# Microhydration of Biomolecules: Revealing the Native Structures by Cold Ion IR Spectroscopy

*Erik Saparbaev, Viktoriia Aladinskaia, Andrei Zviagin, and Oleg V. Boyarkin**

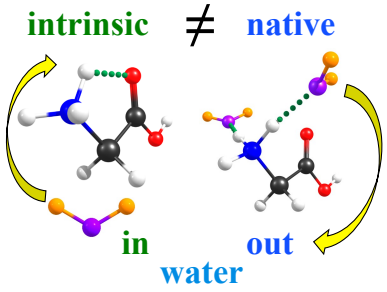
Laboratoire de Chimie Physique Moléculaire, École Polytechnique Fédérale de Lausanne, Station-6, 1015 Lausanne, Switzerland.

*Corresponding author: oleg.boiarkin@epfl.ch

ABSTRACT

The native-like structures of protonated glycine and peptide $Gly_3H^+$ were elucidated using cold ion IR spectroscopy of these biomolecules hydrated by a controlled number of water molecules. The complexes were generated directly from an aqueous solution using gentle electrospray ionization. Already with a single retained water molecule, $GlyH^+$ exhibits the native-like structure characterized by a lack of intramolecular hydrogen bonds. We use our spectra to calibrate the available data for the same complexes, but produced by cryogenic condensation of water onto the gas-phase glycine. Only in some conformers of these complexes $GlyH^+$ adopts the native-like structure, while in the others remains "kinetically" trapped in the intrinsic state. Only upon condensation of 4-5 water molecules the embedded amino acid fully adopts its native-like structure. Similar, condensation of one water molecule onto the tripeptide is insufficient to fully eliminate its kinetically trapped intrinsic states.

TOC GRAPHICS



Three-dimensional structures of biological molecules isolated in the gas phase may differ drastically from the geometries that they adopt in aqueous solutions. These native conformations, which are stabilized by hydrogen bonds with surrounding water, define the functionality of the biological molecules in living organisms. Many biological processes are mediated by water, including molecular recognition, protein folding, proton transfer, etc. All this makes the determination of the 3D structures of biomolecules in their native environment a pivot objective for many fields of life science.

IR spectroscopy of cryogenically cold ions isolated in the gas phase is a recent approach that allows for solving the intrinsic structures of small to midsize biological molecules with high accuracy and conformational resolution. The cooling sharpens vibrational transitions, thus providing detailed spectroscopic signatures that can be used for stringent validation of the computed molecular structures. While, potentially, high in the gas phase, the spectroscopic resolution degrades drastically in aqueous solutions, where biomolecules naturally adopt their native structures. The inhomogeneity and dynamical nature of non-covalent interactions with surrounding water make IR spectra unacceptable for validating structural calculations. As a compromise between the solution phase, where the structures are native but high resolution is difficult to achieve, and the gas phase, where spectral resolution can be high but structures are intrinsic, one can interrogate in the gas phase microhydrated biomolecular ions. Already a few water molecules may lock the main features of native structures of the ions in such complexes, while their limited size yet enables the use of cold ion spectroscopy.[1,2] One of the challenges of this approach is in producing a high number of hydrated complexes, where water molecules are only weakly coupled to a parent ion via hydrogen bonds. The most straightforward method to produce hydrated ions is to transfer them directly from a solution to the gas phase using gentle conditions in an electrospray ionization (ESI) source and subsequently cool the ions down in a cold trap.[1,3-5] There is a narrow interval of ESI conditions, where such complexes

can survive while keeping the ionization efficiency high. The use of soft ESI has been demonstrated on many occasions for generating microsolvated complexes of organic[1,5,6] and inorganic[7-11] ions, as well as for transferring to the gas phase intact protein-ligand complexes.[12]

An alternative and, perhaps, a more universal way of microhydration uses a gas-phase condensation of water molecules onto cryogenically pre-cooled ions.[13-15] Recently, this approach was employed by the group of Garand for spectroscopic study of microhydrated protonated amino acid glycine[16] and a small peptide triglycine.[17] Here we use IR spectroscopy to investigate the same hydrated molecules, but extracted directly from aqueous solution to the gas phase by soft ESI and subsequently cooled to cryogenic temperature. We address the questions of whether the structure of the ions depends on the method of their hydration.

Apart from the modified ion source, our experimental setup has been described elsewhere[18] (see SI for details). Briefly, the hydrated ions are produced from solution by a nano- ESI source and transferred through a metal capillary and three consecutive inline molecular skimmers to a room temperature octupole ion trap for accumulation and thermalization. The thermalized ionic complexes are mass-selected by a quadrupole mass-filter and then guided into a cold (T=6K) octupole trap,[19] where they undergo photofragmentation by an IR laser pulse. The appearing fragments that lost one or two waters are detected by the second quadrupole MS.

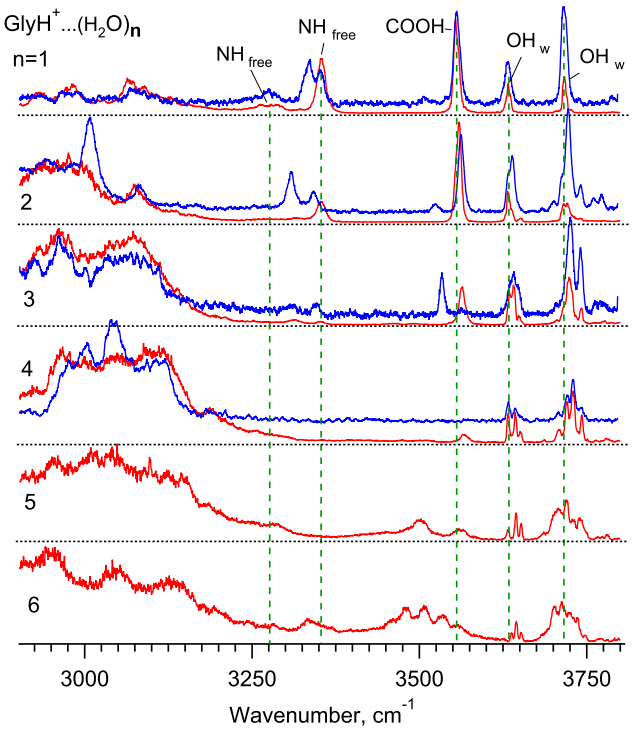



Figure 1. IR spectra of cold $GlyH^+\ldots(H_2O)_n$ ($1 \leq n \leq 6$), produced directly from aqueous solution (red trace; this work) and by cryogenic condensation (blue trace, ref.[16]). The reproduced original blue traces were shifted by -4 $cm^{-1}$ to match to the position of free OH-stretches of waters around 3700 $cm^{-1}$ measured herein (red traces) with the accuracy of ±0.2 $cm^{-1}$. The blue traces are offset up for graphical clarity.

Figure 1 (red traces) shows the IR spectra of $GlyH^+$ hydrated by the well-defined numbers (1 to 6) of water molecules that were retained on the amino acid. For comparison, the blue traces in the same figure reproduce the IR spectra earlier measured by the group of Garand for the same complexes (n=1 to 4), but produced in the gas phase by the cryogenic condensation.[16] Note, that the relative intensities of the transitions in each of the two "action" spectra may differ due to the difference in the used methods of the measurement. The assignment of the peaks in these spectra was based on the comparison with the computed spectra of the low-energy ionic structures that were calculated for the complexes by the same group.[16] For n=1 the spectra of

the complexes extracted from the solution and the complexes prepared by the condensation exhibit all the same transitions, except one extra peak at 3336 $cm^{-1}$ in the spectrum of the latter. The structure associated with this peak was assigned to the most stable conformer of the singly hydrated $GlyH^+$. This conformer is characterized by an intramolecular H-bond between the $NH_3^+$ and C=O groups. The same strong H-bond is present in the most stable gas-phase structure of $GlyH^+$, which was the only conformer observed experimentally.[17] Consistently, the gas-phase hydration by condensation conserves its main structural features. In an aqueous solution, glycine does not have this intramolecular H-bond, because the surrounding water molecules fully solvate the charge and the hydrophilic functional groups of the ion.[20] We thus are to conclude that, the structure assigned to the peak at 3336 $cm^{-1}$ in the spectrum of $GlyH^+…H_2O$ complex produced in the gas phase is relevant not to the native solution phase, but rather to the intrinsic gas-phase structure of the amino acid. The lack of this conformer for the complexes produced from the solution provides direct evidence of that. The peak at 3355 $cm^{-1}$, which appears in the spectra of the complexes produced by both methods of hydration, was earlier assigned to the conformer that has no intramolecular H-bond between the $NH_3^+$ and C=O groups.[16] Competing with this bond, the first condensed or the last retained water molecule non-covalently binds to the protonated N-terminus, partially solvating its charge. Consistently, this is the only structure detected for the complexes extracted from the solution in our experiments (Figure 1, n=1, red trace). We may speculate, that at the mild cryogenic temperature of ~80 K used for condensation,[15] a fraction of complexes residing in the gas-phase structures overcomes the barrier toward the solution-like structure of $GlyH^+$. The rest of the complexes remain "kinetically" trapped in the gas-phase conformational state. This is a mirroring of the process, where a molecule isolated in the gas phase from solution remains "kinetically" trapped in a solution-like state.[6,11,21,22] In opposite, retaining a single water molecule on $GlyH^+$ conserves the main features of the solution-phase structure of the ion in the

gas phase. This single lowest-energy structure remains stable despite the thermalization at room temperature for ~45 ms prior to the cryogenic cooling.

Retaining one more water molecule on $GlyH^+…H_2O$ complex almost does not change the resolved intense peaks assigned to weakly bound NH- and OH-stretches of glycine (compare the red traces for n=1 and 2 in Figure 1). A close inspection of the spectra reveals two spectroscopic manifestations of binding the second water to the N-terminus. The broad weak peak at ~ 3280 $cm^{-1}$, which for n=1 was earlier assigned to one of the two "free" NH stretches,[16] disappears (for n=2). It, likely, shifts to the red due to the H-bond between this NH group and the second water molecule. In opposite, the solvation of the charge weakens the long-distance non-covalent interaction of the two termini. This becomes evident from a slight, but well detected 5.6 $cm^{-1}$ blue shift of the OH-stretch transition of the C–terminus. The spectroscopic signature of the two additional OH-stretches of the second water molecule in the spectrum for n=2 (Figure 1, n=2, red trace) is the slightly split broaden peak at 3722 $cm^{-1}$ and the small peak at 3652 $cm^{-1}$. In comparison with this, the spectrum of the same complex, $GlyH^+…(H_2O)_2$, but produced in the gas phase by condensation[16] (blue trace in Fig. 1, n=2), exhibits 4-5 additional peaks in the region of the absorption by free OH-stretches of water (3750-3780 $cm^{-1}$), as well as 3 peaks earlier assigned to the NH-stretches of $GlyH^+$. This larger number of transitions reflects the existence of more than one conformer of the $GlyH^+…(H_2O)_2$. The calculations and IR-IR double resonance spectroscopy, indeed, revealed two main conformers, which do not have the intramolecular H-bond between the two termini of $GlyH^+$, as well as, at least, one minor conformer with this bond.[16] The most abundant conformer of $GlyH^+…(H_2O)_2$ produced in the gas-phase is associated with the two characteristic intense NH-stretch spectral transitions at 3007 and 3308 $cm^{-1}$ (Figure 1, n=2, blue trace).[17] These peaks are not present, however, in the spectrum of the same complex, but produced from solution (red trace in Figure 1, n=2).

This observation clearly demonstrates once more that, the gas-phase hydration "kinetically" traps the embedded molecules in their intrinsic structures.

The IR transitions of the second main conformer[16] of the gas-phase produced $GlyH^+…(H_2O)_2$ look similar to the spectrum of this complex prepared from solution, except the 14 $cm^{-1}$ redshift of the characteristic NH-stretch peak at 3353 $cm^{-1}$. We may suggest that the structure of this conformer is similar to the structure of the only conformer (conformational family) of the complexes prepared from solution, provided the shift can be explained by the influence of the $D_2$ tag, which was attached to the complexes (produced by the condensation) for their detection.[23,24]

The influence of the $D_2$ tag was also invoked by the authors[16] to explain the 30 $cm^{-1}$ redshift of the (CO)O-H stretch transition in the spectrum of $GlyH^+…(H_2O)_3$ relative to $GlyH^+…(H_2O)_2$ (blue trace in Figure 1, n=2, 3). With this reservation, the spectra of $GlyH^+$ triply hydrated by the two methods look very similar. The calculations suggested that the third water molecule energetically almost equally binds either to the C-terminus, inducing a large calculated redshift of the (CO)O-H stretch transition, or to the protonated N-terminus to fully solvate the charge.[16] The former binding results in a reduction of the intensity of the characteristic peak of the (CO)O-H stretch in both spectra for n=3. The continuing charge solvation becomes evident from the frequency of this transition, which shifts to blue upon retaining one more water on $GlyH^+…(H_2O)_2$ (red traces in Figure 1, n=2 and 3). We, thus, may suggest that the structures of $GlyH^+…(H_2O)_3$ complexes prepared by the two methods become similar and are both native-like.

Retaining one more water molecule to generate $GlyH^+…(H_2O)_4$ from solution further decreases the intensity of the peak, which was assigned to the free (CO)O-H stretch,[16] but doesn't change anymore its frequency. The implication of this, confirmed by the calculations, is that the charge has been already fully solvated by the first three waters, such that the fourth

molecule may now attach to the H atom of the C-terminus. The last three retained water molecules, therefore, already form the first solvation shell of the protonated N-terminus. The remaining intensity of the (CO)O-H stretch peak suggests the presence of another conformer, earlier predicted theoretically,[25,26] in which the fourth water molecule does not bind to the C-terminus, but to the first solvation shell. This conformer is absent, however, for the same $GlyH^+…(H_2O)_4$ complex prepared by the condensation in the gas phase.

The spectra of $GlyH^+$ hydrated by 5 and 6 waters become similar for both methods of their production (Figure S2, n=5 and 6). In all the calculated conformers, the fifth and the sixth water molecules belong to the second solvation shell. We thus may propose that the hydration of $GlyH^+$ by 5-6 water molecules using any of the two methods results in the native structures of this small biomolecule. The identification of such structures, which lack an intramolecular H-bond, becomes evident already for the singly-hydrated $GlyH^+$ released from solution, but remains ambiguous in the case of the gas-phase prepared complexes until as many as four to five water molecules were condensed onto the amino acid.

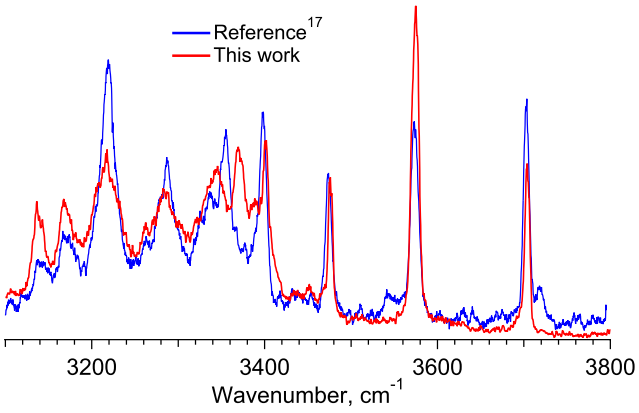



Figure 2. IR vibrational spectra of $Gly_3H^+…(H_2O)$. The red and blue traces correspond to the complexes produced from an aqueous solution and in the gas phase, respectively. The red trace was recorded by detecting $Gly_3H^+$ as a photofragment. The blue spectrum was measured by the group of Garand.[17]

$Gly_3H^+$ is an example of a larger molecule, for which the hydration with the two different methods results in different structures of this tripeptide. In the gas the peptide appears in two conformational states, in which the proton is either on N-terminus or shared between the two amide C=O groups.[17] The former conformer is characterized by the *cis* configuration of the first peptide bond, while the latter conformer is in the *trans* configuration of both peptide bonds. Such peptide conformers are unlikely for aqueous solution, where the charge is usually solvated by water molecules, unless it is strongly protected by the surrounding hydrophobic groups (e.g., aromatic rings[27]). The *cis* configuration of the amide group is also atypical for peptides, except those that contain proline residues. Condensation of a single water molecule to the gas phase conformers, indeed, made the *trans* conformer energetically favorable and reduced the population of the *cis* conformer to 10%. This conformer is naturally not present at all for the same $Gly_3H^+…H_2O$ complex extracted from the solution. The only detected conformer with one retained water molecule exhibits all the same characteristic IR transitions (Figure 2) that were assigned to the main *trans* conformer for the complexes prepared by the condensation. The characteristic IR transitions at 3718 $cm^{-1}$ and 3541 $cm^{-1}$ (Figure 2, blue trace), which were earlier assigned to the *cis* conformer prepared by the condensation, are clearly absent in our IR spectrum (Figure 2, red trace). Apart from this, the only difference in the position of the peaks in the spectra of the differently prepared complexes is a distinct peak at 3369.8 $cm^{-1}$ that appears only in the case of solution-prepared $Gly_3H^+…H_2O$. We may speculate that this feature can be attributed either to the certain structural difference of the differently prepared complexes or to the influence of tagging in the case of the gas-phase hydration.

In conclusion, retaining a single water molecule on $GlyH^+$ already allows for the determination of its native-like structure, which changes only little upon retaining more waters (Table S1). In comparison, the cryogenic condensation may produce some additional conformers of the amino acid, where the intrinsic structures of the biomolecule are

"kinetically" trapped. This makes the identification of the native structures among all the available conformers ambiguous. Upon increasing the number of the attached water molecules, the method of condensation produces, essentially, the same conformers as the structures extracted directly from the solution. Because of the large number of vibrational degrees of freedom and increasing spectral congestion, the structures of these larger complexes are more difficult to solve and to validate, however.

Finally, for larger gas-phase protonated peptides, the increased number of hydrophilic groups will increase the probability for a charge to be self-solvated;[13] more H-bonds between these groups will be deeply buried into the molecule. These strong intramolecular interactions may compete with a subsequent microhydration in the gas phase, protecting peptides in their "kinetically" trapped intrinsic states. We may speculate that the difference between the structures of microsolvated peptides prepared from solution and formed in the gas will increase for larger molecules.

ACKNOWLEDGMENT

This paper is in memory of prof. D. Gerlich, who for years guided us in the use of cold traps. The authors thank prof. E. Garand for the generously supplied originals of the published IR spectra and computed structures. The work was supported by the Swiss National Science Foundation (grant 200020_172522) and by EPFL.

SUPPORTING INFORMATION. Experimental details; mass and IR spectra of the complexes; comparison of structures.

REFERENCES

(1) Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V. Interplay of Intra- and Intermolecular H-Bonding in a Progressively Solvated Macrocyclic Peptide. *Science* **2012,** *336*, 320-323.
(2) Roy, T. K.; Nagornova, N. S.; Boyarkin, O. V.; Gerber, R. B. A Decapeptide Hydrated by Two Waters: Conformers Determined by Theory and Validated by Cold Ion Spectroscopy. *J. Phys. Chem. A* **2017,** *121*, 9401-9408.
(3) Mercier, S. R.; Boyarkin, O. V.; Kamariotis, A.; Guglielmi, M.; Tavernelli, I.; Cascella, M.; Rothlisberger, U.; Rizzo, T. R. Microsolvation Effects on the Excited-State Dynamics of Protonated Tryptophan. *J. Am. Chem. Soc.* **2006,** *128*, 16938-16943.
(4) Spieler, S.; Duong, C. H.; Kaiser, A.; Duensing, F.; Geistlinger, K.; Fischer, M.; Yang, N.; Kumar, S. S.; Johnson, M. A.; Wester, R. Vibrational Predissociation Spectroscopy of Cold Protonated Tryptophan with Different Messenger Tags. *J. Phys. Chem. A* **2018,** *122*, 8037-8046.
(5) Chang, T. M.; Chakrabarty, S.; Williams, E. R. Hydration of Gaseous m-Aminobenzoic Acid: Ionic vs Neutral Hydrogen Bonding and Water Bridges. *J. Am. Chem. Soc.* **2014,** *136*, 10440-10449.
(6) Silveira, J. A.; Fort, K. L.; Kim, D.; Servage, K. A.; Pierson, N. A.; Clemmer, D. E.; Russell, D. H. From Solution o the Gas Phase: Stepwise Dehydration and Kinetic Trapping of Substance P Reveals the Origin of Peptide Conformations. *J. Am. Chem. Soc.* **2013,** *135*, 19147-19153.
(7) Asmis, K. R.; Neumark, D. M. Vibrational Spectroscopy of Microhydrated Conjugate Base Anions. *Accounts Chem. Res.* **2012,** *45*, 43-52.
(8) Inokuchi, Y.; Ebata, T.; Rizzo, T. R.; Boyarkin, O. V. Microhydration Effects on the Encapsulation of Potassium Ion by Dibenzo-18-Crown-6. *J. Am. Chem. Soc.* **2014,** *136*, 1815-1824.
(9) Sun, S. T.; Jiang, L.; Liu, J. W.; Heine, N.; Yacovitch, T. I.; Wende, T.; Asmis, K. R.; Neumark, D. M.; Liu, Z. F. Microhydrated Dihydrogen Phosphate Clusters Probed by Gas Phase Vibrational Spectroscopy and First Principles Calculations. *Phys. Chem. Chem. Phys.* **2015,** *17*, 25714-25724.
(10) Mitra, S.; Duong, C. H.; McCaslin, L. M.; Gerber, R. B.; Johnson, M. A. Isomer-Specific Cryogenic Ion Vibrational Spectroscopy of the $D_2$ Tagged $Cs^+(HNO_3)(H_2O)_{n=0–2}$ Complexes: Ion-Driven Enhancement of the Acidic H-bond to Water. *Phys. Chem. Chem. Phys.* **2020,** *22*, 4501-4507.
(11) Hebert, M. J.; Russell, D. H. Tracking the Structural Evolution of 4-Aminobenzoic Acid in the Transition from Solution to the Gas Phase. *J. Phys. Chem. B* **2020,** *124*, 2081-2087.
(12) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013,** *85*, 2-9.
(13) Liu, D.; Wyttenbach, T.; Barran, P. E.; Bowers, M. T. Sequential Hydration of Small Protonated Peptides. *J. Am. Chem. Soc.* **2003,** *125*, 8458-8464.
(14) Xu, S.; Smith, J. E. T.; Weber, J. M. Hydration of a Binding Site with Restricted Solvent Access: Solvatochromic Shift of the Electronic Spectrum of a Ruthenium Polypyridine Complex, One Molecule at a Time. *J. Phys. Chem A* **2016,** *120*, 7650-7658.
(15) Voss, J. M.; Fischer, K. C.; Garand, E. Accessing the Vibrational Signatures of Amino Acid Ions Embedded in Water Clusters. *J. Phys. Chem. Lett.* **2018,** *9*, 2246-2250.
(16) Fischer, K. C.; Sherman, S. L.; Voss, J. M.; Zhou, J.; Garand, E. Microsolvation Structures of Protonated Glycine and l-Alanine. *J. Phys. Chem. A* **2019,** *123*, 3355-3366.
(17) Fischer, K. C.; Voss, J. M.; Zhou, J.; Garand, E. Probing Solvation-Induced Structural Changes in Conformationally Flexible Peptides: IR Spectroscopy of $Gly_3H^+ \cdot (H_2O)$. *J. Phys. Chem A* **2018,** *122*, 8213-8221.

(18) Boyarkin, O. V. Cold Ion Spectroscopy for Structural Identifications of Biomolecules. *Int. Rev. Phys. Chem.* **2018,** *37*, 559-606.
(19) Boyarkin, O. V.; Kopysov, V. Cryogenically Cooled Octupole Ion Trap for Spectroscopy of Biomolecular Ions. *Rev. Sci. Instrum.* **2014,** *85*, 033105.
(20) Kayi, H.; Kaiser, R. I.; Head, J. D. A Theoretical Investigation of the Relative Stability of Hydrated Glycine and Methylcarbamic Acid—from Water Clusters to Interstellar Ices. *Phys. Chem. Chem. Phys.* **2012,** *14*, 4942-4958.
(21) Patrick, A. L.; Cismesia, A. P.; Tesler, L. F.; Polfer, N. C. Effects of ESI Conditions on Kinetic Trapping of the Solution-Phase Protonation Isomer of p-Aminobenzoic Acid in the Gas Phase. *Int. J. Mass Spectrom.* **2017,** *418*, 148-155.
(22) Zhang, H.; Cao, W.; Yuan, Q.; Zhou, X.; Valiev, M.; Kass, S. R.; Wang, X.-B., Cryogenic "Iodide-Tagging" Photoelectron Spectroscopy: A Sensitive Probe for Specific Binding Sites of Amino Acids. *J. Phys. Chem. Lett.* **2020,** *11*, 4346-4352.
(22) Kamrath, M. Z.; Garand, E.; Jordan, P. A.; Leavitt, C. M.; Wolk, A. B.; Van Stipdonk, M. J.; Miller, S. J.; Johnson, M. A. Vibrational Characterization of Simple Peptides Using Cryogenic Infrared Photodissociation of $H_2$-Tagged, Mass-Selected Ions. *J. Am. Chem. Soc.* **2011,** *133*, 6440-6448.
(24) Masson, A.; Williams, E. R.; Rizzo, T. R. Molecular Hydrogen Messengers Can Lead to Structural Infidelity: A Cautionary Tale of Protonated Glycine. *J. Chem. Phys.* **2015,** *104313*.
(25) Wei, Z.; Chen, D.; Zhao, H.; Li, Y.; Zhu, J.; Liu, B. Ab Initio Investigation of the First Hydration Shell of Protonated Glycine. *J. Chem. Phys.* **2014,** *140*, 085103.
(26) Michaux, C.; Wouters, J.; Perpète, E. A.; Jacquemin, D. Microhydration of Protonated Glycine: An ab initio Family Tree. *J. Phys. Chem. B* **2008,** *112*, 2430-2438.
(27) Pereverzev, A. Y.; Szabó, I.; Kopysov, V. N.; Rosta, E.; Boyarkin, O. V. Gas-Phase Structures Reflect the Pain-Relief Potency of Enkephalin Peptides. *Phys. Chem. Chem. Phys.* **2019,** *21*, 22700-22703.