II}g^- < I_{II}a \)). However, for serine the enthalpy gap between \( S_{II}g^- \) and \( S_{II}c^+ \) is equal to 1.2 kJ mol\(^{-1}\) while for threonine, the analogous difference attains 2.4 kJ mol\(^{-1}\) (G4 calculations). Since the entropies of type II conformers are lower than the entropies of conformers of type I, the \( G_{298}^0 \) of both types of conformers will be affected differently. Roughly, relative \( G_{298}^0 \) are shifted downward by ca. 1.8 kJ mol\(^{-1}\). As a consequence, the order of stability is changed for \( S_{II}g^- \) and \( S_{II}c^+ \) but not for \( T_{II}g^- \) and \( T_{II}c^+ \).

As indicated above, conformers \( T_{II}g^-/T_{II}c^+ + T_{II}a^-/T_{II}c^+ \) represent more than \( 90\% \) of the population of conformers at 298 K (G4 calculations). The exact ratio of \( T_{II}g^-/T_{II}c^+ + T_{II}a^-/T_{II}c^+ \) is 38.0/26.7/18.8/11.4/5.1\%. Comparable results, based on MP2/6-31+G(2df,p) calculations, have been reported by Lin and coworkers.\(^39\) From an experimental point of view, Fourier transform microwave spectroscopy was applied to \( l \)-threonine vaporized by laser ablation.\(^40\) Analysis of the resulting rotational spectra leads the authors to conclude that seven conformers were present in their experimental conditions, the most abundant of which being \( T_{II}c^+ \), \( T_{II}a^- \) and \( T_{II}a^- \) in excellent agreement with our expectations.

Relative enthalpies and free energies at 298 K of the four most stable structures of protonated threonine, among the 12 identified stable structures at the B3LYP/6-31+G(d,p) level, are reported in Table T2 and Fig. S9 and S10 of ESI.\(^1\) Not surprisingly, as for serine, essentially two conformers, \( T_{II}g^- \) and \( T_{II}g^+ \), are predicted to represent the equilibrium population at 298 K. These two structures present indeed particularly strong intramolecular hydrogen bonds between the \( NH_3^+ \) moiety and (i) the oxygen of the carbonyl group, and (ii) the hydroxyl oxygen (Fig. 7). Two other structures, \( T_{II}l^+g \) and \( T_{II}l^-g \), characterized by a less efficient \( NH_3^+ \cdot \cdot \cdot O(H) \cdot C=O \) bonding are situated more than 10 kJ mol\(^{-1}\) above \( T_{II}g^- \) and \( T_{II}g^+ \). This situation is reminiscent of that encountered with serine and similar comments can be made. In particular the small compression of the energy scale observed between \( H_{II}g^0 \) and \( G_{298}^0 \) (Fig. S9 of ESI\(^1\)) is in perfect agreement with the difference in strength of the internal hydrogen bonds and the resulting difference in entropy. A comparable situation has been described before with protonated serine (see Fig. 3 and accompanying comments). Finally, it is evident that the description of the 298 K population of protonated threonine is limited to the participation of conformers \( T_{II}g^- \) and \( T_{II}g^+ \) only. Accordingly, using the G4 computed \( G_{298}^0 \), a \( T_{II}l^+g/T_{II}l^+g + T_{II}l^+g/T_{II}l^-g \) ratio of 61.2/38.1/0.4/0.3\% is evaluated.

Concerning deprotonated threonine, we located 11 stable conformers, the four most stable being similar to those obtained for serine, \( T_{II}a, T_{II}a, T_{II}g^- \), and \( T_{II}a \) (Table T2 and Fig. S11 and S12 of ESI).\(^1\) Their relative \( H_{II}g^0 \) and \( G_{298}^0 \) (Fig. 7) present similar values and the choice of the method seemingly does not alter the corresponding stability orders. A noticeable exception however is the B97-D functional which doesn’t predict \( T_{II}a \) as a stable species but invariably converges toward conformer \( T_{II}a \). The structural difference between \( T_{II}a \) and \( T_{II}a \) is only lying on the value of the HNC—O dihedral angle (Fig. 7). The slightly lower stability of \( T_{II}a \) is probably due to a steric repulsion between the methyl group and the H atom of the amino group not involved in a hydrogen bond. This argument is supported by the fact that, in serine, the two equivalent conformers \( S_{II}a \) and \( S_{II}a \) present quasi identical \( H_{II}g^0 \) and \( G_{298}^0 \) (Fig. 4).

---

Fig. 6 Optimized geometries (B3LYP/6-31G(2df,p)) of the five most stable conformers of neutral \( l \)-threonine (in brackets, relative \( H_{II}g^0 \) and \( G_{298}^0 \) in kJ mol\(^{-1}\) calculated at the G4 level).
Using the G4 free energies at 298 K, we estimate the population of deprotonated threonine, T-H₃a/T-H₉a/T-H₁g⁻/T-H₁g⁺, to be 62.9/31.2/5.6/0.3%.

**Cysteine.** Among the 87 stable conformers of (2R)-L-cysteine identified at the DFT level, the 13 most stable have been investigated with the composite computational methods. The relative \( H_{\text{298}} \) and \( G_{\text{298}} \) and the optimized structures of these conformers are presented in Table T3 and Fig. S13 and S14 of ESL† As a general observation for this system, relative 298 K enthalpies of the ten most stable conformers are close together whatever the computational method used. More significant fluctuations are however observed on the relative free energies. However, all the methods give the type \( \text{II} \) conformer \( \text{CIHg}^- \) (Fig. 8) as the most stable in both \( H_{\text{298}} \) and \( G_{\text{298}} \).

At the G4 level, the order of stability of the cysteine conformers in terms of \( H_{\text{298}} \) is: \( \text{CIHg}^- < \text{CIcg}^- < \text{CIcg}^- < \text{CIcg}^+ < \text{CIII CG}^- < \text{CIII CG}^+ < \text{CII CG}^- < \text{CII CG}^+ < \text{CII CG}^+ \) (Fig. 8). It is noteworthy that the enthalpy difference between \( \text{CIHg}^- \) and the second conformer \( \text{CIcg}^- \) is as large as 5.4 kJ mol\(^{-1}\) at the G4 level. No such large energy gap has been observed for serine and threonine for which conformer \( \text{IIHg}^- \) has been also identified as the most stable in the enthalpy scale, but where the closest conformers were situated only 1.2 and 2.4 kJ mol\(^{-1}\) above. This difference in behaviour may be understood by examining the network of internal hydrogen bonds stabilizing the closest conformers. As repeatedly noted, conformers of type \( \text{II} \) are strongly stabilized by the \( \text{H}_2\text{N} \cdot \cdot \cdot \text{O} \cdot \cdot \cdot \text{H} \) bonding. A secondary stabilization is brought by a NH···OHC(3) interaction for \( \text{SIHg}^- \) and \( \text{THlg}^- \). In the case of \( \text{CIHg}^- \), the homologous interaction, i.e. NH···SHC(3) is reinforced since S is more basic than O (proton affinities of methanol and methanethiol are 754 and 773 kJ mol\(^{-1}\), respectively). Conformers of type \( \text{I} \) are characterized by a NH···O···COH(syn) interaction. In addition, \( \text{Icg}^- \) and \( \text{Icg}^+ \) of serine and threonine enjoy very strong C(3)OH···NH₂ interactions. This result in close enthalpy proximity for conformers \( \text{SIHg}^- \), \( \text{SIcg}^- \) and \( \text{SIcg}^+ \) on one hand and of \( \text{THlg}^- \), \( \text{TIcg}^- \) and \( \text{TIcg}^+ \) on the other. In the case of cysteine, the stabilization of conformers \( \text{CIcg}^- \) and \( \text{CIcg}^+ \) is not observed because the C(3)SH···NH₂ interaction becomes less efficient due to the lower electronegativity of S.

The difference in stability between \( \text{CIHg}^- \) and the other conformers is significantly reduced in the free energy scale. The order of stability given by the G4 calculated \( G_{\text{298}} \): \( \text{CIHg}^- < \text{CIcg}^- < \text{CIcg}^- < \text{CIcg}^+ < \text{CIHg}^- < \text{CIIHg}^- < \text{CIIHg}^+ \) is however not significantly changed with respect to the enthalpy scale. What is noteworthy is the shift of the conformers of type \( \text{I} \) (\( \text{CIcg}^- \), \( \text{CIcg}^+ \), \( \text{CIcg}^+ \), \( \text{CIcg}^+ \)) toward low \( G_{\text{298}} \) as expected from their high third law entropies (near 380 J mol\(^{-1}\) K\(^{-1}\), as compared to 367 J mol\(^{-1}\) K\(^{-1}\) for \( \text{CIHg}^- \)). This entropy difference is obviously related to the strong character of the two major internal hydrogen bonds occurring in \( \text{CIHg}^- \) discussed above. The six conformers presented in Fig. 8 represent ca. 85% of the total population of conformers (Table T3 in ESL). At 298 K the overall ratio \( \text{CIHg}^--/\text{CIcg}^-/\text{CIcg}^--/\text{CIcg}^+/\text{CIcg}^+/\text{CIIHg}^-/\text{CIIHg}^+ \) is predicted to be 35.8/23.9/11.9/7.4/5.9/5.9%.

Previous theoretical works on neutral cysteine reported results obtained at the HF,24 B3LYP26,29,41,43 and MP244,45 levels of theory. Some of these studies however are concerned by the non-natural 2S configuration of cysteine.24,26,45 There is a general consensus in placing conformer \( \text{CIHg}^- \) as the most stable species. Depending upon the theoretical level, the conformer situated immediately above \( \text{CIHg}^- \) is \( \text{CIcg}^- \) or \( \text{CIcg}^+ \) as expected for a structure differing only by few kJ mol\(^{-1}\). Experimentally, no less than six conformers have been claimed to be identified.44 According-ly, microwave spectra of vapor phase cysteine have been interpreted by the major presence of conformers \( \text{CIHg}^- \), \( \text{CIcg}^- \) and \( \text{CIcg}^+ \) beside the minor contributions of conformers \( \text{CIHg}^+ \), \( \text{CIII Hg}^- \), \( \text{CIII Hg}^+ \), \( \text{CIHg}^- \), \( \text{CIHg}^+ \), \( \text{CIIHg}^- \), \( \text{CIIHg}^+ \). Surprisingly enough, conformers \( \text{CIcg}^- \) and \( \text{CIcg}^- \) are not mentioned despite the fact that they are predicted to be more stable than \( \text{CIcg}^- \), \( \text{CIII Hg}^- \) or \( \text{CIII Hg}^+ \). It may be underlined however that rotational constants of \( \text{CIcg}^- \) and \( \text{CIcg}^+ \) are very close to that of \( \text{CIcg}^- \) probably rendering a clear structural assignment uneasy.

**Fig. 7** Optimized geometries (B3LYP/6-31G(2df,p)) of the four most stable conformers of protonated and deprotonated l-threonine (in brackets, relative \( H_{\text{298}} \) and \( G_{\text{298}} \) in kJ mol\(^{-1}\) calculated at the G4 level).
Among the 21 conformers identified for protonated cysteine, four were predicted to be situated in a $\sim 6$ kJ mol$^{-1}$ range in both $H^0_{298}$ and $G^0_{298}$ (Table T3 and Fig. S15, S16 in ESI† Fig. 8). These four structures, $\text{CHI}^-g$, $\text{CHI}^+g$, $\text{CHI}'^-g$, and $\text{CHI}'^+g$, are stabilized by $\text{NH}_3^+\cdots\text{O}=\text{COH}(\text{syn})$ and $\text{NH}_3^+\cdots\text{S}$ hydrogen bonding interactions. The difference between conformers $\text{CHI}^-g$ and $\text{CHI}^0g$, or $\text{CHI}^+g$ and $\text{CHI}'^0g$, lies on a rotation around the C(3)–S bond. For $\text{CHI}^-g$ and $\text{CHI}^0g$ the S–H bond is staggered with respect to the C(3)H$_2$,C(2) group (Fig. 8), whereas the conformation is eclipsed for conformers $\text{CHI}^0g$ and $\text{CHI}'^0g$. This conformational change is associated with a difference in energy of $\sim 5$ kJ mol$^{-1}$. The $g^-/g^+$ conformational change corresponds to an energy difference of $\sim 2$ kJ mol$^{-1}$ seemingly induced by a slight weakening of the $\text{NH}_3^+\cdots\text{O}=\text{COH}(\text{syn})$ interaction in $\text{CHI}^0g$ and $\text{CHI}'^0g$. The four conformers described above represent the essential components of the equilibrium population at 298 K. The distribution of conformers $\text{CHI}^-g$ and $\text{CHI}^+g$, or $\text{CHI}^0g$ and $\text{CHI}'^0g$, is equal to 54.2/28.1/11.4/6.3%. The present results are in qualitative agreement with a previous study, done, as indicated previously, on the $2R$ enantiomeric form of cysteine.$^{26}$

Deprotonated cysteine may exist in either its carboxylate or its thiolate forms (Scheme 2). Recently, mass spectrometry$^{45}$ and photoelectron spectroscopy$^{46}$ experiments were interpreted by the preferential formation of the thiolate form of deprotonated cysteine. By contrast, tentative characterization of the thiolate structure by gas phase IR multi photon dissociation failed but, rather, the results were interpreted by the occurrence of a carboxylate form.$^{36}$ These studies were supported by theoretical computations on several deprotonated cysteine conformers where the thiolate structure is lower in energy than the carboxylate structure. However only a limited number of conformers seems to have been considered in these studies. Moreover, most of the time, the investigations were done using the B3LYP functional. A systematic investigation of the conformational space of deprotonated cysteine up to the G4 level is thus of interest in order to elucidate these observations.

In the present study we identified 25 carboxylate, $\text{C-H}_0$, and 16 thiolate, $\text{C-H}_s$, stable structures, the five most stable are situated in a $15$ kJ mol$^{-1}$ $G^0_{298}$ range (Table T3 and Fig. S17, S18 in ESI† Fig. 9). All the composite methods (CBS-QB3, G3B3, G4MP2 and G4) confirm that the thiolate form is more stable than the carboxylate form of deprotonated cysteine. In fact, four thiolate conformers, $\text{C-H}_s\text{ch}_g$, $\text{C-H}_s\text{ch}_a$, $\text{C-H}_s\text{sa}_a$, $\text{C-H}_s\text{sa}_g$, are predicted to be situated below the lowest
carboxylate form C-HOCg+ (Fig. 9). The four thiolate structures are all characterized by a strong (anti)OCOH⋯S interaction. Secondary favourable interactions involving NH2⋯O=COH or NH2⋯S- hydrogen bonding are observed for the a (C-HSCg and C-HSAb) or g- (C-HSag- and C-HSbg- ) types of conformers, respectively. These four structures are situated in a 6 kJ mol\(^{-1}\) range of \(H_{298}^0\) and appear to be similarly stabilized by their networks of internal hydrogen bonds since they present identical third law entropies of 355 J K\(^{-1}\) mol\(^{-1}\). As a consequence, the relative \(G_{298}^0\) values mimic the \(H_{298}^0\) results.

The carboxylate structure C-HOCg+ is stabilized by a NH2⋯O=CO internal hydrogen bond, accompanied by a cooperative SH⋯NH2 interaction. This structure is situated 12.8 kJ mol\(^{-1}\) above the most stable thiolate conformer C-HScg- in the \(H_{298}^0\) scale. This enthalpy gap is drastically reduced when considering the \(G_{298}^0\). Accordingly, since the entropy of C-HOCg+ is 14 J K\(^{-1}\) mol\(^{-1}\) larger than that of C-HScg-, the difference in free energy is reduced to 8.7 kJ mol\(^{-1}\) at the G4 level. It is confirmed however that the carboxylate form is not the most stable, even on the 298 K free energy scale. The predicted population of conformers C-HScg-/C-HScb/C-HScab/C-HSag- C-HOcg+ is 57.2/28.7/8.1/4.2/1.8% (G4 calculations at 298 K). It is noteworthy that only C-HScb and C-HOCg+ are considered in ref. 36 and 43. It would be consequently interesting to re-examine the photoelectron and IRMPD results in light of the present results which demonstrate the presence of conformers C-HScg-, C-HScb or C-HSag-.

**Methionine.** The neutral methionine potential energy surface has been examined by several groups at various levels of theory.\(^{30,33,48-51}\) Most of the time, a limited number of conformers has been considered. The most recent investigation reports results obtained at the B3LYP, B3P86 and MP2 levels using a 6-311+G(2df,2p) basis set and conclude that six conformers are lying in the first 5 kJ mol\(^{-1}\) of the \(H_{298}^0\) range.\(^{51}\)

In the present study, the 32 most stable conformers obtained at the B3LYP/6-31+G(d,p) level were further investigated with the whole panel of methods used here (Table T4 and Fig. S19 and S20 in ESI).\(^{\dagger}\) Fig. 10 presents the six most stable conformations in the \(G_{298}^0\) scale identified in the present work at the G4 level. These structures are unambiguously of type I: MetI g+g+g+, MetI g+ag-, MetI g+g+a, MetI g+ag+, MetI g+aa and MetI g+g+g- (Fig. 10). These structures are also among the most stable in the \(H_{298}^0\) scale but three type II conformers fall in the same 5 kJ mol\(^{-1}\) energy range: MetII ag-g+g+, MetII ag-g+a and MetII ag-g+g- (Fig. S20).\(^{33}\) All the computational methods agree in locating MetI g+g+g+ as the most stable conformer (Table T4 in ESI).\(^{\dagger}\) It is remarkable that this finding has been reported only in one previous study.\(^{33}\) The stability of structures MetI g+g+g+, MetI g+ag-, MetI g+g+a, MetII g+ag+, MetII g+aa and MetII g+g+g- is due to the NH2⋯O=C interaction (type I conformers) and also to a C(4)H⋯NH2 favourable interaction, attested by a small interatomic distance (\(d\approx 2.6\) Å), and due to the occurrence of a positive charge of ca. +0.2 on the hydrogen atom and a negative charge of ca. −0.7 on the nitrogen atom. The largest stability of MetI g+g+g+ may be attributed to an additional interaction between the C(2)H hydrogen (charge +0.22) and the sulfur atom (charge −0.22) their separating distance being only 2.9 Å while it attains more than 4 Å for the other MetI g+ conformers. Conformers of type II, MetII ag-g+a, MetII ag-g+g- and MetII ag-g+g+, present a NH⋯S hydrogen bonding type interaction (interatomic distance ~2.5 Å) which significantly reduces the backbone flexibility and thus the absolute entropy. The \(S_{298}^0\) values of MetII ag-g+a, MetII ag-g+g- and MetII ag-g+g+ are indeed close to 435 J mol\(^{-1}\) K\(^{-1}\) i.e. 15 J mol\(^{-1}\) K\(^{-1}\) less than the \(S_{298}^0\) values of type I conformers thus explaining why these latter are shifted.

![Fig. 9 Optimized geometries (B3LYP/6-31G(2df,p)) of the five most stable conformers of deprotonated l-cysteine (in brackets, relative \(H_{298}^0\) and \(G_{298}^0\) in kJ mol\(^{-1}\) calculated at the G4 level).](image-url)
to the low $G_{298}^*$ values. The theoretical population of neutral methionine conformers at 298 K calculated at the G4 level is Met$_{1}$g+g+g+/Met$_{1}$g+g-/Met$_{1}$g+g+a/Met$_{1}$g+ag+/Met$_{1}$g+aa/Met$_{1}$g+g+g+: 51.2/13.8/9.4/8.3/6.5%.

Neutral methionine has been studied experimentally by valence core photoelectron spectroscopy in the VUV and soft X-ray regions and by gas phase Fourier transform IR spectroscopy. However, observations were interpreted by computations on conformers situated more than 10–15 kJ mol$^{-1}$ above the most stable form Met$_{1}$g+g+g+. Re-examination of the data in the light of the present results is clearly due to the difference in basicity of both oxygens of the acidic group. Clearly, conformers MetH$_{1}$g+g–g– and MetH$_{1}$g–g–g+ are of identical stabilities and should be hardly distinguishable (Fig. 11). At the G4 level, the calculated relative population of conformers at 298 K, MetH$_{1}$g+g–g–/MetH$_{1}$g–g–g+/MetH$_{1}$g–g+g–/MetH$_{1}$g–g+a/MetH$_{1}$g+g–g+ is equal to 33.8/29.8/22.1/14.3%.

The 24 most stable conformers of deprotonated methionine were investigated with the total panoply of computational methods (Table T4 and Fig. S23 and S24 of ESI). At the G4 level, six structures were found to fall in the 0–5 kJ mol$^{-1}$ $G_{298}^*$ range (Met-H$_{1}$ag–g–, Met-H$_{1}$g–ag+, Met-H$_{1}$ag+g+, Met-H$_{1}$ag–ag, Met-H$_{1}$ag–ag–, Met-H$_{1}$ag–ag–, Met-H$_{1}$ag–ag–). Two additional structures (Met-H$_{1}$ag–g+, Met-H$_{1}$ag–g–) were found to be situated in the 0–5 kJ mol$^{-1}$ range of the $H_{298}^*$ scale (Fig. 12). All the investigated conformers present the NH–…–O=CO interaction with an interatomic distance of ~2.0 Å. It is generally observed that this distance is shorter for conformers of type Met-H$_{1}$ which, as a consequence, are more stable than conformers of type Met-H$_{2}$ (by 1 to 5 kJ mol$^{-1}$) (Fig. 12). Additional stabilizations are provided by interactions of the types (i) C(4)H–…–O=CO (Met-H$_{3}$ag–ag+, Met-H$_{4}$ag–ag+, Met-H$_{5}$ag–g–, Met-H$_{6}$ag–ag–, Met-H$_{7}$ag–ag–, Met-H$_{8}$ag–ag–, Met-H$_{9}$ag–ag–, Met-H$_{10}$ag+g+), (ii) C(4)H–…–NH$_{2}$ (Met-H$_{11}$ag–ag+, Met-H$_{12}$ag+g+, Met-H$_{13}$ag–ag–, Met-H$_{14}$ag–ag–, Met-H$_{15}$ag–ag–, Met-H$_{16}$ag–ag–).
Experimental. The general principle of measurement of gas-phase protonation, and deprotonation, thermochemistry is to determine the Gibbs free energy change associated with a proton transfer between the molecule M of interest and a reference base B_{ref}. Three general methods have been developed, (i) the equilibrium method, (ii) the kinetic methods (“simple” and “extended”) and (iii) the thermokinetic method.1–3

The first determinations of the gas phase basicities of Ser, Thr, Cys and Met came from the measurement of proton transfer equilibrium constants by ion cyclotron resonance mass spectrometry.52 In these experiments, the amino acid molecules M have been volatilized by heating the sample in a direct insertion probe close to the reacting region. No measurement of the temperature has been done but an estimate of 350 K has been proposed by Hunter & Lias.53 The derived GB values quoted in Table 1 take into account this temperature correction. Serine and methionine were also examined by the thermokinetic54 and the extended kinetic49,50 methods. These procedures lead to GB values in good agreement with those obtained by the equilibrium method. Average GB values are indeed associated with reasonable standard deviation (less than 3 kJ mol$^{-1}$).

The “simple” kinetic method is expected to provide an apparent proton affinity value given by PA_{app}(M) = \left[\frac{\text{PA}(M)}{T_{\text{eff}}{\Delta p^S_0(M) - \Delta p^S_0(B_{\text{ref}})}}\right]$ where T_{eff} is an “effective temperature” and \Delta p^S_0(X) = S^E(XH^+) - S^E(X) the “protonation entropy” of the species X (M or B_{ref}).3,53 In the “simple” kinetic method it is assumed that the experimentally determined

---

Fig. 12 Optimized geometries (B3LYP/6-31G(2df,p)) of the six most stable conformers + two conformers of deprotonated l-methionine (in brackets, relative $H^0_{298}$ and $G^0_{298}$ in kJ mol$^{-1}$ calculated at the G4 level).
### Table 1  Experimental protonation thermochernistry of the studied amino acids

<table>
<thead>
<tr>
<th>M</th>
<th>Method</th>
<th>GB(M) (kJ mol(^{-1}))</th>
<th>PA(M) (kJ mol(^{-1}))</th>
<th>(\Delta_p S(M)) (J K(^{-1}) mol(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serine</td>
<td>Equilibrium</td>
<td>880.3</td>
<td>910.4</td>
<td>911.4</td>
</tr>
<tr>
<td></td>
<td>Thermokinetic</td>
<td>876.2 ± 4.3(^d)</td>
<td>910.4</td>
<td>911.4</td>
</tr>
<tr>
<td></td>
<td>Simple kinetic</td>
<td></td>
<td>910.4</td>
<td>911.4</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>878.3 ± 2.9</td>
<td>911.6 ± 1.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evaluated</td>
<td>880.7(^{f}) – 874.3(^{f})</td>
<td>914.6 – 906.4(^{f}) – 5(^{f}) – 0(^{f})</td>
<td></td>
</tr>
<tr>
<td>Threonine</td>
<td>Equilibrium</td>
<td>888.5(^{e})</td>
<td>918.7(^{r})</td>
<td>922.9(^{e})</td>
</tr>
<tr>
<td></td>
<td>Simple kinetic</td>
<td></td>
<td>918.7(^{r})</td>
<td>922.9(^{e})</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>888.5</td>
<td>921.0 ± 2.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evaluated</td>
<td>888.5(^{i}) – 893.5(^{i})</td>
<td>922.5 – 925.9(^{i}) – 5(^{i}) – 0(^{i})</td>
<td></td>
</tr>
<tr>
<td>Cysteine</td>
<td>Equilibrium</td>
<td>868.8(^{s})</td>
<td>904.0(^{j})</td>
<td>901.9(^{j})</td>
</tr>
<tr>
<td></td>
<td>Simple kinetic</td>
<td></td>
<td>904.0(^{j})</td>
<td>901.9(^{j})</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>868.8</td>
<td>903.0 ± 1.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evaluated</td>
<td>869.3(^{k}) – 868.9(^{k})</td>
<td>903.2 – 901.4(^{k}) – 5(^{k}) – 0(^{k})</td>
<td></td>
</tr>
<tr>
<td>Methionine</td>
<td>Equilibrium</td>
<td>901.3(^{l})</td>
<td>928.4(^{m})</td>
<td>931.6(^{o})</td>
</tr>
<tr>
<td></td>
<td>Bracketing</td>
<td>899.0</td>
<td>927.4(^{n})</td>
<td>936.5(^{p})</td>
</tr>
<tr>
<td></td>
<td>Simple kinetic</td>
<td></td>
<td>927.4(^{n})</td>
<td>936.5(^{p})</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>898.5 ± 3.2(^{q})</td>
<td>937.5 ± 2.9(^{r}) – 22 ± 5(^{r})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evaluated</td>
<td>899.6 ± 1.5</td>
<td>931.7 ± 4.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Extended</td>
<td>898.5 ± 3.2(^{q})</td>
<td>937.5 ± 2.9(^{r}) – 22 ± 5(^{r})</td>
<td></td>
</tr>
</tbody>
</table>

*Data anchored to GB(NH\(_3\)) = 819.0 kJ mol\(^{-1}\) and PA(NH\(_3\)) = 853.6 kJ mol\(^{-1}\), and corrected to the Hunter & Lias\(^{55}\) basicity scale using linear correlation with the original data. \(^a\) \(\Delta_p S(M) = S'(MH^+) – S'(M)\). \(^b\) Ref. 52 (original values and, in brackets, as adapted by Hunter & Lias\(^{55}\) by assuming a temperature of 350 K rather than 320 K). \(^c\) Ref. 54. \(^d\) Ref. 57. \(^e\) Ref. 58. \(^f\) Ref. 59. \(^g\) Ref. 49. \(^h\) Ref. 50. \(^i\) Evaluated by Hunter & Lias. \(^j\) Ref. 60.

PA\(_{app}\)(M) is equal to the true PA(M). It consequently supposes that the term \(T_{eff}(\Delta_p S(M) – \Delta_p S(B_{ref}))\) can be neglected. In situations where this quantity is not negligible a more elaborated method such as the “extended kinetic method” is needed. The latter allows the determination of both parameters PA(M) and \(\Delta_p S(M) – \Delta_p S(B_{ref})\) by considering several sets of experiments corresponding to different effective temperatures \(T_{eff}\) i.e. to different ion activation conditions.\(^1\)

This latter method has been applied only to methionine and revealed a negative \(\Delta_p S(M)\text{(Met)}\) of ca. –20 J K\(^{-1}\) mol\(^{-1}\). It may be noted that the PA\(_{app}\) values obtained by the “simple” kinetic method for methionine from five different laboratories (Table 1) are ranging from 928 to 936 kJ mol\(^{-1}\) and present a standard deviation of 4.4 kJ mol\(^{-1}\). By comparison, the standard deviations associated to the PA\(_{app}\) values of Ser, Thr and Cys are situated between 1.3 and 2.1 kJ mol\(^{-1}\). The spread of PA\(_{app}\)(Met) values observed in Table 1 may be interpreted by different effective temperatures \(T_{eff}\) associated with the various experiments. The existence of a non negligible \(\Delta_p S(M)\text{(Met)}\) term would consequently lead to different PA\(_{app}\) values. By the same reasoning one would expect negligible \(\Delta_p S(M)\) values for the three amino acids Ser, Thr and Cys.

Gas phase acidities, \(\Delta_{acid} G\), of serine, threonine, cysteine and methionine have been determined by using the kinetic methods by O’Hair \(et\ al.\)\(^{55}\) and by Poutsma and coworkers.\(^{31}\) They anchor their results to the \(\Delta_{acid} G\) of reference molecules including substituted benzoic acids. In the most recent compilation of thermochemical data concerning negative ions,\(^{56}\) the \(\Delta_{acid} G\) values of the reference acids used by O’Hair \(et\ al.\)\(^{55}\) have changed. We thus reconsider their original experimental data and adjust the resulting \(\Delta_{acid} G\) to the new gas phase acidity scale.\(^{56}\) \(\Delta_{acid} H\) reported in Table 3 are obtained either from the sets of experimental \(\Delta_{acid} G\) obtained by the simple kinetic method\(^{31,55}\) assuming a constant \(TAS\) term of 29.5 kJ mol\(^{-1}\), or from data derived from the extended kinetic method.\(^{31}\) Comparison of the various \(\Delta_{acid} G\) and \(\Delta_{acid} H\) estimates (Table 3) shows a clear agreement from one method to the other with an average deviation equal to 2.7 kJ mol\(^{-1}\).

**Theoretical.** Monoconformer proton affinities and gas phase basicities, PA\(_{msc}\) and GB\(_{msc}\), of Ser, Thr, Cys and Met computed using the most stable conformers in the enthalpy scale and in the free energy scale, respectively are compared in Table 2. Consideration of the overall population of conformers at 298 K for both neutral and protonated species leads to averaged proton affinities, PA\(_{average}\), protonation entropies, \(\Delta_p S_{average}\), and gas phase basicities, GB\(_{average}\). The two latter quantities include the entropy of mixing (eqn (11)) in the “average” values. It should be underlined that “average” values of the thermochemical parameters are dependent on the number of conformers and consequently on the Gibbs free energy range considered. We adopt uniformly a cutoff of \(\Delta G = 6\) kJ mol\(^{-1}\) since, using eqn (7) at 298 K, any conformer more than 6 kJ mol\(^{-1}\) will possess a relative abundance less than 10% that of the most stable conformer.

It is generally observed that PA\(_{msc}\) constitutes a lower limit for the overall, averaged, quantities PA\(_{average}\). This is due to the larger number of conformers for the neutral species with respect to the protonated form in a given energy range. It results in a difference [\(\Delta S_{average}\)\(_{GB}\) – \(\Delta S_{average}\)\(_{PA}\)]\(_{msc}\) larger than [\(\Delta S_{average}\)\(_{GB}\) \(\Delta S_{average}\)\(_{PA}\)]\(_{msc}\) and consequently to PA\(_{average}\) > PA\(_{msc}\). The difference PA\(_{average}\) – PA\(_{msc}\) observed in the present study is equal to 2.2 ± 0.6 kJ mol\(^{-1}\) at the G4 level (Table 2). When looking at the gas phase basicity, the data presented in Table 2 show that GB\(_{average}\) is lower than GB\(_{msc}\) by 1.5 kJ mol\(^{-1}\) for Ser, Thr and Cys at the G4 level. It is noteworthy that the \(\Delta_p S_{average}\) of these three amino acids reduces to the entropy of mixing contribution: the intrinsic \(\Delta_p S\) differences calculated either in the monoconformer approximation or by considering the full population of conformers are, on average, close to zero. In the case of methionine the situation is different since a \(\Delta_p S\) of ca. –20 J K\(^{-1}\) mol\(^{-1}\) is calculated (both using the msc or average methods). As a consequence, GB\(_{average}\) is higher than GB\(_{msc}\) (by 1.4 kJ mol\(^{-1}\) at the G4 level).

It is now interesting to compare the thermochemical results obtained at the various levels of theory investigated here using the G4 method as a benchmark. Signed average deviations (SAD) on GB(M) and PA(M) are reported in Table T5 of the ESL.\(^{f}\) The largest average deviations are obtained using B97-D and M06-2X functionals. Accordingly, PA\(_{msc}\) calculated at the B97-D/6-311++G(3df,2p) level are always overestimated,
mainly for the sulfur containing molecules (average deviation: 10.4 kJ mol\(^{-1}\)) with a maximum deviation equal to 19 kJ mol\(^{-1}\) for methionine), while M06-2X/6-31+G(3df,2p) PA\(_{\text{msc}}\) are systematically underestimated by about 10 kJ mol\(^{-1}\). Interestingly enough, B3LYP/6-31+G(d,p) gives the correct PA\(_{\text{msc}}\) (within less than 2 kJ mol\(^{-1}\)). Excellent agreement is found between G3B3, G4MP2 and G4 with a maximum SAD of 4.4 kJ mol\(^{-1}\). Note that CBS-QB3 leads to PA\(_{\text{msc}}\) values slightly below that given by the G4 method (by ~2.5 kJ mol\(^{-1}\)). Similar conclusion arises from examination of the monomer gas phase basicities, GB\(_{\text{msc}}\).

Computed monomers and averaged \(\Delta a_{\text{acid}}G\), \(\Delta a_{\text{acid}}S\) and \(\Delta a_{\text{acid}}H\) are presented in Table 4. As done for the protonation thermochemistry, a cutoff of \(\Delta G_{\text{msc}}^{298} = 6\) kJ mol\(^{-1}\) has been applied to limit the number of conformers considered in the 298 K population averaging. Averaged \(\Delta a_{\text{acid}}H_{\text{average}}\) appears to be slightly lower than monomeric \(\Delta a_{\text{acid}}H_{\text{msc}}\) by ca. 2 kJ mol\(^{-1}\) (Table 4). As noted for proton affinities, this shift finds its origin in the larger number of conformers of the neutral molecules with respect to their ionized forms in the 298 K populations. The resulting increase in enthalpy of one mol of conformational mixture of neutral amino acid consequently reduces \(\Delta a_{\text{acid}}H\). Turning now to \(\Delta a_{\text{acid}}G\), the data presented in Table 4 show that \(\Delta a_{\text{acid}}G_{\text{average}}\) is slightly higher than \(\Delta a_{\text{acid}}G_{\text{msc}}\), the shift being however less than 1 kJ mol\(^{-1}\).

Taking the G4 results as benchmark, the maximum signed average deviation SAD on \(\Delta a_{\text{acid}}H\) is obtained with the B97-D/6-311++G(3df,2p) data (\(\Delta a_{\text{acid}}H_{\text{msc}} \sim +7\) kJ mol\(^{-1}\)) (Table T6 of ES1). A systematic underestimate of \(\Delta a_{\text{acid}}H_{\text{msc}}\)
Table 3 Experimental deprotonation thermochromy of aliphatic \( \varepsilon \)-amino acids

<table>
<thead>
<tr>
<th>M</th>
<th>Method</th>
<th>( \Delta_{\text{acid}}G(M) )</th>
<th>( \Delta_{\text{acid}}H(M) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serine</td>
<td>Simple kinetic( ^d )</td>
<td>1364.5</td>
<td>1394.0</td>
</tr>
<tr>
<td></td>
<td>Simple kinetic( ^c )</td>
<td>1365.0</td>
<td>1393.4</td>
</tr>
<tr>
<td></td>
<td>Extended kinetic( ^c )</td>
<td>1360.9</td>
<td>1391.0</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>1363.5 \pm 2.2</td>
<td>1393.1 \pm 1.8</td>
</tr>
<tr>
<td>Threonine</td>
<td>Simple kinetic( ^c )</td>
<td>1361.8</td>
<td>1391.3</td>
</tr>
<tr>
<td></td>
<td>Simple kinetic( ^c )</td>
<td>1363.8</td>
<td>1393.4</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>1360.2</td>
<td>1388.1</td>
</tr>
<tr>
<td></td>
<td>Extended kinetic( ^c )</td>
<td>1361.9 \pm 1.8</td>
<td>1390.9 \pm 2.7</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Equilibrium( ^c )</td>
<td>1370.3 \pm 8.8</td>
<td>1391.9 \pm 9.2</td>
</tr>
<tr>
<td></td>
<td>Simple kinetic( ^c )</td>
<td>1365.5</td>
<td>1395.0</td>
</tr>
<tr>
<td></td>
<td>Simple kinetic( ^c )</td>
<td>1369.8</td>
<td>1399.3</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>1366.5( ^{55} )</td>
<td>1395( ^{55} )</td>
</tr>
<tr>
<td></td>
<td>Extended kinetic( ^c )</td>
<td>1364.6(\pm 14)( ^{c} )</td>
<td>1392.9(\pm 14)( ^{c} )</td>
</tr>
<tr>
<td>Methionine</td>
<td>Simple kinetic( ^c )</td>
<td>1379.4</td>
<td>1408.9</td>
</tr>
<tr>
<td></td>
<td>Simple kinetic( ^c )</td>
<td>1384.8</td>
<td>1414.3</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>1380.7 \pm 3.6</td>
<td>1410.1 \pm 3.8</td>
</tr>
</tbody>
</table>

\( ^a \) Data anchored to \( \Delta_{\text{acid}}G(\text{benzoic acid}) = 1394.0 \) \( \text{kJ mol}^{-1} \) and \( \Delta_{\text{acid}}H(\text{benzoic acid}) = 1423.5 \) \( \text{kJ mol}^{-1} \) \( ^{15} \) (ref. 56). \( ^b \) Ref. 55, \( \Delta_{\text{acid}}G \) recalculated using the most recent \( \Delta_{\text{acid}}G \) of the reference acids. \( \Delta_{\text{acid}}H(M) \) were obtained by adding a uniform \( \Delta_{\text{H}} \) term equal to 29.5 \( \text{kJ mol}^{-1} \). \( ^c \) Using the data of ref. 31 obtained at 20\% attenuation. \( ^d \) \( \Delta_{\text{acid}}H(M) \) from ref. 31, \( \Delta_{\text{acid}}G(M) \) are calculated using the \( \Delta_{\text{p}}\text{p}(\text{H} - \text{H}) = S^0(\text{M}) - S^0(\text{M} - \text{H}) \) \( = 8 \pm 35, 16 \pm 20, 13 \pm 19 \) and \( 12 \pm 9 \) \( \text{kJ mol}^{-1} \) \( K^{-1} \) for Ser, Thr, Cys and Met, respectively (personal communication from JC Poutsma). \( ^{15} \) Ref. 43.

**Comparison experiment/theory.** Computed thermochemical data reported in Tables 2 and 4, either in the “msc” or the “average” approximations, are ideal quantities which should be compared carefully with experiment. It is a custom to compare data obtained by mass spectrometry techniques to computed “msc” values. However, the population of ions sampled experimentally does not necessarily consist of a pure single conformer or even to a thermal distribution of structures. It depends on how the ions are formed and handled during the experiments. In equilibrium or kinetic methods, proton exchange is occurring into proton bonded complexes stabilized by ca. 50–100 \( \text{kJ mol}^{-1} \). This deep energy well allows, in principle, the interconversion of conformers through rotational barriers along \( \sigma \)-bonds and consequently all thermodynamically accessible isomers to be present. Recently, evidence was presented on the validity of the “average” approach (including entropy of mixing) when considering proton transfer equilibrium experiments.\( ^{61} \) No such evidence has been reported for data obtained by using the kinetic method. A second point to consider is the ability of a given theoretical method to reproduce the thermochemical parameters of the standards to which the protonation/deprotonation thermochromy is anchored. As indicated in Tables 1 and 3, the reference standards are ammonia for the basicity scale and benzoic acid for the acidity scale. Computed \( \text{PA(NH}_3) \), \( \text{GB(NH}_3) \), \( \Delta_{\text{acid}}H(\text{benzoic acid}) \) and \( \Delta_{\text{acid}}G(\text{benzoic acid}) \) using the panel of methods used in the present work are reported in Table T9 of ESI together with the presently recognized reference values.\( ^{53,56} \) Deviations larger than 12 \( \text{kJ mol}^{-1} \) are observed on \( \text{PA(NH}_3) \) and \( \text{GB(NH}_3) \) whereas maximum deviations less than 6 \( \text{kJ mol}^{-1} \) are obtained for \( \Delta_{\text{acid}}H(\text{benzoic acid}) \) and \( \Delta_{\text{acid}}G(\text{benzoic acid}) \). Correction based on these deviations are applied to the computed \( \text{PA, GB, } \Delta_{\text{acid}}H \) and \( \Delta_{\text{acid}}G \) values and are indicated in parentheses in the data reported in Tables 2 and 4.

When comparing experimental and \( G^4 \) computed proton affinities reported in Table 2, a correct agreement appears using both monoconformer \( \text{PA}_{\text{msc}} \) or averaged \( \text{PA}_{\text{average}} \) values (\( \text{SAD}_{\text{msc}} = -1.4 \pm 1.7 \text{ kJ mol}^{-1} \) against \( \text{SAD}_{\text{average}} = 0.1 \pm 1.6 \text{ kJ mol}^{-1} \)). A clear illustration is given by Fig. 13 where \( G^4 \) calculated proton affinities are plotted against the experimental values. As evidenced in Fig. 13, the maximum deviation is observed for threonine, suggesting that the experimental value is too high by ca. 3 \( \text{kJ mol}^{-1} \). In the case of methionine, only the experimental \( \text{PA} \) value obtained by the extended kinetic method is meaningful since a significant protonation entropy is demonstrated for this molecule. Indeed, the calculated \( G^4 \) values (\( \text{PA}_{\text{msc}} = 937.2 \) and \( \text{PA}_{\text{average}} = 938.2 \) \( \text{kJ mol}^{-1} \)) agree nicely with experiment (\( \text{PA}_{\text{exp}} = 937.5 \pm 2.9 \) \( \text{kJ mol}^{-1} \)). Comparison of gas phase basicities is also correct when considering \( G^4 \) calculated \( \text{GB}_{\text{msc}} \) or \( \text{GB}_{\text{average}} \) and experimental data (\( \text{SAD}_{\text{msc}} = -0.6 \pm 2.0 \text{ kJ mol}^{-1} \) against \( \text{SAD}_{\text{average}} = -1.3 \pm 2.2 \text{ kJ mol}^{-1} \)). Again, the maximum deviation is observed for threonine and one may suggest that the experimental GB(threonine) is too high by ca. 3 \( \text{kJ mol}^{-1} \). Comparison between experimental and \( G^4 \) computed protonation entropy can be done only for methionine. It is observed (Table 2) that the \( \text{SAD}^0 \) value obtained by the extended kinetic method (\( \pm 5 \) \( \text{kJ mol}^{-1} \) \( K^{-1} \)) in excellent agreement with predictions based on the computations either using the msc or the average results (\( \pm 20 \text{kJ mol}^{-1} K^{-1} \) Table 2). No experimental information is available on the protonation entropy of serine, threonine and cysteine but, as recalled above, it is generally assumed negligible.\( ^{53} \) In agreement with this expectation, computed \( \text{SAD}^0 \) are generally limited to less than \( \pm 10 \) \( \text{kJ mol}^{-1} \) \( K^{-1} \); the mean msc and average values are 0.3 \( \pm 4.7 \) and \( -7.7 \pm 1.6 \) \( \text{kJ mol}^{-1} K^{-1} \), respectively.

Experimental \( \Delta_{\text{acid}}H \) of Ser, Thr, Cys and Met are generally slightly closer to the computed monoconformer \( \Delta_{\text{acid}}H_{\text{msc}} \) values (\( \text{SAD}_{\text{msc}} = 0.5 \pm 2.0 \) and \( \text{SAD}_{\text{average}} = -0.8 \pm 3.2 \text{ kJ mol}^{-1} \)) (Table 4 and Table T7 of ESI).\( ^{\dagger} \) Similar conclusions may be drawn when considering \( \Delta_{\text{acid}}G \) (\( \text{SAD}_{\text{msc}} = 0.5 \pm 1.6 \) and \( \text{SAD}_{\text{average}} = 1.2 \pm 0.9 \text{ kJ mol}^{-1} \)). In both cases, the maximum deviation is observed for methionine indicating that the experimental \( \Delta_{\text{acid}}H \) and \( \Delta_{\text{acid}}G \) values are probably too low by \( \pm 3 \text{ kJ mol}^{-1} \). Considering protonation entropy \( \text{SAD}^0 \) associated with gas phase acidity, values situated between 8 and 16 \( \text{kJ mol}^{-1} K^{-1} \) were obtained from extended kinetic method plots,\( ^{31} \) the corresponding 95% error was however of the same order of magnitude (see footnote c...
The present work was supported by an extensive search of the most stable conformers of neutral, protonated and deprotonated serine, threonine, cysteine and methionine. From a large investigation of more than around two thousand trial geometries based on systematic dihedral angle changes completed by Monte-Carlo and molecular dynamics simulations procedures using AMOEBA force field, no less than 131 structures were fully examined at various theoretical levels. These explorations involved three DFT (B3LYP, B97-D and M06-2X functionals using 6-31+G(d,p) optimized geometries and 6-31+G(3df,2p) single points) and four composite (CBS-QB3, G3B3, G4MP2 and G4) methods.

Comparison between these various computational methods leads to the following observations:

(a) Conformers’ relative energies calculated by the composite methods CBS-QB3, G3B3, G4MP2 and G4 by the M06-2X functional are generally nearly identical (within ~2 kJ mol⁻¹). Larger discrepancies are observed when using B97-D or B3LYP functionals, particularly for the neutral species.
The deviation with average PA and $\Delta_{\text{acid}}H$ is however limited to 1.5 kJ mol$^{-1}$.

(b) Computed $\Delta_S S'(M)$ are generally limited to less than $\pm 10$ J mol$^{-1}$ K$^{-1}$ except for methionine for which a value close to $-20$ J mol$^{-1}$ K$^{-1}$ is calculated (both in the msc and the average estimates) in agreement with experiment.

(c) Assignment of the structure(s) (or mixtures of structures) of the most stable neutral conformer(s) is in correct agreement with experiments based on microwave spectra as attested for serine$^{27}$, threonine$^{40}$ and cysteine$^{44}$. By contrast, interpretation of the experimental IR$^{47}$ and VUV photoelectron$^{48}$ spectra of neutral methionine should be reconsidered by taking into account the structure of the most stable conformers identified here. A similar remark applies to the IRMPD spectra of protonated serine and deprotonated cysteine. In the former case, only one structure (namely SH$^-$) has been considered to interpret the spectrum$^{35}$, our results show that a second conformer should also participate (namely SH$^+$). Its impact on the observed absorption bands is however expected to be negligible since both SH$^-$ and SH$^+$ exhibit the same characteristic frequencies in the considered spectral zone. Concerning deprotonated cysteine, our computational results confirm that the thiolate form is the most stable$^{36,43,46}$.

Acknowledgements

The authors would like to thank Bhawani Singh Inda for preliminary calculations, Ashwani Sharma and Carine Clavaguéra for their help during the setup of the computational procedures. This work was performed using HPC resources from GENCI-CINES (grant 2011-c2011085107).

References
