PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Serine *O*-sulfation probed by IRMPD spectroscopy

Roberto Paciotti,[1] Cecilia Coletti,[1] Nazzareno Re,[1] Debora Scuderi,[2] Barbara Chiavarino,[3] Simonetta Fornarini,[3,*] and Maria Elisa Crestoni[3,*]

[1] *Dipartimento di Farmacia, Università G. D'Annunzio, Via dei Vestini 31, I-66100 Chieti, Italy*
[2] *Laboratoire de Chimie Physique d'Orsay, Universitè Paris Sud, UMR8000 CNRS, Facultè des Sciences, Bât 350, 91405 Orsay Cedex, France*
[3] *Dipartimento di Chimica e Tecnologie del Farmaco, Università degli Studi di Roma La Sapienza, P.le A. Moro 5, I-00185 Roma, Italy*

Corresponding authors:
*Tel: +39 06 4991 3596. Fax: +39 06 4991 3602.
Email: mariaelisa.crestoni@uniroma1.it
*Tel: +39 06 4991 3510. Fax: +39 06 4991 3602.
Email: simonetta.fornarini@uniroma1.it

Keywords:
Post-translational modification
Sulfoserine
IR spectroscopy
Density functional theory calculations
Molecular Structure

Abstract

Sulfation of amino acids is a frequent post-translational modification. It is highly labile, though, and characterizing it by mass spectrometry, an otherwise powerful and widely exploited tool in analytical proteomics, is a challenge. The presently reported study is aimed at revealing the O-sulfation of L-serine and elucidating the effects of protonation and deprotonation on the structure and stability of the ensuing ionic species, $[sSer+H]^+$ and $[sSer-H]^-$. These ions are obtained as gaseous, isolated species by electrospray ionization, trapped in a Paul ion-trap, and sampled by IR multiple photon dissociation (IRMPD) spectroscopy in either the 750-1900 cm^{-1} fingerprint range, or the 2900 and 3700 cm^{-1} range encompassing the N-H and O-H stretching modes. The recorded IRMPD spectra present diagnostic signatures of the sulfate modification that are missing in the spectra of the native serine ions, $[Ser+H]^+$ and $[Ser-H]^-$. The experimental IRMPD features have been interpreted by comparison with the linear IR spectra of the lowest energy structures that are likely candidates for the sampled ions, calculated at the M06-2X/6-311+G(d,p) level of theory. Evidence is gathered that the most stable conformations of $[sSer+H]^+$ are stabilized by hydrogen bonding interactions between the protonated amino group and both the carbonyl and sulfate oxygens. $[sSer-H]^-$ ions possess a negatively charged sulfate group involved in either a $S{=}O\cdots HN$ or

a S=O···HO hydrogen bond. The experimental IRMPD spectra are consistent with the presence of multiple low-lying structures in a thermally equilibrated population of several species particularly in the case of $[sSer-H]^-$ ions, where the high structural flexibility combined with the presence of a negative charge favors the co-existence of several different H-bonding motifs.

Introduction

Sulfation is a widespread post-translational modification (PTM) which involves covalent attachment of a sulfate group. This PTM may affect several endogenous acceptors, such as proteins, steroids, sugars as well as xenobiotics.[1] Consequently, in addition to the most common O-linkage involving alcoholic or phenolic functions (O-sulfation)[2-4], also other sites may be sulfated, including amides (N-sulfation) and thiols (S-sulfation).[5] Sulfation events are thought to modulate protein-protein interactions, in relevant processes like hemostasis and leukocyte adhesion in inflammatory response,[6-8] and, similarly to phosphorylation, growth regulation in cancer.[9] Considerable efforts have been focused on identifying, quantifying, and localizing sulfation sites in peptides and proteins, an issue of crucial relevance for the biological function.[5,10] Varied spectroscopic approaches, including (heteronuclear) NMR, UV, and FT-IR spectroscopy have been widely used and proven valuable for the characterization and quantification of purified sulfopeptides.[11-12] Mass spectrometry, which in recent years has become the method of choice for the analysis of protein PMTs, by virtue of its high sensitivity and ability to cope with the heterogeneity of biological samples,[5] has been exploited for differentiating sulfation and phosphorylation events on side chains, either present individually[13] or simultaneously.[14] The latter result, obtained only by electrospray ionization (ESI) coupled to Fourier transform ion cyclotron resonance (FT-ICR) mass analysis, has allowed to directly distinguish sulfopeptides from the isobaric phosphopeptides (Δm = 9.5 mDa) with baseline resolution in a complex proteolytic digest.

Also conventional tandem mass spectrometry (MS/MS) has proven useful for differentiating phosphorylation from sulfation, yielding positively charged fragments in low-energy collision induced dissociation (CID) experiments involving a characteristic neutral loss of 98 Da [H_3PO_4] and of 80 Da (SO_3), respectively.[4] However, the high sensitivity of sulfopeptides to acid hydrolysis, combined with the presence of very acidic sites and the labile nature of the sulfoester bond severely complicate the characterization of sulfation sites by MS. As a result, cleavage of the sulfation PTM before backbone fragmentation is observed in most cases, yielding CID spectra basically coincident

with those of the native species, and so preventing one to assign the modification site.[15] Partial retention of the sulfoester group in fragment ions may nevertheless be achieved in negative ion mode CID assays, once again producing similar PTM losses from phospho- and sulfopeptides. To circumvent the poorly informative neutral loss exhibited by positive ion CID MS/MS experiments, the application of gentler activation techniques, like metal-assisted electron-capture dissociation (ECD)[13] or electron transfer dissociation (ETD)[16] has been explored and found to afford an improved retention of the sulfation motif. However, these techniques work only with multiply charged positive ions.[17] More successfully, recently established negative ion mode activation techniques (niECD and NETD) are reported to display complete retention of PTMs and extensive sequence characterization for singly and doubly deprotonated sulfopeptides and, in general, acidic biomolecules.[18,19] Besides, spectroscopic approaches have been recently applied to identify sulfation vs phosphorylation in tyrosine containing peptide ions providing diagnostic sequence information in anions using 193 nm ultraviolet photodissociation.[20] Alternatively, the presence of characteristic side-chain OH stretches could be revealed by means of infrared multiple photon dissociation (IRMPD) spectroscopy.[21]

IRMPD spectroscopy is indeed a sensitive structural probe which relies on the coupling of a tunable source of IR light with tandem mass spectrometry. It is an "action" spectroscopy based on the photofragmentation process triggered by the sequential absorption of resonant IR photons.[22-26] In recent years, the potential of IRMPD spectroscopy to complement and support important inferences obtained by conventional mass spectrometry has allowed to shed light on inherent molecular properties and structural features of a variety of mass-selected (bio)molecular ions in an isolated environment. In this medium bulk solvent, matrix and counterion interferences are absent and one can take advantage of the direct comparative analysis between experimental IRMPD and calculated IR spectra. Implemented in a number of investigations, this powerful tool has been exploited to address issues such as the characterization of synthetic and natural (modified) amino acids and peptides[27-37] and of their fragmentation pathways,[38,39] (de)protonation of (hetero) aromatic molecules,[40-43] metal cation binding patterns,[44-51] as well as the discrimination between isomers and conformers of a variety of biomolecular ions.[34,36,52-54] A clue into the conformational landscape of fairly large and flexible molecular ions has also been attained, disclosing the groups entangled in intramolecular hydrogen bond networks and the impact of PTMs on the higher order structure of the native species.[21,54-56] Also, the sulfate motif is not novel to IRMPD inquiry because simple inorganic ions and clusters such as protonated sulfuric acid, hydrated sulfate dianion, and hydrogen

sulfate complexes, have been sampled for their vibrational features, most significant modes being associated to S-O and O-H stretching and SOH bending.[57-60]

In the present contribution, a study is reported that combines MS coupled to IRMPD spectroscopy with extensive density functional theory (DFT) calculations. This approach is intended to evaluate the influence brought about by the sulfation PTM on the geometric determinants of electrosprayed (de)protonated O-sulfoserine. The (de)protonated serine ions, already assayed by IRMPD spectroscopy, are re-examined here as native species to offer a useful guide in identifying distinctive signatures revealing the presence of this modification. These diagnostic features may serve as a probe for O-sulfation in more complex serine-containing biomolecules, providing a differentiation from the related phosphorylation variants.

Experimental and theoretical methods

L-serine, L-serine O-sulfate potassium salt, and all the solvents used in this work were commercial products (Sigma-Aldrich), used as received. The experimental setup and general procedure of IRMPD experiments, based on the coupling of a tunable IR laser source with a modified ion trap mass spectrometer, has been described in detail previously.[61] Briefly, deprotonated L-serine, $[Ser-H]^-$ (m/z 104), and L-serine-O-sulfate, $[sSer-H]^-$ (m/z 184), ions were generated by ESI of a 3 μM solution of L-serine and L-serine O-sulfate potassium salt, respectively, in water/methanol (1:1), whereas protonated L-serine, $[Ser+H]^+$ (m/z 106), and L-Ser-O-sulfate ions, $[sSer+H]^+$ (m/z 186), were obtained by ESI of a 2 μM solution of the former precursors in water/methanol (1:1), with 2% formic acid to assist protonation.

IRMPD spectroscopy

The IRMPD spectra of mass-selected (de)protonated ions from L-serine and L-serine-O-sulfate have been recorded employing two different experimental setups. The fingerprint range (750-1900 cm^{-1}) was explored using a commercial Paul-ion trap mass spectrometer (BrukerEsquire 3000+) coupled to the bright and tunable IR radiation (5-25 μm) of the free electron laser (FEL) beamline at the Centre Laser Infrarouge d'Orsay (CLIO),[62] which is generated by a 10–50 MeV electron linear accelerator and delivered in 8μs-long macropulses at a repetition rate of 25 Hz, each containing 500 micropulses (1–2 ps long). Typical macropulse energies are 40 mJ.

For the present study, the electron energy was set at 40 MeV and at 44 MeV in order to optimize the laser power in the frequency region of interest and a fairly stable power around 800–900 mW was ensured, so that no correction for laser power variation has been considered. In order to avoid saturation of the most pronounced bands, IRMPD spectra were also obtained using an attenuator that decreases the power by a factor of three. The IR-FEL spectral width (fwhm) was less than 0.5% of the central wavelength. In the trap, mass-selected ions were typically accumulated for 5 ms, and irradiated with 10 macropulses.

A tabletop IR optical parametric oscillator/amplifier (OPO/OPA; LaserVision, Bellevue, WA) laser system coupled to a modified Paul type quadrupole ion trap mass spectrometer (Bruker Esquire 6000+)[63], assembled at the Università di Roma "La Sapienza", has been used to investigate the 2900-3700 cm^{-1} wavenumber range for mass-selected $[Ser+H]^+$, $[sSer+H]^+$ and $[sSer-H]^-$ ions. The parametric converter is pumped by a Nd:YAG laser (Continuum Surelite II) running at 10 Hz and delivering 600 mJ/pulse (4-6 ns long). In the investigated spectral region the typical output pulse energy from the OPO/OPA laser source was 20-24 mJ/pulse, with a spectral bandwidth of about 3-4 cm^{-1}. In the ion trap, ions are accumulated for 20 ms and mass selected prior to IR irradiation for 0.5-1.0 s. The laser wavelength was continuously moved at a speed of 0.1 cm^{-1} s^{-1}, and the mass spectrum was typically obtained from an accumulation over four scans.

When the IR beam is resonant with an active vibrational transition, sequential absorption of multiple photons can activate a photofragmentation process, yielding a structurally diagnostic "action" spectroscopy. The ensuing IRMPD spectrum is obtained by plotting the photofragmentation yield R ($R = -\ln[I_P/(I_P + \sum I_F)]$, where I_P and I_F are the abundances of the parent ion and of a fragment ion) as a function of the radiation wavenumber.[64] In all acquired spectra, a quantitative consistency is verified between the decrease of the parent ion signal and the rise of the product ions signals.

The good inter-setup reproducibility is shown by the comparison of the IRMPD spectra of the protonated O-phosphotyrosine (m/z 262), used as calibrant for the optimization of the experimental conditions, recorded on the two mentioned experimental platforms over the range of 3250-3700 cm^{-1} (Fig S1).

Computational details

The identification of the most stable isomers and conformers of (de)protonated L-serine-O-sulfate ions was carried out through a preliminary conformational analysis. For each possible isomer a

conformational search was performed with Macromodel 9.6[65] by using a Multiple Minimum MonteCarlo method with AMBER as force field. 1000 conformations for each isomer were minimized, and only those within 25 kJ/mol from the global minimum were recorded. Similar conformations were grouped by using XCluster[65] and visually analyzed to devise additional conformers escaped from the conformational search. All conformations obtained from this analysis were then fully optimized by means of density functional theory calculations using the G09 suite of program.[66] Preliminary calculations performed with two different functionals which have demonstrated to perform well in the calculation of harmonic frequencies,[67] namely B3LYP and M06-2X,[68] show that the latter functional is able to better reproduce the experimental peaks (Fig. S2), most probably because it describes more accurately non-covalent interactions, which play a crucial role in these structures. In addition, among several tested basis sets (also shown in Fig. S2), 6-311+G(d,p) was found to reproduce the experimental behavior and no improvement was observed by adding diffuse functions on the hydrogen atoms, as in 6-311++G(d,p), so that the former set was employed due to its lower computational cost. Thus, the M06-2X/6-311+G(d,p) level of theory was used for the simulation of all spectra.

Harmonic vibrational frequencies, performed also to verify the stationary points as local minima, were calculated within the double harmonic approximation for the most stable conformations of each isomer, namely the ones comprised within 10 kJ/ mol relative to the global minimum structure. Enthalpy corrections and entropies were evaluated at the same level of theory. Single point energy calculations at MP2/aug-cc-pVTZ level were also carried out to obtain more accurate values of electronic energies. The relative energies at 0 K, enthalpies and free energies at 298 K were thus evaluated for each conformer.

Harmonic vibrational frequencies were scaled by 0.940, the best fit scale factor obtained to quantitatively reproduce the experimental frequencies. In fact, this is the same factor determined by Truhlar and co-workers[69] in their database for the very close M06-2X/6-31+G(d,p) level of theory. The frequencies involving the S-O and S=O stretching modes of the sulfate group, however, were left unscaled, as suggested by previous evidence.[32, 33, 37, 70, 71]

For consistency with experimental spectral resolution, the calculated IR bands were convoluted with a Gaussian profile with an associated width (fwhm) of 15 cm^{-1} (5 cm^{-1}) in the 750-1900 cm^{-1} (2900-3700 cm^{-1}) frequency range.

Results and Discussion

Photodissociation and CID mass spectra of sulfoserine ions

Electrospray ionization of L-serine-O-sulfate in aqueous methanol solution yields (de)protonated $[sSer-H]^-$ and $[sSer+H]^+$ ions that have been mass-selected, trapped and exposed to IR radiation from two distinct, tunable laser sources. The sequential absorption of multiple, resonant photons and intramolecular vibrational redistribution provide a slow heating of the sampled species, ultimately promoting dissociation along the lowest energy pathway. The photofragmentation of $[sSer-H]^-$ ions at m/z 184 in the IR fingerprint region involves the loss of $[C_3,H_5,N,O_2]$, which leads exclusively to hydrogen sulphate ion at m/z 97 (a fragment highly informative of the PTM), without any contribution of a skeletal cleavage by CH_2O loss, as detected for $[Ser-H]^-$ions.[31] Fig. S3 compares exemplary mass spectra recorded upon mass selection and storage of $[sSer-H]^-$ ions either prior to (upper trace) or following (bottom trace) irradiation with CLIO FEL light on resonance at 1206 cm^{-1}. The same fragmentation behavior (with formation of hydrogen sulfate at m/z 97, as the only observed fragment) is found when $[sSer-H]^-$ ions are assayed by low-energy CID, as commonly observed when deprotonated sulfopeptides are probed by tandem mass spectrometry.

IRMPD spectroscopy has been also performed on protonated L-serine-O-sulfate ion, $[sSer+H]^+$, in both the 900-1900 cm^{-1} and the 2900-3700 cm^{-1} ranges, using two different laser sources and experimental setups. The dissociation pathways observed upon resonant IR excitation lead to the formation of fragment ions at m/z 88 and m/z 106, as major channels arising from neutral PTM losses of H_2SO_4, or $H_2O + SO_3$, and SO_3, respectively. Albeit in minor amount, a fragment at m/z 140, by loss of $H_2O + CO$, is also found. Fig. S4 shows the mass spectrum when selected $[sSer+H]^+$ ions are stored in the Paul ion trap either prior to or following irradiation by OPO/OPA IR photons tuned at 3572 cm^{-1}. Recently, IRMPD spectroscopy has been used to elucidate the structure of the sulfate-loss product ions at m/z 88 and m/z 106, fragment ions of potential sulfoproteomic interest.[72] While the fragment ion due to loss of sulfuric acid (m/z 88) has been identified as an aziridine ion, so finding a counterpart in the kinetically controlled photodissociation of the isobaric protonated phosphoserine parent ion,[39] the fragment ion at m/z 106 corresponds to protonated serine. The same dissociation products arising from removal of sulfate modification are consistently generated when $[sSer+H]^+$ ions are probed by low-energy CID, conforming to the well-established propensity of protonated sulfotyrosine-containing peptides to lose the sulfoester group.[10]

For comparison purposes, the parent amino acid lacking the sulfation PTM has been also probed by IRMPD spectroscopy. The fingerprint range in the case of both protonated, $[Ser+H]^+$ (m/z 106),[28, 72] and deprotonated, $[Ser-H]^-$ (m/z 104),[31] ions, as well as the IRMPD spectrum of $[Ser+H]^+$ in the

NH/OH stretching region[72] were already explored. These IRMPD spectra are re-examined here to help assigning the IR absorptions due to the sulfation PTM. While the most abundant photofragment ion of $[Ser+H]^+$ appears at m/z 88 and involves loss of water, the dominant fragmentation pathway of $[Ser-H]^-$ leads to a product ion at m/z 74 by elimination of CH_2O.

IRMPD spectroscopy of deprotonated sulfoserine

A structural characterization of $[sSer-H]^-$ ions was sought by IRMPD spectroscopy in the so-called fingerprint region (750-1900 cm^{-1}), where the powerful FEL beamline at CLIO may reveal the sulfate modification through the presence of characteristic features, such as the symmetric and antisymmetric SO_3 stretches.

The dependence of the IRMPD yield of $[sSer-H]^-$ on the wavenumber of the incident radiation is presented in Fig. 1 (black profile). It shows two prominent, only partially resolved absorbances centred at 1207, and 1297 cm^{-1}, along with a band at 1407 cm^{-1}, two sharp peaks at 1008, and 1732 cm^{-1}, and a weak resonance at 1057 cm^{-1}. Interestingly, when compared with the spectrum of the deprotonated native amino acid, previously reported in the 600-1650 cm^{-1} range,[31] and re-examined here (Fig. 1, green profile), notable differences emerge. The IRMPD spectrum of the native amino acid, $[sSer-H]^-$, is in fact dominated by the strong absorbances due to the combined NH_2 scissor mode combined with the antisymmetric (1610 cm^{-1}) and the symmetric carboxylate (1330 cm^{-1}) stretching modes.

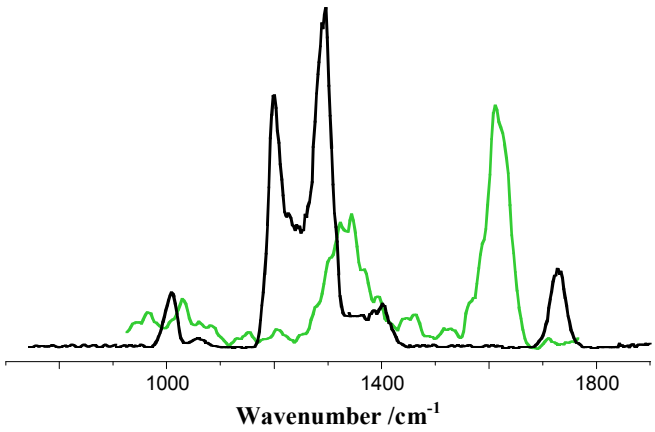

Fig 1 Experimental IRMPD spectra of $[sSer-H]^-$ (black profile) and $[Ser-H]^-$ (green profile) recorded in the fingerprint range.

Conversely, in the IRMPD spectrum of $[sSer-H]^-$, the bands centered at 1207, 1297, and 1732 cm^{-1} may be ascribed to the antisymmetric SO_3 and carboxylic C=O stretches. The latter C=O mode at 1732 cm^{-1} indicates that deprotonation occurs at the sulfate site, although its frequency is markedly red shifted with respect to a typical free carbonyl stretch, which commonly occurs around 1800 cm^{-1}. The observed frequency thus points to a strong intramolecular hydrogen bond that weakens the carbonyl double bond in $[sSer-H]^-$. The high energy portion of the spectrum was also carefully explored in the 2900-3700 cm^{-1} range. However, the search for diagnostic features in the structurally informative NH and OH stretching region was unsuccessful, because the IRMPD spectrum of $[sSer-H]^-$ ions appeared remarkably flat and no activity was observed throughout the whole range explored (Fig. S5). This result may be traced to the low intensity (less than 30 km/mol) of the free NH stretches above 3300 cm^{-1}, as shown in the calculated spectra in Fig. S5. In the wavenumber region below 3300 cm^{-1}, highly active bands are predicted. However, the related stretching modes pertain to NH bonds engaged in strong hydrogen bonding which is known to adversely affect the IR signal, causing dramatic broadening and weakening effects in IRMPD spectroscopy.

Computed Structures of [sSer-H]⁻

In order to gain insight into the IRMPD spectrum, which mainly reflects the absorption of the first resonant IR photon, an extensive search has been performed to identify the most stable structures of $[sSer-H]^-$ ions. Several isomers can be envisaged arising from deprotonation at either the sulfate or the carboxylic groups, and a large array of conformers result from rotations about single bonds in the highly flexible skeleton involving a variety of many possible, different hydrogen bonding motifs. The relevant thermodynamic data, including the relative enthalpy (ΔH_{rel}°) and free energy (ΔG_{rel}°) values at 298 (kJ mol^{-1}), calculated at MP2/aug-cc-pVTZ//M06-2X/6311+G(d,p), are collected in Table S1. Henceforth, the relative enthalpy values are given in plain text, whereas the relative free energies (kJ mol^{-1}) at 298 K are given in parentheses. According to theoretical calculations, the sulfate group is the most favorable deprotonation site and twelve low-energy conformers have been identified within a 12.6 (10.8) kJ mol^{-1} window. A few representative optimized geometries determined at the M06-2X/6-311+G(d,p) level are illustrated in Fig. 2 (the six most stable structures **S*t*-1, S*t*-2**, **S*t*-3, S*t*-4**, **S*t*-5, S*t*-6**) and in Fig. S6 (the following six **S*t*-7**, **S*t*-8**, **S*t*-9**, **S*t*-10**, **S*t*-11**, **S*t*-12**), together with the relative entalpy and free energy value at 298 K (kJ mol^{-1}). The values of some selected dihedral angles are listed in Table S2.

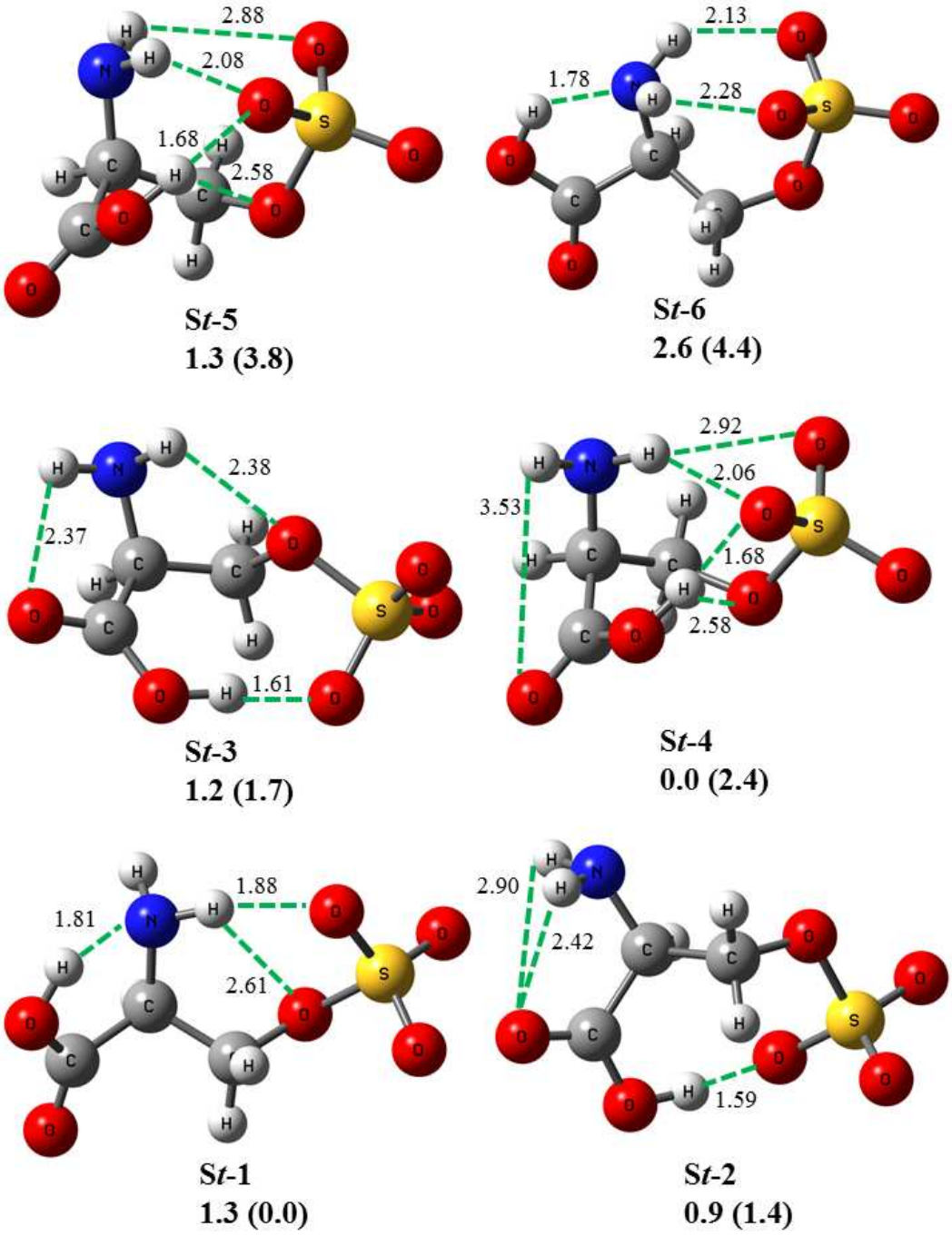

Fig 2 Optimized geometries of the most stable conformers (**S*t*-1, S*t*-2**, **S*t*-3, S*t*-4**, **S*t*-5, S*t*-6**) of deprotonated L-serine-O-sulfate [sSer-H]$^-$ determined at the M06-2X/6-311+G(d,p) level. Relative enthalpy and free energy (in parentheses) values (kJ mol^{-1}, in parentheses) at 298 K are computed at the MP2/aug-cc-pVTZ level. Interatomic bond distances, marked by dashed lines, are given in Å.

The fact that rotamers **S*t*-1, S*t*-2**, **S*t*-3, S*t*-4**, **S*t*-5**, and **S*t*-6** are very close in free energy (lying within 5 kJ mol^{-1}), with differences well within the accuracy of the method (4 kJ mol^{-1} at best) and that their ordering in terms of relative level is different if we consider enthalpy or free energy, suggests that several species should contribute to the Maxwell-Boltzmann averaged ion population at 298 K, and that their weight could be comparable.

The six most stable structures **S*t*-1, S*t*-2**, **S*t*-3, S*t*-4**, **S*t*-5**, and **S*t*-6**, comprised within 2.6 (4.4) kJ mol^{-1} window, present the sulfate group rotated so as to benefit from the stabilization of the excess negative charge via formation of an intramolecular hydrogen bond with either the amine NH (in **S*t*-1**, and **S*t*-6**), the carboxylic OH (in **S*t*-2**, and **S*t*-3**) or both (in **S*t*-4**, and **S*t*-5**) groups. In the carboxylic group, in a consistently *trans* configuration, the carbonyl can be almost free (in **S*t*-1**, **S*t*-4**, **S*t*-5**, and **S*t*-6**) or directed towards HN (in **S*t*-2**, and **S*t*-3**), thus establishing a C=O···H-N hydrogen bond. In addition, some structures may benefit from a stabilizing interaction of the sulfoester oxygen with a hydrogen atom belonging either to the amine (in **S*t*-1**, and **S*t*-3**) or the carboxylic (in **S*t*-4**, and **S*t*-5**) functionality. The lowest energy conformer **S*t*-1** of $[sSer-H]^-$ adopts a rather unfolded conformation stabilized by three hydrogen bonds formed between the amino hydrogen with both a negatively charged oxygen of the sulfate ($r_{SO...H-N}$ =1.88 Å) and the sulfoester oxygen ($r_{C3O...H-N}$ =2.61 Å), and between the COH donor and the amino nitrogen acceptor ($r_{N...H-O}$ =1.81Å). The latter binding motif results in a small value of the ∠O(H)C1C2N dihedral angle (-6.8° for **S*t*-1**), in marked contrast with rotamers **S*t*-2**, **S*t*-3**, **S*t*-4**, and **S*t*-5**, where the OH group interacts by hydrogen bonding with the sulfate group (Table S2). Comparison between the most stable conformer **S*t*-1** and conformer **S*t*-4**, lying 2.4 kJ mol^{-1} higher in free energy, suggests that the rather compact arrangement of the latter species may likely contribute to its favorable enthalpy, 1.3 kJ mol^{-1} lower with respect to **S*t*-1**.

As previously noted, other species, including isomer **S*c*-1** with a *cis* geometry of the carboxylic group, or deprotonated at alternative sites, such as the zwitterionic isomer **ZW-1**, bearing a proton at the amino group and deprotonated at both the carboxylic and sulfate sites, and the carboxylate structure **C-1**, lie 31.0, 14.4, and 149.2 kJ mol^{-1} higher in free energy at 298 K relative to the global minimum **S*t*-1**, respectively (Fig. S7).

Spectral assignment of [sSer-H]⁻

For comparative purposes, the IRMPD spectrum of $[sSer-H]^-$ is presented along with the calculated IR spectra of rotamers **S*t*-1**, **S*t*-2**, **S*t*-3**, **S*t*-4**, and **S*t*-5**, the most stable structures in free energy and enthalpy (Fig. 3).

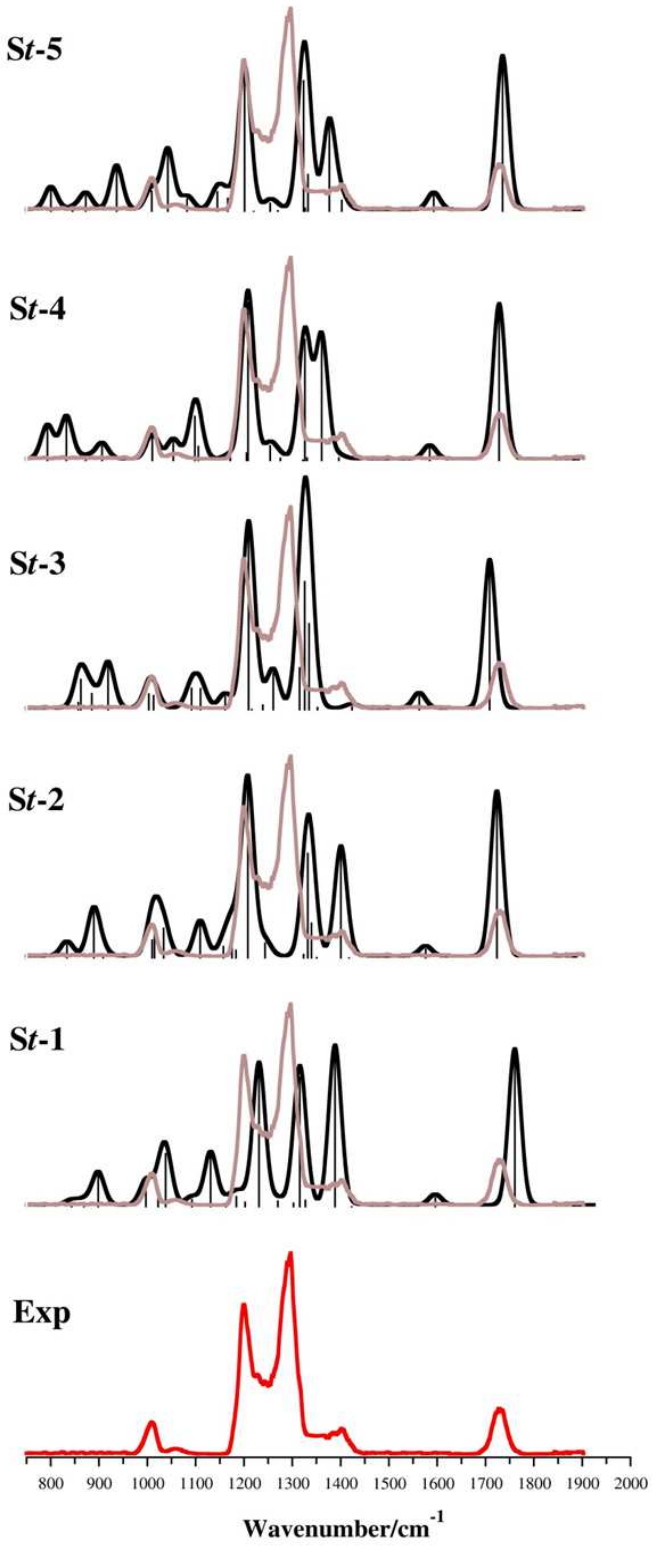

Fig. 3 Experimental IRMPD spectrum of deprotonated L-serine-O-sulfate [sSer-H]$^-$ (bottom), and computed IR spectra (at the M06-2X/6-311+G(d,p)level) of conformers **S*t*-1, S*t*-2**, **S*t*-3, S*t*-4**, and **S*t*-5**. The experimental spectrum is also reported as a pale trace over the calculated spectra to enable immediate comparison.

The position of observed IRMPD bands and the computed transitions of the lowest lying structure **S*t*-1** and **S*t*-4**, respectively, together with a concise description of the assigned vibrational modes are given in Table 1. The calculated IR absorption spectra of the optimized structures **S*t*-6**, **S*t*-7**, **S*t*-8**, and **S*t*-9**, as well as **S*t*-10**, **S*t*-11**, **S*t*-12** are reported in Fig. S8 and S9, respectively, together with the experimental IRMPD spectrum. Inspection of Fig. 3 reveals that the main experimental features are rather well reproduced by IR spectra of all low-lying conformers **S*t*-1**, **S*t*-2**, **S*t*-3**, **S*t*-4**, and **S*t*-5**, with small variations from one conformer to another to be ascribed to the different hydrogen bonding patterns. Here, frequencies pertaining to conformer **S*t*-1** (the lowest in free energy) and conformer **S*t*-4** (the lowest in enthalpy), as representatives of two different hydrogen bonding motifs, will be explicitly discussed.

The band observed at 1732 cm^{-1} presents a slight tailing on the blue side and can be attributed to the C=O stretching mode, predicted at 1740 cm^{-1} for structure **S*t*-1**, and at 1728 cm^{-1} for **S*t*-4** and **S*t*-5**, all bearing an almost free CO group. However, when the CO acceptor is involved in a hydrogen bond with the amino donor, the carbonyl stretch is expected as low as 1708 cm^{-1} (conf **S*t*-2** and **S*t*-3**). In fact, a closer look at the IRMPD band reveals that the peak is structured, a feature that could be ascribed to the presence of different conformers in the probed ion population, presenting slight variations for the absorption relative to the C=O stretching mode. Furthermore, the latter experimental frequency, is in very good agreement with a related band previously observed at 1728 cm^{-1} for deprotonated O-phosphoserine.[33] This finding is a clear indication of sulfate deprotonation in $[sSer\text{-}H]^-$, at variance with the carboxylate structure of deprotonated native serine which displays a strong antisymmetric OCO stretch at 1610 cm^{-1} (Fig. 1). Moreover, any significant contribution of the less stable isomers **ZW-1** and **C-1** can be ruled out. In fact these isomers present highly active antisymmetric carboxylate stretches below 1700 cm^{-1} (Fig. S10) where $[sSer\text{-}H]^-$ shows hardly any vibrational activity. The multiplet of peaks in the frequency range from 1200 to 1300 cm^{-1} correlates with the intense antisymmetric SO_3 stretch coupled to NH_2 twisting, the weak C-OH and the strongly active antisymmetric SO_2 stretches, predicted for **S*t*-1** at 1231, 1270, and 1315 cm^{-1}, and at 1208, 1254, and 1326 cm^{-1} for **S*t*-4**, respectively. Vibrational modes at similar values are observed in the calculated IR spectra of **S*t*-2**, **S*t*-3**, and **S*t*-5**. All these conformers, with small differences, reproduce quite well the experimental frequencies and the separation of about 30 cm^{-1} between the main features of the IRMPD broad band, with **S*t*-1** presenting more pronounced variations and a somewhat smaller separation between the two prominent bands at lower frequency. Assuming, from the close energy spacing discussed in the previous section, that conformers **S*t*-1**, **S*t*-2**, **S*t*-3**, **S*t*-4** and **S*t*-5**, should all contribute to the Boltzmann population of the sampled ion, the

envelope of absorbances in this region of the IRMPD spectrum, giving rise to the experimentally observed broad band, is accounted for.

On the blue side of the envelope, the shoulder at 1407 cm^{-1} finds its counterpart in a comparatively intense OH bending mode expected at 1388 cm^{-1} in S*t*-**1**, where theory predicts the formation of a proton bridge between the carboxylic hydroxyl group and the amine site and at 1361 cm^{-1} in S*t*-**4**, red-shifted due to the presence of a hydrogen bond between the carboxylic hydroxyl group and the sulfate oxygen. Indeed, as already pointed out, this mode may be significantly sensitive to the hydrogen bonding network, and in the considered conformers its frequency varies from 1332 cm^{-1} (conformer S*t*-**5**) to 1400 (conformer S*t*-**2**). This variation might be also related to the strongly anharmonic nature of this feature, well demonstrated when the proton is shared between two carboxylate groups,[31] as well as in (modified) deprotonated serine.[31, 33]

At still lower frequency, the weak resonances at 1008 and 1057 cm^{-1} are accounted for mainly by the symmetric SO_3 stretch coupled with the C2-O(SO_3) stretch and the NH_2 wagging, calculated at 1010 and 1054 cm^{-1}, respectively, for S*t*-**4**, and 1022, 1038 and 1131 cm^{-1} for S*t*-**1**.

Interestingly, the aforementioned symmetric and antisymmetric SO_3 stretching modes agree well in terms of frequency with those previously observed in hydrated HSO_4^- clusters.[58] In this region, a much stronger activity, due to additional contributions of different vibrational modes, is conversely expected for the less stable salt-bridge **ZW-1** isomer (NH bending coupled with C2-C3 stretching), S*c*-**1** isomer (COH bending and SOH bending), in contrast with the experimental evidence.

The above analysis suggests that the IRMPD spectrum for $[sSer-H]^-$, showing relatively wide bands and structured peaks, is in fact the result of the simultaneous presence of several low energy conformers, whose vibrational modes are somewhat spread in frequency due to their different H-bonding networks. However, the structural motif present in S*t*-**4**, endowed with the lowest enthalpy value in the series (Table S1) shows a better agreement with the experimental frequency values. In particular, compared to S*t*-**1**, the resolved antisymmetric SO_2 streching and COH bending are merged in S*t*-**4**, while the carbonyl stretch is slightly red-shifted at 1728 cm^{-1}, in very good agreement with the experimental peak observed at 1732 cm^{-1}. Such considerations are confirmed by comparing the thermally averaged spectrum to the experimental one (Fig. 4).

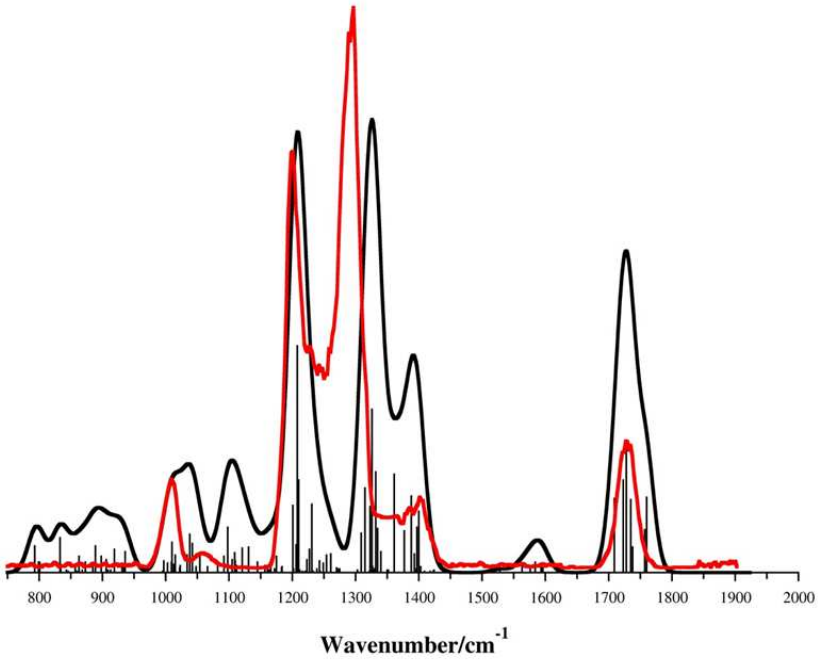

Fig. 4 Averaged spectrum of deprotonated L-serine-O-sulfate [sSer-H]$^-$, computed by considering contributions from all thermally accessible conformers according to their relative enthalpy. The experimental spectrum (red line) is also reported to enable immediate comparison.

Here, the thermally averaged spectrum was built by considering contributions of each conformer to the Boltzmann population weighed according to its relative ΔH value, rather than ΔG value: the discussion above highlighted that conformer **S*t*-4**, which is the lowest in enthalpy, seems to better reproduce the experimental frequencies. In any case, the spectrum averaged by taking into account ΔG values is also reported in Fig. S11, where it can be seen that differences mostly concern band shapes, rather than their position. Fig. 4 shows that the experimental frequencies are generally well reproduced, the main discrepancy arising in the position of the experimental band at 1297 cm^{-1}, whose corresponding calculated value (1326 cm^{-1}) gives however a percent deviation within 2.5 %. The contribution of several conformers in the 1200-1400 cm^{-1} region yields a broader band with respect to the single structures somewhat reminiscent of the experimental spectrum.

IRMPD spectroscopy of protonated sulfoserine

The IRMPD spectrum of mass-selected $[sSer+H]^+$ ions exhibits (Fig. 5, black profile) several distinct features in the explored "fingerprint" region (800-1900 cm^{-1}) at 908, 1022, 1446 and 1784 cm^{-1}, and two largely overlapping bands with maxima at 1161 and 1206 cm^{-1}.

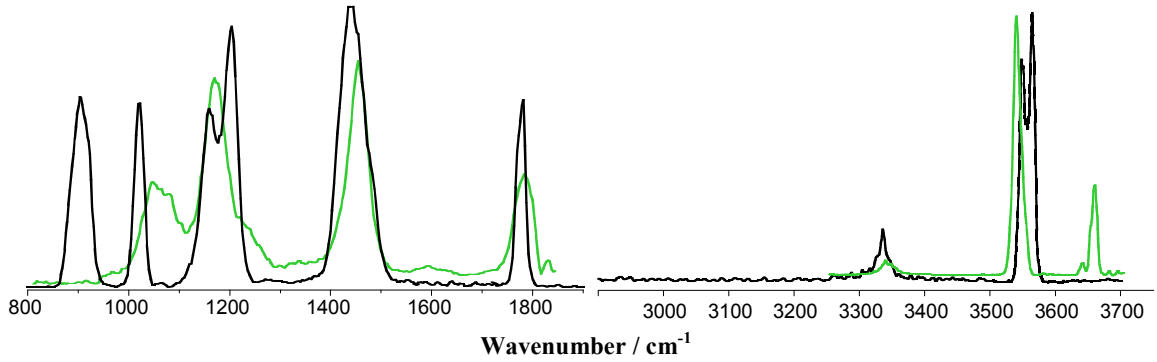

Fig. 5 Experimental IRMPD spectra of $[sSer+H]^+$ (black profile) and $[Ser+H]^+$ (green profile) recorded in the fingerprint and in the NH/OH stretch ranges.

In the same panel of Fig. 5, the IRMPD spectrum of the protonated native amino acid, $[Ser+H]^+$, is shown as a green profile. The latter spectrum is characterized by absorbances at 1050, 1167, 1456 and 1791 cm^{-1} which all find a counterpart in the IRMPD spectrum of the sulfo-derivative $[sSer+H]^+$. However, diagnostic sulfate signatures seem to appear in the 800-1300 cm^{-1} range, where $[sSer+H]^+$ presents four features to be compared with the only two of $[Ser+H]^+$. So, two additional absorptions appear in the IRMPD spectrum of $[sSer+H]^+$, namely a wide (fwhm = 48 cm^{-1}) band at 908 cm^{-1}, and a sharp (fwhm= 12 cm^{-1}) one at 1022 cm^{-1}, offering a distinct experimental signature for the sulfo-modification.

Vibrational features have been also recorded in the 2900-3700 cm^{-1} range, where the NH/OH stretching transitions can be explored. In this region, the IRMPD spectrum of $[sSer+H]^+$ presents two prominent, overlapping bands centered at 3548 and 3562 cm^{-1}, together with a relatively weaker resonance at 3335 cm^{-1} (Fig. 5, black profile), whereas $[Ser+H]^+$ absorbs at 3340, 3541 and 3660 cm^{-1} (Fig. 5, green profile). However, in the lower-energy range, the intense peaks associated to H-bonded N+-H group/interactions predicted at 2900 and 3150 cm^{-1} are notably absent in the IRMPD spectrum. In general, stretching modes involved in extensive hydrogen-bond network are not well represented in IRMPD spectroscopy. Significant red-shift and broadening will affect the signatures of the related modes, as already documented.

Computed Structures of [sSer+H]$^+$

According to calculations, the amino group is the most favourable protonation site of sulfoserine, as previously described for protonated serine in the native[28] and phosphorylated form.[33] Stable conformers of [sSer+H]$^+$ ions, **N*c*-1, N*c*-2, N*c*-3,** and **N*c*-4**, all bearing an ammonium group involved in multiple intramolecular interactions and a *cis* configuration of the carboxylic group, lie within a 11.2 (11.7) kJ mol^{-1} enthalpy (free energy at 298 K) window, and are depicted in Fig. 6.

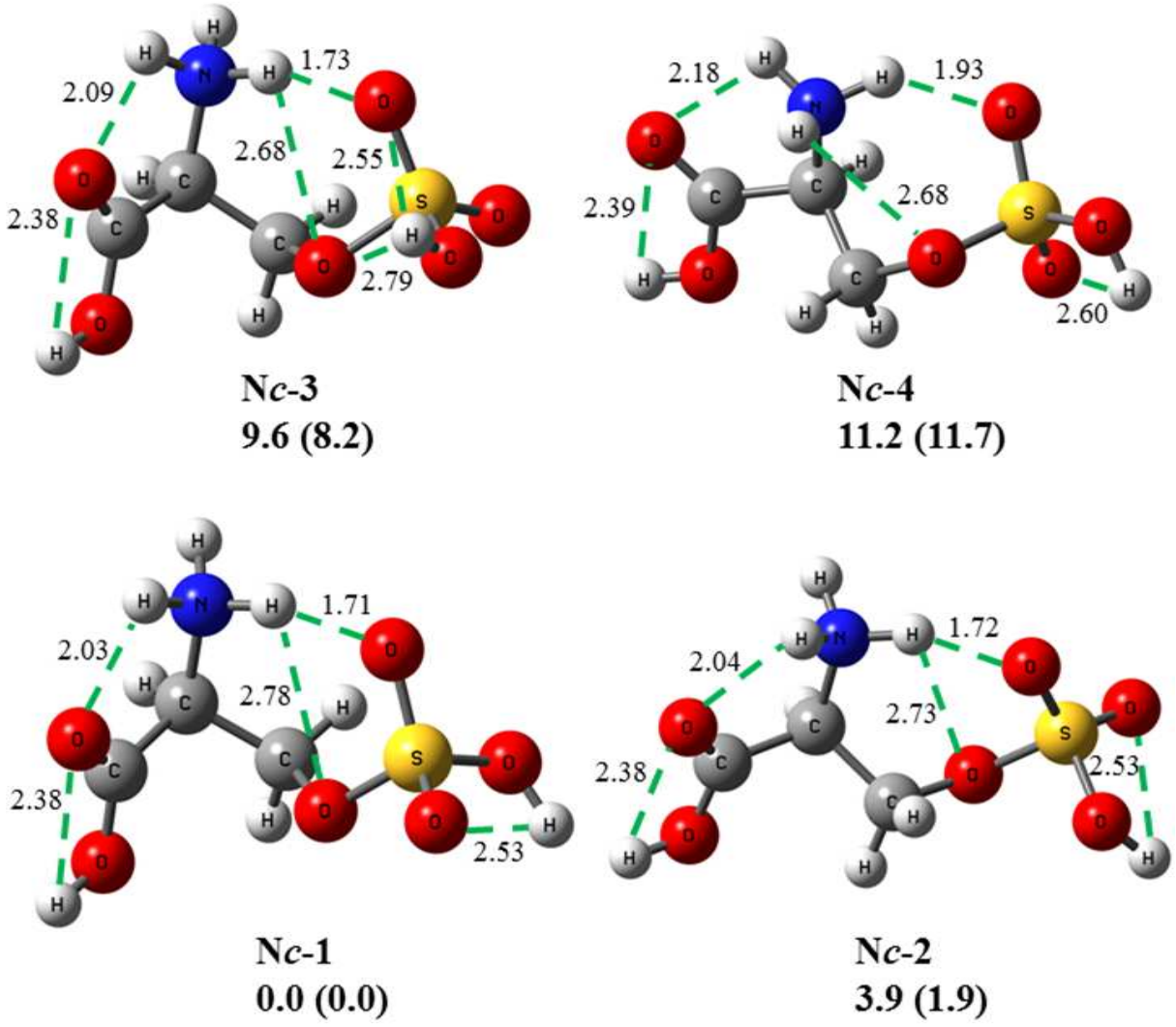

Fig 6 Optimized geometries of the most stable conformers (**N*c*-1, N*c*-2**, **N*c*-3**, and **N*c*-4**) of protonated L-serine-O-sulfate [sSer+H]$^+$ determined at the M06-2X/6-311+G(d,p) level. Relative enthalpy and free energy (in parentheses) values (kJ mol^{-1}, in parentheses) at 298 K are computed at the MP2/aug-cc-pVTZ level. Interatomic bond distances, marked by dashed lines, are given in Å.

The considerable number of low-lying structures for protonated sulfoserine found in this study is matched by similar findings from a previous IRMPD investigation of protonated serine.[28] Still, the protonation event was found to drastically decrease the number of coexisting conformers lying within about 8 kJ mol^{-1}, with respect to gaseous neutral serine, as observed by Fourier transform microwave spectroscopy.[73]

The relevant thermodynamic data are presented in Table S1. Stepwise rotation about the C-O(S), (C)O-S, S-O(H) and C-C bonds of the flexible side chain leads to conformational changes described by dihedral angles listed in Table S3. The ∠C2C3OS, ∠C1C2C3O, ∠C3OSO, ∠OSOH dihedral angles of the ground-state conformer **N*c*-1**, equal to -100.5°, -43.8°, 43.3°, and -92.0°, respectively, favor an arrangement whereby two ammonium protons are engaged in a hydrogen bond with the carbonyl oxygen ($r_{CO...HN}$ = 2.03 Å) and in a bifurcated H-bond motif with both the sulfate ($r_{SO...HN}$ = 1.71 Å) and the sulfoester oxygen ($r_{SO..HN}$ = 2.78 Å), respectively. As can be noted in Fig. 6, the four lowest lying conformers present very similar patterns of hydrogen bonding interactions. Rotamer **N*c*-2,** 3.87 (1.94) kJ mol^{-1} less stable, though bearing a completely different orientation of dihedral angles, presents to a hydrogen bonding motif (in terms of number, kind and strength of the interactions) very much alike to the global minimum **N*c*-1**, also explaining the proximity in energy of the two conformers**.** The rotation around the (C3)O-S bond of **N*c*-1**, changing the C3OSO dihedral angle to 71.9°, allows the sulfate hydroxyl to establish hydrogen bonds with both sulfate ($r_{SO...HO(S)}$ = 2.55 Å) and sulfoester ($r_{CO(S)..HO(S)}$ = 2.79 Å) oxygen atoms and yields rotamer **N*c*-3** which lies at 9.6 (8.2) kJ mol^{-1} relative to **N*c*-1**. In the case of **N*c*-4**, at 11.2 (11.7) kJ mol^{-1} relative to **N*c*-1**, the ∠C2C3OS dihedral angle is found equal to -64.0°, resulting in all three ammonium hydrogens involved in hydrogen bonding. The H bond distances are either longer, in the case of carbonyl ($r_{CO..HN}$ = 2.18 Å and) and sulfate ($r_{SO..HN}$ = 1.93 Å) acceptor groups, or shorter, in the case of the sulfoester oxygen ($r_{SO..HN}$ = 2.68 Å), if compared with the corresponding distances within **N*c*-1**. Interestingly, all structures adopt a ∠C1C2C3O dihedral angle of the opposite sign as the ∠C3OSO dihedral, which facilitates the formation of intramolecular hydrogen bonds by the ammonium group with the carbonyl and sulfate oxygens.

Other structures have been examined, including the N-protonated rotamer **N*t*-1** with a *trans* configuration of the carboxylic group, and isomers bearing a proton on either the carbonyl (**CO-1**) or the sulfate (**SOH-1**) oxygen atoms (Fig. S12). The **N*t*-1** isomer differs from **N*c*-1** mainly for the *trans* setting of the carboxylic group, keeping the same pattern of hydrogen bonding. This variation

causes a destabilization by 40.3 (39.2) kJ mol^{-1} (Table S1). In the **CO-1** and **SOH-1** isomers, the sulfoester oxygen is engaged in hydrogen bonding with both an NH and either the charged C(OH) ($r_{SO..HO(C)}$ = 1.71 Å) or S(OH) ($r_{SO..HO(S)}$ = 2.85 Å) group. Both structures, however, lie much higher in enthalpy (free energy) relative to **N*c*-1**, by 123.0 (121.4) and 177.8 (168.7) kJ mol^{-1} respectively, which makes their contribution to the probed ion population extremely unlikely.

Calculations were also performed on the aziridium ion and sulphuric acid, the prevalent fragments obtained upon both soft CID and IRMPD. The results show that the threshold energy associated to this fragmentation process amounts to 128.9 kJ mol^{-1} (E_0 at 0 K). This relatively high endothermicity requires the absorption of multiple (more than 5) photons to obtain a sizeable photofragmentation yield, considering that in the 900–1900 cm^{-1} range the photon energy is 12–23 kJ mol^{-1}.

Spectral assignment of [sSer+H]$^+$

Fig. 7 compares the experimental IRMPD spectrum of [sSer+H]$^+$, covering both the fingerprint and the NH/OH stretching regions, with the calculated IR absorption spectra of the two most stable rotamers **N*c*-1** and **N*c*-2** among the ones characterized by the favored protonation at the amino group. Indeed, these two conformers lie sufficiently close in free energy (within 5 kJ mol^{-1}) to likely contribute to the thermally averaged population of sampled ions at room temperature.

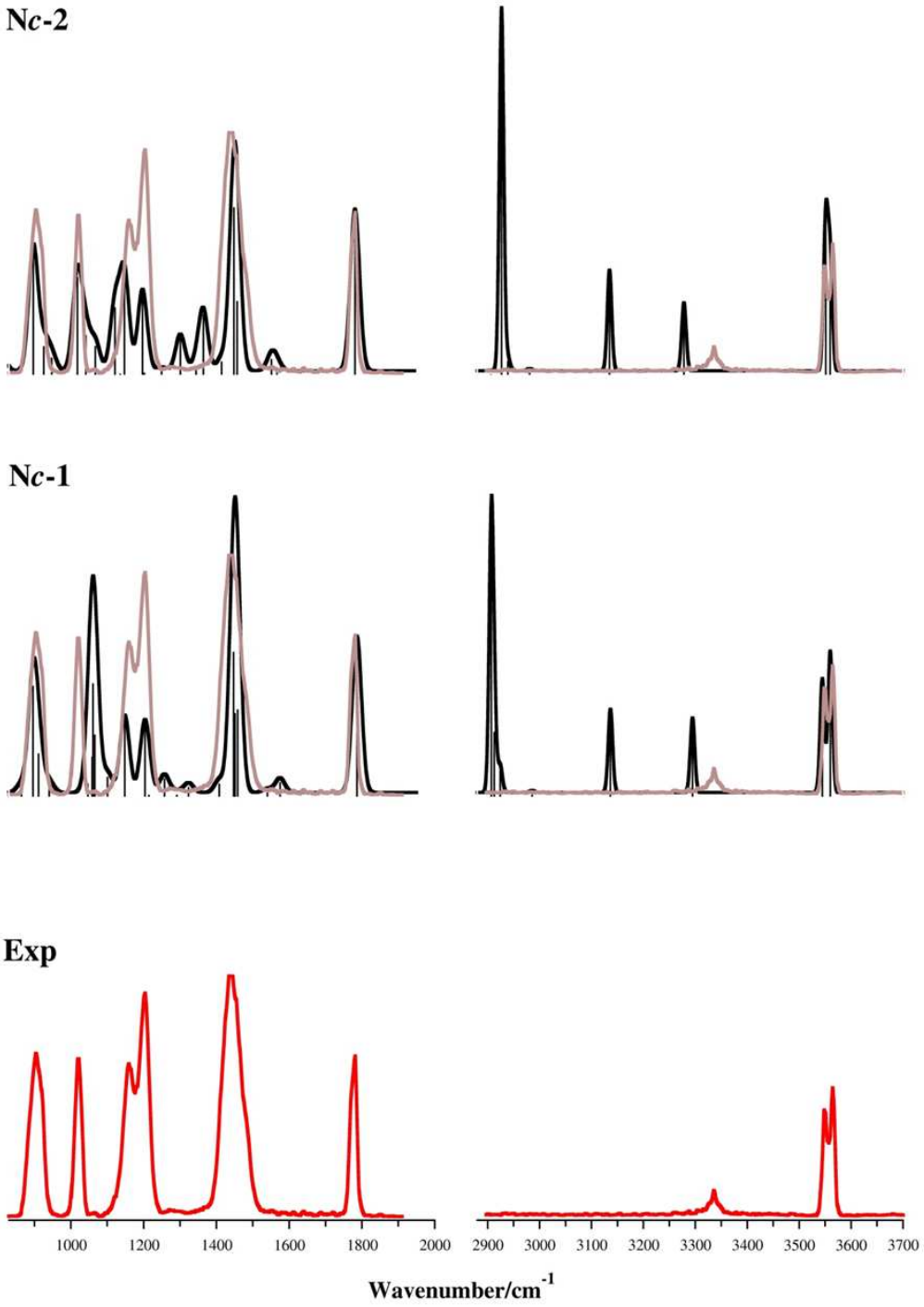

Fig. 7 Experimental IRMPD spectrum of protonated L-serine-O-sulfate $[sSer+H]^+$ (bottom), and computed IR spectra (at the M06-2X/6-311+G(d,p) level) of conformers **Nc-1,** and **Nc-2**. The experimental spectrum is also reported as a pale trace over the calculated spectra to enable immediate comparison.

For this reason, in Fig. 8 the thermally averaged spectra obtained, as in the case of deprotonated solfoserine, by considering contributions of the two conformers to the Boltzmann population weighed according to their relative ΔH value, is also reported for comparison.

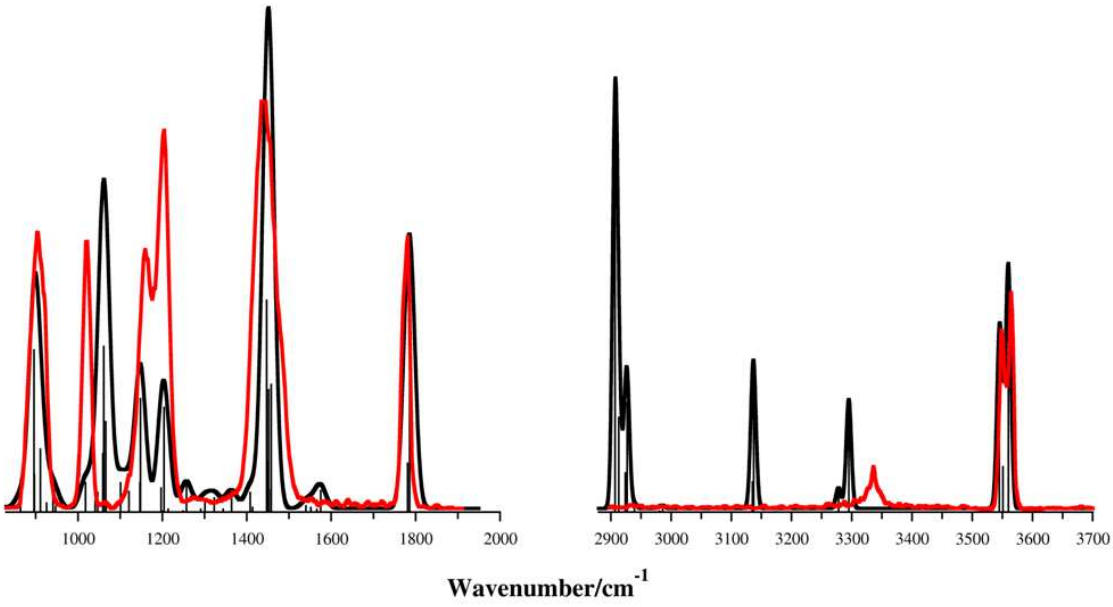

Fig. 8 Averaged spectrum of of protonated L-serine-O-sulfate $[sSer+H]^+$, computed by considering contributions from all thermally accessible conformers, according to their relative enthalpy. The experimental spectrum (red line) is also reported to enable immediate comparison.

The spectrum averaged by taking into account ΔG values is also reported in Fig. S13.

A comprehensive presentation of the linear IR spectra of all structures protonated at either the amino site (**N*c*-1**, **N*c*-2**, **N*c*-3**, **N*c*-4,** and **N*t*-1**), the carbonyl (**CO-1**) or the sulfate (**SOH-1**) group is provided in Fig. S14 and S15, respectively. As can be seen in Fig. S14, all the low energy N-protonated, *cis* structures present very similar IR spectra. This finding is not surprising given their close structural resemblance. However, certain differences between the calculated spectra appear in the low wavenumber region, mainly associated to the presence of (poorly resolved) features around 950 cm^{-1} (**N*c*-3**), 1100 cm^{-1} (**N*c*-4**), and 1370 cm^{-1} (**N*c*-2**). Some difference also appears in the three peaks between 3100 and 3300 cm^{-1} (**N*c*-4**). A concise description of the main assigned IRMPD

absorptions is presented in Table 2, along with the experimental data and the computed frequencies and intensities of the ground state structure **N*c*-1**, whose well resolved IR active modes presents the best overall match for the entire IRMPD spectrum. This result provides evidence that the protonated $[sSer+H]^+$ ions is reasonably described by the lowest energy structure.

It might be worth noting that this situation is remarkably different from the behavior of deprotonated sulfoserine $[sSer-H]^-$ examined in the previous sections, where several comparably stable conformers contribute to the overall spectrum. This finding suggests that the presence of a deprotonated site (with a relatively mobile sulfate group) increases significantly the possibility of multiple H-bonding. As a consequence, an explanation of the different behavior in the comparison between the IRMPD spectra of the parent amino acid with the sulfated derivatives might be put forward. In the case of the more rigid protonated ions, only additional features due to sulfation are observed in the IRMPD spectrum of the modified species with other modes otherwise unchanged from the native amino acid ion. In the case of the negatively charged $[sSer-H]^-$ ion, the IRMPD spectrum presents different features spread throughout the frequency range, with respect to $[Ser-H]^-$, that are not solely ascribable to sulfation.

Among the most active features in the fingerprint range, the first band at 908 cm^{-1} matches with a band comprising coupled antisymmetric S-OH stretching and SOH bending modes with the C1-C2, and C2-N stretches predicted at 896 cm^{-1}; the narrow (fwhm= 24 cm^{-1}) peak at 1022 cm^{-1} can be attributed to the (C3)O-S and C3-O(S) stretches, and the N-H bending calculated at 1061 cm^{-1}, besides the molecular deformation and the NH2 scissoring expected at 1065 cm^{-1}; the envelope of absorptions centered at 1161 and 1206 cm^{-1} are associated to the COH bending modes of the carboxylic group, calculated at 1148 cm^{-1}, and to the symmetric SO_2 stretch predicted at 1204 cm^{-1}, which is conversely lacking in isomer **SOH-1** (Fig. S15). At higher frequency, the quite broad (fwhm= 80 cm^{-1}) absorption at 1446 cm^{-1} comprises the antisymmetric SO_3 stretch, predicted at 1447 cm^{-1}, merged with the NH_3^+ umbrella mode calculated at 1451 cm^{-1}, which marks a clear signature of the protonation at the amino group. On the other hand, the latter band is missing in less stable isomers such as **CO-1** and **SOH-1** (Fig. S15). Any appreciable contribution of these species can be thus ruled out. In addition, the sharp peak (fwhm= 22 cm^{-1}) at 1784 cm^{-1} corresponds to the C=O stretch expected at 1786 cm^{-1}, which is obviously missing in the IR spectrum of **CO-1**. Although some weaker bands find a fairly good correspondence in the computed IR absorptions, the IRMPD spectrum appears remarkably flat in the region where only a weak peak at 1560 cm^{-1} (not reported in Table 2 since its intensity is lower than 50 km mol^{-1}) is expected. Interestingly,

comparative inspection of the IRMPD spectrum of $[sSer+H]^+$ ions recorded here with a previous report of the IRMPD spectrum of protonated O-phosphoserine reveals very similar absorptions.[32]

The high-energy portion of the spectrum (2900-3700 cm^{-1}) presents a prominent unresolved absorbance with maxima at 3548 and 3562 cm^{-1}, accompanied by a weak band at 3335 cm^{-1}, faithfully reproduced by the carboxylic and sulfuric O-H stretches and by the N^+-H stretch, when not engaged in strong hydrogen bonding. In fact, the theoretically predicted values are 3545, 3560, and 3295 cm^{-1}, respectively, for the lowest energy structure **N*c*-1**. Noteworthily, this vibrational pattern is diagnostic of the major contribution of N-protonated structures, in contrast with the calculated SO-H stretches for **N*t*-1**, **CO-1**, and **SOH-1** isomers which are instead significantly parted from the CO-H stretch mode and in marked contrast with the experiment.

Although the experimental band positions and (partially) intensities occurring at wavenumber above 3300 cm^{-1} are faithfully reproduced by **N*c*-1**, inspecting the range below 3300 cm^{-1} the experimental spectrum appears notably flat. In this range, though, intense bands for NH/OH stretches involved in hydrogen bonding are expected, including a highly active mode for the carbonyl bound N^+-H stretching expected at 3136 cm^{-1}. However, as previously reported for related systems, such as protonated (S-nitroso)cysteine, 4-hydroxyproline, and the native serine itself,[35, 36, 72] a noticeably weak intensity or even absence of detectable peaks has been observed in this part of the spectrum, where weaker laser power, red-shift and broadening due to anharmonicity effects, incomplete vibrational relaxation, and pronounced nonlinear effects may result in what has been termed IRMPD "transparency".[59, 74] Anyhow, it is worth noting that also in this portion of the IRMPD spectrum a signature of the serine O-sulfation may be still recognized when comparing the sulfate OH stretch of $[sSer+H]^+$ with the alcohol OH stretch of $[Ser+H]^+$,[72] and the phosphate OH stretch of phosphopeptides,[21, 55] which are active at 3564, 3650, and ca. 3670 cm^{-1}, respectively.

Conclusions

Vibrational signatures of both protonated and deprotonated L-serine-O-sulfate ions, $[sSer-H]^-$ and $[sSer+H]^+$, have been obtained by IRMPD spectroscopy assisted by quantum chemical calculations. Two IR spectral regions have been explored, namely the fingerprint (750-1900 cm^{-1}) and the NH/OH stretching (2900-3700 cm^{-1}) ranges, using either the CLIO FEL light source, or a tabletop OPO/OPA laser source, respectively, coupled with tandem mass spectrometry. The experimental features have been assigned by comparison with the calculated IR spectra of the low energy conformers of $[sSer-H]^-$ and $[sSer+H]^+$ which contribute to the sampled ion population.

Experimental and theoretical results indicate that the low energy structures of $[sSer+H]^+$ bear a proton on the amino group and two hydrogen atoms of the ammonium group establish a C=O…HN and a S=O…HN hydrogen bonding interaction. The O-sulfation event in $[sSer+H]^+$ has been found responsible for the appearance of highly active vibrational modes in the 900-1300 cm^{-1} energy range, due to S-OH, C-OS, and CO-S stretches, SOH bending and symmetric SO_2 stretching modes, at frequencies where the native protonated serine $[Ser+H]^+$ displays scant or anyway distinctly different activity. Conversely, the intense antisymmetric SO_2 stretch at 1446 cm^{-1} is not diagnostic, being coupled with the NH_3 umbrella mode so that a pronounced band is present in both the sulfo and native ions. In the OH stretching region, the SO-H stretch at 3562 cm^{-1} can be recognized from carboxylic and alcoholic O-H stretches, providing a signature for the presence of this PTM.

For the conjugate base of sulfoserine, $[sSer-H]^-$, in the lowest energy structures deprotonation occurs at the sulfate group which forms intramolecular hydrogen bonds with either the amino hydrogen atoms or the carboxylic hydroxyl group. Here the negative charge together with the higher flexibility imparted by the sulfate group, facilitate the formation of many different hydrogen bond patterns, leading to rotamers very close in energy. As a result, the experimental IRMPD spectrum presents broad and structured bands which are compatible with contributions coming from different conformations. Distinctive vibrational features of $[sSer-H]^-$, associated with the symmetric SO_3 (1008 cm^{-1}) and antisymmetric SO_3 (1204 cm^{-1}) and SO_2 (1297 cm^{-1}) stretches, and lacking in the unmodified amino acid, $[Ser-H]^-$, may provide a useful reference for the assignment of sulfate IR modes of O-sulfated amino acids in their natural deprotonated state. Besides these evident features due to sulfonation, $[sSer-H]^-$ and its parent deprotonated amino acid spectra differ for the position of other vibrational modes, again to be related to the increased flexibility of the sulfonate species and its ability to form intramolecular hydrogen bonds. This is in contrast with the behavior of the comparatively more rigid protonated species, where only additional features due to sulfonation are found for $[sSer+H]^+$ as compared to $[Ser+H]^+$.

Acknowledgements

We thank Philippe Maitre, Jean-Michel Ortega and Vincent Steinmetz, for the support of the CLIO team and are grateful to Annito Di Marzio for experiments at the OPO/OPA laser. This work has been funded by the Università degli Studi di Roma La Sapienza and by the European Commission (CLIO project IC009-11).

References

1 C. A. Strott, *Endocr. Rev.*, 2002, **23**, 703-732.
2 W. B. Huttner, *Nature*, 1982, **299**, 273-276.
3 C. Niehrs, and W. B. Huttner, *EMBO J.,* 1990, **9**, 35-42.
4 K. F. Medzihradszky, Z. Darula E. Perlson, M. Fainzilber, R. J. Chalkley, H. Ball, D. Greenbaum, M. Bogyo, D. R. Tyson, R. A. Bradshaw, and A. L. Burlingame, *Mol. Cell. Proteomics*, 2004, **3**, 429-440.
5 F. Monigatti, B. Hekking, and H. Steen, *Biochim. Biophys. Acta*, 2006, **1764**, 1904-1913.
6 J. W. Kehoe, and C. R. Bertozzi, *Chem. Biol.*, 2000, **7**, R57-R61.
7 K. L. Moore, *J. Biol. Chem.*, 2003, **278**, 24243-24246.
8 P. P. Wilkins, K. L. Moore, R. P. McEver, and R. D. Cummings, *J. Biol. Chem.*, 1995*,* **270**, 22677-22680.
9 J. Zhang, P. L. Yang, and N. S. Gray, *Nat. Rev. Cancer*, 2009, **9**, 28-39.
10 C. Seibert, and T. P. Sakmar, *Pept. Sci.*, 2008, **90**, 459-477.
11 C. T. Veldkamp, C. Seibert, F. C. Peterson, T. P. Sakmar, and B. F. Volkman, *J. Mol. Biol.*, 2006, **359**, 1400-1409.
12 C. M. Ward, R. K. Andrews, A. I. Smith, and M. C. Berndt, *Biochem.*, 1996, **35**, 4929-4938.
13 N. L. Kelleher, R. A. Zubarev, K. Bush, B. Furie, B. C. Furie, F. W. McLafferty, and C. T. Walsh, *Anal. Chem*., 1999, **71**, 4250-4253.
14 R. E. Bossio, and A. G. Marshall, *Anal. Chem.*, 2002, **74**, 1674-1679.
15 J. F. Nemeth-Cawley, S. Karnik, and J. C. Rouse, *J. Mass Spectrom*., 2001, **36**, 1301-1311.
16 K. F. Medzihradszky, S. Guan, D. A. Maltby, and A. L. Burlingame, *J. Am. Soc. Mass Spectrom.*, 2007, **18**, 1617-1624.
17 H. Liu, and K. Hakansson, *Anal. Chem*., 2006, **78**, 7570-7576.
18 H. J. Yoo, N. Wang, S. Zhuang, H. Song, and K. Hakansson, *J. Am. Chem. Soc.,* 2011*,* **133**, 16790-16793.
19 K. E. Hersberger, and K. Hakansson, *Anal. Chem.*, 2012, **84**, 6370-6377.
20 M. R. Robinson, K. L. Moore, and J. S. Brodbelt, *J. Am. Soc. Mass Spectrom*., 2014, **25**, 1461-1471.
21 A. L. Patrick, C. N. Stedwell, and N. C. Polfer, *Anal. Chem*., 2014, **86**, 5547-5552.
22 J. Oomens, B. G. Sartakov, G. Meijer, and G. von Helden, *Int. J. Mass Spectrom*., 2006, **254**, 1-19.
23 J. R. Eyler, *Mass Spectrom. Rev*., 2009, **28**, 448−467.
24 T. D. Fridgen, *Mass Spectrom. Rev*., 2009, **28**, 586-607.
25 N. C. Polfer, and J. Oomens, *Mass Spectrom. Rev*., 2009, **28**, 468-494.
26 J. Roithovà, *Chem. Soc. Rev*., 2012, **41**, 547-559.
27 A. Simon, L. MacAleese, P. Maitre, and T. B. McMahon, *J. Am. Chem. Soc*., 2007, **129**, 2829-2840.
28 R. H. Wu, and T. B. McMahon, *ChemPhysChem*, 2008, **9**, 2826-2835.
29 T. D. Vaden, S. A. N. Gowers, T. S. J. A. de Boer, J. D. Steill, J. Oomens, L. C. Snoek, *J. Am. Chem. Soc*., 2008, **130**, 14640-14650.
30 M. Citir, E. M. S. Stennet, J. Oomens, J. D. Steill, M. T. Rodgers, and P. B. Armentrout, *Int. J. Mass Spectrom*., 2010, **297**, 9-17.
31 J. Oomens, J. D. Steill, and B. Redlich, *J. Am. Chem. Soc*., 2009, **131**, 4310-4319.
32 C. F. Correia, P. O. Balaj, D. Scuderi, P. Maitre, and G. Ohanessian, *J. Am. Chem. Soc*., 2008, **130**, 3359-3370.
33 D. Scuderi, C. F. Correia, O. P. Balaj, G. Ohanessian, J. Lemaire, and P. Maitre, *ChemPhysChem*, 2009, **10**, 1630-1641.

34 C. Coletti, N. Re, D. Scuderi, P. Maitre, B. Chiavarino, S. Fornarini, F. Lanucara, R. K. Sinha, and M. E. Crestoni, *Phys. Chem. Chem. Phys*., 2010, **12**, 13455-13467.
35 F. Lanucara, B. Chiavarino, M. E. Crestoni, D. Scuderi, R. K. Sinha, P. Maitre, and S. Fornarini, *Int. J. Mass Spectrom.*, 2012, **330-332**, 160-167.
36 M. E. Crestoni, B. Chiavarino, D. Scuderi, A. Di Marzio, and S. Fornarini, *J. Phys. Chem. B*, 2012, **116**, 8771−8779.
37 M. Ignasiak, D. Scuderi, P. de Oliveira, T. Pedzinski, Y. Rayah, and C. Houee Levin, *Chem. Phys. Lett.*, 2011, **502**, 29–36.
38 B. J. Byhtell, P. Maitre, and B. Paizs, *J. Am. Chem. Soc*., 2010, **132**, 14766-14779.
39 F. Lanucara, B. Chiavarino, D. Scuderi, P. Maitre, S. Fornarini, and M. E. Crestoni, *Chem. Comm.*, 2014, **50**, 3845-3848.
40 B. Chiavarino, M. E. Crestoni, O. Dopfer, P. Maitre, and S. Fornarini, *Angew. Chem., Int. Ed. Engl.*, 2012, **51**, 4947-4949.
41 B. Chiavarino, M. E. Crestoni, M. Schütz, A. Bouchet, S. Piccirillo, V. Steinmetz, O. Dopfer, and S. Fornarini, *J. Phys. Chem. A*, 2014, **118**, 7130-7138.
42 B. Chiavarino, M. E. Crestoni, S. Fornarini, L. Lanucara, J. Lemaire, P. Maitre, and D. Scuderi, *Chem. - Eur. J*., 2009, **15**, 8185-8195.
43 B. Chiavarino, P. Maitre, S. Fornarini, and M. E. Crestoni, *J. Am Soc. Mass Spectrom*., 2013, **24**, 1603-1607.
44 C. Kapota, J. Lemaire, P. Maitre, G. Ohanessian, *J. Am. Chem. Soc*., 2004, **126**, 1836−1842.
45 N. C. Polfer, J. Oomens, D. T. Moore, G. von Helden, G. Meijer, and R. C. Dunbar, *J. Am. Chem. Soc.*, 2006, **128**, 517-525.
46 P. B. Armentrout, M. T. Rodgers, J. Oomens, and J. D. Steill, *J. Phys. Chem. A*, 2008, **112**, 2248-2257.
47 B. Chiavarino, M. E. Crestoni, S. Fornarini, S. Taioli, I. Mancini, P. Tosi, *J. Chem. Phys*., 2012, **137**, 024307/1- 024307/9.
48 F. Lanucara, D. Scuderi, B. Chiavarino, S. Fornarini, P. Maitre, and M. E. Crestoni, *J. Phys. Chem. Lett*., 2013, **4**, 2414-2417.
49 B. Chiavarino, M. E. Crestoni, S. Fornarini, D. Scuderi, and J. Y. Salpin, *J. Am. Chem. Soc*., 2013, **135**, 1445−1455.
50 A. Gholami and T. D. Fridgen, *J. Phys. Chem. B*, 2013, **117**, 8447-8456.
51 A. De Petris, A. Ciavardini, C. Coletti, N. Re, B. Chiavarino, M. E. Crestoni, and S. Fornarini, *J. Phys. Chem. Lett*., 2013, **4**, 3631−3635.
52 R. C. Dunbar, J. D. Steill, and J. Oomens, *J. Am. Chem. Soc*., 2011, 133, 1212-1215.
53 F. Lanucara, M. E. Crestoni, B. Chiavarino, S. Fornarini, O. Hernandez, D. Scuderi, and P. Maitre, *RSC Adv.*, 2013, **3**, 12711-12720.
54 B. Schindler, J. Joshi, A.-R. Allouche, D. Simon, S. Chambert, V. Brites, M.-P. Gaigeot, and I. Compagnon, *PhysChemChemPhys*, 2014, **16**, 22131-22138.
55 F. Turecek, C. L. Moss, I. Pikalov, R. Pepin, K. Gulyuz, N. Polfer, M. F. Bush, J. Brown, J. Williams, and K. Richardson, Int. J. Mass Spectrom., 2013, **354-355**, 249-256.
56 B. Gregori, L. Guidoni, B. Chiavarino, D. Scuderi, E. Nicol, G. Frison, S. Fornarini, and M. E. Crestoni, *J. Phys. Chem. B*, 2014, **118**, 12371-12382.
57 M. F. Bush, R. J. Saykally, and E. R. Williams, *J. Am. Chem. Soc.*, 2007, **129**, 2220-2221.
58 T. I. Yacovitch, T. Wende, L. Jiang, N. Heine, G. Meijer, D. M. Neumark, and K. R. Asmis, *J. Phys Chem. Lett.*, 2011, **2**, 2135-2140.
59 T. I. Yacovitch, N. Heine, C. Brieger, T. Wende, C. Hock, D. M. Neumark, and K. R. Asmis *J. Phys. Chem. A*, 2013, **117**, 7081-7090.
60 R. K. Sinha, B. Chiavarino, S. Fornarini, J. Lemaire, P. Maitre, and M. E. Crestoni, *J. Phys. Chem. Lett.*, 2010, **1**, 1721-1724.

61 L. Mac Aleese, A. Simon, T. B. McMahon, J.-M. Ortega, D. Scuderi, J. Lemaire, and P. Maitre, *Int. J. Mass Spectrom*., 2006, **249/250**, 14-20.
62 J. Lemaire, P. Boissel, M. Heninger, G. Mauclaire, G. Bellec, H. Mestdagh, A. Simon, S. Le Caer, J. M. Ortega, F. Glotin, and P. Maitre, *Phys. Rev. Lett*., 2002, **89**, 273002-273004.
63 R. K. Sinha, P. Maitre, S. Piccirillo, B. Chiavarino, M. E. Crestoni, and S. Fornarini, *PhysChemChemPhys*, 2010, **12**, 9794-9800.
64 J. S. Prell, J. T. O'Brien, and E. R. Williams, *J. Am. Soc. Mass Spectrom*., 2010, **21**, 800-809.
65 MACROMODEL, Version 9.6, Schrödinger, LLC: New York, NY, 2008.
66 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Peterson, et al. *GAUSSIAN 09* (Revison A.01); Gausssian, Inc.: Wallington, CT, 2009.
67 M. M. Quesada-Moreno, J. R. Avilés-Moreno, A. Á. Márquez-García,F. Partal-Ureña, and J. J. López González *J. Mol. Structure,* 2013, **1046**, 136–146.
68 Y. Zhao, and D. G. Truhlar, *Theor Chem Acc,* 2008*,* **120**, 215–241.
69 I. M. Alecu, J. Zheng, Y. Zhao, and D. G. Truhlar, *J. Chem. Theory Comput,* 2010, **6**, 2872-2887.
70 R. K. Sinha, B. Chiavarino, S. Fornarini, J. Lemaire, P. Maitre, and M. E. Crestoni, *J. Phys. Chem. Lett*., 2010, **1**, 1721-1724.
71 Y.-w. Nei, N. Hallowita, J. D. Steill, J. Oomens, and M. T. Rodgers, *J. Phys. Chem. A*, 2013, **117**, 1319-1335.
72 A. L. Patrick, C. N. Stedwell, B. Schindler, I. Compagnon, G. Berden, J. Oomens, and N. C. Polfer, *Int. J. Mass Spectrom.* 2015, **379**, 26-32.
73 S. Blanco, M. E. Sanz, J. C. Lopez, and J. L. Alonso, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 20183-20188.
74 J. P. Beck, J. M. Lisy, *J. Chem. Phys.*, 2011, **135**, 044302/1-044302/6.

Table 1 Observed IRMPD resonances and calculated vibrational frequencies for the low-lying rotamers S***t*****-1**, and S***t*****-4** of deprotonated L-serine-O-sulfate [sSer-H]$^-$.

IRMPD[a]	Calculated[a,b]		vibrational mode
	S*t*-1	**S*t*-4**	
	997 (68)		δ_{wag} -NH_2, δ_{rock} -CH_2
1008 (0.04)	1022[c] (43)	1010[c] (69)	ν_{sym} SO_3, δ NH_2
	1038[c] (154)		ν_{sym} SO_3, δ_{wag} NH_2, molecule deformation
1057 (0.01)	1131[c] (149)	1054[c] (56)	ν C-O(SO_3), ν_{sym} SO_3, δ_{twist} NH_2
		1098[c] (128)	δ NH_2, ν (C)O-S, ν SO_3
1207 (0.17)	1231[c] (399)	1208[c] (448)	ν_{asym} SO_3, δ_{twist} NH_2
1297 (0.24)	1315[c] (368)	1326[c] (344)	ν_{asym} O=S=O
1407 (0.03)	1388 (346)	1361 (346)	δ_{scis} (C=O)O-H
1732 (0.05)	1740 (439)	1728 (435)	ν C=O
-	3111 (619)	3042 (1067)	ν O-H
-	3176 (220)	3285 (119)	ν N-H

[a] In cm^{-1}. [b] The reported intensities given in parentheses are in km mol^{-1}. Bands with an intensity lower than 50 km mol^{-1} are not included. Frequencies are scaled by a factor of 0.940. [c] Frequencies are scaled by a factor of 1.00.

TABLE 2 Experimental IRMPD resonances and calculated vibrational frequencies for the **N*c*-1** isomer of $[sSer+H]^+$ ions (at M06-2X/6-311+G(d,p) level).

IRMPD[a]	Calculated[a,b]	vibrational mode
908 (0.60)	896 (221)[c]	ν S-OH, ν C_α-N, ν C_α-C(=O)
	911 (86)[c]	ν_{asym} (C_β)O-S-OH, δ H-C_β-H
1022 (0.59)	1061 (225)[c]	ν (C_β)O-S, ν C_β-OS, δ N-H
	1058 (79)	$\delta_{sciss.}$ S-O-H
1065 (0.57)	1066 (123)	Molecule deformation
1161 (0.62)	1148 (154)	δ_{sciss} (C=O)-O-H
1206 (0.78)	1204 (142)[c]	ν O=S=O, δ N-H
1446 (0.89)	1447 (288)[c]	ν_{asym} O=S=O, molecule deformation
	1458 (173)[c]	ν (C=O)-OH, ν_{asym} O=S=O
	1451 (166)	δ NH_3^+ umbrella mode
1784 (0.60)	1786 (312)	ν C=O
-	2907 (534)	ν N-H
-	2913 (129)	ν_{sym} CH_2
-	2924 (53)	ν C-H
-	3136 (165)	ν N-H
3335[d]	3295 (148)	ν N-H
3548[d]	3545 (226)	ν COO-H
3562[d]	3560 (281)	ν SO-H

[a] In cm^{-1}. [b] The reported intensities given in parentheses are in km mol^{-1}. Bands with an intensity lower than 50 km mol^{-1} are not included. Frequencies are scaled by a factor of 0.940. [c] Scaling factor=1.00. [d] The relative intensity of this feature is not available since this wavelength range (2900-3700 cm^{-1}) was explored with a different experimental set-up.

Table of Contents

The gas phase structures of (de)protonated O-sulfoserine ions have been revealed by infrared multiple photon dissociation (IRMPD) spectroscopy and computations.