Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Probing Local Electrostatics of Glycine in Aqueous Solution by THz Spectroscopy

Federico Sebastiani, Chun Yu Ma, Sarah Funke, Alexander Bäumer, Dominique Decka, Claudius Hoberg, Alexander Esser, Harald Forbert, Gerhard Schwaab, Dominik Marx, and Martina Havenith**

anie_202014133_sm_miscellaneous_information.pdf

Supporting Information

Probing local electrostatics of glycine in aqueous solution by THz spectroscopy

Federico Sebastiani[a], Chun Yu Ma[a], Sarah Funke[a], Alexander Bäumer[a], Dominique Decka[a], Claudius Hoberg[a], Alexander Esser[b], Harald Forbert[c], Gerhard Schwaab[a], Dominik Marx[b], and Martina Havenith*[a]

[a]Lehrstuhl für Physikalische Chemie II, Ruhr-Universität Bochum, 44780 Bochum, Germany. E-mail: martina.havenith@rub.de
[b]Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany
[c]Center for Solvation Science ZEMOS, Ruhr-Universität Bochum, 44780 Bochum, Germany

Materials and Methods

L-glycine, L-valine and diglycine with a nominal purity higher than 99% were purchased from Sigma Aldrich. The samples were lyophilized overnight and then dissolved without further purification in HPLC grade ultrapure water or in NaOH/HCl aqueous solutions at different concentrations. The concentrations of the glycine, valine and diglycine solutions were 1 M, 0.3 M and 0.5 M, respectively. The mass density of the solutions at 20°C was measured by an Anton-Paar DMA58 density meter to calculate accurately the concentration of solutes and solvent in all the samples. The amount of hydrolysis was determined by pH measurements. For the 1 M glycine aqueous solution before NaOH/HCl titration, the pH was about 6.

Broadband THz-Far Infrared Fourier Transform absorption measurements were performed using a Bruker Vertex 80V FTIR spectrometer with an average aperture of 4 mm and equipped with a liquid He cooled bolometer (Infrared Laboratories) as detector and a mercury arc lamp as light source and a Mylar multilayer beamsplitter. The sample compartment was continuously purged with N_2 gas to minimize atmospheric water vapor lines and kept at temperature controlled conditions (20±0.2°C), while the interferometer compartment was evacuated. The samples were measured in a Bruker liquid cell with chemical vapor deposition-grown diamond windows (Diamond Materials, GmbH). The sample layer thickness is fixed by Kapton spacers inbetween the windows with a nominal thickness of 25 μm. The exact layer thickness was determined by recording the etalon fringes of the empty cell by mid-infrared absorption spectroscopy. Each sample was measured several times in the frequency range from 50 to 400 $^{-1}$ as an average over at least 128 scans with a typical resolution of 1 cm^{-1}.

For the **THz time-domain spectrometer** (THz-TDS), a pulsed 800 nm Ti:Sa laser with a repetition rate of 80 MHz and a pulse length of 70 fs (MaiTai DeepSee eHP, SpectraPhysics) was employed as a light source, which is split into a pump and a probe beam. The pump beam is focused on a Tera-SED3 antenna (Gigaoptics), biased with 20 V and a modulation frequency of 92 kHz. The emitted THz radiation passes a 100 μm-thick Bruker liquid cell with z-cut quartz windows and overlaps with the probe beam on a 0.5 mm ZnTe crystal. The electric field of the THz radiation is detected via electric-optical sampling. A Lock-In amplifier (Zürich Instruments) with a 3 dB filter and a time constant of 1 μs was employed. By using an additional fast delay stage (5 Hz, ScanDelay 150, APE) in the pump beam path, a time window of 8 ps was covered. To correct the signal for fluctuations in temperature, humidity and laser power, we used a sample-changer to

record transmission in a sample and a reference cell. Within 30 s we reach a dynamic range of 50 dB for aqueous solutions. All measurements were carried out at 20°C and low humidity (<5%). The recorded THz pulses from the THz-TDS were Fourier transformed into the power spectrum. We obtained the absorption spectrum via an iterative standing wave correction approach as reported in Ref.[1]. For all recorded spectra we determine the density corrected effective absorption $\alpha^{\text{eff}}_{Sample}(\nu, c)$:[2–4]

$$\alpha^{\text{eff}}_{Sample}(\nu, c) = \alpha_{solution}(\nu, c) - \frac{c_w}{c^0_w}\alpha_{water}(\nu) \tag{S.1}$$

with $\alpha_{solution}$ and α_{water} being the absorption coefficients of the aqueous solution and of bulk water, and c_w and c^0_w being the concentrations of water in the solution as obtained from density measurements of the solutions and of bulk water at room temperature, respectively. Using water as a substance with a very similar index of refraction, artifacts due to reflections at the cell windows can be minimized. In addition, this procedure allows us to remove nonlinearities caused by a solute(s) concentration dependent change of the apparent molar volume. In fact, $\alpha^{\text{eff}}_{Sample}$ contains now only the contributions from the solvated aminoacid or peptide as well as any change in the water absorption induced by the solute itself compared to bulk water.

Principal Component Analysis (PCA)

For the analysis of the measured THz absorption spectra we applied a principal component analysis (PCA) using a singular value decomposition approach as described previously.[2,3] This approach allows us to orthogonalize spectral information, i.e. frequency dependency of the important components, and chemical information (importance of individual components as function of concentration).
In summary, an input matrix M consisting of individual spectra as row vectors is decomposed into a score matrix U, a singular value matrix Σ and a loadings matrix V:

$$M = U \Sigma V^T \tag{S.2}$$

Principal components (PCs) are obtained by the product $PC_i = \sigma_{i,i}\, l_i$, where $\sigma_{i,i}$ is a diagonal element of the singular value matrix Σ and l_i is a row vector of the transposed loadings matrix V^T. The entries (s) of the score matrix U quantify the concentration dependent contribution of the corresponding principal component to a spectrum. In our case, the total number of principal components is limited by the number of measured spectra. PCs are ordered according to their explained variance. While the most important PCs provide physically meaningful information, the less important PCs contain contributions from statistical noise and systematic errors like changes in residual water vapor absorption. For all spectra discussed in this work, two principal components were found to be sufficient to reconstruct all the measured spectra within experimental uncertainty.

PCA was performed by using Wolfram Mathematica 11 without any preprocessing of the data. As a remark, we did not mean-center our data, since the first principal component resulting from our analysis is very close to the average spectrum obtained during our measurements.

For the titration with HCl we found two components with an explained variance of 99.8 % and 0.2 % for the low and 99.6 % and 0.4 % for the high frequency regime (Fig. S1). For the titration with NaOH we found two components with an explained variance of 96.7 % and 3.2 % for the low and 99.0 % and 0.9 % for the high frequency regime (Fig. S2).

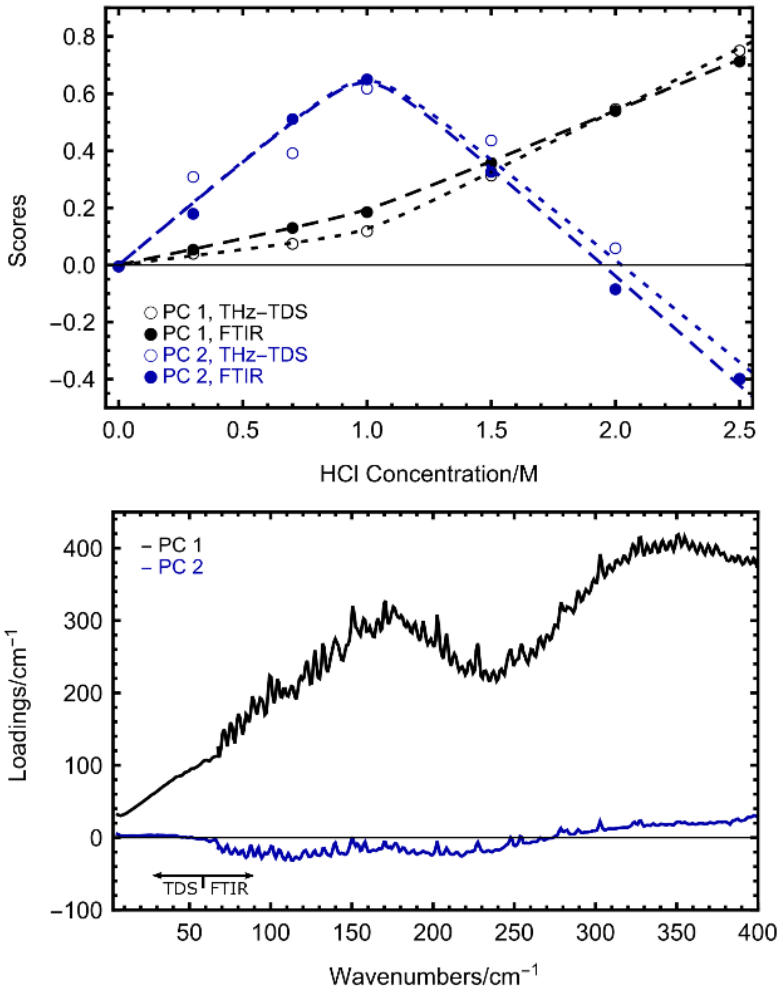

Figure S1. Results of the PCA performed on the difference spectra of the titration of glycine with HCl ($\alpha^{\mathrm{eff}}_{\mathrm{Gly+HCl}} - \alpha^{\mathrm{eff}}_{\mathrm{Gly}}$), see Fig. 1 of the main text. The scores are displayed in the upper panel and the corresponding principal components are located in the lower panel. A physical model (see text for details) was used to fit the scores as a function of concentration. Best fit models for the THz-TDS and FTIR measurements are represented as dotted and dashed lines, respectively.

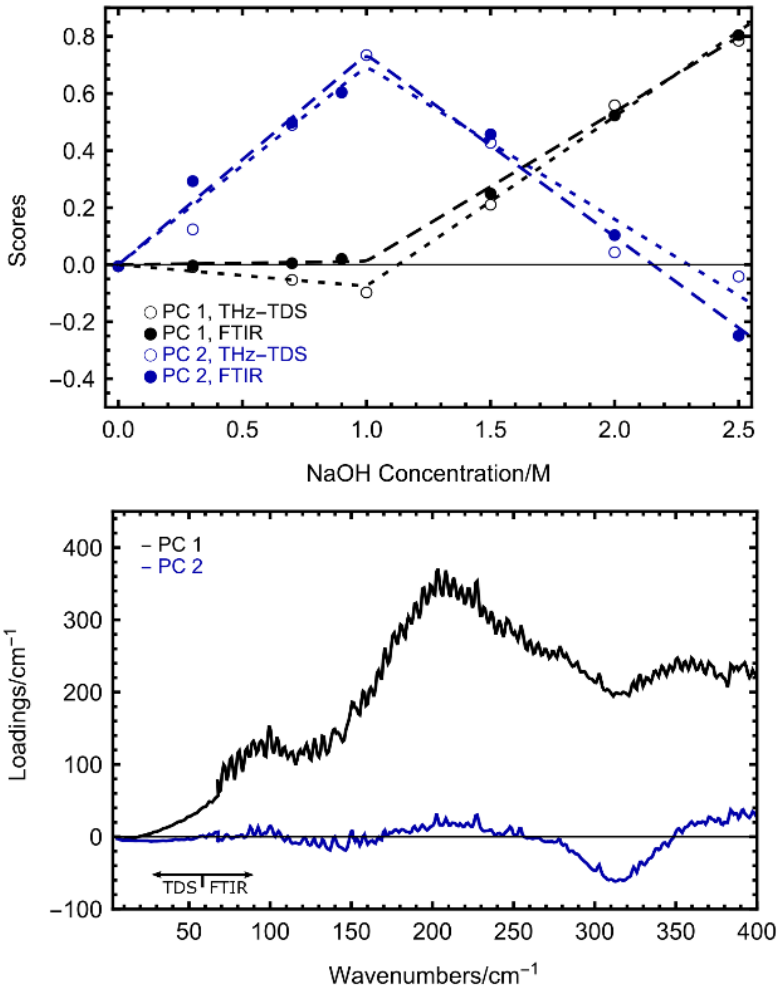

Figure S2. Results of the PCA performed on the difference spectra of the titration of glycine with NaOH ($\alpha^{\mathrm{eff}}_{\mathrm{Gly+NaOH}} - \alpha^{\mathrm{eff}}_{\mathrm{Gly}}$). The scores are displayed in the upper panel and the corresponding principal components are located in the lower panel. A physical model (see text for details) was used to fit the scores as a function of concentration. Best fit models for the THz-TDS and FTIR measurements are represented as dotted and dashed lines, respectively.

The scores of the low and the high frequency regimes follow the same trend, revealing the same physical background. Interestingly, all scores show a turning point at 1 M HCl/NaOH (i.e. at 1 equimolar with respect to glycine). This indicates that the low frequency modes are strongly coupled to the peaks observed by FTIR measurements and simulations, although none of those modes is evidently present in

the THz-TDS spectra. To receive the dissociation degree β of 1 M glycine we calculated the pH at different HCl (Eqs. S.3a, S.4a) and NaOH (Eqs. S.3b, S.4b) concentrations c assuming full dissociation:

$$\mathrm{pH} = -\log\,(c - x) \tag{S.3a}$$

$$\mathrm{pH} = 14 + \log\,(c - x) \tag{S.3b}$$

with

$$x = \frac{(1\mathrm{M} + c + 10^{-\mathrm{p}K_A(\mathrm{COOH})}) \pm \sqrt{(1\mathrm{M} + c + 10^{-\mathrm{p}K_A(\mathrm{COOH})})^2 - 4\mathrm{M}\cdot c}}{2} \tag{S.4a}$$

$$x = \frac{(1\mathrm{M} + c + 10^{-\mathrm{p}K_B(\mathrm{NH_2})}) \pm \sqrt{(1\mathrm{M} + c + 10^{-\mathrm{p}K_B(\mathrm{NH_2})})^2 - 4\mathrm{M}\cdot c}}{2}\,. \tag{S.4b}$$

Here, for the carboxyl group $\mathrm{p}K_A(\mathrm{COOH})$ is 2.34 and for the amino group $\mathrm{p}K_B(\mathrm{NH_2})$ is 4.40. We subsequently calculated the dissociation degree β for the titration with HCl (Eq. S.5a) and NaOH (Eq. S.5b) via the Henderson-Hasselbalch equation and $\mathrm{p}K_A(\mathrm{NH_2})$ = 9.60 for the amino group. The resulting dissociation degrees are displayed in Fig. S3.

$$\beta = \frac{10^{\mathrm{pH}}}{10^{\mathrm{pH}} + 10^{\mathrm{p}K_A C}} \tag{S.5a}$$

$$\beta = \frac{10^{\mathrm{pH}}}{10^{\mathrm{pH}} + 10^{\mathrm{p}K_A N}} \tag{S.5b}$$

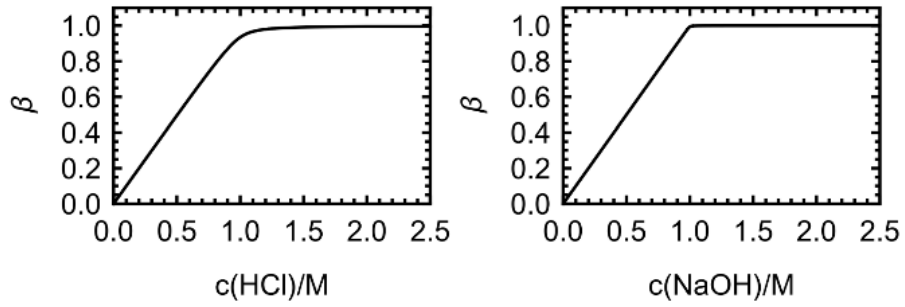

Figure S3. Calculated dissociation degree β of 1 M glycine as a function of the HCl/NaOH concentration.

The calculation of the dissociation degree reveals no zwitterionic glycine ($Gly^{\pm}$) at around 1 M HCl/NaOH. Subsequently, the scores below 1 M HCl/NaOH are dominated by spectral changes due to the protonation/deprotonation of the glycine molecules ($Gly^{+}\cdot Cl^{-}$/$Gly^{-}\cdot Na^{+}$), whereas the scores above 1 M are dominated by spectral changes due to solvation of HCl and NaOH. We used Eqs. S.4 and S.5 to fit the scores s (Fig. S1 and S2), whereby the percentage of the different glycine states is determined by $\beta(c)$:

$$s_{PC1}(c) = [q_{11}c + q_{12}\beta(c)] \tag{S.6}$$

$$s_{PC2}(c) = [q_{21}c + q_{22}\beta(c)] \tag{S.7}$$

The spectrum can be reconstructed by multiplying the scores with the weighted loadings for the selected principal components:

$$\alpha^{\mathrm{eff}}_{\mathrm{Gly+HCl/NaOH}} - \alpha^{\mathrm{eff}}_{\mathrm{Gly}}(c,\nu) = s_{PC1}(c)\cdot l_{PC1}(\nu) + s_{PC2}(c)\cdot l_{PC2}(\nu) \tag{S.8}$$

Insertion of Eqs. S.6 and S.8 into Eq. S.8 yields:

$$\alpha^{\mathrm{eff}}_{\mathrm{Gly+HCl/NaOH}} - \alpha^{\mathrm{eff}}_{\mathrm{Gly}} = [q_{11}c + q_{12}\beta(c)]\cdot l_{PC1}(\nu) + [q_{21}c + q_{22}\beta(c)]\cdot l_{PC2}(\nu) \tag{S.9}$$

Eq. S.9 can be transposed to Eq. S.10, so that the first part is a function of the HCl/NaOH concentration and the second part is a function of the percentage of the different glycine states.

$$\alpha^{\mathrm{eff}}_{\mathrm{Gly+HCl/NaOH}} - \alpha^{\mathrm{eff}}_{\mathrm{Gly}} = [q_{11}l_{PC1}(\nu) + q_{21}l_{PC2}(\nu)]\cdot c + [q_{12}l_{PC1}(\nu) + q_{22}l_{PC2}(\nu)]\cdot \beta(c) \tag{S.10}$$

Our physical model describes the effective THz absorption α_s^{eff} as a superposition of effective solute specific extinctions $\varepsilon^{\mathrm{eff}}$ shown in the following equations:

$$\alpha^{eff}_{\mathrm{Gly+HCl/NaOH}} - \alpha^{\mathrm{eff}}_{\mathrm{Gly}} = c_{H_3O^+} \cdot \varepsilon^{\mathrm{eff}}_{H_3O^+} + c_{Cl^-} \cdot \varepsilon^{\mathrm{eff}}_{Cl^-} + c_{Gly^{zi}} \cdot \varepsilon^{\mathrm{eff}}_{Gly^{\pm}} + c_{Gly^+} \cdot \varepsilon^{\mathrm{eff}}_{Gly^+} - c_{Gly} \cdot \varepsilon^{\mathrm{eff}}_{Gly^{\pm}} \quad \text{(S.11)}$$

where $c_{Gly} = 1\ M$ is the glycine concentration and $Gly^{\pm}$ represents the zwitterion. With the concentrations of the solutes defined below, we can deduce the effective absorption α_s^{eff} (Eq. S.16):

$$c_{H_3O^+} = c_{HCl} - c_{Gly} \cdot \beta \quad \text{(S.12)}$$

$$c_{Cl^-} = c_{HCl} \quad \text{(S.13)}$$

$$c_{Gly^{zi}} = c_{Gly} \cdot (1 - \beta) \quad \text{(S.14)}$$

$$c_{Gly^+} = c_{Gly} \cdot \beta \quad \text{(S.15)}$$

$$\alpha^{eff}_{\mathrm{Gly+HCl}} - \alpha^{\mathrm{eff}}_{\mathrm{Gly}} = c_{HCl} \cdot \left(\varepsilon^{\mathrm{eff}}_{H_3O^+} + \varepsilon^{\mathrm{eff}}_{Cl^-}\right) + c_{Gly} \cdot \beta \cdot (\varepsilon^{\mathrm{eff}}_{Gly^+} - \varepsilon^{\mathrm{eff}}_{Gly^{\pm}} - \varepsilon^{\mathrm{eff}}_{H_3O^+}) \quad \text{(S.16)}$$

Therefore, when inserting the fit parameters q and the loadings corresponding to the fitted scores, we are able to deduce the HCl absorption spectrum $\left(\varepsilon^{\mathrm{eff}}_{H_3O^+} + \varepsilon^{\mathrm{eff}}_{Cl^-}\right)$, as well as the difference absorption spectra $(\varepsilon^{\mathrm{eff}}_{Gly^+} - \varepsilon^{\mathrm{eff}}_{Gly^{\pm}} - \varepsilon^{\mathrm{eff}}_{H_3O^+})$.

For the titration with NaOH we can follow a similar procedure.

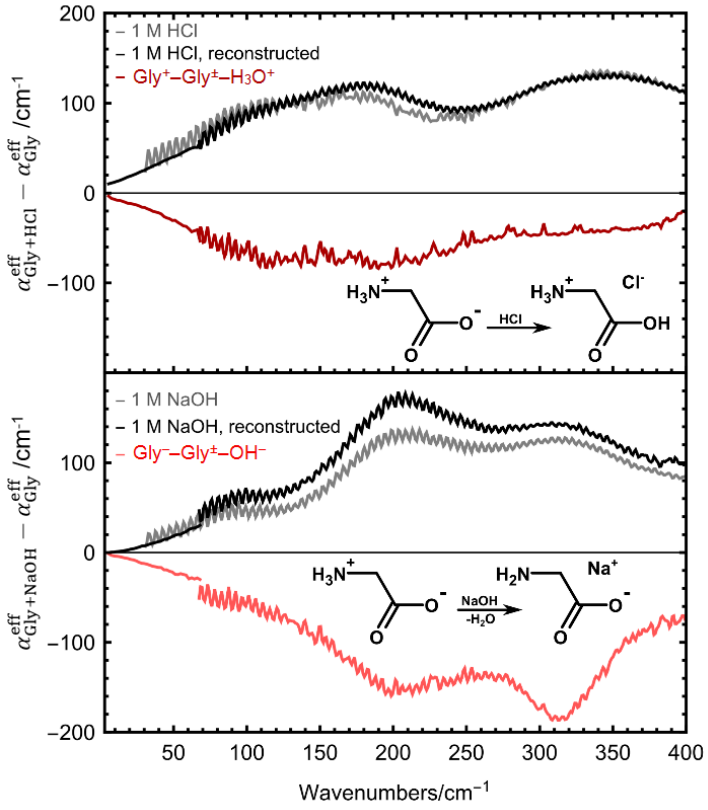

Figure S4. Difference spectra reconstructed by our PCA analysis. The reconstructed spectra of HCl/NaOH are displayed in black and the measured spectra are displayed in gray for comparison. The difference absorption spectra are displayed in dark red (upper panel) and light red (lower panel), respectively.

Based on the titration spectra we can reconstruct the spectrum of HCl and NaOH in the glycine solution and compare these spectra with the experimentally measured spectra of 1 M aqueous solution of HCl and NaOH (Fig. S4). The reconstructed spectrum of HCl resembles very well the typical Cl^- rattling mode at 180 cm^{-1} and the maxima at 140 cm^{-1} and 340 cm^{-1}.[3] The reconstructed NaOH spectrum consists of two pronounced bands at 100 cm^{-1} and 200 cm^{-1} and a broad band at 310 cm^{-1}.[5] In Ref.[4] we assigned the rattling mode of Na^+ peaks observed around 100 cm^{-1} and 150 cm^{-1}. The reconstructed spectra and the bulk spectra of solvated HCl and NaOH agree very well, thus supporting our additive approach; the subtle difference might reflect minor additional non-linear or/and counter ion effects, which cannot be ruled out completely.

To further validate our analysis, we applied the method described above to the combined THz-TDS and FTIR spectra. The results of the corresponding PCA for the titration of glycine with HCl and NaOH, respectively, are shown in Fig. S5 and S6. The results are in excellent agreement with those from the separate analysis of the low and high frequency regimes. In detail, for the titration with HCl we found two components with an explained variance of 92.9 % and 5.6 % (Fig. S5). For the titration with NaOH we found two components with an explained variance of 87.3 % and 8.3 % (Fig. S6).

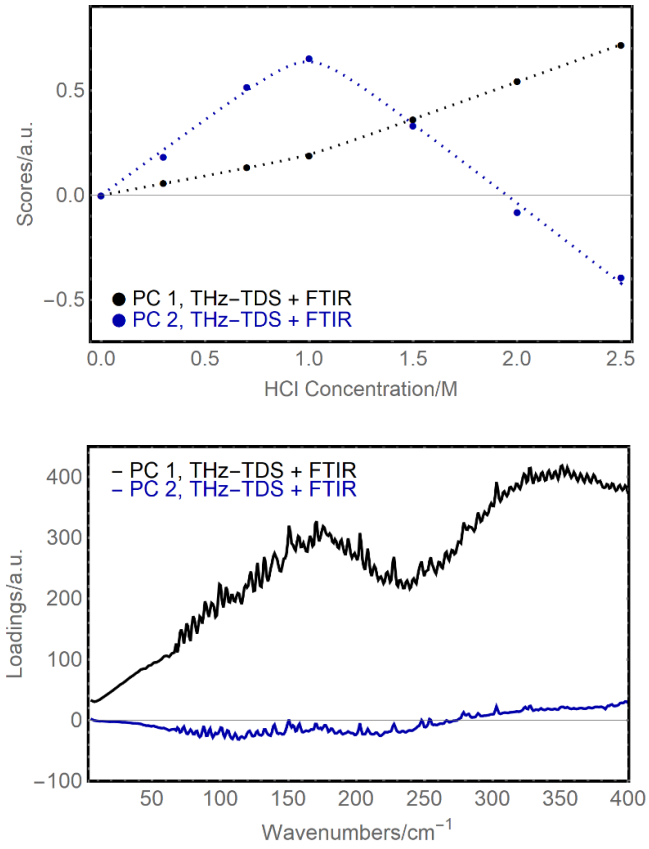

Figure S5. Results of the PCA performed on the combined difference spectra of the titration of glycine with HCl measured by both THz-TDS and FTIR spectrometers. The scores are displayed in the upper panel and the corresponding principal components in the lower panel. Best fit models are represented as by the dotted lines (see text for details).

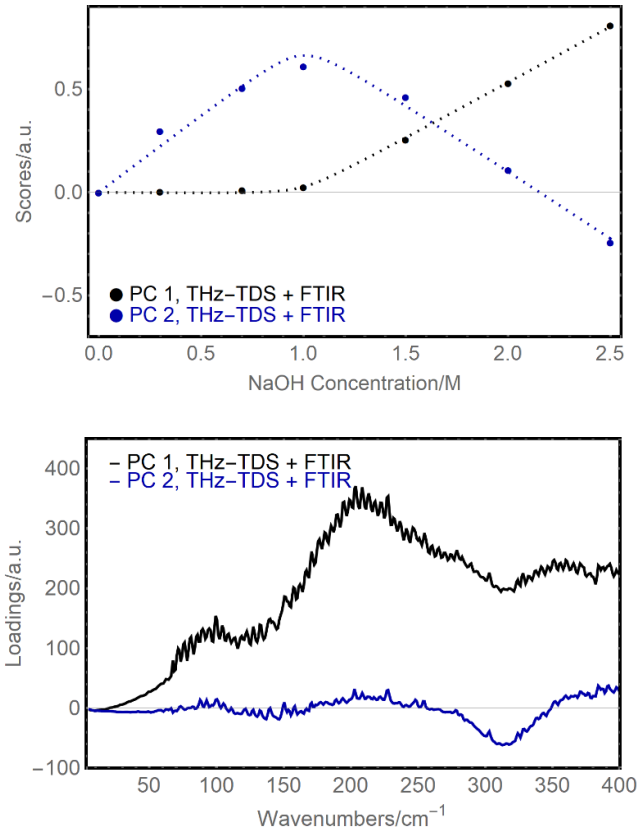

Figure S6. Results of the PCA performed on the combined difference spectra of the titration of glycine with NaOH measured by both THz-TDS and FTIR spectrometers. The scores are displayed in the upper panel and the corresponding principal components in the lower panel. Best fit models are represented as by the dotted lines (see text for details).

Glycine in aqueous electrolyte solutions

In order to test whether the observed intensity changes reflect only the protonation state or whether these are also sensitive to the cation/anion identity, we have recorded THz spectra of ternary glycine salt mixtures (including $NaNO_3$, NaCl, CsBr). Chemicals with a purity higher than 99.5% were purchased from Alfa Aesar for NaCl and CsBr and from Sigma-Aldrich for $NaNO_3$, respectively. Without further purification, aqueous salt solutions at 1 M and aqueous glycine solutions at the same concentration with and without electrolytes were prepared. The pH of these solutions was found to be not affected upon salt addition. We obtain a difference extinction spectrum:

$$\epsilon^{\text{eff}}_{\text{Gly+salt}} - \epsilon^{\text{eff}}_{\text{salt}} \tag{S.17}$$

for 1 M glycine solutions from Equations 1 and 2 of the main text by subtracting the effective absorption coefficients of glycine/salt aqueous solution, with 1 M glycine to 1 M salt ratio. The results are shown in Fig. S7. In contrast to the large intensity changes upon addition of HCl and NaOH, the addition of these salts results only in small changes, none in the frequency range up to 220 cm^{-1} within our experimental uncertainty. In the high frequency range above 220 cm^{-1}, glycine shows slightly less absorption when CsBr and $NaNO_3$ are added. However, a detailed analysis of these small changes is beyond the scope of this manuscript.

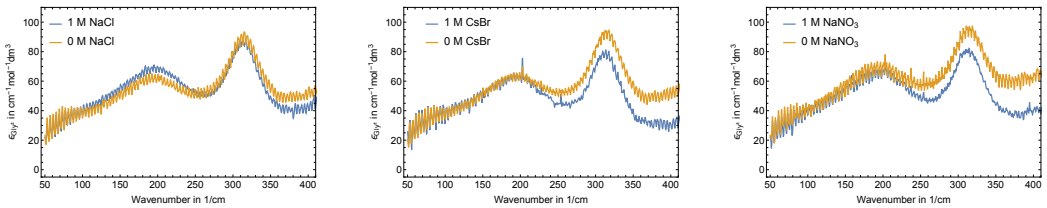

Figure S7: The extinction coefficient $\epsilon^{\text{eff}}_s = \epsilon^{\text{eff}}_{\text{Gly+salt}} - \epsilon^{\text{eff}}_{\text{salt}}$ of 1 M glycine aqueous solution with NaCl (left), CsBr (middle) and $NaNO_3$ (right).

Spectroscopic pH study of valine and diglycine aqueous solutions

In order to test whether the observed protonation dependence of the THz spectra holds also for other amino acids, we have carried out a titration study for valine. For valine aqueous solutions, the concentration-dependent absorption coefficient has already been reported in Ref.[5]. The N-C-C-O open/close mode is centered at about 350 cm^{-1} as predicted by simulation and observed experiments.

The difference THz absorption spectra of 0.3 M valine as a function of HCl/NaOH concentrations ($\alpha^{\text{eff}}_{\text{valine+HCl/NaOH}} - \alpha^{\text{eff}}_{\text{valine}}$) are displayed in Fig. S8. For the titration with HCl we deduce two principle components with an explained variance of 86.4 % and 8.2 %, respectively (Fig. S9, left panels). For the titration with NaOH we find two principle components with an explained variance of 91.3 % and 3.6 %, respectively (Fig. S9, right panels).

The difference THz absorption spectra of 0.5 M diglycine as a function of HCl/NaOH concentrations ($\alpha^{\text{eff}}_{\text{diglycine+HCl/NaOH}} - \alpha^{\text{eff}}_{\text{diglycine}}$) are displayed in Fig. S10. For the titration with HCl we deduce two principle components with an explained variance of 89.3 % and 7.1 %, respectively (Fig. S11, left panels). For the titration with NaOH we deduce two principle components with an explained variance of 85.6 % and 7.4 %, respectively (Fig. S11, right panels). Qualitatevely, the relative increase in intensity of the spectra is similar to what is observed for glycine.

Interestingly, by applying PCA analysis and Henderson-Hasselbalch equation for valine as well as for diglycine, we obtain the same results as for glycine. The scores of the first and second principal components of both valine and diglycine aqueous solution upon titration with HCl or NaOH display a turning point at an equimolar concentration of the aminoacid/peptide, i.e. at the point where valine and diglycine is expected to be fully protonated.

We determine pK_a (and pK_b) by choosing the concentration which corresponds to half of the concentration of the fitted turning point. The result is listed in Table 1 in the main text for solvated glycine, valine and diglycine in water and compared with literature values. Based on the comparison we conclude that the combination of THz-TDS and FTIR spectroscopy on glycine allows a quantitative evaluation of the pK_a value. This result underlines that THz spectra provide a sensitive probe for protonation/deprotonation.

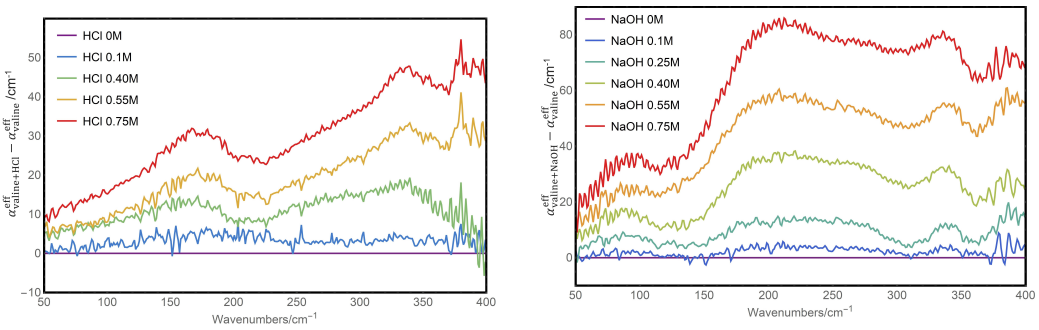

Figure S8: Difference absorption spectra of 0.3 M valine aqueous solutions upon titration with HCl (left) and NaOH (right) measured by far-infrared FTIR spectroscopy.

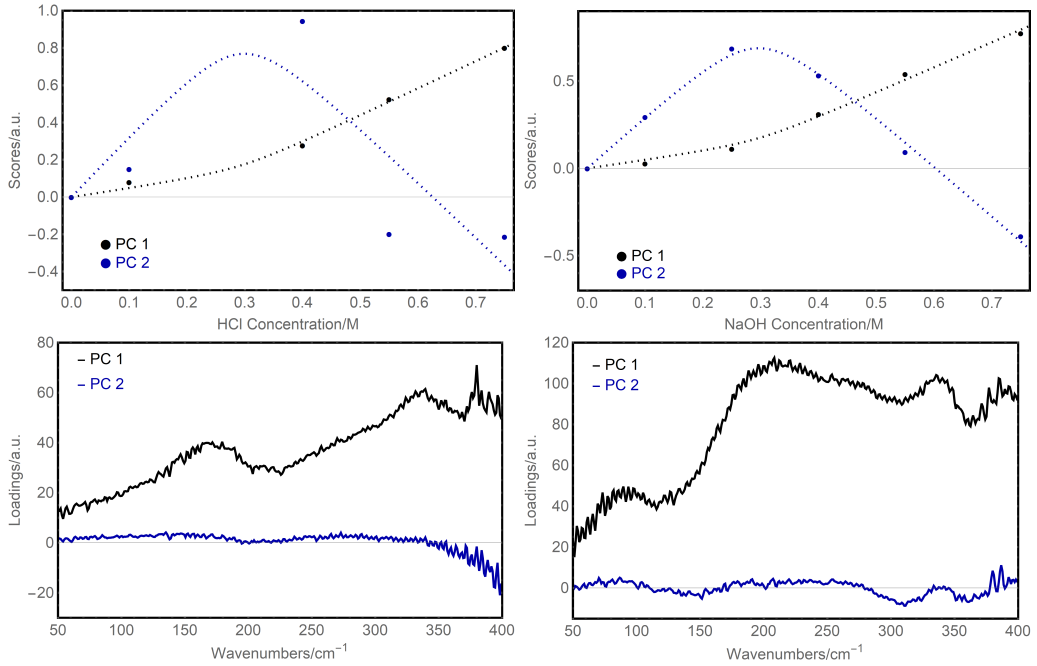

Figure S9: Results of the PCA performed on the difference absorption spectra of 0.3 M valine aqueous solutions upon titration with HCl (left panels) and NaOH (right panels), as shown in Fig. S8. The corresponding scores and principal components are displayed in the top and bottom panels, respectively. The turning point corresponds to the concentration where the amino acid is fully protonated.

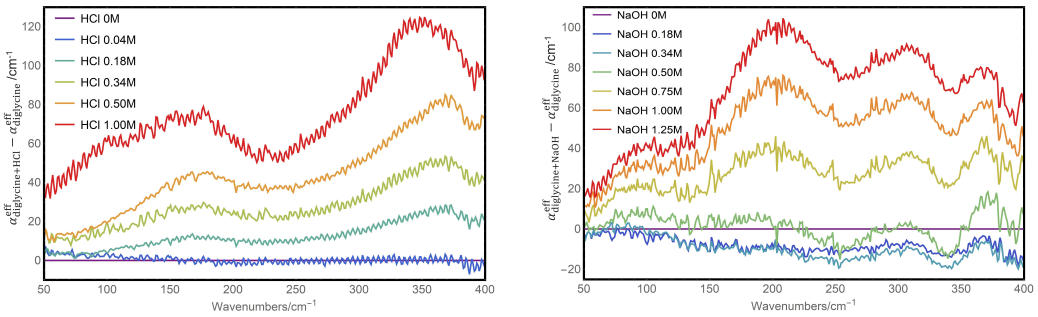

Figure S10: Difference absorption spectra of 0.5 M diglycine aqueous solutions upon titration with HCl (left) and NaOH (right) measured by far-infrared FTIR spectroscopy.

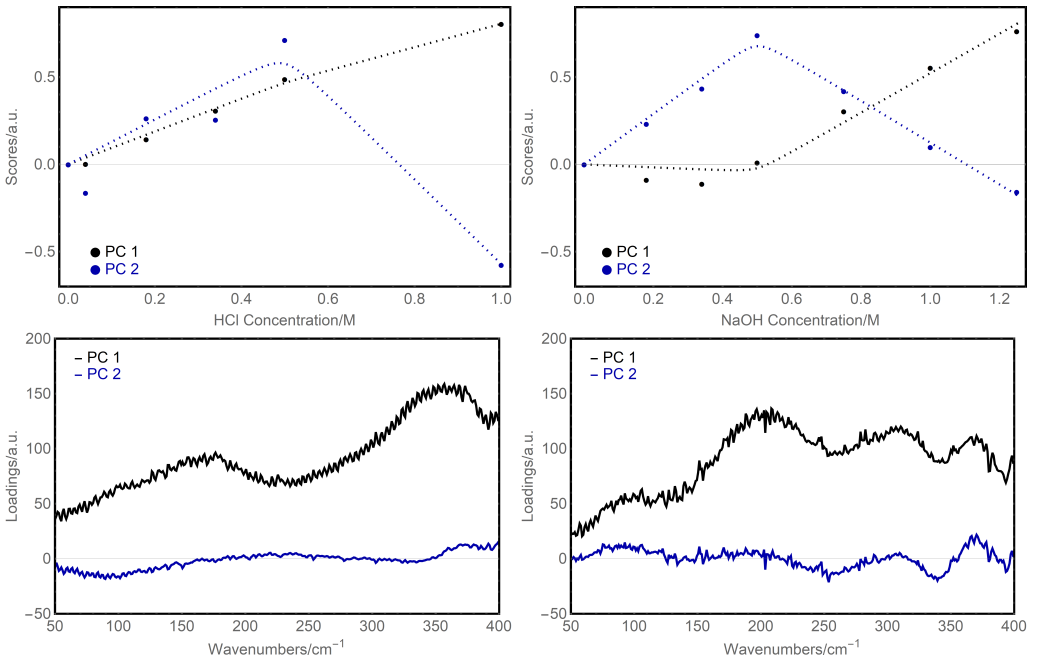

Figure S11: Results of the PCA performed on the difference absorption spectra of 0.5 M diglycine aqueous solutions upon titration with HCl (left panels) and NaOH (right panels), as shown in Fig. S10. The corresponding scores and principal components are displayed in the top and bottom panels, respectively.

Simulation Details

Ab initio Born-Oppenheimer molecular dynamics simulations[6] were carried out using the Quickstep module[7] in the CP2k program package[8]. The electronic structure was obtained with the RPBE functional[9] as implementated in the libxc library[10] with a combination of triple-ζ quality TZV2P Gaussian basis set with polarization functions[11] for the atomic orbitals and an auxiliary plane wave basis for the electron density[12] with a charge density cutoff of 500 Ry together with the NN50 smoothing both for the charge density and its derivatives. Core electrons were replaced by the corresponding Goedecker-Teter-Hutter norm-conserving pseudopotentials.[13,14] London dispersion interaction was taking into account using the so-called D3 correction (with zero-damping).[15] Since the influence of its three-body terms has been shown to be negligible for liquid water,[16] we included only the two-body terms.

All systems were simulated in cubic supercells containing one glycine molecule in one of the three different charged states surrounded by 128 water molecules. The sidelength of the periodic supercells were 15.76 Å in all three cases. For the simulation of the protonated glycine species an additional soft harmonic potential between the carboxyl oxygen and the proton was used during NVT equilibration to prevent dissociation.

After initial equilibration, long NVT trajectories were generated for all three cases using massive Nosé-Hoover chain thermostats at a temperature of 300 K. These simulations sample independent NVE starting configurations for the subsequent simulations which were used to compute the spectra. A total of 27 NVE trajectories were generated for the zwitterionic species of a total length of 2.9 ns while 27 NVE trajectories of total length of 1.9 ns were obtained for the deprotonated system and 38 NVE trajectories of total length of 2.1 ns were produced for the protonated system. Integration time steps were set to 0.5 fs while every 1 fs maximally localized Wannier functions were computed to allow the computation of molecular dipoles.

The mode decomposition analysis is carried out using the supermolecular solvation complex (SSC) methodology described in Ref.[17]. The molecular cross-correlation analysis (CCA) is carried out following the description in Ref.[4] and Ref.[5]. Contributions to the reconstructed effective absorption cross sections from the sodium and chloride counter ions have been obtained from Ref.[4]. For the glycine contributions to the reconstructed effective absorption cross sections, the glycine dipole auto-correlations and cross-correlations between glycine and the first two solvation shells were included. The solvation shells were defined via

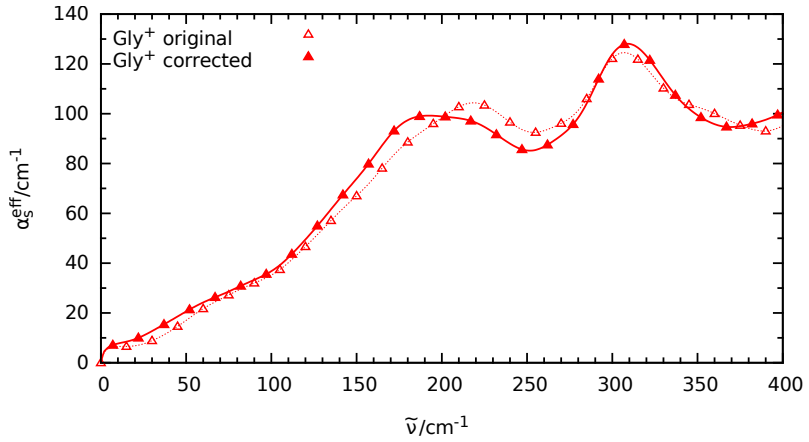

Figure S12: Computed effective absorption spectrum of $Gly^+ + Cl^-$ using the original THz contribution of Cl^- from Ref.[4] (empty triangles/dotted line) depicted as the red spectrum in Fig. 2(B) of the main text and using a corrected Cl^- contribution (filled triangles/solid line) as explained in the text.

geometrical criteria to describe hydrogen-bonds, that is first solvation shell water molecules are defined as being hydrogen-bonded to either the $-COO^-$ or $-COOH$ group or $-NH_3^+$ or $-NH_2$ group. The second solvation shell water molecules are defined as being hydrogen-bonded to first solvation shell water molecules as in Ref.[18]. Additionally, the change of the dipole auto-correlations of first solvation shell water molecules compared to bulk water molecules and the change of the dipole cross-correlations of first solvation shell water molecules with their hydrogen-bonded neighbor water molecules (part of the second solvation shell of glycine) entered the reconstruction of the effective absorption cross sections. Further possible contributions, like change of auto-correlations of second solvation shell water molecules, were not significant and would have only added noise.
Since the computational chloride counter ion contribution of Cl^- from Ref.[4] is blue-shifted by about $28\,cm^{-1}$ compared to the experimental peak of the aqueous NaCl solution as provided in the same reference, the influence of this shift is investigated in Fig. S12. The reconstructed $Gly^+ + Cl^-$ spectrum is shown therein for the case of using the original chloride contribution (depicted as the red spectrum in Fig. 2(B) of the main text) as well as for a corrected case in which the original chloride contribution was red-shifted by $28\,cm^{-1}$ in order to counterbalance the slight mismatch of that THz peak position with respect to experiment.[4] Applying

this correction, the maximum in the total computational spectrum of $Gly^+ + Cl^-$ at roughly $220\,cm^{-1}$ is found to red-shift and thus to move closer to the experimentally observed peak in Fig.2(A) of the main text, while the good agreement found for the most prominent resonance at higher frequency is preserved. This systematic improvement results from a combination of two phenomena, namely that the lower-frequency peak is indeed largely due to the chloride counter ion, whereas the rest of the spectrum, in particular the main resonance at $310\,cm^{-1}$ which is due to the independent Gly^+ N−C−C−O open/close motion, remains largely uneffected.

References

[1] K. Aoki, K. Shiraki, T. Hattori, *Appl. Phys. Lett.* **2013**, *103*, 173704.

[2] G. Schwaab, F. Sebastiani, M. Havenith, *Angew. Chem. Intl. Ed.* **2019**, *58*, 3000–3013.

[3] D. Decka, G. Schwaab, M. Havenith, *Phys. Chem. Chem. Phys.* **2015**, *17*, 11898–11907.

[4] P. Schienbein, G. Schwaab, H. Forbert, M. Havenith, D. Marx, *J. Phys. Chem. Lett.* **2017**, *8*, 2373–2380.

[5] A. Esser, H. Forbert, F. Sebastiani, G. Schwaab, M. Havenith, D. Marx, *J. Phys. Chem. B* **2018**, *122*, 1453–1459.

[6] D. Marx, J. Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*, Cambridge University Press, **2009**.

[7] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, J. Hutter, *Comput. Phys. Commun.* **2005**, *167*, 103–128.

[8] J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 15–25.

[9] B. Hammer, L. B. Hansen, J. k. Nørskov, *Phys. Rev. B* **1999**, *59*, 7413.

[10] M. A. L. Marques, M. J. T. Oliveira, T. Burnus, *Compt. Phys. Commun.* **2012**, *183*, 2272–2281.

[11] J. VandeVondele, J. Hutter, *J. Chem. Phys.* **2007**, *127*, 114105.

[12] G. Lippert, J. Hutter, M. Parrinello, *Mol. Phys.* **1997**, *92*, 477–488.

[13] S. Goedecker, M. Teter, J. Hutter, *Phys. Rev. B* **1996**, *54*, 1703.

[14] C. Hartwigsen, S. Goedecker, J. Hutter, *Phys. Rev. B* **1998**, *58*, 3641.

[15] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.

[16] R. Jonchiere, A. P. Seitsonen, G. Ferlat, A. M. Saitta, R. Vuilleumier, *J. Chem. Phys.* **2011**, *135*, 154503.

[17] J. Sun, G. Niehues, H. Forbert, D. Decka, G. Schwaab, D. Marx, M. Havenith, *J. Am. Chem. Soc.* **2014**, *136*, 5031–5038.

[18] S. Imoto, H. Forbert, D. Marx, *Phys. Chem. Chem. Phys.* **2015**, *17*, 24224–24237.