

Observation of Conformational Simplification upon *N*-Methylation on Amino Acid Iodide Clusters

Wenjin Cao,[#] Hanhui Zhang,[#] Qinqin Yuan, Xiaoguo Zhou,^{*} Steven R. Kass,^{*} and Xue-Bin Wang^{*}



Cite This: *J. Phys. Chem. Lett.* 2021, 12, 2780–2787



Read Online

ACCESS |



Metrics & More

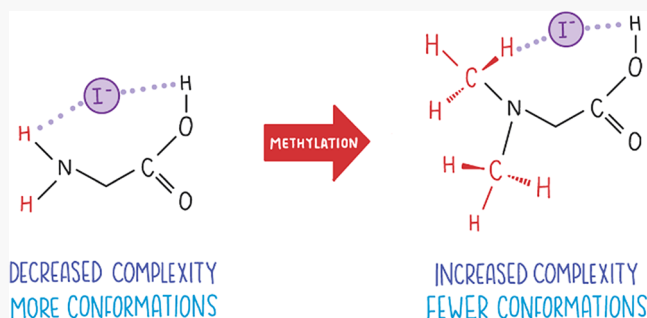


Article Recommendations



Supporting Information

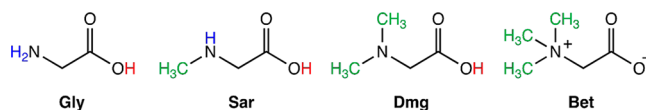
ABSTRACT: This Letter reports a counterintuitive observation that methylation of the glycine-iodide cluster leads to fewer conformations and spectroscopic simplicity. Cryogenic “iodide-tagging” negative ion photoelectron spectroscopy (NIPES) is used to probe specific binding sites of three *N*-methylated glycine derivatives, i.e., *N*-methylglycine (sarcosine), *N,N*-dimethylglycine, and *N,N,N*-trimethylglycine (glycine betaine). NIPES reveals a progressive spectral simplification of the iodide clusters with increasing methylation due to fewer contributing structures. Low energy conformers and tautomers of each cluster are computationally identified, and those observed in the experiments are assigned based on excellent agreement between the NIPES spectra and theoretical simulations. Zwitterionic cluster structures are found to be less stable than their canonical forms and do not contribute to the observed spectra. This work demonstrates the power of iodide-tagging NIPES in probing conformations of amino acid-iodide clusters and provides a molecular level understanding on the effect of methyl substitution on amino acid binding sites.



Amino acids (AAs) are the building blocks of proteins and play a crucial role in living cells and biological processes.^{1,2} All 20 proteinogenic AAs adopt zwitterionic structures with the N-terminus protonated and the C-terminus deprotonated under typical physiological conditions, whereas they exist in their canonical forms in the gas phase.^{3–5} Previous studies have shown that the gaseous zwitterionic forms of AAs can be stabilized upon solvation and exist in AA-ion clusters, illustrating the remarkable molecular binding forces exerted by solvents and ions on AA structures.^{6–23} AA-ion complexes are tractable both by first principle calculations and detailed spectroscopic characterizations, making them attractive models for probing fundamental intermolecular interactions important to the understanding of protein and ion function under biological settings.

AAs, like all other biomolecules, are usually bound to anions at electrophilic sites via hydrogen bonds. Amino and carboxyl groups are common N–H and O–H hydrogen bond donors, respectively, found in AAs. Recent catalytic studies revealed that methyl- and methylene-H can act as hydrogen bond donors even though C–H...X[–] interactions are not as strong as N–H...X[–] or O–H...X[–] hydrogen bonds.^{24,25} Taken together, these results indicate that a diverse range of binding motifs between amino acids and anions maybe accessible. This was indeed found to be the case for the simplest amino acid, glycine, as revealed in our recent in-depth study of the glycine-iodide cluster anion (Gly-I[–]) in which five different structures were identified.²³ Glycine has a series of *N*-methylated derivatives (Scheme 1), namely, *N*-methylglycine (sarcosine,

Scheme 1. Molecular Structures of Gly, Sar, Dmg, and Bet



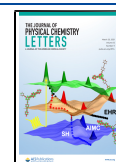
Sar), *N,N*-dimethylglycine (Dmg), and trimethylglycine (glycine betaine, Bet), that are commonly classified as unnatural AAs even though they all are widely found in nature. Incorporation of these AAs into proteins could alter and broaden their biological functions,^{26–29} and they offer numerous medicinal and pharmacological opportunities^{30–32} as well as applications in biolabeling.^{33,34} Such functional modifications presumably stem from alterations in intermolecular interactions due to changes in protein binding sites induced by including unnatural AAs.

The conformers of Gly, Sar, and Dmg have been well studied in the gas phase, revealing how steric effects brought about by *N*-methyl substitution affects their populations and relative energies.^{35–44} The number of low-lying conformers were identified as 6, 8, and 7 for Gly, Sar, and Dmg,

Received: January 12, 2021

Accepted: March 5, 2021

Published: March 12, 2021



ACS Publications

© 2021 American Chemical Society

2780

<https://dx.doi.org/10.1021/acs.jpclett.1c00125>
J. Phys. Chem. Lett. 2021, 12, 2780–2787

respectively, in a recent study.⁴⁴ No conclusions were drawn on the effects of replacing amine hydrogens with methyl groups on the available binding sites or subsequent intermolecular interactions. Sequential replacement of the amino hydrogen atoms in glycine by methyl groups, however, is known to increase the basicity of the resulting AA, to reduce the number of N–H hydrogen bond donating sites from two in Gly to zero in Dmg and Bet, and also to introduce methyl–H binding sites. As a result, four types of hydrogen bond donors (one N–H, one O–H, two CH–H, and three CH₂–H) exist in Sar, three types (one O–H, two CH–H, and six CH₂–H) in Dmg, and two types (two CH–H and nine CH₂–H) in Bet. Since all of these kinds of hydrogen can form hydrogen bonds with anions, this leads to additional challenges in identifying the active binding sites in anionic clusters of these AAs. An effective method for addressing this issue is to determine the molecular structures of these complexes, and by comparing differences between glycine and its mono-, di-, and trimethylated derivatives, the consequences of methyl group substitution on glycine binding sites can be deduced. This should shed light toward understanding how certain functional group changes affect biological processes.

Gas phase ion spectroscopy accompanied by high-level *ab initio* computations has proven to be a powerful means for determining biomolecule-ion structures,^{4–23} but there are significant challenges regarding (1) how to experimentally populate energized isomers and (2) how to obtain spectroscopic signatures for distinguishing different but similar structures. We recently demonstrated in the Gly-I[−] case study that cryogenic “iodide-tagging” negative ion photoelectron spectroscopy (NIPES) coupled with an electrospray ionization (ESI) source is an effective approach for this purpose.²³ Multiple AA-iodide isomers are accessible from room temperature micro-sized droplets using this approach, and since they undergo rapid solvent evaporative cooling, the cluster anion structures are kinetically trapped in the adjustable low temperature (≥ 20 K) ion trap. This method is different from the widely used infrared multiple photon dissociation technique that characterizes different ion geometries by identifying signature vibrational modes,^{6–19} in that “iodide-tagging” NIPES, as the name indicates, uses iodide as a messenger to interpret electronic and structural information on its cluster ions. This methodology takes advantage of the fact that the NIPE spectra of these complexes are expected to be dominated by atomic iodide transitions shifted to various electron binding energies (EBEs), which exhibit two distinct bands separated by ~ 0.9 eV due to the $^2P_{3/2}$ and $^2P_{1/2}$ spin-orbit states of the iodine atom.^{45–47} Consequently, different conformations and tautomers (henceforth referred to as isomers) can be identified from resolved peaks based on small EBE differences. For Gly-I[−], five isomers all with glycine in its canonical form were characterized,²³ revealing the presence of multiple binding motifs for this complex. However, whether this method can be effectively applied to more complex clusters remains to be seen. In this work, various structures of the Sar-I[−], Dmg-I[−], and Bet-I[−] cluster anions were generated and kinetically trapped in a 3-D cryogenic ion trap and were characterized with iodide-tagging NIPES. In combination with quantum chemical calculations, isomer identifications for each ion complex were achieved based on excellent agreement between calculated vertical detachment energies (VDEs) and measured band positions in addition to well-matched spectral simulations. A progressive simplification

of the spectral profiles along with fewer computationally located isomers suggests that fewer structures are produced upon methylation in this series of amino acid cluster ions.

Figure 1 exhibits the 20 K NIPE spectra for Gly-I[−], Sar-I[−], Dmg-I[−], and Bet-I[−] anions collected at 193 nm (left panel) and

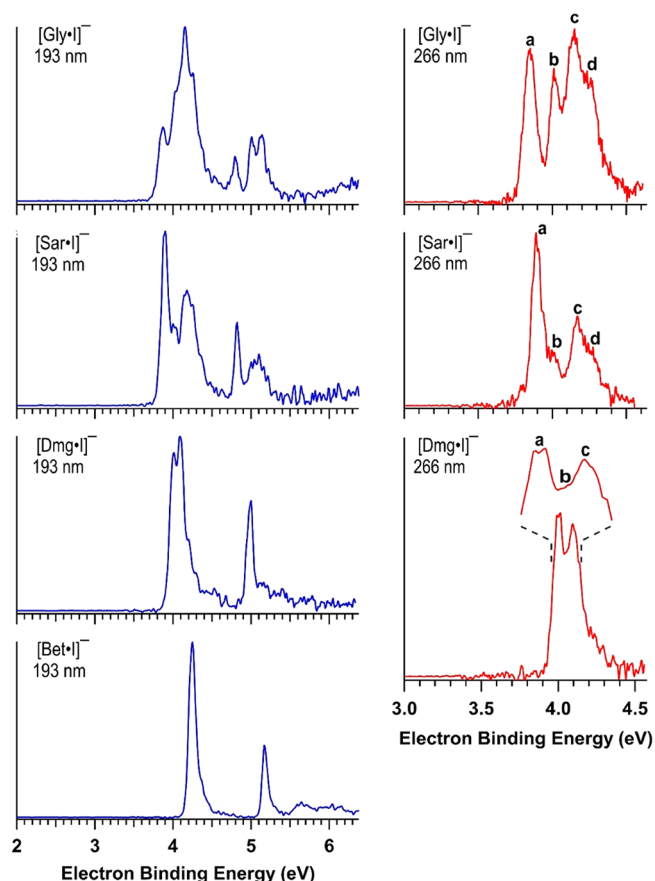


Figure 1. The 20 K NIPE spectra of Gly-I[−], Sar-I[−], Dmg-I[−], and Bet-I[−] anions obtained at 193 nm (6.424 eV, left panel) and those for Gly-I[−], Sar-I[−], and Dmg-I[−] at 266 nm (4.661 eV, right panel). The Gly-I[−] spectra are adapted from ref 23.

266 nm (right panel), obtained using the PNNL cryogenic temperature-dependent NIPE spectrometer.⁴⁸ These 193 nm spectra share a common feature in that they are all composed of two band systems separated by ~ 0.9 eV and are characterized with nearly identical band profiles but less intensity for the higher EBE feature. Based on a previous Gly-I[−] study, these two band systems correspond to formation of the neutral complex with iodine in $^2P_{3/2}$ ($\Omega = 3/2$) and $^2P_{1/2}$ ($\Omega = 1/2$) spin-orbit states, respectively and are composed of multiple features due to the coexistence of different isomers with slightly different EBEs.²³ Interestingly, as the molecular complexity increases from Gly-I[−] to Bet-I[−] the spectral profiles of these cluster ions become simpler, consistent with the presence of fewer isomers. The least complex Gly-I[−] anion possesses the most complicated spectra, with the $\Omega = 3/2$ band system consisting of four bands (labeled a, b, c, and d) corresponding to five different isomers, g1–g5 (Figure 2) based on the good agreement between observed band positions and calculated VDEs, as well as well-matched Franck–Condon factor (FCF) simulations; in this regard it is worth noting that the EBEs for g1 and g3 overlap so both conformers correspond to band d. The 266 nm Sar-I[−]

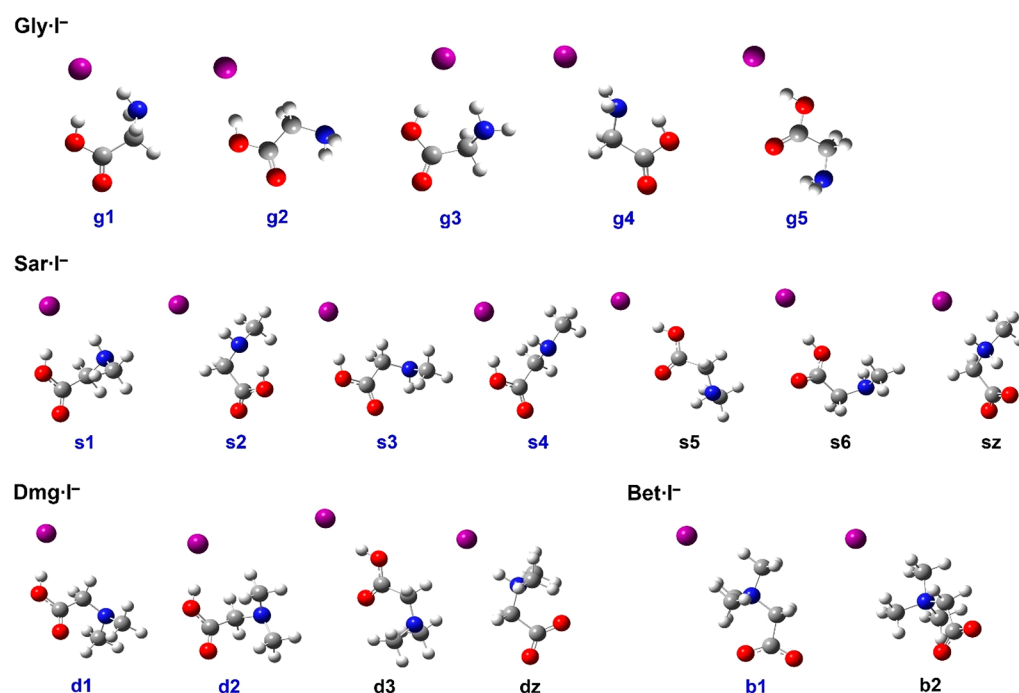


Figure 2. Observed isomeric structures of Gly-I[−] anion (adapted from ref 23) and low-lying M06-2X structures of Sar-I[−], Dmg-I[−], and Bet-I[−] anions. The ones labeled in blue are confirmed with NIPES.

spectrum consists of at least two bands at 3.89 and 4.16 eV (labeled **a** and **c**), both of which have a shoulder at slightly higher EBEs (labeled **b** and **d**), located at 4.01 and 4.26 eV, respectively. With one more amino-H substituted by $-\text{CH}_3$, the Dmg-I[−] anion exhibits a significantly simpler spectrum at 266 nm for the $\Omega = 3/2$ band with only two major features at 4.01 (**a**) and 4.10 (**c**) eV along with a discernible shoulder between them at ~ 4.06 eV (**b**). However, it is worth noting that in the 193 nm spectrum **a** and **c** are resolved in the $\Omega = 3/2$ band system but exhibit only a single transition in the $\Omega = 1/2$ band system despite better spectral resolution expected at higher EBE. As for the spectrum of the Bet-I[−] anion, it is the simplest one despite this complex being larger than the others. Only one sharp transition for each band system located at 4.25 and 5.17 eV is visible. Due to this extremely simple spectral profile and the resulting spectral assignment, the 266 nm spectrum was not needed or recorded.

Due to a good correlation between the $\Omega = 3/2$ and $\Omega = 1/2$ band systems, structural assignments for the Sar-I[−] complex were based on the 266 nm spectrum, taking advantage of its better spectral resolution. A total of seven isomers labeled **s1**–**s6** and **sz** (the letter ‘z’ indicating a zwitterionic structure) were located by calculations at the M06-2X/maug-cc-pVTZ(-PP) level of theory as shown in Figure 2 with their corresponding neutral radical structures summarized in Figure S1. To assign the obtained NIPES spectrum, calculated relative energies and VDEs of the low-lying isomers of the Sar-I[−] complex at the CCSD(T) level are summarized in Table 1. In addition, the 266 nm NIPES spectrum is compared with each isomer-specific FCF simulated spectrum in Figure S2. Similar to the previously studied Gly-I[−] cluster,²³ the Sar-I[−] spectrum is also attributed to multiple isomers since none of the simulated spectra are able to individually reproduce the NIPES spectrum that spans ~ 0.8 eV. The simulated spectra of **s1**–**s6** are all dominated by vertical transitions, with those for **s1** and **s2** both displaying a close-lying band 0.46 and 0.10 eV higher in energy,

Table 1. Calculated Relative Energies (ΔE , kcal mol^{−1}) and Both Experimental (expt) and Calculated (calcd) VDEs (eV) for Different Gly-I[−], Sar-I[−], Dmg-I[−], and Bet-I[−] Isomers^a

		VDE				VDE	
	ΔE	calcd	expt		ΔE	calcd	expt
Gly-I [−]							
g1	0	4.30	4.26	g4	1.57	3.96	4.01
g2	0.45	4.16	4.15	g5	3.93	3.83	3.85
g3	1.34	4.30	4.26	gz	7.49	4.86	—
Sar-I [−]							
s1	0	4.23	4.26	s5	4.48	3.74	—
s2	0.40	3.90	3.89	s6	5.08	3.77	—
s3	0.99	4.07	4.16	sz	2.79	4.59	—
s4	1.00	4.24	4.26				
Dmg-I [−]							
d1	0	4.04	4.01	d3	3.63	3.71	—
d2	0.53	4.08	4.06	dz	2.94	4.73	—
Bet-I [−]							
b1	0	4.29	4.25	b2	4.79	4.05	—

^aComputed values correspond to CCSD(T)//M06-2X energies except for the Gly-I[−] clusters which come from ref 23 and are CCSD(T)//B3LYP values.

respectively. Those for **s3** and **s4** have two extra bands with larger EBEs by 0.12 and 0.38 eV for **s3** and 0.13 and 0.38 eV for **s4**. The first higher EBE bands for **s1**–**s4** are attributed to vibrational excitations of combination modes consisting of an O–H stretch for **s1**, a N–H rock for **s2**, an O–H rock for **s3**, and an O–H rock with a CH₂ twist for **s4**, whereas the second additional band for **s3** and **s4** involve a C–H and N–H stretch for **s3** and **s4**, respectively. Even though all six isomers are energetically accessible from the ESI source, the presence of **s5** and **s6** are excluded based on the differences between the observed band positions and the calculated VDE values as well

as their relatively high energies. The presence of **sz** is also discounted for the same reasons. That is, there are no bands in either the 266 nm or the 193 nm NIPE spectra near the wide-spanning simulated spectrum with a calculated VDE of 4.59 eV, and the zwitterion complex is computed to be less stable than the canonical isomers **s1**–**s4**. As a result, the observed spectral bands are assigned to the remaining four isomers, **s1**–**s4**.

Band **a** at 3.89 eV is assigned to **s2** according to its calculated VDE of 3.90 eV, and band **d** at 4.26 eV is attributed to **s1** and **s4**, which have nearly identical calculated VDEs of 4.23 and 4.24 eV, respectively. Isomer **s3** with a calculated VDE of 4.07 eV is attributed to band **c** at 4.16 eV instead of band **b** at 4.01 eV, and the latter spectral feature is attributed to vibrational combination bands of **s2** as predicted by FCF simulations (Figure 3). These assignments are supported by

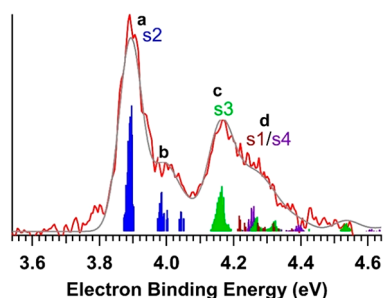


Figure 3. Simulated NIPE spectrum of the Sar-I[−] anion (gray) using calculated stick spectra (dark brown, blue, green, and violet for **s1**–**s4**, respectively) convoluted with Gaussian line broadening (where the fwhm for each Gaussian was set to 60 meV) superimposed onto the experimental 266 nm spectrum (red). The intensities of **s3** and sum of **s1** + **s4** are scaled by 0.38 and 0.59 respectively relative to that of **s2** to give the best fit of the experimental spectrum. The simulated spectra for each individual structure compared with the experimental spectrum are provided in Figure S2.

comparisons between the 20 K spectrum and the spectra collected at higher temperatures (120 and 200 K, Figure S3). Similar **a/b** intensity ratios in all three spectra suggest that these two features arise from the contribution of a single isomer. As it is harder to resolve the four distinct bands at higher temperatures, we simply compare the relative intensities of bands **a** and **c** (Figure S3) which are attributed to individual isomers (i.e., **s2** and **s3**, respectively). There is a significant increase in the **c/a** intensity ratio from 120 to 200 K, which qualitatively agrees with the calculated relative Gibbs free energies of **s2** and **s3** and their theoretical populations based on the Boltzmann distribution law (Table S1). The **c/a** intensity ratios at 20 and 120 K, however, are nearly identical despite a predicted thermal population difference of more than a factor of 10^4 (Table S1). This is probably due to rapid cooling of the cluster ions in the low temperature regime which prevents them from overcoming energy barriers and thermally equilibrating. Based upon these assignments, an excellent match was obtained between the measured and simulated spectra, which were convoluted from combinations of stick spectra that were normalized on the basis of the experimental peak height for **a** (Figure 3).

It is worth noting that the intensity of band **d**, which is composed of both **s1** and **s4**, is significantly weaker than that for band **a** assigned to **s2**. This intensity ratio at first glance seems to violate the Boltzmann distribution law since a larger

population is expected for a more thermally stable species. However, the actual band intensity is related not only to the population of the corresponding isomer but also to its photodetachment cross section, which is inherently related to the interaction strength between the two cluster ion components. Being the most thermally stable isomer, **s1** has the highest predicted thermal population but also the highest calculated binding energy (Table S3), along with a large expected anion-to-neutral geometry change upon photodetachment (Figure S1). As shown in Figure S2, the main band for **s1** is much broader than that of the other isomers, leading to a less pronounced band profile in the spectrum. This phenomenon that the most thermally stable isomer appears to have a small photodetachment cross section was also observed in previous studies.^{23,46}

Substitution of the amino-H in Sar-I[−] by a methyl group affords Dmg-I[−] and leads to a further decrease in the number of located isomers. Four structures, **d1**–**d3** and **dz**, were identified (Figure 2); similar to Sar-I[−], all of their FCF simulated spectra are dominated by vertical transitions, although **d1** also has an additional band due to a vibrational combination mainly associated with a C–H stretching mode that is ~ 0.4 eV higher in energy (Figure S4). Despite the simpler spectrum with fewer observed bands than for Gly-I[−] and Sar-I[−], and a smaller number of structures to consider, the assignment for Dmg-I[−] was initially confusing due to the inconsistency between the number of peaks observed in the $\Omega = 3/2$ and $\Omega = 1/2$ bands. Both **d3** and **dz** were eliminated due to their high relative energies (Table 1) as well as the absence of peaks in the range near their calculated VDEs (3.71 and 4.73 eV, Table 1 and Figure S4). In contrast, the most stable isomers (**d1** and **d2**) have calculated VDEs of 4.04 and 4.08 eV, which are in good accord with the observed bands at 4.01 and 4.10 eV. However, the two most intense bands **a** and **c** are not assigned directly to **d1** and **d2** but instead to the further spin–orbit coupling (SOC) splitting of the $I(^2P_{3/2})$ state of **d1** due to the interaction between the iodine atom and the closed-shell molecule. In Hund's case (a), the $I(^2P_{3/2})$ state further splits into two substates, whereas the $I(^2P_{1/2})$ state has no further splitting using the pseudodiatomic approximation; this results in two transitions in the $\Omega = 3/2$ band system and only one in the $\Omega = 1/2$ band system for one specific isomer (i.e., **d1**).^{47,49–51} The shoulder labeled **b** with an EBE of ~ 4.06 eV is consequently assigned to **d2**. Confirmation of the SOC splitting of **d1** was confirmed by CASPT2/maug-cc-pVTZ(-DK3) calculations which predict the gap between the two lowest states to be 0.12 eV (Table S2), in accord with the observed value of 0.09 eV. Since these gaps for the Sar-I[−] and Bet-I[−] complexes are both too small to be resolved (Table S2), this phenomenon is only observed in the case of the Dmg-I[−] complex. A similar SOC splitting also exists for **d2** (Table S2), but it was not detected since the EBE for this isomer is very close to that for band **c**. The experimental spectrum also can be reproduced by combining **d1**, **d2**, and their SOC peaks (Figure 4).

Our assignment of both **a** and **c** to **d1** and **b** to **d2** is also supported by comparing the 20 and 120 K spectra (Figure S5), which exhibit a clear enhancement of the contribution from **d2** at higher temperature. This trend is in accord with the Boltzmann distribution law since the higher-energy isomer **d2** becomes more populated at the higher temperature (Table S1). It also should be pointed out that the intensity change upon temperature variation for both Sar-I[−] and Dmg-I[−] is less

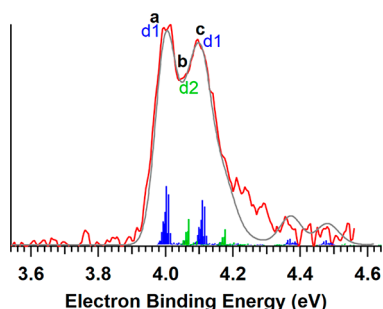


Figure 4. Simulated NIPE spectrum of the Dmg·I[−] anion (gray) using calculated stick spectra (blue and green for **d1** and **d2**, respectively) convoluted with Gaussian line broadening (where the fwhms of each Gaussian were set to 60 meV) superimposed onto the experimental 266 nm spectrum (red). There are two sets of identical stick spectra for each isomer with an EBE separation of 0.09 eV to represent the SOC splitting of each isomer. The intensities of **d2** are scaled by 0.26 relative to that of **d1** to match the experimental spectral profile. The simulated spectra for each individual structure compared with the experimental spectrum are provided in Figure S5.

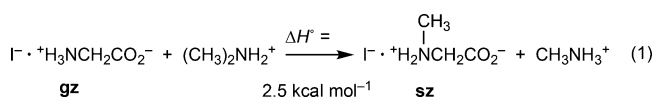
significant compared to the Gly·I[−] complex.²³ This is because in Sar·I[−] and Dmg·I[−] the more energetically favorable isomers (**s2** vs **s3** and **d1** vs **d2**) have larger entropies which to a certain extent negates the population gains for the less stable isomers at higher temperature. In contrast, the opposite is true for Gly·I[−] (i.e., the energetically most stable isomer has a smaller entropy); more details and equations are provided in the Supporting Information.

One transition is observed for both the $\Omega = 3/2$ and $1/2$ bands of Bet·I[−], indicating the formation of only one isomer. Due to the zwitterionic structure of betaine and the absence of amino- and carboxyl-H atoms which could migrate and result in a canonical–zwitterionic isomerization, the Bet·I[−] cluster anion is also predicted only to be in the zwitterionic form. M06-2X calculations located two isomers, **b1** with C_s symmetry and the iodide interacting with one H from each of the three methyl groups and **b2** with C₁ symmetry and the iodide interacting with two methyl-H atoms remotely located from the carboxylate group. Both species give rise to a simulated spectrum dominated by a vertical transition as shown in Figure S6, but **b1** is 4.79 kcal mol^{−1} more stable than **b2** and has a calculated VDE of 4.29 eV, which is close to the observed band position of 4.25 eV (Table 1). In contrast, **b2** has a calculated VDE of 4.05 eV, indicating that this isomer can be excluded and that **b1** is the only isomer formed for the Bet·I[−] cluster anion.

Upon successive *N*-methyl substitution of the Gly·I[−] anion cluster increasingly larger Sar·I[−], Dmg·I[−], and Bet·I[−] complexes with additional molecular complexity are formed. One might assume that this would lead to a greater number of isomers and more poorly resolved and less informative NIPE spectra. In fact, the opposite is observed due to increased steric interactions and the loss of amino-H atoms leading to changes in the preferred iodide anion binding sites. In Gly·I[−], five different canonical isomers that span a 3.93 kcal mol^{−1} range were identified, and the glycine is in a different conformation in each of them.²³ Both **g1** and **g3** have similar nonplanar glycine structures that differ in the orientation about the CO–CH₂NH₂ bond, and neither is among the predicted low-energy forms of the free amino acid.^{35–39} Their calculated and observed VDEs are both 4.30 and 4.26 eV, respectively. These are the largest values for the observed isomers, and the 1.20 eV

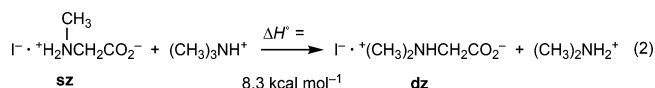
increase in the experimental VDE relative to I[−] is indicative of a strong interaction between I[−] and Gly in their anionic cluster ion. Among **g2**, **g4**, and **g5** with planar glycine heavy-atom (C, N, and O) arrangements, the glycine moiety of **g5** possesses the lowest-energy form for bare glycine based upon several computational studies,^{35–39} but **g5** is the least stable iodide cluster among all of the identified isomers. As for the zwitterionic form of the cluster (**gz**), it is much less stable than the canonical structures (3.56–7.49 kcal mol^{−1}) and does not contribute to the observed NIPE spectra.²³

For the Sar·I[−] complex, **s1**, **s3**, and **s4** are similar both structurally and in terms of their relative energies to **g1**, **g2**, and **g3**, respectively (Figure 2 and Table 1). Isomers **s2** and **g4** also share similar amino acid structures, but they have different iodide binding motifs. That is, there are two CH·I[−] interactions and one NH·I[−] hydrogen bond in **s2**, whereas the iodide anion binds to two amino-H atoms in **g4**. As a result, **s2** is preferentially stabilized and is only slightly less stable than **s1** but more stable than **s3** and **s4** (Table 1). In addition, the methylene group is slightly rotated in **s4** compared to that in **g3** so that the iodide also binds to one methylene-H. In all four observed Sar·I[−] isomers, three XH·I[−] hydrogen bonds are formed (one OH·I[−], NH·I[−], and CH·I[−] in **s1** and **s4**, two CH·I[−] and one NH·I[−] in **s2**, and two CH·I[−] and one OH·I[−] in **s3**). Isomer **s5** is related to **g5** and is a high-energy structure (+4.48 kcal mol^{−1} relative to **s1**), so it is not surprising that it is not observed in the NIPES measurements. Overall, the binding motifs in Gly·I[−] are largely retained in the Sar·I[−] complexes; but, one less isomer is observed, and the relative energies of the zwitterionic forms of these clusters drops from 7.49 (Gly·I[−]) to 2.79 kcal/mol (Sar·I[−]). This presumably is due to their primary versus secondary amine structures, respectively, as it is well-known that the basicity of amines increases upon alkyl substitution (i.e., PA = 204.0 (NH₃), 214.9 (MeNH₂), 222.2 (Me₂NH), and 226.8 (Me₃N) kcal mol^{−1}).⁵² In accord with this explanation, an isodesmic reaction using CCSD(T) energies that attempts to account for the different amine substitution in Gly and Sar (eq 1) is nearly thermoneutral.

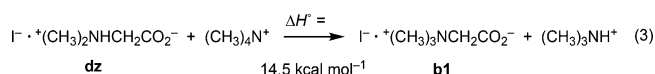


Upon incorporation of a *N*-methyl group into Sar·I[−] significant changes are observed in the resulting Dmg·I[−] cluster anion. This is because there are no N–H bonds in the canonical form of the latter complex, and only one traditional hydrogen bond donor (i.e., the CO₂H group) is available to interact with the iodide anion. As a consequence, only two Dmg·I[−] isomers (**d1** and **d2**) are present, and further spectral simplification is observed. Neither **d1** or **d2** correspond to the two most stable structures of Sar·I[−]; instead, they are related to **s3** and **s4**, respectively. In addition, the relative energy of the zwitterionic isomer of Dmg·I[−] (**dz**, 2.94 kcal mol^{−1}) is similar to that for Sar·I[−] (**sz**) even though Dmg is a tertiary amine and is expected to be more basic than Sar (which is a secondary amine). This surprising result is presumably due to the presence of only one N–H bond in **dz** versus two in **sz**. As a result, the latter ion can form N–H hydrogen bonds to both the I[−] and CO₂[−], whereas this is not the case in the former complex. To assess this further, an analogous isodesmic reaction to eq 1 was computed, and the more substituted amino acid cluster (**dz**) is disfavored relative

to the **gz/sz** difference by 5.8 kcal mol⁻¹. This can be largely attributed to the disparity between NH...O₂C and CH...O₂C hydrogen bonds.



N,N,N-Trimethylglycine in its neutral form is commonly referred to as glycine betaine or just betaine, and all of its hydrogens are attached to carbon atoms. Consequently, Bet-I⁻ can form only CH...I⁻ hydrogen bonds with iodide. A C_s structure (**b1**) with three nontraditional CH...I⁻ hydrogen bonds is found to be 4.79 kcal mol⁻¹ more stable than a C₁ isomer with two such interactions, and it is the only observed species. This cluster ion is predicted to be much less stable than **dz** as indicated in eq 3 due to the absence of a NH hydrogen bond donor and the consequent formation of both CH...I⁻ and CH...O₂C interactions.



In conclusion, iodide-tagging NIPES has been successfully applied to three *N*-methylated glycine-iodide cluster anions, Sar-I⁻, Dmg-I⁻, and Bet-I⁻, to investigate their isomeric structures. With the help of high-level calculations, four isomers of the Sar-I⁻ complex, two isomers of the Dmg-I⁻ cluster anion, and one isomer of Bet-I⁻ have been observed and identified. This reveals a clear trend of decreasing isomeric diversity upon increasing *N*-methyl substitution due to the loss of NH and OH hydrogen bond donating sites and an increase in the basicity of the amino group. Zwitterionic structures of Gly-I⁻, Sar-I⁻, and Dmg-I⁻ are less favorable than their canonical forms and do not contribute to the observed NIPES spectra. This study marks an important step forward in validating the effectiveness of iodide-tagging NIPES in probing conformations and tautomeric structures of iodide tagged clusters and lays the foundation for future studies of larger clusters of biological interest.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcllett.1c00125>.

Experimental and theoretical methods; thermal population analyses of different isomers; calculated relative electronic energies, Gibbs free energies, and thermal populations of each cluster isomer (Table S1); calculated CASPT2 relative energies of the three lowest SOC states for observed Sar-I⁻, Dmg-I⁻, and Bet-I⁻ isomers (Table S2); calculated dissociation energies (*D_e*, in kcal mol⁻¹) for each of Gly-I⁻, Sar-I⁻, and Dmg-I⁻ isomer (Table S3); optimized geometries of low-lying Sar-I⁻, Dmg-I⁻, and Bet-I⁻ anions and their corresponding neutral structures (Figure S1); experimental and simulated NIPE spectra of Sar-I⁻ for five low-energy canonical structures along with the most favorable zwitterionic form (Figure S2); *T*-dependent NIPE spectra of Sar-I⁻ (Figure S3); experimental versus simulated NIPE spectrum of Dmg-I⁻ of three low-energy canonical structures along with the most favorable zwitterionic form (Figure S4); *T*-dependent NIPE spectra of Dmg-I⁻ (Figure S5); experimental

versus simulated NIPE spectrum of Bet-I⁻ (Figure S6); and coordinates of identified isomers (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Xue-Bin Wang – Physical Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99352, United States; orcid.org/0000-0001-8326-1780; Email: xuebin.wang@pnnl.gov

Steven R. Kass – Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455, United States; orcid.org/0000-0001-7007-9322; Email: kass@umn.edu

Xiaoguo Zhou – Hefei National Laboratory for Physical Sciences at the Microscale, Department of Chemical Physics, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China; orcid.org/0000-0002-0264-0146; Email: xzhou@ustc.edu.cn

Authors

Wenjin Cao – Physical Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99352, United States; orcid.org/0000-0002-2852-4047

Hanhui Zhang – Physical Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99352, United States; Hefei National Laboratory for Physical Sciences at the Microscale, Department of Chemical Physics, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China

Qinqin Yuan – Physical Sciences Division, Pacific Northwest National Laboratory, Richland, Washington 99352, United States; orcid.org/0000-0001-5771-6147

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpcllett.1c00125>

Author Contributions

*W.C. and H.Z. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences, and was performed using EMSL, a national scientific user facility sponsored by the DOE's Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory, which is operated by Battelle Memorial Institute for the DOE. The theoretical calculations were conducted on the EMSL Cascade Supercomputer and at the University of Minnesota Supercomputing Institute. S.R.K. also acknowledges support from the National Science Foundation (CHE-1955186). X.Z. appreciates the financial support of the National Natural Science Foundation of China (21873089 and 22073088) and the Ministry of Science and Technology of China (2012YQ220113).

■ REFERENCES

- Wendisch, V. F. *Amino Acid Biosynthesis—Pathways, Regulation and Metabolic Engineering*; Springer-Verlag: Berlin, 2007.
- Fagg, G. E.; Foster, A. C. *Amino Acid Neurotransmitters and Their Pathways in the Mammalian Central Nervous System*. *Neuroscience* **1983**, *9*, 701–719.

- (3) Locke, M. J.; McIver, R. T. Effect of Solvation on the Acid/Base Properties of Glycine. *J. Am. Chem. Soc.* **1983**, *105*, 4226–4232.
- (4) Xu, S.; Nilles, J. M.; Bowen, K. H., Jr. Zwitterion Formation in Hydrated Amino Acid, Dipole Bound Anions: How Many Water Molecules are Required? *J. Chem. Phys.* **2003**, *119*, 10696–10701.
- (5) Bush, M. F.; Prell, J. S.; Saykally, R. J.; Williams, E. R. One Water Molecule Stabilizes the Cationized Arginine Zwitterion. *J. Am. Chem. Soc.* **2007**, *129*, 13544–13553.
- (6) Bush, M. F.; O'Brien, J. T.; Prell, J. S.; Saykally, R. J.; Williams, E. R. Infrared Spectroscopy of Cationized Arginine in the Gas Phase: Direct Evidence for the Transition from Nonzwitterionic to Zwitterionic Structure. *J. Am. Chem. Soc.* **2007**, *129*, 1612–1622.
- (7) Bush, M. F.; Oomens, J.; Saykally, R. J.; Williams, E. R. Effects of Alkaline Earth Metal Ion Complexation on Amino Acid Zwitterion Stability: Results from Infrared Action Spectroscopy. *J. Am. Chem. Soc.* **2008**, *130*, 6463–6471.
- (8) Bush, M. F.; Oomens, J.; Williams, E. R. Proton Affinity and Zwitterion Stability: New Results from Infrared Spectroscopy and Theory of Cationized Lysine and Analogues in the Gas Phase. *J. Phys. Chem. A* **2009**, *113*, 431–438.
- (9) Kapota, C.; Lemaire, J.; Maitre, P.; Ohanessian, G. Vibrational Signature of Charge Solvation vs Salt Bridge Isomers of Sodiater Amino Acids in the Gas Phase. *J. Am. Chem. Soc.* **2004**, *126*, 1836–1842.
- (10) Dunbar, R. C.; Hopkinson, A. C.; Oomens, J.; Siu, C.-K.; Siu, K. W. M.; Steill, J. D.; Verkerk, U. H.; Zhao, J. Conformation Switching in Gas-Phase Complexes of Histidine with Alkaline Earth Ions. *J. Phys. Chem. B* **2009**, *113*, 10403–10408.
- (11) Armentrout, P.; Rodgers, M.; Oomens, J.; Steill, J. Infrared Multiphoton Dissociation Spectroscopy of Cationized Serine: Effects of Alkali-Metal Cation Size on Gas-Phase Conformation. *J. Phys. Chem. A* **2008**, *112*, 2248–2257.
- (12) Polfer, N. C.; Oomens, J.; Moore, D. T.; von Helden, G.; Meijer, G.; Dunbar, R. C. Infrared Spectroscopy of Phenylalanine Ag(I) and Zn(II) Complexes in the Gas Phase. *J. Am. Chem. Soc.* **2006**, *128*, 517–525.
- (13) Coates, R. A.; Boles, G. C.; McNary, C. P.; Berden, G.; Oomens, J.; Armentrout, P. B. Zn²⁺ and Cd²⁺ Cationized Serine Complexes: Infrared Multiple Photon Dissociation Spectroscopy and Density Functional Theory Investigations. *Phys. Chem. Chem. Phys.* **2016**, *18*, 22434–22445.
- (14) Boles, G. C.; Hightower, R. L.; Berden, G.; Oomens, J.; Armentrout, P. B. Zinc and Cadmium Complexation of L-Threonine: An Infrared Multiple Photon Dissociation Spectroscopy and Theoretical Study. *J. Phys. Chem. B* **2019**, *123*, 9343–9354.
- (15) Lemoff, A. S.; Bush, M. F.; Williams, E. R. Binding Energies of Water to Sodiater Valine and Structural Isomers in the Gas Phase: The Effect of Proton Affinity on Zwitterion Stability. *J. Am. Chem. Soc.* **2003**, *125*, 13576–13584.
- (16) O'Brien, J. T.; Prell, J. S.; Berden, G.; Oomens, J.; Williams, E. R. Effects of Anions on the Zwitterion Stability of Glu, His and Arg Investigated by IRMPD Spectroscopy and Theory. *Int. J. Mass Spectrom.* **2010**, *297*, 116–123.
- (17) Schmidt, J.; Kass, S. R. Zwitterion vs Neutral Structures of Amino Acids Stabilized by a Negatively Charged Site: Infrared Photodissociation and Computations of Proline–Chloride Anion. *J. Phys. Chem. A* **2013**, *117*, 4863–4869.
- (18) Burt, M.; Wilson, K.; Marta, R.; Hasan, M.; Scott Hopkins, W.; McMahon, T. Assessing the Impact of Anion– π Effects on Phenylalanine Ion Structures Using IRMPD Spectroscopy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 24223–24234.
- (19) Corinti, D.; Gregori, B.; Guidoni, L.; Scuderi, D.; McMahon, T. B.; Chiavarino, B.; Fornarini, S.; Crestoni, M. E. Complexation of Halide Ions to Tyrosine: Role of Non-Covalent Interactions Evidenced by IRMPD Spectroscopy. *Phys. Chem. Chem. Phys.* **2018**, *20*, 4429–4441.
- (20) Milner, E. M.; Nix, M. G. D.; Dessent, C. E. H. Collision-Induced Dissociation of Halide Ion–Arginine Complexes: Evidence for Anion-Induced Zwitterion Formation in Gas-Phase Arginine. *J. Phys. Chem. A* **2012**, *116*, 801–809.
- (21) Walker, M.; Harvey, A. J. A.; Sen, A.; Dessent, C. E. H. Performance of M06, M06-2X, and M06-HF Density Functionals for Conformationally Flexible Anionic Clusters: M06 Functionals Perform Better than B3LYP for a Model System with Dispersion and Ionic Hydrogen-Bonding Interactions. *J. Phys. Chem. A* **2013**, *117*, 12590–12600.
- (22) Wytenbach, T.; Witt, M.; Bowers, M. T. On the Stability of Amino Acid Zwitterions in the Gas Phase: The Influence of Derivatization, Proton Affinity, and Alkali Ion Addition. *J. Am. Chem. Soc.* **2000**, *122*, 3458–3464.
- (23) 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.
- (24) Shirakawa, S.; Liu, S.; Kaneko, S.; Kumatabara, Y.; Fukuda, A.; Omagari, Y.; Maruoka, K. Tetraalkylammonium Salts as Hydrogen-Bonding Catalysts. *Angew. Chem., Int. Ed.* **2015**, *54*, 15767–15770.
- (25) Payne, C.; Kass, S. R. Structural Considerations for Charge-Enhanced Brønsted Acid Catalysts. *J. Phys. Org. Chem.* **2020**, *33*, No. e4069.
- (26) Noren, C.; Anthony-Cahill, S.; Griffith, M.; Schultz, P. A General Method for Site-Specific Incorporation of Unnatural Amino Acids into Proteins. *Science* **1989**, *244*, 182–188.
- (27) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. Expanding the Genetic Code of *Escherichia coli*. *Science* **2001**, *292*, 498–500.
- (28) Hendrickson, T. L.; de Crécy-Lagard, V. d.; Schimmel, P. Incorporation of Nonnatural Amino Acids Into Proteins. *Annu. Rev. Biochem.* **2004**, *73*, 147–176.
- (29) Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. Encoding Multiple Unnatural Amino Acids via Evolution of a Quadruplet-Decoding Ribosome. *Nature* **2010**, *464*, 441–444.
- (30) Axup, J. Y.; Bajjuri, K. M.; Ritland, M.; Hutchins, B. M.; Kim, C. H.; Kazane, S. A.; Halder, R.; Forsyth, J. S.; Santidrian, A. F.; Stafin, K.; et al. Synthesis of Site-Specific Antibody-Drug Conjugates Using Unnatural Amino Acids. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 16101.
- (31) Zimmerman, E. S.; Heibeck, T. H.; Gill, A.; Li, X.; Murray, C. J.; Madlansacay, M. R.; Tran, C.; Uter, N. T.; Yin, G.; Rivers, P. J.; et al. Production of Site-Specific Antibody–Drug Conjugates Using Optimized Non-Natural Amino Acids in a Cell-Free Expression System. *Bioconjugate Chem.* **2014**, *25*, 351–361.
- (32) Yin, G.; Stephenson, H. T.; Yang, J.; Li, X.; Armstrong, S. M.; Heibeck, T. H.; Tran, C.; Masikat, M. R.; Zhou, S.; Stafford, R. L.; et al. RF1 Attenuation Enables Efficient Non-Natural Amino Acid Incorporation for Production of Homogeneous Antibody Drug Conjugates. *Sci. Rep.* **2017**, *7*, 3026.
- (33) Wang, L.; Xie, J.; Deniz, A. A.; Schultz, P. G. Unnatural Amino Acid Mutagenesis of Green Fluorescent Protein. *J. Org. Chem.* **2003**, *68*, 174–176.
- (34) Cohen, B. E.; McAnaney, T. B.; Park, E. S.; Jan, Y. N.; Boxer, S. G.; Jan, L. Y. Probing Protein Electrostatics with a Synthetic Fluorescent Amino Acid. *Science* **2002**, *296*, 1700–1703.
- (35) Iijima, K.; Tanaka, K.; Onuma, S. Main Conformer of Gaseous Glycine: Molecular Structure and Rotational Barrier from Electron Diffraction Data and Rotational Constants. *J. Mol. Struct.* **1991**, *246*, 257–266.
- (36) Csaszar, A. G. Conformers of Gaseous Glycine. *J. Am. Chem. Soc.* **1992**, *114*, 9568–9575.
- (37) Hu, C. H.; Shen, M.; Schaefer, H. F. Glycine Conformational Analysis. *J. Am. Chem. Soc.* **1993**, *115*, 2923–2929.
- (38) Godfrey, P. D.; Brown, R. D. Shape of Glycine. *J. Am. Chem. Soc.* **1995**, *117*, 2019–2023.
- (39) Kasalová, V.; Allen, W. D.; Schaefer, H. F., III; Czinki, E.; Császár, A. G. Molecular Structures of the Two Most Stable Conformers of Free Glycine. *J. Comput. Chem.* **2007**, *28*, 1373–1383.

- (40) Balabin, R. M. Conformational Equilibrium in Glycine: Experimental Jet-Cooled Raman Spectrum. *J. Phys. Chem. Lett.* **2010**, *1*, 20–23.
- (41) Gómez-Zavaglia, A.; Fausto, R. Conformational Study of Sarcosine as Probed by Matrix-Isolation FT-IR Spectroscopy and Molecular Orbital Calculations. *Vib. Spectrosc.* **2003**, *33*, 105–126.
- (42) Cocinero, E. J.; Villanueva, P.; Lesarri, A.; Sanz, M. E.; Blanco, S.; Mata, S.; López, J. C.; Alonso, J. L. The Shape of Neutral Sarcosine in Gas Phase. *Chem. Phys. Lett.* **2007**, *435*, 336–341.
- (43) Lesarri, A.; Cocinero, E. J.; López, J. C.; Alonso, J. L. Gas-Phase Structure of N, N-Dimethylglycine. *ChemPhysChem* **2005**, *6*, 1559–1566.
- (44) Cormanich, R. A.; Ducati, L. C.; Tormena, C. F.; Rittner, R. A. Theoretical Investigation of the Dictating Forces in Small Amino Acid Conformational Preferences: The Case of Glycine, Sarcosine and N, N-dimethylglycine. *Chem. Phys.* **2013**, *421*, 32–38.
- (45) Hou, G.-L.; Wang, X.-B. Spectroscopic Signature of Proton Location in Proton Bound $\text{HSO}_4^- \cdot \text{H}^+ \cdot \text{X}^-$ ($\text{X} = \text{F}, \text{Cl}, \text{Br}, \text{and I}$) Clusters. *J. Phys. Chem. Lett.* **2019**, *10*, 6714–6719.
- (46) Wang, L.; Yuan, Q.; Cao, W.; Han, J.; Zhou, X.; Liu, S.; Wang, X.-B. Probing Orientation-Specific Charge-Dipole Interactions between Hexafluoroisopropanol and Halides: A Joint Photoelectron Spectroscopy and Theoretical Study. *J. Phys. Chem. A* **2020**, *124*, 2036–2045.
- (47) Zhang, H.; Cao, W.; Yuan, Q.; Wang, L.; Zhou, X.; Liu, S.; Wang, X.-B. Spectroscopic Evidence for Intact Carbonic Acid Stabilized by Halide Anions in the Gas Phase. *Phys. Chem. Chem. Phys.* **2020**, *22*, 19459–19467.
- (48) Yuan, Q.; Cao, W.; Wang, X.-B. Cryogenic and Temperature-Dependent Photoelectron Spectroscopy of Metal Complexes. *Int. Rev. Phys. Chem.* **2020**, *39*, 83–108.
- (49) Haberland, H. On the Spin-Orbit Splitting of the Rare Gas-Monohalide Molecular Ground State. *Z. Phys. A: At. Nucl.* **1982**, *307*, 35–39.
- (50) Zhao, Y.; Arnold, C. C.; Neumark, D. M. Study of the $\text{I} \cdot \text{CO}_2$ van der Waals Complex by Threshold Photodetachment Spectroscopy of $\text{I}^- \text{CO}_2$. *J. Chem. Soc., Faraday Trans.* **1993**, *89*, 1449–1456.
- (51) Arnold, D. W.; Bradforth, S. E.; Kim, E. H.; Neumark, D. M. Study of Halogen–Carbon Dioxide Clusters and the Fluoroformyloxyl Radical by Photodetachment of $\text{X}^-(\text{CO}_2)$ ($\text{X} = \text{I}, \text{Cl}, \text{Br}$) and FCO_2^- . *J. Chem. Phys.* **1995**, *102*, 3493–3509.
- (52) Hunter, E. P. L.; Lias, S. G. Evaluated Gas Phase Basicities and Proton Affinities of Molecules: An Update. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413–656.