Vibrational Spectroscopy 137 (2025) 103783

Contents lists available at ScienceDirect

Vibrational Spectroscopy

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

Uncovering the vibrational modes of zwitterion glycine in aqueous solution

Mark Christie [a], Mozhdeh Mohammadpour [b,c], Jan Sefcik [b,d], Karen Faulds [a], Karen Johnston [b,*]

[a] Centre for Molecular Nanometrology, WestCHEM, Department of Pure and Applied Chemistry, Technology and Innovation Centre, University of Strathclyde, 99 George Street, Glasgow G1 1RD, United Kingdom
[b] Department of Chemical and Process Engineering, University of Strathclyde, 75 Montrose Street, Glasgow G1 1XJ, United Kingdom
[c] School of Engineering, Lancaster University, Lancaster LA1 4YW, United Kingdom
[d] EPSRC Future Manufacturing Research Hub in Continuous Manufacturing and Advanced Crystallisation, University of Strathclyde, Glasgow G1 1RD, United Kingdom

ARTICLE INFO

Keywords:
Spectroscopy
Vibrational frequencies
Glycine
Zwitterion

ABSTRACT

Vibrational spectroscopy is widely employed to probe and characterise chemical, biological and biomedical samples. Glycine solutions are relevant in a variety of biological and chemical systems, yet the reported experimental vibrational wavenumbers of the glycine zwitterion, which is the dominant species in aqueous solution, are inconsistent and incomplete. This study presents a procedure that obtained a complete set of vibrational frequencies for the glycine zwitterion in aqueous solution, apart from the two lowest wavenumber modes which are available from a previous THz study. Vibrational spectra were measured using IR and Raman spectroscopy, to obtain both IR and Raman-active modes for a range of different glycine solution concentrations using four different instruments. Insight from a literature survey of density functional theory calculations in implicit and explicit water was used to guide the deconvolution of the experimental spectra into vibrational modes, giving 22 out of 24 vibrational wavenumbers with a standard error of less than 3 cm^{-1}. This thorough analysis of the glycine vibrational spectra has enabled missing and erroneous wavenumbers in literature to be identified, and the systematic procedure for determining vibrational modes will pave the way for deeper quantitative analysis of glycine systems, and serve as a benchmark for computational method development.

1. Introduction

Glycine is the smallest amino acid and has been used as a model system for the study of proteins in electrolyte systems [1,2] and as a lixiviant for metal extraction [3]. Glycine also exhibits interesting physical phenomena, including polymorphism and mesoscale clusters [4,5] and it predominantly exists in zwitterionic form in aqueous solution. Vibrational analysis plays a pivotal role in chemical and materials science and processing, and biological and clinical sciences, as it helps to elucidate the fundamental properties of molecules and solids [6–8], such as identifying functional groups, distinguishing between different crystalline polymorphs, and studying adsorption structures. However, despite extensive investigations [9–12–15,16,17], a complete set of individual vibrational wavenumbers for the glycine zwitterion is not available and some reported wavenumbers are inconsistent.

A detailed investigation of THz spectra of glycine aqueous zwitterion solutions [18] identified a number of modes in the low wavenumber region corresponding to both intramolecular and solvent interaction vibrations. For higher wavenumbers of glycine in aqueous solution, Raman spectra [9–12,13,14] and IR spectra [15–17] are available but only one study used both Raman and IR spectroscopy [16], even though certain vibrational modes are difficult to identify from just a single technique. The absence of a complete and reliable set of glycine zwitterion vibrational wavenumbers limits the interpretation of spectroscopic studies of more complex systems containing glycine.

For many systems, including glycine, obtaining a comprehensive and trustworthy set of individual vibrational frequencies presents a formidable challenge. Inconsistencies and missing peaks can arise from differences in experimental techniques and instrumentation used to acquire vibrational frequency data and it can be difficult to distinguish weak peaks from noise, which can result in missing peaks or inclusion of spurious peaks. Varying the concentration is a frequently used technique for verifying peaks, and for determining which peaks correspond to which molecular species [19]. In addition, utilizing both IR and Raman spectroscopy can help to obtain reliable peak positions for both IR active and Raman active modes. Another challenge in acquiring vibrational frequencies from literature is that deconvolution of spectral peaks is often neglected, i.e. studies often report the peak position as the

* Corresponding author.
E-mail address: karen.johnston@strath.ac.uk (K. Johnston).

https://doi.org/10.1016/j.vibspec.2025.103783
Received 5 December 2024; Accepted 13 February 2025
Available online 18 February 2025

vibrational frequency but the peak may be comprised of multiple, overlapping modes. However, without having information about the number of vibrational modes that contribute to a given peak, reliable deconvolution can be difficult to achieve.

A number of density functional theory (DFT) studies calculated the vibrational frequencies of zwitterionic glycine. The zwitterionic species is not stable in vacuum, and calculations have accounted for water implicitly using a polarisable continuum model (PCM) [20,21], or explicitly by including water molecules in the calculated system [11,12,14,17,22]. However, DFT calculations are very sensitive to methodological choices, such as exchange and correlation functional, harmonic or anharmonic approximations, and even the most accurate methods still yield wavenumbers that differ significantly from experimental values. Nevertheless, calculated wavenumbers can still be used to provide insight into the systems vibrational properties.

In this work, we aimed to acquire a comprehensive experimental dataset of vibrational frequencies for glycine zwitterion in aqueous solution. After describing the methodology, we present a review of DFT calculations of vibrational wavenumbers and identify ranges for distinct vibrational wavenumbers. To obtain reliable experimental vibrational mode frequencies, we performed a combined IR and Raman spectroscopy study on glycine solutions over a range of concentrations. The spectra were deconvoluted into 22 individual vibrational modes guided by the DFT calculations, where only the lowest two modes were outside the range of our IR and Raman instruments. Finally we combined these 22 wavenumbers, and compared it to experimental literature determining missing and erroneous modes. Combining these with the two lowest wavenumbers from a THz study [18], we produced a benchmark set of 24 vibrational wavenumbers, which we used to compare DFT results from literature.

2. Methodology

2.1. Experimental

Glycine solutions were prepared through the dissolution of glycine, (≥98.5 %) purchased from Sigma-Aldrich, as 1 mL solutions using doubly distilled deionised water. Concentrations of 1.00, 2.00, 3.01, and 4.04 mol dm^{-3} (indicating moles of solute per cubic decimeter of solvent) glycine solutions were prepared using 75, 150, 226, and 303 mg of glycine dissolved in 1 mL of double-distilled water. The 0.5 mol dm^{-3} solutions were prepared by diluting a 1 mol dm^{-3} stock solution. To ensure complete dissolution of glycine solid, each solution was stirred at 600 rpm and 90˚C for 1 h before analysis. It is worth noting that the 4 mol dm^{-3} solution, although supersaturated, did not form crystals during the spectroscopic measurements.

Solutions of glycine were analysed using both Raman and IR spectroscopy and four sets of data were obtained, referred to as Raman 1, Raman 2, IR 1, and IR 2. Raman 1 set contains Raman measurements between 200–3200 cm^{-1}, performed using a Renishaw inVia Raman Microscope system with a laser wavelength of 532 nm, 5× objective lens, laser power of 18 mW at the sample, an acquisition time of 10 s, and a spectral resolution of 1 cm^{-1}. The Raman 2 set contains Raman spectroscopic measurements between 200–2000 cm^{-1}, were performed using a Snowy Range SnRI spectrometer with a laser wavelength of 785 nm, laser power of 75 mW at the sample, an acquisition time of 10 s, with a spectral resolution of 4 cm^{-1}. All Raman measurements were performed on solutions contained within glass vials. The IR 1 dataset consists of IR measurements obtained between 500 and 4000 cm^{-1} obtained using a Nicolet iS5 FTIR Spectrometer with an iD5 ATR accessory with a spectral resolution of 4 cm^{-1}. The IR 2 set contains IR measurements from 650–4000 cm^{-1} performed using an Agilent ATR-IR 5500 Series Spectrometer with a spectral resolution of 4 cm^{-1}. For IR measurements, glycine solutions were pipetted directly onto the ATR crystal upon complete dissolution of the glycine solid, and measurements were taken rapidly, well before any significant evaporation from the deposited droplets.

The baseline corrections, which used the “adaptive” baseline setting, and the scaled background subtraction for all spectral data were performed using Spectragryph v.1.2.14. All spectra presented in Figures have had respective H_2O background spectra subtracted followed by individual baseline correction. All peak fitting analyses and deconvolution were performed using OriginPro 2022 with a mixture of Gaussian and Lorentzian functions. Boundaries of 0–1 were set with respect to peak shape to enable varying degrees of Gaussian/Lorentzian character, thus also enabling wholly Gaussian or Lorentzian peak shapes. Peak positions and amplitudes were not fixed and allowed to automatically adjust to acquire the best fit.

2.2. Density functional calculations

Density functional theory calculations for the glycine zwitterion were performed using the Gaussian package (g09-D0.1) [23]. Calculations were performed in two levels of theory using a generalized gradient functional (PBE) [24] and a hybrid density functional (B3LYP) [25,26]. The all-electron correlation consistent aug-ccPVTZ basis set [27,28] is used for all atoms. Optimization of zwitterionic glycine is performed in an implicit water environment using the Polarization Continuum Model (PCM) [29] implemented in the Gaussian package.

Harmonic vibrational frequencies for the relaxed structure were calculated analytically within the Gaussian package. The anharmonic correction was performed using the second order vibrational perturbation theory method implemented in Gaussian 09 [30–32]. No scaling factors were applied to the wavenumbers obtained. The vibAnalysis software [33] was employed to obtain the potential energy distribution of normal modes. In our analysis, we utilize the following notation for the vibrational modes: ν for stretch, δ for bending, and τ for torsion.

3. Results and discussion

In this section we first present DFT results and review published glycine vibrational modes in the literature to identify distinct regions in the zwitterion spectra, and determine the number of vibration modes in each region. Next we present IR and Raman spectra, and use information from DFT calculations to guide the deconvolution of the experimental spectra.

3.1. Calculated vibrational modes

In the geometry-optimised zwitterion molecule, the NH_3^+ group is rotated so that one of the hydrogen atoms point towards one of the carboxyl oxygen atom, with H-bond length of 1.66 Å for PBE and 1.79 Å for B3LYP, which is the same minimum energy structure found in other studies in implicit solvent. The atomic coordinates of the optimised zwitterion molecules for each method are reported in the provided open access dataset (see Data Statement). Computed vibrational modes using the PBE and B3LYP functionals with the harmonic approximation, and anharmonic B3LYP calculations are shown in Table 1 along with intensities, and the computational assignment based on the B3LYP harmonic calculation. As anticipated, the PBE results mainly yield lower wavenumbers compared to the B3LYP results but are generally in good agreement. The anharmonic B3LYP wavenumbers are mostly lower than the harmonic B3LYP wavenumbers. We note that our first two anharmonic wavenumbers were negative and have been omitted, and it has been previously noted that anharmonic calculations are more sensitive to the energy minimisation of the molecular structure [17].

Our B3LYP harmonic results are in good agreement with similar calculations in implicit water carried out by Gontrani et al. [21] and Tortonda et al. [20], with root mean square differences (RMSD) across all 24 modes of 51 and 20 cm^{-1}, respectively. They are also in reasonable agreement with the PCM results of Tiwari et al. [22], although the

Table 1
Calculated vibrational wavenumbers, ν (cm^{-1}), and IR intensities, I (km mol^{-1}), using PBE and B3LYP functionals.

Mode	PBE harmonic		B3LYP harmonic		B3LYP anharmonic		Assignment based on B3LYP harmonic
	ν (cm^{-1})	I (km mol^{-1})	ν (cm^{-1})	I (km mol^{-1})	ν (cm^{-1})	I (km mol^{-1})	
1	104	1	95	1	–	–	67 % $\tau_{OCCH+HCNH}$, 32 % δ_{HCN}
2	302	159	288	0	–	–	62 % $\tau_{OCCN+OCCH+HCNH+CCNH}$, 38 % δ_{NCH}
3	333	0	306	103	229	110	75 % $\delta_{CCN+CCO}$, 24 % ν_{CC}
4	493	97	500	71	496	51	58 % $\nu_{CN+CO+CC}$, 42 % $\delta_{CCN+CCO}$
5	551	2	575	4	531	1	90 % $\tau_{HCNH+CCNH}$
6	670	12	679	13	653	2	49 % $\nu_{CN+CO+CC}$, 49 % $\delta_{CCO+OCO}$
7	840	103	864	112	798	142	69 % ν_{CN+CO}, 29 % δ_{OCO}
8	895	19	929	24	816	30	71 % $\tau_{CCNH+HCNH}$, 29 % δ_{NCH}
9	969	14	979	16	884	11	70 % $\nu_{CN+CO+CC}$, 16 % τ_{OCCH}, 14 % δ_{CCO}
10	1059	1	1100	1	927	55	56 % $\tau_{HCNH+CCNH}$, 36 % $\delta_{CCH+CNH}$
11	1071	196	1105	125	982	84	80 % $\delta_{CCN+CNH+HNH}$, 20 % ν_{CO}
12	1250	110	1306	7	1236	8	48 % $\tau_{NCCO+HCNH}$, 45 % δ_{CNH}
13	1259	5	1325	30	1238	114	55 % δ_{CCH}, 45 % ν_{CO}
14	1284	286	1358	286	1245	184	65 % ν_{CO}, 35 % δ_{OCO}
15	1344	507	1427	533	1260	503	56 % $\delta_{CNH+HNH}$, 38 % ν_{CO}
16	1424	17	1482	16	1331	3	63 % $\delta_{HCH+NCH+CCN}$, 22 % ν_{CO}
17	1574	29	1632	161	1388	57	49 % ν_{CO}, 48 % δ_{HNH}
18	1592	37	1660	42	1459	28	82 % δ_{HNH}, 10 % τ_{OCCN}
19	1635	636	1665	643	1637	532	67 % ν_{CO}, 33 % δ_{HNH}
20	2651	720	3025	478	2626	81	87 % ν_{NH}, 13 % δ_{CNH}
21	3046	5	3115	8	2974	5	100 % ν_{CH}
22	3103	1	3169	1	3008	3	100 % ν_{CH}
23	3413	66	3486	94	3228	60	100 % ν_{NH}
24	3478	99	3543	111	3291	77	100 % ν_{NH}

authors did not report all frequencies. Chowdhry et al. [11] reported a complete set of vibrational frequencies for B3LYP PCM calculations, however, they used an empirical scaling factor to match their frequencies to experiment. A table summarising our calculated values compared to literature is presented in Supporting Information. Our lowest frequency of 95 cm^{-1} is in reasonable agreement with the wavenumber of 70.54 cm^{-1} reported in a recent Hartree-Fock/MP2 study of glycine in water using the conductor-like PCM [34].

We note that the intramolecular hydrogen bond may be less likely to occur in the presence of explicit water molecules, where the H and O molecules may interact with neighbouring water molecules, to give several different glycine conformations. Several studies have calculated vibrational modes for complexes of zwitterionic glycine with various numbers of explicit water molecules [11,12,22]. However, only one of these studies reported the full set of 24 vibrational modes [11] but used empirical scaling. Other studies reported only 19 modes [22], 10 modes [12], 13 modes [17], and 9 modes [14]. The reasons for reporting only a subset of glycine frequencies may be due to the fact that it is difficult to differentiate glycine modes from water modes. Several studies studied systems with increasing numbers of water molecules [12,14,22], however, there was no clear systematic convergence of the reported wavenumbers with increasing number of water molecules. The Supporting Information presents reported literature values for studies with explicit water.

While the presence of water and molecular motion will perturb and smear out the glycine vibrational frequencies, compared to a gas phase molecule, we still expect the experimental spectra to be centred around the wavenumbers corresponding to the global minimum zwitterion structure. To tease the individual wavenumbers out of the experimental spectra we need information about where we expect the peaks to be, and thus we expect that, although the conformations of glycine may be somewhat different in explicit vs implicit water, the calculated DFT modes can be used to guide the analysis of the experimental spectra. Despite variations in the different methods, in all studies using implicit or explicit water, we can identify the same distinct regions of vibrational modes, indicated by horizontal lines in Table 1. There is a clear low wavenumber region up to 700 cm^{-1} that contains six vibrational modes. The next region begins at 800 cm^{-1} with a gap of around 100 cm^{-1} after five modes, which ends at approximately 1100–1200 cm^{-1}. The next region contains eight modes between 1200–1800 cm^{-1}. The highest region contains five modes over 2600 cm^{-1}.

3.2. Identification of vibrational frequencies and modes

The spectroscopic study of glycine in solution was performed over a range of concentrations from 0.0001 mol dm^{-3} to 4 mol dm^{-3} utilising both Raman and IR spectroscopy. Of the studied concentrations, the five most concentrated solutions (0.5 mol dm^{-3} to 4 mol dm^{-3}) were carried forward for further comparison because of a sufficiently strong vibrational response from glycine and better signal-to-noise ratio. Spectra for all concentrations studied are presented in SI and included as a dataset (details in Data Statement). In addition, the 2 mol dm^{-3} solution was discounted in the Raman 2 analysis due to the apparent presence of impurities within this sample and the 4 mol dm^{-3} solution Raman 1 spectrum is not displayed due to sections of the Raman spectrum saturating the detector at this concentration, thus giving an incomplete spectrum for further analysis. The Raman and IR spectra for different concentrations and different instruments are shown in Fig. 1. The peak positions are very consistent between the different concentrations, showing that there are no significant structural or speciation changes at higher concentrations. There is also good agreement between the two Raman data sets, and between the two IR data sets.

We then utilised the knowledge of the calculated glycine vibrational modes to guide the peak deconvolution and identify the individual modes from the experimental spectra. Upon inspection of the experimental spectra in Fig. 1 (see Supporting Information for more detailed spectra in each region), we can identify the corresponding spectral ranges for the regions discussed in the previous subsection. These regions are labelled and presented in Table 2 along with the wavenumber range and the expected number of modes in each region. Knowing the expected number of modes in each region, we then proceeded with deconvolution by analysing the Raman and IR spectra in the different spectral regions.

In region I, only the Raman 1, Raman 2, and IR 1 spectra have data available in this region, as IR 2 has measurements only above 650 cm^{-1}. Below 300 cm^{-1} there are no peaks consistent across the two Raman datasets and the different concentrations. The first distinct peak can be seen in the Raman 2 spectra at around 300 cm^{-1}, which appears as a

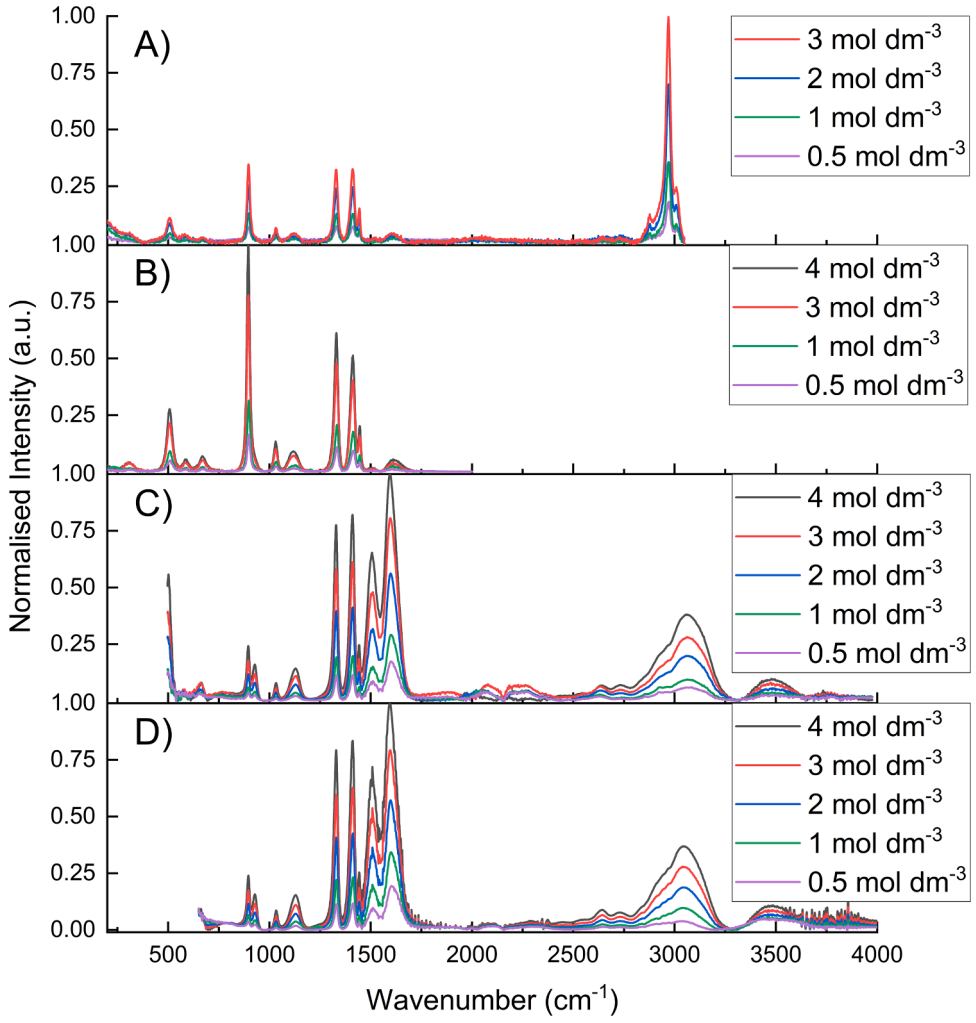

Fig. 1. (A) Raman 1, (B) Raman 2, (C) IR 1, and (d) IR 2 spectra of glycine aqueous solutions ranging from 0.5 mol dm^{-3} to 4 mol dm^{-3}.

Table 2
Expected number of vibrational modes in defined spectral regions.

Region	Range (cm^{-1})	Number of modes
I	0–800	6
II	800–1200	5
III	1200–1800	8
IV	2500–3500	5
All	0–3500	24

shoulder in the Raman 1 spectra. This is likely to correspond to the peak at 305 cm^{-1} identified using THz spectroscopy [18]. There are three identifiable symmetric peaks in the Raman spectra between 500 and 700 cm^{-1}. We can also identify these three peaks in the IR 1 spectra, although only the two higher wavenumber peaks are available. We have identified these four peaks as modes 3–6, while the lowest two modes at 137 cm^{-1} and 170 cm^{-1} were identified in the THz study mentioned above.

In region II, five modes are expected, but only three peaks are clearly visible in the Raman spectra and four in the IR spectra. The strong Raman peak close to 900 cm^{-1} has a weak shoulder at the higher wavenumber side, and corresponds to a second peak in the IR spectra. Thus we have deconvoluted this Raman peak into two modes, which correspond to modes 7 and 8. Of the remaining two peaks in region II, the second peak at around 1150 cm^{-1} shows as a broad peak in Raman and an asymmetric peak in IR. The DFT calculations predict that the upper two modes in this region have very similar wavenumbers, so this peak is also deconvoluted into two modes, 10 and 11, to give the expected five modes in this region.

In region III, eight modes are expected but only five peaks are observed in the IR spectra. The Raman spectra show three strong peaks up to around 1450 cm^{-1}, a very weak peak at around 1500 cm^{-1}, and a weak peak at around 1600 cm^{-1}. We note first that the lowest wavenumber peak is slightly asymmetric with a shoulder on the lower wavenumber side. The calculations show that the first two modes in this region have wavenumbers typically separated by less than 20 cm^{-1}, and thus this peak was deconvoluted into two modes, 12 and 13. The deconvoluted modes are separated by only around 10 cm^{-1}, which is in good agreement with the DFT predictions. The next peak is also asymmetric with a broadening on the lower wavenumber side, which is distinct from the adjacent peak. The DFT harmonic calculations show an expected difference of around 60 cm^{-1} between modes 14 and 15, although the anharmonic calculations predict only a small difference of 15 cm^{-1} between these modes. This peak was deconvoluted into the two modes, 14 and 15. The next small sharp peak at around 1450 cm^{-1} was taken as a single mode, 16, which corresponds to a weak but distinct peak in all DFT calculations. The peak just above 1500 cm^{-1} is slightly asymmetric in the IR spectra, which is likely due to the overlap with the adjacent peaks, and this was thus treated as a single mode, 17. The broad peak above 1600 cm^{-1} is asymmetric with a broadening on the high wavenumber side, suggesting two modes, and has thus been deconvoluted into two modes, 18 and 19.

Finally, in region IV we expect five modes and we observe more peaks in this region than predicted. There are two peaks visible in both Raman and IR at around 2650 cm^{-1} and 2750 cm^{-1}. These peaks were not selected for peak fitting as this region is known to commonly relate to overtones and combination bands and therefore were not considered to be fundamental vibrations of glycine. Therefore the region from 2850–3100 cm^{-1} was selected for deconvolution. The IR spectra were not considered within this region as, even with water background subtraction, there remained too much ambiguity in the remaining signal. This range is outside the range of the spectrometer used to collect the Raman 2 spectra and, therefore, only the Raman 1 spectra are available for this region. The first two peaks are asymmetric and thus they were each deconvoluted into two modes, corresponding to modes 20–23. The third peak was treated as a single mode due to its lack of asymmetry and sharpness and was assigned as mode 24.

The mode wavenumbers and standard errors for each instrument in both IR and Raman spectroscopies, averaged across all concentrations, are presented in Table 3. The largest errors across concentrations are for modes 17–19 from Raman spectra, which is a consequence of fitting very weak peaks. As can be seen from the data, the agreement across instruments and spectroscopies is excellent. The resulting wavenumbers and standard error for the 22 vibrational modes uncovered in this work averaged across all instruments and all concentrations is presented in Table 4.

In Fig. 2 we show the differences between the mode wavenumber of each individual experimental measurement and the average experimental mode wavenumbers shown in Table 4. It can be seen that the individually measured mode wavenumbers are approximately evenly distributed around the average values. There is no obvious bias with respect to measurement technique or solution concentration. This data illustrates that individual measurements are routinely 5–10 cm^{-1} different from the true value. One also has to consider that the instrument resolution is 4 cm^{-1} except for Raman 1, which is 1 cm^{-1}.

3.3. Comparison and reconciliation with literature results

Utilising the newly established dataset following deconvolution, a comparison to all available IR and Raman literature of zwitterionic glycine was performed to establish similarities to the literature as well as discovery of previously unidentified modes or peaks. This comparison is presented in Table 4. We have attempted to align modes, but where wavenumbers significantly differ from other results we have included the modes on separate rows. The available Raman and IR peak positions in this work are, generally, in good agreement with the present work with a few exceptions. We note that although the literature investigated different concentrations of solutions, based on our results, as expected, we did not see a variation in mode wavenumber with concentration.

Table 3
Average deconvoluted mode wavenumbers and standard errors (cm^{-1}) for each data set.

Mode	Raman 1		Raman 2		IR 1		IR 2	
3	307	±3.0	306	±1.8	–	–	–	–
4	507	±0.6	508	±0.4	498	±1.4	–	–
5	584	±1.2	586	±0.3	576	±1.1	–	–
6	673	±0.8	671	±0.3	658	±0.7	–	–
* 7	897	±0.1	897	±0.7	895	±0.1	896	±0.0
* 8	929	±2.1	933	±1.0	928	±0.1	929	±0.2
9	1032	±0.2	1031	±0.5	1032	±0.1	1033	±0.1
* 10	1111	±1.1	1110	±4.0	1109	±1.1	1117	±0.7
* 11	1134	±1.7	1131	±3.0	1135	±0.6	1139	±1.4
* 12	1314	±3.1	1322	±1.3	1321	±0.2	1319	±0.7
* 13	1330	±0.3	1332	±0.7	1332	±0.2	1331	±0.2
* 14	1392	±2.1	1402	±1.6	1406	±1.1	1406	±0.5
* 15	1411	±0.3	1413	±0.8	1415	±0.2	1414	±0.2
16	1444	±0.1	1446	±0.1	1445	±0.1	1443	±0.1
17	1516	±2.7	1492	±4.8	1509	±0.4	1508	±0.9
* 18	1600	±4.6	1600	±3.5	1596	±1.8	1593	±1.7
* 19	1636	±8.3	1634	±6.3	1625	±3.0	1623	±2.5
* 20	2863	±2.9	–	–	–	–	–	–
* 21	2876	±0.6	–	–	–	–	–	–
* 22	2953	±1.0	–	–	–	–	–	–
* 23	2970	±0.1	–	–	–	–	–	–
24	3013	±0.2	–	–	–	–	–	–

An asterisk indicates the mode is from a deconvoluted peak.

In region I, the two lowest modes were reported previously using THz spectroscopy [18]. Mode 3 at 305 cm^{-1} is in agreement with the mode reported in the same THz study. The modes at 505 cm^{-1}, 583 cm^{-1} and 671 cm^{-1} have been previously identified in Raman spectra [9,11–13]. The Raman study by Krauklis et al. [14] also observed three modes, but their peaks at 603 cm^{-1} and 698 cm^{-1} correspond to those of α-glycine [10] and it is, therefore, likely that the spectra are from a slurry containing solid particles rather than fully dissolved glycine. Similarly, the IR study by Kumar et al. [16] (values taken from their Fig. 3 for IR spectra of glycine in water) appear to show α-glycine peaks. The study by Shi et al. [10] investigated a glycine solution using Raman spectroscopy, but did not detect any modes for the zwitterion below 510 cm^{-1}.

In region II, the Raman and IR literature reported two or three peaks, but not always the same peaks. Only the IR study by Max et al. reported five modes [15]. All Raman studies reported a single mode from the strong peak between 893–900 cm^{-1}, but did not identify the small shoulder corresponding to mode 8. The IR study of Max et al. [15] reported two modes and the agreement with our results is excellent, with mode 8 only reported in the present study and by Max et al. There is more variation in modes 10 and 11, which we deconvoluted from a broad peak. Several studies reported a single wavenumber between 1109 cm^{-1} and 1130 cm^{-1}. The Raman study by Chowdhry et al. reported two modes at 1110 cm^{-1} and 1130 cm^{-1} and the IR study by Max et al. reported modes at 1110 cm^{-1} and 1132 cm^{-1}, which are in excellent agreement with our results of 1114 cm^{-1} and 1137 cm^{-1}.

In region III, most studies identified between two and four peaks. Six peaks (12, 13, 15, 16, 17 and 19) were identified by both Furic et al. [9], using Raman, and by Lutz et al., using IR. All studies found a peak at around 1330 cm^{-1} and Max et al., and Furic et al. reported a peak at 1320 cm^{-1}. This supports our deconvolution of the asymmetric peak at around 1330 cm^{-1} into modes 12 and 13.

Max et al. reported an additional peak at 1598 cm^{-1} in their IR spectra, supporting the deconvolution of the broad, asymmetric peak at around 1600 cm^{-1} into modes 18 and 19. This region commonly represents vibrations due to water, however, it was found here that the peak intensity, once again, increases and broadens slightly with an increase in glycine concentration, therefore signifying some contribution from glycine.

Within region IV, five previous works provide peak positions [9–11,16,17]. There is general agreement about the most intense peak at 2970 cm^{-1}, and the second most intense at 3013 cm^{-1}. Furic observed a peak at 2880 cm^{-1}, which is close to the lower intensity peak. Kumar et al. listed peaks at 2898 cm^{-1} and 3170 cm^{-1}, but again, this could be affected by undissolved solid glycine. Aside from this, there is no good agreement. As discussed previously, it is known that five vibrational modes should exist within this region, and thus we proposed that the additional vibrations occur at 2863 cm^{-1}, 2876 cm^{-1}, and 2953 cm^{-1}. Furic et al. [9] attributed their peaks above 3235 cm^{-1} to water.

Overall, all 22 vibrational modes above 300 cm^{-1} were resolved, making this study the most complete and reliable set of vibrational modes for the glycine zwitterion in aqueous solution. By combining our results with the two lowest THz wavenumbers by Sun et al., we have defined a complete set of 24 wavenumbers from experiment, which we now use for benchmarking the DFT calculations. In Fig. 3 we show the differences between the calculated wavenumbers for various DFT methods and the average experimental mode wavenumbers shown in Table 4. We have used ascending order for calculated and experimental wavenumbers rather than aligning wavenumbers using the vibrational

Table 4
Present average experimental wavenumbers and standard error across all instruments and concentrations for zwitterion glycine in aqueous solution, alongside existing literature data.

			Literature									
			Raman						IR			THz
Mode	Present		[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]
1	–		–	–	–	–	–	–	–	–	–	137
2	–		–	–	–	–	–	–	–	–	–	170
3	306	±1.5	–	–	–	–	–	–	–	–	–	305
4	504	±1.5	507	510	506	506	507	492	–	503	–	–
5	582	±1.5	585	–	585	584	593	–	–	–	–	–
–		–	–	–	–	–	–	603	–	608	–	–
6	667	±2.1	671	–	669	670	675	–	–	–	–	–
–		–	–	–	–	–	–	698	–	697	–	–
* 7	896	±0.2	897	899	898	899	900	893	897	893	898	–
* 8	930	±0.7	–	–	–	–	–	–	929	–	–	–
9	1032	±0.2	1031	1034	–	1033	1037	1036	1034	1032	1032	–
* 10	1112	±1.2	–	–	1110	–	–	1109	1110	–	–	–
* 11	1135	±1.0	1121	–	1130	1123	–	–	1132	–	1130	–
* 12	1319	±1.0	1320	–	–	–	–	–	–	–	1320	–
* 13	1331	±0.3	1330	1334	1331	1329	1338	1326	1331	1332	1330	–
* 14	1403	±1.5	–	–	–	–	–	–	–	–	–	–
* 15	1414	±0.4	1412	1408	1412	1413	1419	1410	1412	1412	1412	–
16	1444	±0.3	1444	1443	1445	1433	1445	–	1444	–	1444	–
17	1508	±1.3	1514	–	1511	–	–	–	1510	1510	1510	–
* 18	1597	±1.5	–	–	–	–	–	–	1598	–	–	–
* 19	1629	±2.7	1620	–	–	1618	–	–	–	1615	1608	–
–		–	–	–	–	–	1639	–	–	–	–	–
–		–	–	–	–	–	–	1670	–	–	–	–
		–	–	–	–	–	–	–	–	2124	–	–
* 20	2863	±2.9	–	–	–	–	–	–	–	–	–	–
* 21	2876	±0.6	2880	–	–	–	–	–	–	2898	–	–
* 22	2953	±1.0	–	–	–	–	–	–	–	–	–	–
* 23	2970	±0.1	2972	2976	2971	–	–	–	–	–	2968	–
24	3013	±0.2	3016	3012	3013	–	–	–	–	–	3034	–
–		–	–	3128–3528	–	–	–	–	–	3170	–	–
–		–	3235	–	–	–	–	–	–	–	–	–
–		–	3415	–	–	–	–	–	–	–	–	–
–		–	3635	–	–	–	–	–	–	–	–	–

An asterisk indicates the mode is from a deconvoluted peak.

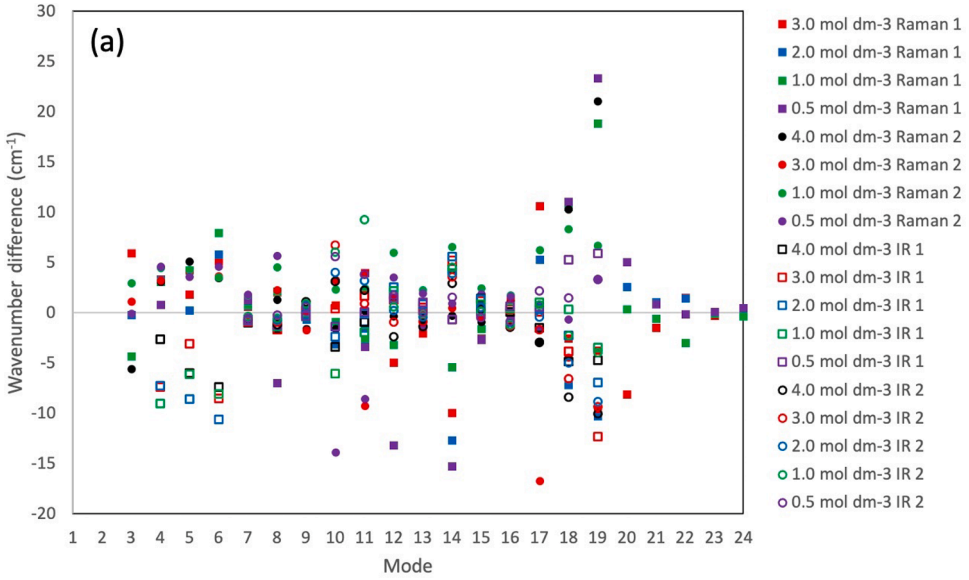

Fig. 2. Wavenumber difference between individual experimental datasets and the average experimental wavenumbers.

modes, as is often done in literature. We note that the decomposition of the vibrational mode into components often gives mixed modes, and it is unclear how reliable the mode decomposition is. By using ascending order, the root mean square differences (RMSDs) used here to compare the different methods will be lower than if the modes were ordered based on the mode assignment.

In region IV, it is clearly seen that there is a significant overestimation of wavenumbers for N–H and C–H stretching modes, which is typical for DFT calculations of isolated and solvated molecules. The RMSDs for implicit water calculations range from 368 cm^{-1} (present),

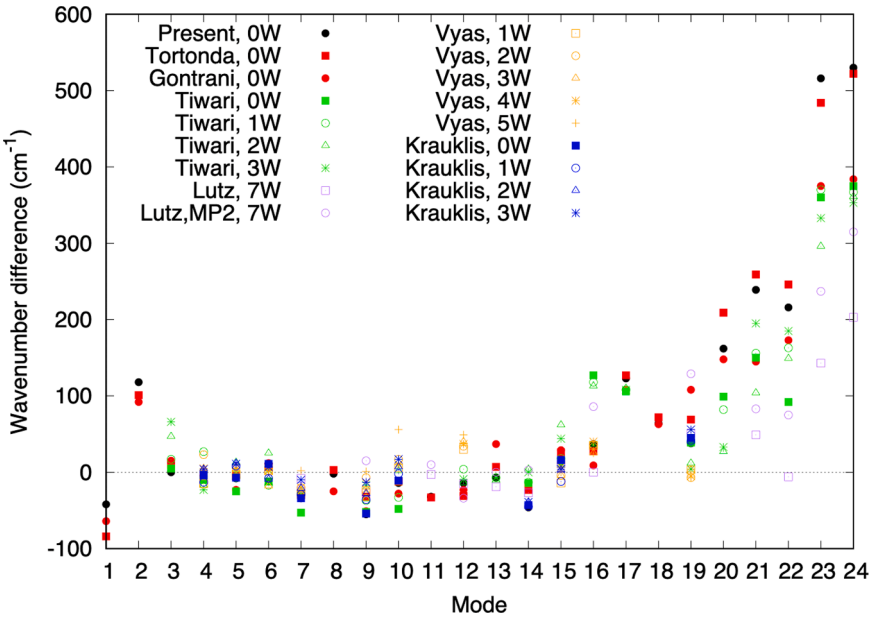

Fig. 3. Wavenumber difference between unscaled DFT calculations using implicit or explicit water (where *i*W denotes the number *i* of explicit water molecules) and benchmark experimental wavenumbers. Key refers to calculations by Tortonda et al. [20], Gontrani et al. [21], Tiwari et al. [22], Lutz et al. [17], Vyas et al. [12], and Krauklis et al. [14].

269 cm^{-1} [21] and 249 cm^{-1} [22] for B3LYP, and 368 cm^{-1} [20] for B3PW91 functionals. The study by Tiwari et al. added explicit water molecules but showed no systematic improvement, with RMSD values of 249 cm^{-1} for implicit water and 257 cm^{-1}, 224 cm^{-1}, and 249 cm^{-1} for one to three water molecules, respectively [22]. Including an anharmonic approximation gives good results in calculations with seven explicit water molecules, yielding RMSDs of 127 cm^{-1} and 205 cm^{-1} for the four reported wavenumbers for B3LYP and MP2, respectively [17].

In regions II and III (modes 7–11 and 12–19) the agreement of the B3LYP/B3PW91 harmonic implicit water calculations with experiment is reasonably good with RMSDs ranging from 25–65 cm^{-1} in each region. The anharmonic results of Lutz et al. with seven explicit water gave a lower RMSD for the modes reported (three out of five in region II and six out of eight in region III), however, for region III they did not report wavenumbers for modes 17 and 18, which are the modes with higher differences in other studies. In general, in all regions there is little systematic improvement gained by increasing the number of explicit water molecules [12,14,22].

Overall, the RMSDs between the DFT wavenumbers and the experimental wavenumbers for the present studies are 55 cm^{-1}, 42 cm^{-1} and 114 cm^{-1} for PBE, B3LYP and B3LYP anharmonic calculations, respectively. The anharmonic approximation gives a significant improvement in the high wavenumber region. Based on the literature results, the addition of explicit water molecules does not appear to systematically improve agreement with experiment. The exchange and correlation functional has the most significant effect on the results, as evidenced by present results, and also the B3LYP vs MP2 calculations by Lutz et al. [17]. The experimental dataset presented here will thus provide a useful benchmark for future calculations of vibrational modes of zwitterionic glycine.

4. Conclusion

Glycine is widely used in pharmaceutical industries and biological processes, yet reported vibrational modes for the glycine zwitterion, which is the dominant form in aqueous solution, have remained incompletely reported. This work obtained IR and Raman spectra for a range of different glycine solution concentrations to enable vibrational modes to be distinguished. The IR and Raman measurements were each performed using two different instruments to check consistency and to maximise the spectral range. A review of DFT calculations of glycine zwitterion vibrational wavenumbers in implicit water and with explicit water molecules were used to help interpret the experimental spectra. Despite variations in wavenumbers between different calculation methods, the DFT modes fell into four distinct regions. The DFT calculations were used to determine how many modes were expected in each region and helped to deconvolute the spectra into the correct number of vibrational modes.

Excellent agreement was found between the deconvoluted experimental datasets across different instruments and solution concentrations as expected. Overall, 22 out of 24 vibrational modes were resolved, making this study the most complete and reliable set of vibrational modes for the glycine zwitterion in aqueous solution. Only the lowest two vibration modes below 300 cm^{-1} could not be measured in this study as these were below the cut off of the edge filter in Raman instruments. This technique for obtaining reliable vibrational modes is particularly useful in resolving uncertainty in the vibrational frequencies presented in current literature, and we were able to determine missing and erroneous wavenumbers reported in the literature. Combining the present 22 wavenumbers with the two lowest frequencies from a previous THz study [18] has thus enabled us to present the first complete set of all 24 wavenumbers for the glycine zwitterion in aqueous solution.

This benchmark experimental data set provides a strong basis from which further vibrational calculations of zwitterionic glycine in the solution can be performed. A comparison of DFT calculations to these benchmark experimental results has shown that the exchange and correlation function has a significant effect on the accuracy of the predicted wavenumbers and that anharmonic corrections give significant improvement to results in the high wavenumber region. However, the addition of explicit water molecules to DFT calculations does not appear to provide any systematic improvement in the accuracy of DFT calculated wavenumbers.

In summary, this complete set of vibrational wavenumbers for the glycine zwitterion in aqueous solution provides a useful benchmark for future validation and refinement of vibrational calculation methodologies, as well as enabling a deeper interpretation and analysis for future spectroscopic experiments of complex glycine aqueous systems.

CRediT authorship contribution statement

Jan Sefcik: Conceptualization, Formal analysis, Funding acquisition, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Karen Johnston:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. **Karen Faulds:** Funding acquisition, Methodology, Supervision, Writing – original draft. **Mozhdeh Mohammadpour:** Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Mark Christie:** Data curation, Formal analysis, Investigation, Methodology, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data underpinning this publication, DFT input and output files and raw and processed experimental spectra, are openly available from the University of Strathclyde KnowledgeBase at https://doi.org/10.15129/6e1b685e-6044-47de-ad9d-60166a22b1f2.

Acknowledgements

This project was supported by a Leverhulme Trust Research Project Grant (RPG-2019-197). Results were obtained using the ARCHIE-WeSt High-Performance Computer (www.Archie-west.ac.uk) based at the University of Strathclyde.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.vibspec.2025.103783.

References

[1] A. Rimola, M. Corno, C.M. Zicovich-Wilson, P. Ugliengo, Ab initio modeling of protein/biomaterial interactions: glycine adsorption at hydroxyapatite surfaces, J. Am. Chem. Soc. 130 (48) (2008) 16181–16183.
[2] M.V. Fedotova, S.E. Kruchinin, Ion-binding of glycine zwitterion with inorganic ions in biologically relevant aqueous electrolyte solutions, Biophys. Chem. 190-191 (2014) 25–31.
[3] H. Li, E. Oraby, J. Eksteen, Extraction of copper and the co-leaching behaviour of other metals from waste printed circuit boards using alkaline glycine solutions, Resour., Conserv. Recycl. 154 (2020) 104624.
[4] A. Jawor-Baczynska, J. Sefcik, B.D. Moore, 250 nm glycine-rich nanodroplets are formed on dissolution of glycine crystals but are too small to provide productive nucleation sites, Cryst. Growth Des. 13 (2) (2013) 470–478.
[5] G. Zimbitas, A. Jawor-Baczynska, M.J. Vesga, N. Javid, B.D. Moore, J. Parkinson, J. Sefcik, Investigation of molecular and mesoscale clusters in undersaturated glycine aqueous solutions, Colloids Surf. A Physicochem. Eng. Asp. 579 (2019) 123633.
[6] P.U. Jepsen, D.G. Cooke, M. Koch, Terahertz spectroscopy and imaging – modern techniques and applications, Laser Photonics Rev. 5 (1) (2011) 124–166.
[7] T. De Beer, A. Burggraeve, M. Fonteyne, L. Saerens, J.P. Remon, C. Vervaet, Near infrared and raman spectroscopy for the in-process monitoring of pharmaceutical production processes, Int. J. Pharm. 417 (1) (2011) 32–47.
[8] M.J. Baker, K. Faulds, Fundamental developments in clinical infrared and raman spectroscopy, Chem. Soc. Rev. 45 (2016) 1792–1793.
[9] K. Furić, V. Mohaček, M. Bonifačić, I. Štefanić, Raman spectroscopic study of H_2O and D_2O water solutions of glycine, J. Mol. Struct. 267 (1992) 39–44.
[10] Y. Shi, L. Wang, Collective vibrational spectra of α-and γ-glycine studied by terahertz and Raman spectroscopy, J. Phys. D Appl. Phys. 38 (19) (2005) 3741.
[11] B.Z. Chowdhry, T.J. Dines, S. Jabeen, R. Withnall, Vibrational spectra of α-amino acids in the Zwitterionic state in aqueous solution and the solid state: DFT calculations and the influence of hydrogen bonding, J. Phys. Chem. A 112 (41) (2008) 10333–10347.
[12] N. Vyas, A.K. Ojha, A. Materny, Simulation of the Raman spectra of zwitterionic glycine+ nH2O (n = 1, 2,…, 5) by means of DFT calculations and comparison to the experimentally observed Raman spectra of glycine in aqueous medium, Vib. Spectrosc. 55 (1) (2011) 69–76.
[13] Y. Numata, M. Otsuka, K. Yamagishi, H. Tanaka, Quantitative determination of glycine, alanine, aspartic acid, glutamic acid, phenylalanine, and tryptophan by Raman spectroscopy, Anal. Lett. 50 (4) (2017) 651–662.
[14] I.V. Krauklis, A.V. Tulub, A.V. Golovin, V.P. Chelibanov, Raman spectra of glycine and their modeling in terms of the discrete–continuum model of their water solvation shell, Opt. Spectrosc. 128 (10) (2020) 1598–1601.
[15] J.-J. Max, M. Trudel, C. Chapados, Infrared titration of aqueous glycine, Appl. Spectrosc. 52 (2) (1998) 226–233.
[16] S. Kumar, A.K. Rai, V.B. Singh, S.B. Rai, Vibrational spectrum of glycine molecule, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 61 (11-12) (2005) 2741–2746.
[17] O.M.D. Lutz, C.B. Messner, T.S. Hofer, L.R. Canaval, G.K. Bonn, C.W. Huck, Computational vibrational spectroscopy of glycine in aqueous solution - fundamental considerations towards feasible methodologies, Chem. Phys. 435 (2014) 21–28.
[18] J. Sun, G. Niehues, H. Forbert, D. Decka, G. Schwaab, D. Marx, M. Havenith, Understanding THz spectra of aqueous solutions: glycine in light and heavy water, J. Am. Chem. Soc. 136 (13) (2014) 5031–5038.
[19] K.Z. Gaca-Zajac, B.R. Smith, A. Nordon, A.J. Fletcher, K. Johnston, J. Sefcik, Investigation of IR and Raman spectra of species present in formaldehyde-water-methanol systems, Vib. Spectrosc. 97 (2018) 44–54.
[20] F.R. Tortonda, J.-L. Pascual-Ahuir, E. Silla, I. Tunon, F.J. Ramirez, Aminoacid zwitterions in solution: geometric, energetic, and vibrational analysis using density functional theory-continuum model calculations, J. Chem. Phys. 109 (2) (1998) 592–603.
[21] L. Gontrani, B. Mennucci, J. Tomasi, Glycine and alanine: a theoretical study of solvent effects upon energetics and molecular response properties, J. Mol. Struct. THEOCHEM 500 (1-3) (2000) 113–127.
[22] S. Tiwari, P.C. Mishra, S. Suhai, Solvent effect of aqueous media on properties of glycine: significance of specific and bulk solvent effects, and geometry optimization in aqueous media, Int. J. Quantum Chem. 108 (5) (2008) 1004–1016.
[23] 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. Petersson, et al., Gaussian, Inc 09, Revision D. 01, Gaussian, Inc., Wallingford CT, 2009. ⟨http://www.gaussian.com⟩.
[24] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (18) (1996) 3865.
[25] A.D. Becke, A new mixing of Hartree-Fock and local density-functional theories, J. Chem. Phys. 98 (2) (1993) 1372–1377.
[26] R.H. Hertwig, W. Koch, On the parameterization of the local correlation functional. What is Becke-3-LYP? Chem. Phys. Lett. 268 (5-6) (1997) 345–351.
[27] T.H. Dunning Jr., Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen, J. Chem. Phys. 90 (2) (1989) 1007–1023.
[28] R.A. Kendall, T.H. Dunning Jr, R.J. Harrison, Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions, J. Chem. Phys. 96 (9) (1992) 6796–6806.
[29] J. Tomasi, B. Mennucci, R. Cammi, Quantum mechanical continuum solvation models, Chem. Rev. 105 (8) (2005) 2999–3094.
[30] V. Barone, Vibrational zero-point energies and thermodynamic functions beyond the harmonic approximation, J. Chem. Phys. 120 (7) (2004) 3059–3065.
[31] V. Barone, Anharmonic vibrational properties by a fully automated second-order perturbative approach, J. Chem. Phys. 122 (1) (2004) 014108.
[32] C. Cappelli, F. Lipparini, J. Bloino, V. Barone, Towards an accurate description of anharmonic infrared spectra in solution within the polarizable continuum model: reaction field, cavity field and nonequilibrium effects, J. Chem. Phys. 135 (10) (2011) 104505.
[33] F. Teixeira, M. Natália, D.S. Cordeiro, Improving vibrational mode interpretation using Bayesian regression, J. Chem. Theory Comput. 15 (1) (2019) 456–470.
[34] R. Boča, J. Štofko, R. Imrich, Effect of solvation on glycine molecules: a theoretical study, ACS Omega 8 (31) (2023) 28577–28582.