Side-chain effects on the structures of protonated amino acid dimers: A gas-phase infrared spectroscopy study

Jongcheol Seo,[a] Waldemar Hoffmann,[a,b] Sebastian Malerz,[a,b] Stephan Warnke,[a,‡] Michael T. Bowers,[c] Kevin Pagel,[b] and Gert von Helden[a]*

[a] *Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany*

[b] *Freie Universität Berlin, Institute of Chemistry and Biochemistry, Takustrasse 3, 14195 Berlin, Germany.*

[c] *Department of Chemistry and Biochemistry, University of California Santa Barbara, Santa Barbara, California 93106, United States.*

* e-mail: helden@fhi-berlin.mpg.de

A protonated amino acid can interact in several ways with another uncharged amino acid molecule to form a protonated dimer. In case of amino acids that do not have basic or acidic side chains, the most likely protonation site is the amino group and the then protonated amine can be involved in a pairwise interaction with a neutral amine, a carboxylic acid, a carboxylate group and/or the sidechain of the partner amino acid. Here, we employ gas-phase infrared spectroscopy and density functional theory to identify these pairwise interactions in protonated homodimers of serine, isoleucine, phenylalanine and tyrosine. The results show the influence of the different side-chains on the respective interactions. A charge-solvated structure with pairwise interaction between a protonated amine and a neutral amine is preferred if the side chain can provide additional stabilizing interaction with the positive charge. In contrast, for amino acids where the side chain only interacts weakly with the protonated amine group, a protonated dimer is formed by an interaction between the protonated amine and the neutral carboxylic acid of the second amino acid.

1 Introduction

For biological molecules, protonation is a basic process that can have a profound impact on structure by strongly affecting non-covalent intra- and intermolecular interactions.[1-7] For peptides and proteins, the amine moiety is the most commonly preferred protonation site and protonated amines can engage in various interactions with other functional groups to play key roles in diverse processes such as protein/peptide folding,[2, 4] proton transfer dynamics[1, 3] or peptide backbone dissociation.[8-10] Recently, these interactions have also been recognized to be of importance in the formation of metastable amino acid clusters[11-14] with the possible relevance to aggregation related diseases such as phenylketonuria.[15]

Protonated amino acid (AA) dimers in the gas phase are of great interest because they provide simple models for the pairwise interactions between protonated amines and other functional groups that might be present in peptides and proteins. Due to their small size, such species are amenable to rigorous structural characterization by means of spectroscopy or first-principle theoretical calculations. To this end, gas-phase infrared (IR) spectroscopy, often performed as IR multiple photon dissociation (IRMPD) spectroscopy, has proven to be a powerful tool, since molecular vibrations of the important functional groups are sensitive to their protonation state and their local interactions and show well-resolved diagnostic IR features.[5, 7, 16-21] By comparing experimental IR spectra with theoretical IR spectra accurate structures can then be deduced.

Previously, several protonated AA dimers have been studied, both experimentally as well as by theory.[22-35] O–H and N–H stretching vibrations in Gly_2H^+,[22, 25] Ser_2H^+,[23] Thr_2H^+,[32] Lys_2H^+,[22, 29] Pro_2H^+,[31] and Trp_2H^+ [33] were investigated in the 2700–3600 cm^{-1} range and Gly_2H^+,[24, 26] Ala_2H^+,[26] Val_2H^+,[26] Cys_2H^+,[35] and Pro_2H^+ [24] were studied in the 1000–2000 cm^{-1} region, in which C=O stretching and N–H bending vibrations can be monitored. Heterodimers, such as $Ser{\cdot}PheH^+$,[28, 30] $Ala{\cdot}GlyH^+$,[26] $Gly{\cdot}LysH^+$,[22, 25] and $Gly{\cdot}PheH^+$,[34] were also investigated in both, the 2700–3700 cm^{-1} and 800–2000 cm^{-1} ranges. By combining those spectroscopic results with theoretical calculations, a large amount of structural detail was obtained. The general consensus is that for protonated AA dimers having neutral side chains, charge-solvated structures are preferred in which a protonated amine group engages in an intermolecular interaction with the neutral carboxylic acid or amine group of the partner AA. Dimers in which the second AA is zwitterionic and where the interaction is present in the form of a salt-bridge are only found for dimers of basic AAs.[22, 24, 29, 31]

For charge-solvated species, the pairwise interactions can occur in two ways. One possibility is intermolecular ionic hydrogen bonding between the protonated amine and the carbonyl group ($[H_2N–H]^+\cdots O=COH–$), which is suggested for several protonated dimers, such as, $Gly{\cdot}AlaH^+$, Gly_2H+, Ala_2H^+, Val_2H^+ and Cys_2H^+.[24-26] For species such as $Gly{\cdot}PheH^+$, Thr_2H^+, and Trp_2H^+, on the other hand, IR spectra indicate hydrogen bonding between the protonated amine and the neutral

amine group of the partner AA ($-[H_2N-H]^+\cdots NH_2-$).[32-34] These results suggests that the pairwise interactions in protonated AA dimers are strongly influenced by the AA side chains. The effects of cation-π interactions in the Gly·PheH$^+$ dimers have been investigated by replacing the -$CH_2C_6H_5$ (benzyl-) side chain of Phe with -$CH_2C_6F_5$.[34]

However, in spite of the progress made to date, a general understanding of the influence of the side chains on the dimer binding interactions is lacking. Here, we present a study on protonated AA dimers having polar uncharged, aliphatic as well as aromatic side chains, namely of the AAs serine, isoleucine, phenylalanine and tyrosine. IRMPD spectra in the mid-IR fingerprint region (1000–2000 cm^{-1}) as well as collision cross sections (CCSs) by using ion mobility spectrometry are reported. The data is compared with results from density functional theory calculations and the structure and pairwise interactions as well as the effects of side chains on the formation of the protonated dimer in the gas phase are discussed.

2 Materials and Methods

2.1 Sample preparation

Isoleucine (Ile), serine (Ser), phenylalanine (Phe), and tyrosine (Tyr) are purchased from Sigma-Aldrich (Taufkirchen, Germany), dissolved in pure water to the final concentration of 500 μM and used without further purification.

2.2 Ion mobility-mass spectrometry

Experiments are performed using an earlier described home-built, hybrid drift-tube ion mobility quadrupole time-of-flight (Q-TOF) mass spectrometer,[5, 36-38] equipped with a nanoelectrospray source (see also Figure S1 in ESI). For all experiments 5–10 μL of sample solution is loaded into a Pd/Pt-coated borosilicate capillary and sprayed with an electrospray voltage of 0.8–1.0 kV. Ions are transferred into the instrument via a capillary inlet into an RF ion funnel, from which they are pulsed into a 80 cm long drift tube. The drift tube is filled with helium buffer gas (4.5 mbar) and ions traverse the drift region with the aid of a weak electric field (10–15 V cm^{-1}). The ions exit the drift tube via a second ion funnel, pass through two differentially pumped ion guides, and enter a high vacuum region where *m*/*z* selection is done using a quadrupole mass filter. Ion currents can be recorded as a function of time to obtain an arrival time distribution (ATD). From the ATD, the angle-averaged collision cross-section (CCS, Ω_{exp}) can be determined using the Mason-Schamp equation.[39] In cases where more than one species is present in the ATD, electrostatic gating can be performed to allow for spectroscopic investigations on species that are then at the same time both, *m*/*z* as well as mobility (i.e. geometrical size) selected. Additional details are given in the Electronic Supplementary Information (See ESI).

2.3 Infrared multiple photon dissociation (IRMPD) spectroscopy

The *m*/*z*- as well as mobility-selected protonated AA dimer ions are irradiated with infrared photons provided by the Fritz Haber Institute free electron laser (FHI FEL).[40] When the IR light is resonant with an IR active mode of the complex, multiple photons are absorbed, which can result in the dissociation of the complex. An IRMPD spectrum is then constructed by plotting the fragmentation yield as a function of IR wavenumber in a range of 1000–1900 cm^{-1}. Further details are described in the Electronic Supplementary Information (See ESI).

2.4 Theoretical calculations

Starting geometries of protonated AA dimers are generated by a conformational search, as implemented in the MacroModel 10.3 software[41] and using the OPLS3 force field.[42] Structures in local minima that are within a 30 kJ mol^{-1} from the overall lowest energy structure are kept for further analysis. Several low-energy structures are then further optimized at the B3LYP level using Gaussian 09[43] with the aug-cc-pVDZ basis set with the addition of Grimme D3 dispersion correction.[44] Vibrational frequencies are scaled by 0.971,[45] and convoluted with Gaussian functions (10 cm^{-1} bandwidth) to construct theoretical IR spectra that can be compared with the experimental spectra. Theoretical CCS values (Ω_{theory}) of optimized structures are calculated using the trajectory method as implemented in MOBCAL.[46, 47]

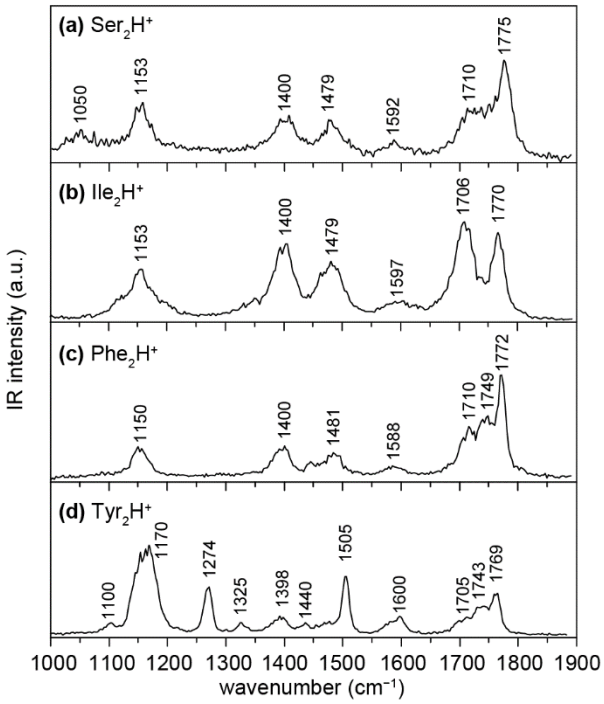

Figure 1. *IRMPD spectra of protonated homodimers of (a) serine (Ser_2H^+), (b) isoleucine (Ile_2H^+), (c) phenylalanine (Phe_2H^+), and (d) tyrosine (Tyr_2H^+).*

3 Results and Discussion

Protonated homodimers of serine, isoleucine, phenylalanine, and tyrosine (Ser_2H^+, Ile_2H^+, Phe_2H^+, and Tyr_2H^+) are abundantly found in the mass spectra as shown in Fig. S1 (See ESI). Arrival time distributions of those protonated dimers are shown in Fig. S2 (See ESI), which show for each dimer only one peak. Thus, only one IRMPD spectrum for each dimer has been recorded.

IRMPD spectra of Ser_2H^+, Ile_2H^+, Phe_2H^+, and Tyr_2H^+ are shown in Fig. 1. The IRMPD spectra of Ser_2H^+, Ile_2H^+, and Phe_2H^+ resemble each other, showing common IR features such as an IR band at ~1150 cm^{-1}, three bands in the range 1400–1600 cm^{-1}, and partly resolved bands around 1700–1780 cm^{-1}. These features are also found in the spectrum of Tyr_2H^+, however, with the presence of several additional intense bands at 1170, 1274, and 1505 cm^{-1}.

The relative intensities of the common IR bands differ between the spectra. This is especially the case in the 1700–1780 cm^{-1} region, where for Ser_2H^+, Phe_2H^+ and Tyr_2H^+ the IR band at ~1775 cm^{-1} is stronger than the features in the range of 1700–1750 cm^{-1}, in contrast to the case of Ile_2H^+ where the lower wavenumber band at 1706 cm^{-1} is more intense.

The IR features that are common in all spectra likely stem from N–H bending modes of neutral or protonated amine groups and from C=O or C–O stretching modes of neutral or deprotonated carboxylic acid groups, which are common for all AAs. Typically in AAs, the 1600–1800 cm^{-1} region is occupied by the stretching vibration of C=O in carboxylic acids and the exact band positions as well as intensities are highly sensitive to the protonation state and hydrogen bonding environment. A free carboxylic acid group has its C=O stretching vibration at ~1780 cm^{-1},[18] but when hydrogen bonded a shift towards 1710–1730 cm^{-1} is observed.[16] In contrast, a carboxylate group has a strong antisymmetric O–C–O stretching vibration at 1620–1640 cm^{-1},[17, 19] and a weaker symmetric stretch vibration around 1300–1400 cm^{-1}.[17, 19]

For a protonated dimer formed by two AAs, which do not have basic or acidic side chains, three different types of interaction are reasonable as schematically shown in Fig. 2. First, there are two different kinds of charge-solvated dimers in which both carboxylic acid groups are neutral and one amine group is protonated. The protonated amine can either interact with a neutral amine (Fig 2, type A) or a carboxylic acid group (Fig 2, type B) of the partner AA. The second possibility involves a zwitterionic partner that forms a salt bridge between the protonated amine group and the carboxylate group (Fig 2, type Z).

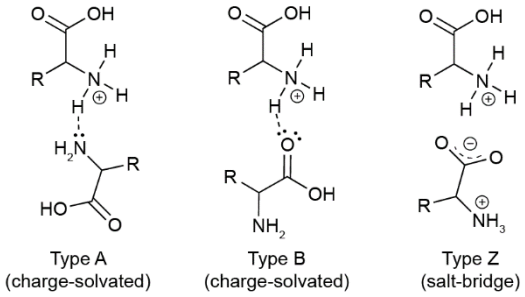

Figure 2. *Possible pairwise interactions in a protonated AA dimer.*

To investigate those structural motifs, *ab initio* calculations were performed and the lowest-energy structures found for type A, B, and Z dimers are shown in Fig. 3. The nomenclature **XA**, **XB**, and **XZ** denotes type A, B, and Z dimers of amino acid X, respectively, where X is the one-letter code of an amino acid. The *xyz*-coordinates of structures in Fig. 3 are given in Table S1–S4 (See ESI). Other several low-energy structures are shown in Fig. S4 (See ESI), but they are at least 5 kJ mol^{-1} less stable than the structures shown in Fig. 3. In the structures **SA**, **IA**, **FA**, and **YA**, the protonated -NH_3^+ group on one AA acts as a hydrogen bond donor to the -NH_2 group on the partner AA. In addition, the -NH_3^+ group is in close vicinity to at least one carboxylic acid -C=O group. Further, for **SA**, **FA**, and **YA,** at least one of the side chains interacts with the positively charged site. In the case of **SA**, the -OH side chain forms a hydrogen bond with the -NH_3^+ moiety. In **FA** and **YA,** both electron-rich aromatic rings are coordinating to the -NH_3^+ group, leading to the formation of a sandwich-like geometry. In the case of **IA**, on the other hand, the aliphatic side chains do not interact with the charged moiety and remain bystanders.

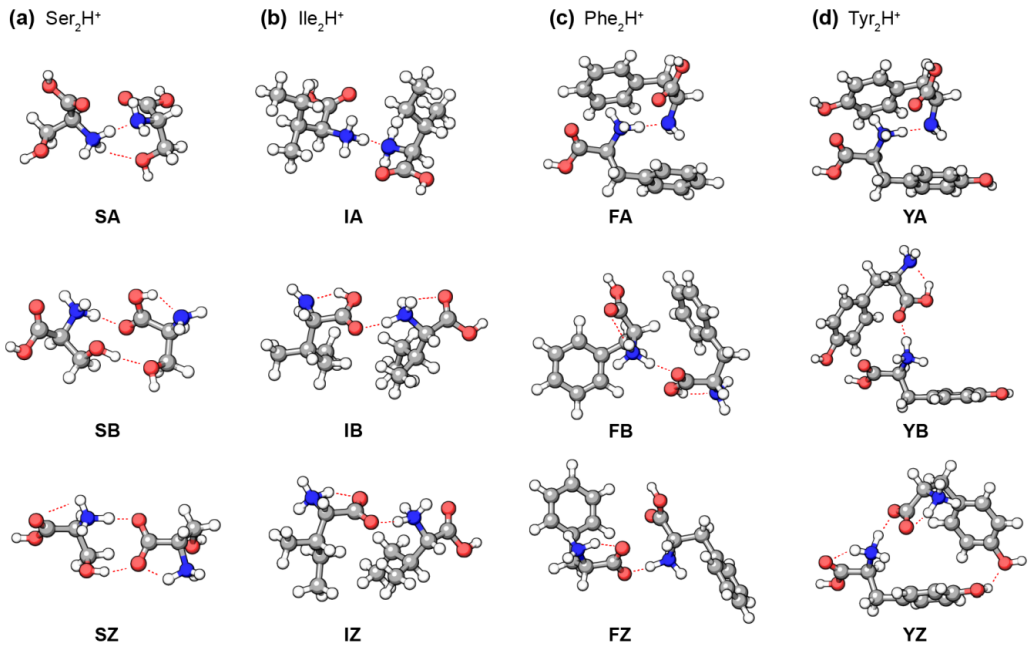

Figure 3. *Optimized geometries for (a) Ser_2H^+, (b) Ile_2H^+, (c) Phe_2H^+, and (d) Tyr_2H^+. Charge-solvated dimers are shown in the first and second rows (Type A and B, respectively), and salt-bridge (zwitterionic) dimers (Type Z) are shown in the third row. For each type, the lowest energy structure is shown.* ***XA***, ***XB***, *and* ***XZ*** *denote type A, B, and Z dimers of amino acid X, respectively, where X is a one-letter code of an amino acid.*

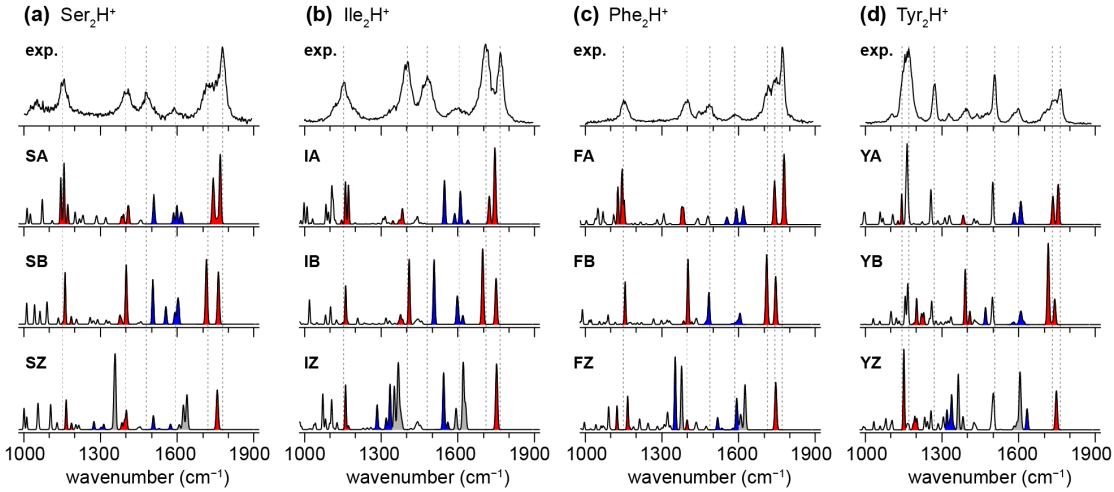

Figure 4. *Comparison between experimental and theoretical IR spectra of (a) Ser_2H^+ (b) Ile_2H^+ (c) Phe_2H^+ and (d) Tyr_2H^+. IR bands coloured in red and grey stem from vibrations of neutral carboxylic acid (-COOH) and carboxylate ($-COO^-$) groups, respectively. The IR bands coloured in blue result from the deformation modes of neutral or protonated amines ($-NH_2$ or NH_3^+).*

In types B dimers, the charged $-NH_3^+$ group not only coordinates to the -C=O group of the other AA but is an integral part of a hydrogen bonding network. Here the protonated amine group is central in an oriented chain of interactions ($H_2N\cdots HO–C=O\cdots[H–NH–H]^+\cdots O=C–OH$) and acts as hydrogen bond donor in both intra- and inter-molecular hydrogen bonds with two carboxylic acids. The -COOH group in the middle is in an *anti* conformation, in which the -OH proton is aligned opposite to C=O, and serves as a donor for an intra-molecular hydrogen bond with the $-NH_2$ group. Further, the side chains in type B dimers are not directly interacting with the positive charge and remain in the periphery (in contrast to the type A dimers).

The structures of type Z dimers are similar to those of type B dimers. The carboxylate groups and the protonated amines form an alternating chain of interactions ($[H_2N–H]^+\cdots[O–C–O]^-\cdots[H–NH–H]^+$). The side chains are mostly not directly involved in these core interactions with the exception of **SZ** where one hydroxyl group forms a hydrogen bond with the central COO^- group.

The experimental spectra shown in Fig. 1 are compared to those calculated for the various type A, B, and Z structures in Fig. 4. A detailed assignment of the vibrational modes is given in the Table S5–S8 (See ESI).

Spectra predicted for type A, B, and Z structures yield clearly distinct spectral signatures. Type A and B structures show two IR bands at 1700–1800 cm^{-1} which can be assigned to C=O stretching vibrations in the two -COOH groups and which are shown as coloured in red in Fig. 4. Type Z structures on the other hand show only one C=O stretching band in the 1700–1800 cm^{-1} range (coloured in red), as one of two carboxylic acid groups is deprotonated. Other bands that stem from the -COOH group are a C–O–H bending mode near 1400 cm^{-1} and a C–OH stretching mode around 1160 cm^{-1}. These are marked in red in Fig. 4. When comparing relative intensities between bands arising from vibrations of -COOH groups, it is interesting to note that in type A structures, the C=O stretching mode with higher wavenumber has a larger IR intensity while the opposite is the case for type B structures. For the bands arising from C–O–H bending modes, it can be noted that this band is particularly strong in type B structures. This can be attributed to a C–O–H$\cdots$$NH_2$ interaction which only occurs in type B structures and causes a strong infrared intensity in the C–O–H bending mode at ~1400 cm^{-1}. Furthermore, the umbrella mode of NH_3^+ at 1480−1500 cm^{-1} can be also indicative for type B structure, because it is typically blue-shifted to ~1540 cm^{-1} for type A or Z structures.

Type Z structures are characterized by an O–C–O antisymmetric stretching mode located at 1600–1650 cm^{-1} and a symmetric stretching mode at 1350–1400 cm^{-1}. These are coloured in grey in Fig. 4 and it can be noted, that the symmetric stretching mode has a slightly higher IR intensity than the antisymmetric stretching mode.

Bands arising from N–H bending modes in -NH_3^+ or -NH_2 are found in the range 1500–1620 cm^{-1} and are coloured in blue in Fig. 4. Their positions and intensities depend on the details of the structure and it appears difficult to assign systematic and consistent trends between type A, B, and Z structures.

All experimental spectra show the presence of double peaks at 1700–1800 cm^{-1}. This, together with absence of bands in the experimental spectra that would correspond to symmetric and antisymmetric O–C–O stretching vibrations, rules out the presence of type Z structures. The splitting and pattern of relative intensities between the two C=O stretch bands between 1700–1800 cm^{-1} can be used as a diagnostic feature to distinguish between type A and B structures. Experimentally, the C=O stretch band at higher frequency has a higher intensity than the one to lower frequency for all species except Ile_2H^+. Together with the good agreement in positions and relative intensities in the 1000–1800 cm^{-1} range, this indicates that Ser_2H^+, Phe_2H^+, and Tyr_2H^+ are present in the experiment as type A structures. For Ile_2H^+, the pattern of the C=O stretching modes

as well as of bands at lower wavenumber, especially of the three characteristic IR bands at ~1400 cm^{-1} (C−O−H bending), ~1480 cm^{-1} (NH_3^+ umbrella mode), and ~1600 cm^{-1} (NH_3^+ scissoring mode) indicate a type B structure.

The predicted gas-phase relative free energies as well as the theoretical CCS values (Ω_{theory}) of the structures in Fig. 3 are given in Table 1. For Ser_2H^+, Phe_2H^+, and Tyr_2H^+, type A structures (**SA**, **FA**, and **YA**) are lowest in energy, respectively. For Ile$_2$H+, on the other hand, a type B structure (**IB**) is predicted to be more stable. This is in agreement with the structure assignment based on the IR spectra given above. For the calculated structures, type A, B and Z structures give similar CCS values with the largest deviation being ~9% between type A and Z for Phe_2H^+. For all other species, the differences are much smaller and it is not possible to unambiguously distinguish structure types based only on CCS values. Nonetheless, the best agreement between experiment and theory is obtained for type A structures for Ser_2H^+, Phe_2H^+, and Tyr_2H^+ and for a type B structure for Ile_2H^+, which is consistent with the results from both spectroscopy and theory.

Table 1. Relative stabilities (ΔG_{298}) and collision cross sections (Ω_{exp} and Ω_{theory}) of protonated AA dimers.

species	experiment Ω_{exp} [a]	theory structure	Ω_{theory} [a]	ΔG_{298K} [b]
Ser_2H^+	84	**SA**	84.6 ± 1.8	0.0
		SB	82.7 ± 0.9	10.9
		SZ	87.5 ± 1.8	12.6
Ile_2H^+	104	**IA**	106.9 ± 1.7	3.9
		IB	104.1 ± 2.4	0.0
		IZ	104.6 ± 1.8	6.3
Phe_2H^+	112	**FA**	112.5 ± 2.1	0.0
		FB	115.4 ± 1.6	25.2
		FZ	122.7 ± 2.3	23.5
Tyr_2H^+	119	**YA**	119.2 ± 1.6	0.0
		YB	124.4 ± 1.3	20.8
		YZ	123.7 ± 1.9	23.1

[a] unit $Å^2$

[b] unit kJ mol^{-1}

The IRMPD spectrum of Ile_2H^+ in the present work is very similar to literature spectra of Gly_2H^+, Ala_2H^+, and Val_2H^+,[24, 26] which show a strong IR band at ~1700 cm^{-1} and a weaker feature at higher wavenumber and for those dimers with aliphatic side chains, type B interactions have been predicted.[24, 26] In the case of Ser_2H^+, the present work can be compared to the previous results on Thr_2H^+.[32] For both Ser_2H^+ and Thr_2H^+, the theoretically most stable conformer is type A; that is charge-solvated interactions with an additional H-bond between the protonated amine and a hydroxyl oxygen. For Phe_2H^+ and Tyr_2H^+, our results show that type A structures are preferred as

well and the calculations indicate that cation-π interactions play a role in stabilizing the positive charge. This result is in line with previous results for Trp_2H^+ and $Gly{\cdot}PheH^+$.[33, 34] We can thus conclude that if the protonated amine can interact with a side chain to gain further stabilization (*via* H-bonding or cation-π interaction), type A interactions ($H_2N\cdots[H–NH_2]^+$) are preferred to yield charge-solvated protonated dimer structures. In contrast, when the side chain cannot provide any significant further stabilization to the charge, the protonated dimers tend to form a charge-solvated structure with type B ($HO–C=O\cdots[H–NH_2]^+$) interaction.

Table 2. Summary of the proposed interaction types of protonated amino acid dimers

species	wavenumber region	proposed type[a]
Gly_2H^+	1000–2000 cm^{-1}	B[24, 26]
	2700–3700 cm^{-1}	A[22], B[25]
Ala_2H^+	1000–2000 cm^{-1}	B[26]
Val_2H^+	1000–2000 cm^{-1}	B[26]
Ile_2H^+	1000–2000 cm^{-1}	B[b]
Phe_2H^+	1000–2000 cm^{-1}	A[b]
Tyr_2H^+	1000–2000 cm^{-1}	A[b]
Trp_2H^+	2700–3700 cm^{-1}	A[33]
Ser_2H^+	1000–2000 cm^{-1}	A[b]
	2700–3700 cm^{-1}	A[23]
Thr_2H^+	2700–3700 cm^{-1}	A[32]
Cys_2H^+	1000–2000 cm^{-1}	B[35]
Pro_2H^+	1000–2000 cm^{-1}	Z[24, 31]
	2700–3700 cm^{-1}	
Lys_2H^+	2700–3700 cm^{-1}	A,[22, 27] Z[29]
$Ala{\cdot}GlyH^+$	1000–2000 cm^{-1}	B[26]
$Gly{\cdot}LysH^+$	2700–3700 cm^{-1}	Z[22]
$Ser{\cdot}PheH^+$	2700–3700 cm^{-1}	-[c, 28, 30]
$Gly{\cdot}PheH^+$	1000–2000 cm^{-1}	A[34]
$Gly{\cdot}Phe(F_5)H^+$ [d]	1000–2000 cm^{-1}	B[34]

[a] A, charge-solvated ($H_2N\cdots[H–NH_2]^+$); B, charge-solvated ($HO\text{-}C=O\cdots[H–NH_2]^+$); Z, salt-bridge (zwitterionic).

[b] This work.

[c] Not clear.

[d] $Phe(F_5)$, pentafluorophenylalanine.

4 Conclusions

Gas-phase IR spectroscopy in the 1000–2000 cm^{-1} fingerprint region combined with density functional theory is used to determine the structures and interactions in protonated homodimers of

serine, isoleucine, phenylalanine, and tyrosine. Three different types of pairwise interactions between the two AA molecules were considered, for which density functional theory predicts IR signatures that are distinguishable from each other. Comparing those predicted IR spectra with experimental spectra allows for the clear identification of the type of pairwise interaction of each protonated dimer. The present results as well as previous results are all summarized in Table 2. The results suggest that the interactions of the side chains with the charged sites play an important role in determining which structural motif is preferred. For AAs with non-interacting aliphatic side chains, charge-solvated dimer structures that are characterized by an interaction between the protonated amine and a carboxylic acid are formed. When the side chain is capable of stabilizing the positive charge either *via* H-bonding or cation-π interaction, charge solvated structures with intermolecular hydrogen bonding between a protonated amine and a neutral amine ($H_2N\cdots[H–NH_2]^+$) are favoured.

5 Acknowledgements

Authors acknowledge the expert technical assistance from S. Gewinner and W. Schöllkopf of the FHI free electron laser facility. M.T.B. gratefully acknowledges the support of the National Science Foundation (USA) for support under grants CHE-1301032 and CHE-1565941 and support from the Alexander von Humboldt foundation.

6 References

1. G. J. Kearley, F. Fillaux, M.-H. Baron, S. Bennington and J. Tomkinson, *Science*, 1994, **264**, 1285-1289.
2. J. M. Sanchez-Ruiz and G. I. Makhatadze, *Trends Biotechnol.*, 2001, **19**, 132-135.
3. C. J. Chang, M. C. Y. Chang, N. H. Damrauer and D. G. Nocera, *Biochim. Biophys. Acta*, 2004, **1655**, 13-28.
4. C. A. Jimenez-Cruz, G. I. Makhatadze and A. E. Garcia, *Phys. Chem. Chem. Phys.*, 2011, **13**, 17056-17063.
5. S. Warnke, J. Seo, J. Boschmans, F. Sobott, J. H. Scrivens, C. Bleiholder, M. T. Bowers, S. Gewinner, W. Schöllkopf, K. Pagel and G. von Helden, *J. Am. Chem. Soc.*, 2015, **137**, 4236-4242.
6. J. Boschmans, S. Jacobs, J. P. Williams, M. Palmer, K. Richardson, K. Giles, C. Lapthorn, W. A. Herrebout, F. Lemiere and F. Sobott, *Analyst*, 2016, **141**, 4044-4054.
7. J. Seo, S. Warnke, S. Gewinner, W. Schöllkopf, M. T. Bowers, K. Pagel and G. von Helden, *Phys. Chem. Chem. Phys.*, 2016, **18**, 25474-25482.
8. A. R. Dongré, J. L. Jones, Á. Somogyi and V. H. Wysocki, *J. Am. Chem. Soc.*, 1996, **118**, 8365-8374.
9. V. H. Wysocki, G. Tsaprailis, L. L. Smith and L. A. Breci, *J. Mass Spectrom.*, 2000, **35**, 1399-1406.
10. B. Paizs and S. Suhai, *Mass Spectrom. Rev.*, 2005, **24**, 508-548.
11. L. Adler-Abramovich, L. Vaks, O. Carny, D. Trudler, A. Magno, A. Caflisch, D. Frenkel and E. Gazit, *Nat. Chem. Biol.*, 2012, **8**, 701-706.

12. H. W. German, S. Uyaver, Hansmann and H. E. Ulrich, *J. Phys. Chem. A*, 2015, **119**, 1609-1615.
13. T. D. Do, W. M. Kincannon and M. T. Bowers, *J. Am. Chem. Soc.*, 2015, **137**, 10080-10083.
14. T. D. Do, N. E. C. de Almeida, N. E. LaPointe, A. Chamas, S. C. Feinstein and M. T. Bowers, *Anal. Chem.*, 2016, **88**, 868-876.
15. W. B. Hanley, *Am. J. Med.*, 2004, **117**, 590-595.
16. C. Kapota, J. Lemaire, P. Maître and G. Ohanessian, *J. Am. Chem. Soc.*, 2004, **126**, 1836-1842.
17. M. N. Blom, I. Compagnon, N. C. Polfer, G. von Helden, G. Meijer, S. Suhai, B. Paizs and J. Oomens, *J. Phys. Chem. A*, 2007, **111**, 7309-7316.
18. R. Linder, K. Seefeld, A. Vavra and K. Kleinermanns, *Chem. Phys. Lett.*, 2008, **453**, 1-6.
19. J. Oomens, J. D. Steill and B. Redlich, *J. Am. Chem. Soc.*, 2009, **131**, 4310-4319.
20. A. I. González Flórez, D.-S. Ahn, S. Gewinner, W. Schöllkopf and G. von Helden, *Phys. Chem. Chem. Phys.*, 2015, **17**, 21902-21911.
21. A. P. Cismesia, G. R. Nicholls and N. C. Polfer, *J. Mol. Spectrosc.*, 2017, **332**, 79-85.
22. H.-B. Oh, C. Lin, H. Y. Hwang, H. Zhai, K. Breuker, V. Zabrouskov, B. K. Carpenter and F. W. McLafferty, *J. Am. Chem. Soc.*, 2005, **127**, 4076-4083.
23. X. Kong, I. A. Tsai, S. Sabu, C.-C. Han, Y. T. Lee, H.-C. Chang, S.-Y. Tu, A. H. Kung and C.-C. Wu, *Angew. Chem. Int. Ed.*, 2006, **45**, 4130-4134.
24. R. Wu and T. B. McMahon, *J. Am. Chem. Soc.*, 2007, **129**, 4864-4865.
25. C. G. Atkins, K. Rajabi, E. A. L. Gillis and T. D. Fridgen, *J. Phys. Chem. A*, 2008, **112**, 10220-10225.
26. K. Rajabi and T. D. Fridgen, *J. Phys. Chem. A*, 2008, **112**, 23-30.
27. R. Wu, R. A. Marta, J. K. Martens, K. R. Eldridge and T. B. McMahon, *J. Am. Soc. Mass. Spectrom.*, 2011, **22**, 1651.
28. U. J. Lorenz and T. R. Rizzo, *J. Am. Chem. Soc.*, 2012, **134**, 11053-11055.
29. X. Kong, *J. Am. Soc. Mass. Spectrom.*, 2014, **25**, 422-426.
30. U. J. Lorenz and T. R. Rizzo, *J. Am. Chem. Soc.*, 2014, **136**, 14974-14980.
31. Y. J. Alahmadi, A. Gholami and T. D. Fridgen, *Phys. Chem. Chem. Phys.*, 2014, **16**, 26855-26863.
32. H. Yin and X. Kong, *J. Am. Soc. Mass. Spectrom.*, 2015, **26**, 1455-1461.
33. R. Feng, H. Yin and X. Kong, *Rapid Commun. Mass Spectrom.*, 2016, **30**, 24-28.
34. W. Fu, J. Xiong, M. J. Lecours, P. J. J. Carr, R. A. Marta, E. Fillion, T. McMahon, V. Steinmetz and W. S. Hopkins, *J. Mol. Spectrosc.*, 2016, **330**, 194-199.
35. C. Ieritano, P. J. J. Carr, M. Hasan, M. Burt, R. A. Marta, V. Steinmetz, E. Fillion, T. B. McMahon and W. Scott Hopkins, *Phys. Chem. Chem. Phys.*, 2016, **18**, 4704-4710.
36. P. R. Kemper, N. F. Dupuis and M. T. Bowers, *Int. J. Mass spectrom.*, 2009, **287**, 46-57.
37. S. Warnke, C. Baldauf, M. T. Bowers, K. Pagel and G. von Helden, *J. Am. Chem. Soc.*, 2014, **136**, 10308-10314.
38. S. Warnke, G. von Helden and K. Pagel, *Proteomics*, 2015, **15**, 2804-2812.
39. E. A. Mason and E. W. McDaniel, *Transport Properties of Ions in Gases*, Wiley-VCH, New York, 2005.
40. W. Schöllkopf, S. Gewinner, H. Junkes, A. Paarmann, G. von Helden, H. Bluem and A. M. M. Todd, *Advances in X-Ray Free-Electron Lasers Instrumentation III, Proceedings of SPIE*, 2015, **9512**, 95121L.
41. Schrödinger Release 2016-4: MacroModel (Schrödinger, LLC, New York, NY, 2016).
42. E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel and R. A. Friesner, *J. Chem. Theory Comput.*, 2016, **12**, 281-296.
43. Gaussian 09, Revision D.01 (Gaussian, Inc., Wallingford CT, USA, 2016).
44. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
45. J. P. Merrick, D. Moran and L. Radom, *J. Phys. Chem. A*, 2007, **111**, 11683-11700.

46. M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz and M. F. Jarrold, *J. Phys. Chem.*, 1996, **100**, 16082-16086.
47. A. A. Shvartsburg and M. F. Jarrold, *Chem. Phys. Lett.*, 1996, **261**, 86-91.

Electronic Supplementary Information

Side-chain effects on the structures of protonated amino acid dimers: A gas-phase infrared spectroscopy study

Jongcheol Seo,[a] Waldemar Hoffmann,[a,b] Sebastian Malerz,[a,b] Stephan Warnke,[a,‡] Michael T. Bowers,[c] Kevin Pagel,[b] and Gert von Helden[a]*

[a] Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4–6, 14195 Berlin, Germany.

[b] Freie Universität Berlin, Takustrasse 3, 14195 Berlin,

[c] Department of Chemistry and Biochemistry, University of California Santa Barbara, Santa Barbara, California 93106, United States.

* E-mail: helden@fhi-berlin.mpg.de

1. Detailed Experimental Procedures

1.1. Instrumental details. A scheme of the instrument is shown in Figure S1. Ions are generated in a nano-electrospray source. Ion mobility separation occurs in the helium filled drift tube through which ions drift under the influence of a weak electric field (10–15 V cm^{-1}) at a buffer gas pressure of ~4.5 mbar. The ions then traverse two differentially pumped ion guides (10^{-2} mbar and 10^{-5} mbar, respectively) and enter the low vacuum region of the quadrupole and TOF analyzers (10^{-8} mbar). Ion mobility selection occurs directly after the drift region, followed by mass selection, IR irradiation and detection in the TOF analyzer.

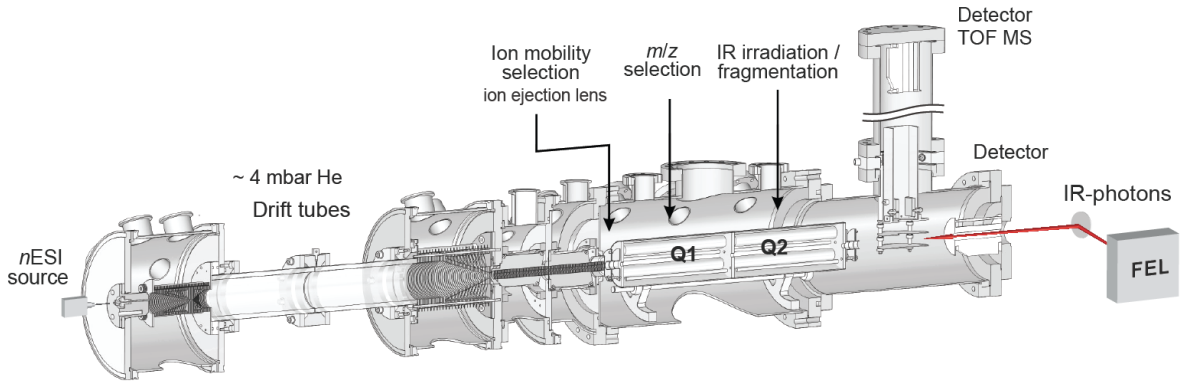

Figure S1. A scheme of the hybrid drift-tube quadrupole time-of-flight (Q-TOF) instrument.

1.2. Collision Cross-Section (CCS) Determination. In the present work, the observed arrival time (t_D) is a sum of a drift time (t_d) in the drift tube and a flight time (t_0) after the drift tube, which can be described as following equation S1 where K is ion mobility, V_d is voltage difference across the drift tube, and L is the length of the drift tube.

$$t_D = t_d + t_0 = \frac{L^2}{K}\frac{1}{V_d} + t_0 \qquad \text{(S1)}$$

Plotting t_D as a function of $1/V_d$ yields a highly linear curve, which allows the determination of the ion mobility K from the slope. The CCS (Ω) is determined from the Mason-Schamp equation (equation S2), where P_0 and T_0 are the standard pressure and temperature, N_0 is the number density of buffer-gas molecules, q is the charge of the ion, μ is the reduced mass of ion-buffer gas molecule system, and k_B is the Boltzmann constant.

$$\Omega = \frac{3q}{16N_0}\sqrt{\frac{2\pi}{\mu k_B T}}\frac{P_0}{P}\frac{T}{T_0}\frac{1}{K} \qquad \text{(S2)}$$

1.3. Infrared Multiple Photon Dissociation (IRMPD) Spectroscopy. A narrow fraction of the drift-time separated ion cloud is selected by electrostatic time gating after the drift tube. Subsequently, *m*/*z*-selection is performed in the quadrupole mass filter. The now ion mobility as well as *m*/*z*-selected ions are then irradiated by a ~7 μs pulse of IR photons with ~20 mJ of pulse energy provided by Fritz

Haber Institute free electron laser (FHI FEL). The relative abundances of precursor and fragment ions are recorded as a function of the IR wavenumber in the 1000–1900 cm^{-1} range in 3 cm^{-1} steps using a time-of-flight (TOF) mass analyzer. The final IR spectrum is obtained by plotting the fragmentation yield as a function of the IR wavenumber. The fragmentation yield (Y) at the specific wavenumber ($\tilde{\nu}$) is determined by equation S3, where I_p and I_f are the wavenumber-dependent abundances of the precursor and the fragment ions, respectively, and P is the IR pulse energy.

$$Y(\tilde{\nu}) = -\log\left\{\frac{I_p(\tilde{\nu})}{I_p(\tilde{\nu}) + \sum I_f(\tilde{\nu})}\right\}\frac{1}{P(\tilde{\nu})} \qquad \text{(S3)}$$

2. Supplementary Data

2.1 Time-of-Flight Mass Spectra of Protonated Dimers

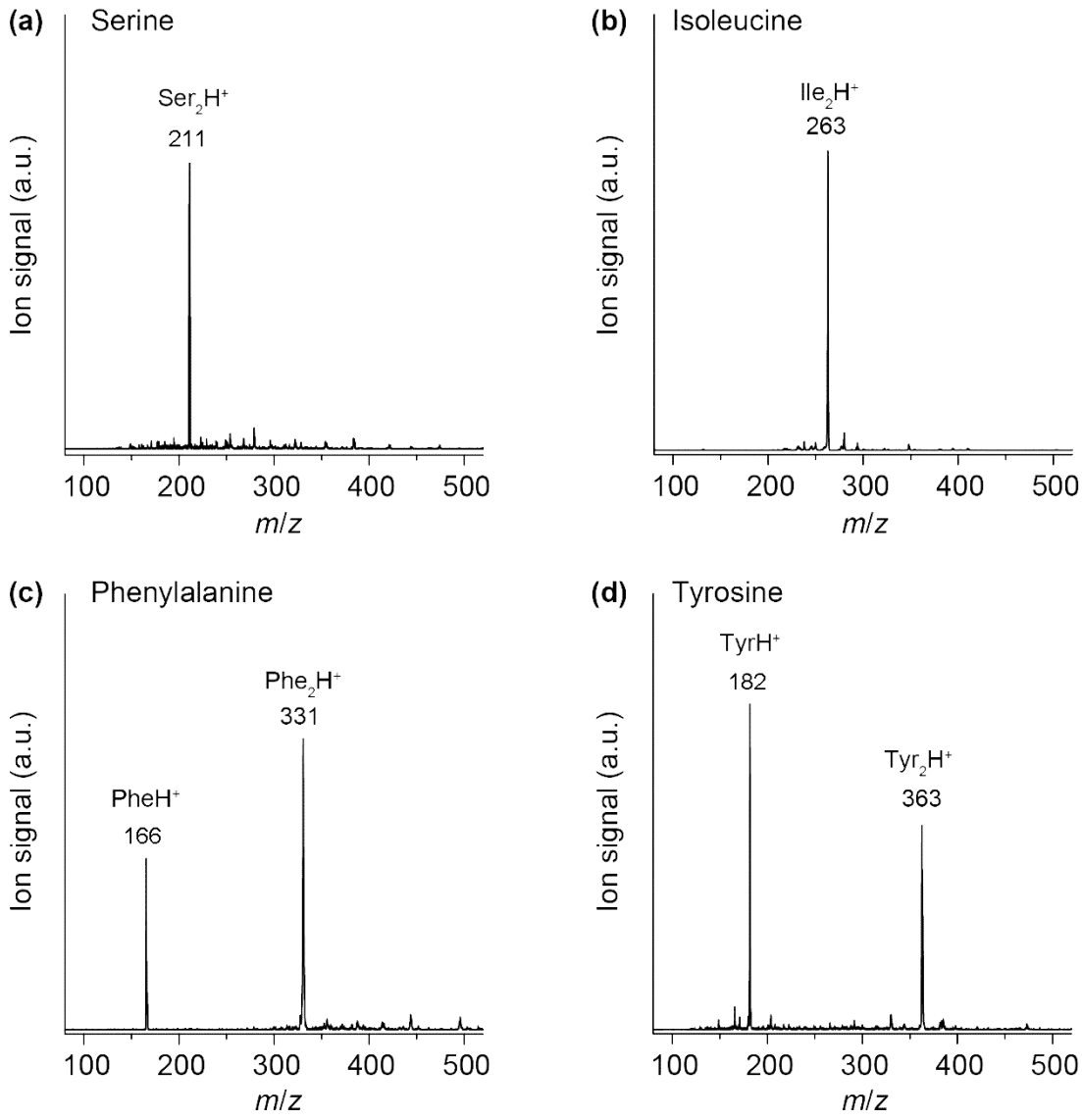

Figure S2. Time-of-flight mass spectra of (a) serine, (b) isoleucine, (c) phenylalanine, and (d) tyrosine electrosprayed from water (500 μM). Signals for $SerH^+$ and $IleH^+$ are not observed because of low transmission for ions with *m/z* below 160 at the present experimental conditions.

2.2 Arrival Time Distributions of Protonated Dimers

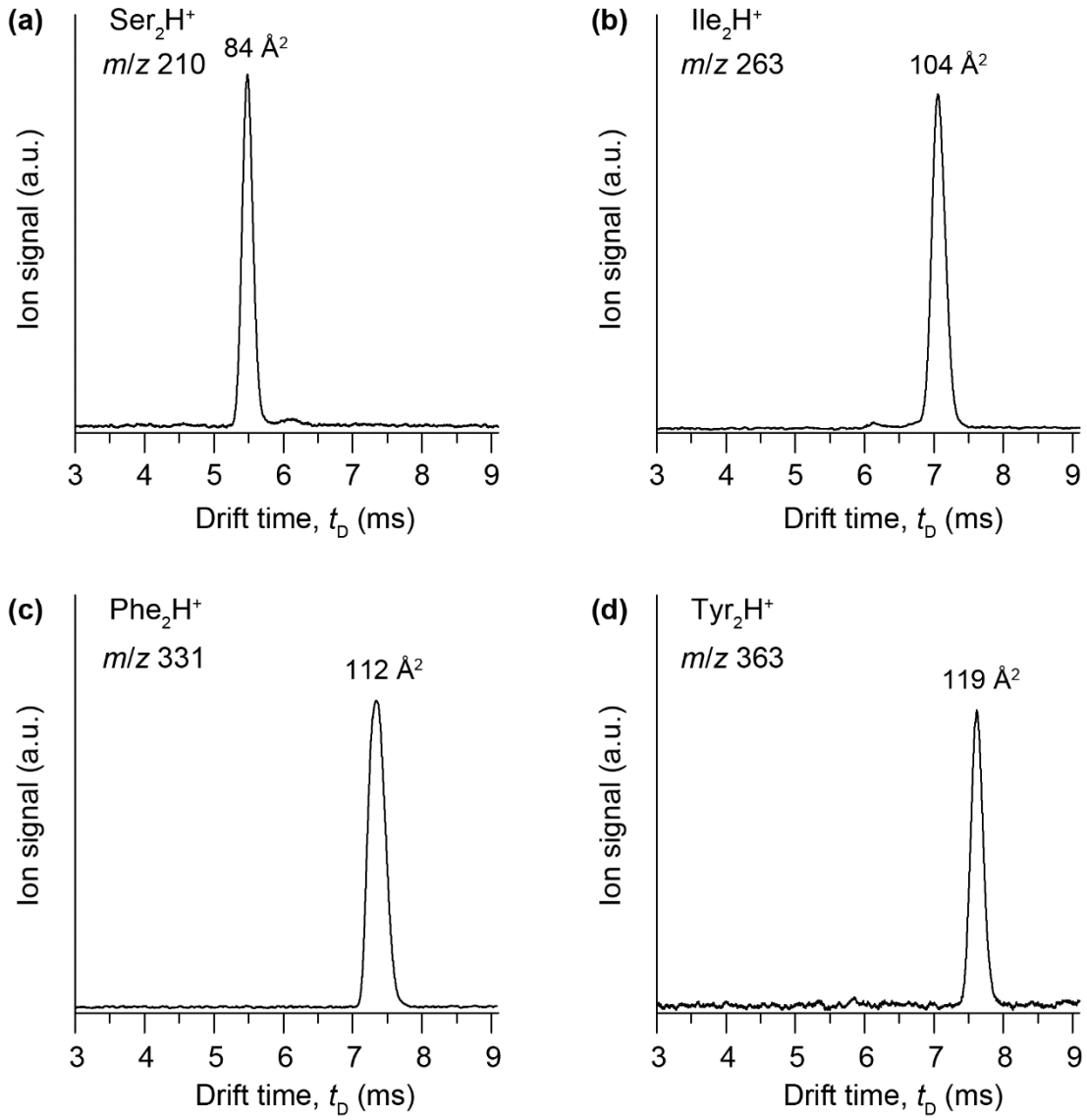

Figure S3. Arrival time distribution of (a) Ser_2H^+, (b) Ile_2H^+, (c) Phe_2H^+, and (d) Tyr_2H^+ and their corresponding collision cross-sections.

2.3 Geometry of the lowest-energy protonated dimer structures

Table S1. The xyz-coordinates of **SA**, **SB**, and **SZ** shown in Figure 3.

SA				**SB**				**SZ**			
H	–1.450279	–0.115430	2.470482	H	3.344758	0.280347	–1.792014	H	4.512129	1.523144	–0.963747
N	–1.279841	–0.423741	1.515330	N	3.508342	–0.518287	–1.183188	C	3.212246	–0.160738	–0.654522
C	–2.515472	–0.349191	0.723829	C	3.089131	–0.238289	0.194136	C	1.666026	0.000876	–0.599800
C	–2.218399	–0.831255	–0.693559	C	1.751746	–0.958012	0.437366	O	1.002383	–1.059844	–0.664083
O	–1.105294	–1.044477	–1.128414	O	1.080707	–0.818225	1.452669	C	3.765521	–1.179176	0.336051
C	–3.044634	1.088266	0.679572	C	2.972408	1.247190	0.537973	O	3.653371	–0.557117	1.618888
O	–2.063926	1.912910	0.029340	O	2.076679	1.840623	–0.424323	H	4.077639	1.224720	0.633160
H	–0.927971	–1.378002	1.573471	H	4.481707	–0.794738	–1.255635	H	3.487924	–0.430889	–1.680353
H	–3.326595	–0.980009	1.122475	H	3.797596	–0.688929	0.904320	H	3.155793	–2.089043	0.258212
H	–3.995827	1.099933	0.130688	H	2.582284	1.364451	1.558029	H	4.814955	–1.419857	0.103308
H	–3.223285	1.438595	1.708731	H	3.965379	1.711628	0.468579	H	3.989748	–1.141809	2.308559
H	–2.447298	2.783557	–0.131076	H	2.211824	2.795922	–0.448722	O	1.256454	1.190618	–0.527344
O	–3.345952	–0.980632	–1.400689	O	1.371436	–1.753422	–0.549257	H	–1.746039	0.022332	–1.550422
H	–3.120862	–1.278555	–2.298620	H	2.091871	–1.645340	–1.237157	C	–2.269482	–0.526149	0.350595
H	1.183405	1.584350	0.763648	H	–1.133163	–1.347639	–0.440802	C	–3.750685	–0.293728	0.072469
C	1.687444	0.091220	–0.575396	C	–2.201749	–0.036242	0.750160	O	–4.236301	–0.379433	–1.032423
C	2.181607	–0.970249	0.407122	C	–3.359392	–0.059796	–0.246689	C	–1.558509	0.690241	0.987137
O	1.714676	–1.135564	1.510552	O	–3.557650	–0.987109	–0.996748	O	–1.418112	1.708168	0.016515
C	2.808011	0.970383	–1.12632	C	–1.249717	1.151737	0.562121	H	–2.054386	–1.593351	–1.443360
O	3.300580	1.704930	–0.007609	O	–0.601900	0.959801	–0.684249	H	–2.162431	–1.407065	0.996380
H	0.110413	1.478864	–0.499473	H	–1.994068	–2.134744	0.728176	H	–2.168201	1.048509	1.825381
H	1.166988	–0.426002	–1.388632	H	–2.606941	–0.034568	1.769728	H	–0.582222	0.366529	1.374259
H	3.575449	0.327325	–1.577102	H	–1.844559	2.076068	0.584544	H	–0.477386	1.731625	–0.248145
H	2.404166	1.641837	–1.902182	H	–0.531336	1.168094	1.394152	O	–4.416495	–0.007556	1.197761
H	4.070828	2.228912	–0.257688	H	0.290360	1.350730	–0.649694	H	–5.354901	0.131188	0.981679
O	3.162616	–1.702129	–0.135193	O	–4.100859	1.047690	–0.150295	N	–1.614063	–0.807901	–0.961021
H	3.439629	–2.380353	0.504688	H	–4.832180	0.997057	–0.790484	H	–0.551803	–0.955657	–0.842762
N	0.689587	0.934681	0.143810	N	–1.429301	–1.303792	0.542141	N	3.762128	1.211945	–0.346131
H	–0.019915	0.363049	0.731242	H	–0.538699	–1.298173	1.101526	H	2.901629	1.814669	–0.428174

Table S2. The xyz-coordinates of **IA**, **IB**, and **IZ** shown in Figure 3.

IA				**IB**				**IC**			
H	-1.585344	0.372242	-2.510753	H	4.003319	-2.247193	-1.219828	H	-3.751224	-2.477231	1.423119
N	-1.600773	-0.080107	-1.598043	N	3.344480	-1.476577	-1.178981	C	-2.812338	-1.321375	-0.120005
C	-2.794933	0.357841	-0.841157	C	2.932405	-1.165139	0.205441	C	-1.276183	-1.505123	-0.186670
C	-2.551902	1.762697	-0.303027	C	1.418744	-1.368886	0.278206	O	-0.737154	-1.307909	-1.302665
O	-1.471246	2.320701	-0.238940	O	0.752889	-1.219592	1.299017	C	-3.257911	0.093976	-0.525886
C	-3.143722	-0.622209	0.315556	C	3.319088	0.238934	0.752792	C	-2.605054	1.179938	0.354081
C	-1.989633	-0.782089	1.324863	C	2.500397	1.410604	0.165433	C	-4.787491	0.200154	-0.576013
C	-3.616443	-1.970703	-0.248691	C	4.833943	0.455612	0.652570	C	-2.726985	2.590918	-0.229784
C	-2.390550	-1.542901	2.593748	C	2.556789	1.616922	-1.353679	H	-2.175560	-1.846519	1.702393
H	-1.661202	-1.080528	-1.777333	H	3.766222	-0.684011	-1.650459	H	-3.291781	-2.069449	-0.760751
H	-3.680792	0.422225	-1.490780	H	3.364280	-1.909264	0.889484	H	-2.863086	0.205574	-1.545327
H	-3.992925	-0.158920	0.837220	H	3.059525	0.194346	1.820034	H	-3.054634	1.175212	1.362360
H	-1.155836	-1.308879	0.836832	H	1.452089	1.298850	0.474433	H	-1.540260	0.947981	0.491588
H	-1.628968	0.217700	1.617600	H	2.856393	2.324287	0.661475	H	-5.099632	1.167981	-0.983428
H	-4.057195	-2.585512	0.543709	H	5.112423	1.386150	1.162787	H	-5.221607	-0.581749	-1.213325
H	-4.383682	-1.834643	-1.022405	H	5.386047	-0.367021	1.127030	H	-5.238181	0.115397	0.426002
H	-2.786073	-2.551542	-0.677647	H	5.182422	0.536966	-0.385602	H	-2.204416	3.317114	0.405282
H	-2.646988	-2.587220	2.377388	H	2.039035	2.543423	-1.628773	H	-2.280212	2.632390	-1.232493
H	-1.565296	-1.551954	3.317254	H	3.588264	1.707808	-1.720183	H	-3.771203	2.916753	-0.308222
H	-3.257148	-1.071888	3.078111	H	2.063468	0.802112	-1.900925	O	-0.710305	-1.803418	0.903484
O	-3.694464	2.321499	0.118485	O	0.853155	-1.709624	-0.872359	H	0.810887	-1.298359	-1.048799
H	-3.501725	3.207137	0.470720	H	1.646991	-1.771332	-1.495275	C	2.709724	-0.224776	-0.791673
H	0.654284	1.587638	-0.163460	H	-0.868341	-1.288074	0.968803	C	3.978589	-0.535054	-0.002281
C	1.972527	0.120696	-0.912447	C	-2.831244	-0.364143	0.732170	O	4.101918	-1.508221	0.707454
C	1.874600	-1.395484	-1.034726	C	-3.927787	-0.605833	-0.301984	C	2.046979	1.110394	-0.369332
O	0.958886	-2.046736	-0.579789	O	-3.817321	-1.396252	-1.213495	C	1.506292	1.042758	1.072577
C	3.241583	0.599266	-0.167531	C	-2.199093	1.044733	0.609552	C	0.998292	1.561479	-1.396502

C	3.329716	0.007255	1.252700	C	-1.427968	1.208688	-0.715695	C	1.037973	2.398290	1.611620
C	3.330493	2.132143	-0.185892	C	-1.364490	1.382480	1.853303	H	1.613492	-1.579523	0.378706
C	4.673031	0.273317	1.940928	C	-1.002900	2.653505	-0.996139	H	2.973172	-0.175695	-1.855857
H	0.628528	0.167547	0.683677	H	-1.658202	-1.598689	-0.473942	H	2.868776	1.839617	-0.387210
H	1.968306	0.541993	-1.926611	H	-3.257153	-0.488146	1.734811	H	0.674548	0.324747	1.126491
H	4.078012	0.198336	-0.756475	H	-3.060695	1.726765	0.586415	H	2.293613	0.652578	1.736043
H	2.521907	0.420240	1.880118	H	-0.535711	0.565473	-0.710018	H	0.652561	2.574459	-1.163462
H	3.166652	-1.080299	1.206601	H	-2.055799	0.858414	-1.548838	H	1.425388	1.589027	-2.407536
H	4.314139	2.463001	0.163388	H	-1.039336	2.427750	1.817347	H	0.125842	0.898144	-1.413203
H	3.192670	2.533217	-1.199342	H	-1.955913	1.258008	2.769832	H	1.834860	3.150706	1.536015
H	2.585289	2.595686	0.478634	H	-0.467529	0.756302	1.933513	H	0.165630	2.776159	1.063633
H	4.830446	1.341963	2.132807	H	-0.550052	2.731290	-1.991926	H	0.755898	2.313837	2.668521
H	4.717673	-0.245094	2.906564	H	-1.867757	3.329904	-0.966765	O	4.919678	0.397897	-0.202350
H	5.506287	-0.090778	1.323987	H	-0.266357	3.012222	-0.266960	H	5.710624	0.158629	0.311110
O	2.918524	-1.897069	-1.704694	O	-4.991083	0.170678	-0.071847	N	1.796911	-1.399169	-0.615093
H	2.825032	-2.863939	-1.760037	H	-5.662913	0.000617	-0.755166	H	2.228913	-2.252930	-0.973923
N	0.714160	0.571374	-0.250735	N	-1.820330	-1.460429	0.531275	N	-3.156586	-1.654777	1.323027
H	-0.181960	0.277259	-0.794759	H	-2.172191	-2.354461	0.880292	H	-3.589102	-0.873477	1.818565

Table S3. The *xyz*-coordinates of **FA**, **FB**, and **FZ** shown in Figure 3.

FA				**FB**				**FZ**			
H	1.589979	1.661074	-0.388533	H	4.133871	-2.215045	-1.467193	H	-2.072363	-1.614455	1.904993
N	0.640634	1.533376	-0.737484	N	3.532902	-2.695021	-0.804684	C	-2.372370	-2.361091	0.071845
C	-0.207305	2.680780	-0.382062	C	3.393959	-1.959916	0.468273	C	-1.032158	-1.578775	0.038515
C	-0.418968	2.677932	1.128244	C	1.901391	-1.861269	0.793313	O	-0.431775	-1.562061	-1.067580
O	-0.102412	1.776433	1.878826	O	1.476664	-1.333429	1.814106	C	-3.386820	-1.902276	-0.995216
C	-1.561680	2.618432	-1.143197	C	4.002965	-0.533008	0.473048	C	-3.943270	-0.528584	-0.684762
C	-2.266440	1.290662	-0.979537	C	3.171837	0.559463	-0.178599	C	-5.212341	-0.390890	-0.096546
C	-3.046320	1.029470	0.158989	C	2.591977	0.412633	-1.449828	C	-3.169462	0.623133	-0.902769
C	-2.104126	0.270271	-1.929254	C	2.993111	1.778865	0.495057	C	-5.699487	0.871446	0.267121
C	-3.623364	-0.229821	0.359010	C	1.861069	1.454037	-2.031637	C	-3.653973	1.881272	-0.540444
C	-2.684615	-0.988944	-1.739053	C	2.270687	2.826126	-0.086648	C	-4.918698	2.009102	0.047179
C	-3.438422	-1.245882	-0.587919	C	1.700106	2.666449	-1.354082	H	-3.712733	-1.511008	1.461807
H	0.719747	1.460544	-1.748904	H	3.917943	-3.624659	-0.666057	H	-2.148618	-3.428122	-0.040739
H	0.262142	3.646852	-0.626544	H	3.854099	-2.521525	1.292605	H	-2.849431	-1.912634	-1.950636
H	-2.189071	3.444506	-0.788671	H	4.175887	-0.258341	1.519315	H	-4.199336	-2.637861	-1.061434
H	-1.350968	2.805400	-2.204864	H	4.995693	-0.593970	0.002199	H	-5.844240	-1.270984	0.044726
H	-3.205493	1.816487	0.896442	H	2.709396	-0.515063	-2.007682	H	-2.184785	0.539890	-1.359296
H	-1.527218	0.463864	-2.834644	H	3.433233	1.915805	1.483237	H	-6.692151	0.963938	0.707680
H	-4.223905	-0.417620	1.248005	H	1.429186	1.320935	-3.023298	H	-3.039735	2.762116	-0.725027
H	-2.560664	-1.763959	-2.495407	H	2.159766	3.768580	0.449231	H	-5.299145	2.992973	0.321413
H	-3.889615	-2.224384	-0.431343	H	1.138820	3.480508	-1.810694	O	-0.680960	-1.040882	1.118748
O	-1.010973	3.811064	1.538202	O	1.093633	-2.352460	-0.133780	H	0.738254	-0.645877	-1.070537
H	-1.120680	3.775234	2.503624	H	1.722936	-2.689111	-0.836695	C	1.658273	0.783910	0.247371
H	-1.256390	-0.743540	0.616400	H	-1.708959	-0.731642	1.526671	C	0.466236	1.718634	0.354291
C	0.300969	-2.040671	0.115571	C	-0.774503	0.843133	0.532679	O	-0.220142	2.042438	-0.591984
C	-0.574477	-3.164556	0.654626	C	-1.556913	2.095497	0.917281	C	2.996896	1.543946	0.388865
O	-1.419155	-3.007923	1.505436	O	-2.004018	2.276773	2.028859	C	4.174275	0.615573	0.183868
C	1.778944	-2.231643	0.521133	C	-1.209899	0.173719	-0.790261	C	4.546770	-0.289006	1.191385
C	2.645530	-1.063216	0.104280	C	-2.568280	-0.482961	-0.680592	C	4.875445	0.603138	-1.031932
C	3.095635	-0.132478	1.049836	C	-3.747606	0.264804	-0.826671	C	5.598318	-1.184996	0.987188
C	2.983937	-0.878796	-1.245895	C	-2.664879	-1.849793	-0.365736	C	5.928814	-0.296059	-1.239259
C	3.868626	0.967149	0.656537	C	-4.995786	-0.339447	-0.658033	C	6.289999	-1.192285	-0.230338
C	3.754204	0.216902	-1.643204	C	-3.915803	-2.454842	-0.193482	H	1.338490	0.780845	-1.804583
C	4.197594	1.145213	-0.691401	C	-5.082404	-1.699042	-0.336374	H	1.564320	0.013287	1.018137
H	0.041968	-0.568790	1.597464	H	-0.028178	-0.689851	1.794405	H	3.024242	2.368715	-0.337453
H	0.208101	-2.012368	-0.976241	H	0.281715	1.132741	0.465473	H	3.011779	1.987986	1.391661
H	1.829542	-2.373624	1.610020	H	-1.194641	0.951021	-1.560751	H	4.018948	-0.283985	2.146678
H	2.123631	-3.163403	0.056967	H	-0.444770	-0.564387	-1.051339	H	4.617379	1.318870	-1.815259
H	2.851852	-0.266003	2.104650	H	-3.688698	1.321489	-1.089114	H	5.885578	-1.874079	1.781348
H	2.658294	-1.606484	-1.991314	H	-1.756355	-2.450352	-0.283422	H	6.472039	-0.286690	-2.184139
H	4.223353	1.675092	1.404966	H	-5.904317	0.249319	-0.783831	H	7.113685	-1.888559	-0.386529
H	4.024820	0.339306	-2.691845	H	-3.977431	-3.517628	0.040006	O	0.313493	2.171810	1.603821
H	4.809850	1.992554	-0.998347	H	-6.057357	-2.168649	-0.208145	H	-0.436849	2.789371	1.620374
O	-0.266426	-4.332351	0.070853	O	-1.642273	2.967663	-0.094062	N	1.557352	0.096761	-1.074241

H	-0.820913	-5.032878	0.455899	H	-2.106980	3.761443	0.223174	H	2.447494	-0.357251	-1.299219
N	-0.229477	-0.748773	0.628358	N	-0.905252	-0.108586	1.685277	N	-2.894383	-2.134687	1.475400
H	0.116934	0.097320	0.059350	H	-1.096623	0.428573	2.540531	H	-3.117213	-2.988598	1.985703

Table S4. The *xyz*-coordinates of **YA**, **YB**, and **YZ** shown in Figure 3.

YA				**YB**				**YZ**			
H	2.098772	-1.078798	0.145951	H	3.796186	4.609503	0.271557	H	2.848178	1.932667	1.514609
N	1.191659	-1.270727	0.570264	N	2.948853	4.323035	-0.209971	C	1.440376	3.161279	0.539508
C	0.757346	-2.647260	0.289109	C	2.932826	2.882767	-0.502921	C	0.075351	2.414175	0.543267
C	0.459368	-2.766199	-1.201120	C	1.485711	2.399417	-0.378487	O	-0.652104	2.537251	-0.470689
O	0.396489	-1.836007	-1.980567	O	1.090971	1.357650	-0.886594	C	2.282555	2.970431	-0.734274
C	-0.490104	-3.016891	1.143018	C	3.837062	2.131240	0.515510	C	2.720862	1.534656	-0.916135
C	-1.606362	-2.003585	1.034362	C	3.698051	0.627969	0.492584	C	1.861169	0.596431	-1.511832
C	-1.731634	-0.964688	1.970854	C	3.043418	-0.041349	1.534499	C	3.963298	1.087054	-0.439442
C	-2.507613	-2.027252	-0.040635	C	4.179160	-0.127177	-0.585684	C	2.219176	-0.743706	-1.608680
C	-2.696285	0.033808	1.833529	C	2.844284	-1.421956	1.495948	C	4.324787	-0.264155	-0.506854
C	-3.474460	-1.028503	-0.203170	C	3.976114	-1.506156	-0.648737	C	3.438970	-1.179867	-1.079276
C	-3.568716	0.011153	0.735978	C	3.294410	-2.144398	0.390237	O	3.691702	-2.531660	-1.138417
O	-4.472263	1.024914	0.643067	O	2.994971	-3.497616	0.364751	H	1.325375	2.125134	2.243376
H	1.312096	-1.139910	1.571643	H	2.861737	4.891231	-1.048365	H	1.236619	4.225847	0.707699
H	1.538582	-3.392095	0.508788	H	3.259586	2.635564	-1.522451	H	1.651864	3.300312	-1.567909
H	-0.827054	-4.013603	0.834811	H	4.872319	2.427372	0.292244	H	3.152871	3.639793	-0.693088
H	-0.158873	-3.099724	2.187106	H	3.598013	2.510588	1.519150	H	0.900749	0.927099	-1.904297
H	-1.065855	-0.935778	2.834409	H	2.682066	0.520280	2.396812	H	4.684298	1.802267	-0.035993
H	-2.455906	-2.832507	-0.773624	H	4.706960	0.365025	-1.403322	H	1.560912	-1.463730	-2.086935
H	-2.799690	0.827300	2.572271	H	2.341171	-1.941991	2.308931	H	5.290897	-0.594247	-0.123146
H	-4.162368	-1.064038	-1.048355	H	4.333884	-2.076428	-1.506764	H	4.610490	-2.715865	-0.901422
H	-4.998631	0.941250	-0.162833	H	3.387125	-3.912546	-0.415166	O	-0.171446	1.764032	1.599977
O	0.264051	-4.046000	-1.557603	O	0.695531	3.176875	0.343866	H	-1.974384	1.691224	-0.108719
H	0.088125	-4.080779	-2.513159	H	1.249650	3.980072	0.538086	C	-3.300840	0.002693	-0.460286
H	0.163671	-0.120374	-0.201014	H	-1.945328	0.222667	-1.294178	C	-4.696675	0.528329	-0.766252
C	-0.534195	1.910980	-0.234881	C	-1.070982	-1.271312	-0.130560	O	-5.221986	1.423401	-0.140936
C	-1.803567	2.579600	-0.744973	C	0.058424	-2.251248	-0.422756	C	-3.362548	-1.381526	0.237999
O	-2.519767	2.100171	-1.594828	O	0.615638	-2.293615	-1.501517	C	-1.975810	-1.901726	0.534625
C	0.736472	2.684193	-0.652931	C	-2.448888	-1.973677	-0.094800	C	-1.373406	-1.703069	1.787189
C	2.000395	1.952365	-0.260391	C	-3.576083	-0.976331	0.044059	C	-1.226267	-2.537403	-0.467253
C	2.746454	1.236180	-1.207801	C	-3.869501	-0.390689	1.285430	C	-0.049545	-2.075455	2.022988
C	2.433612	1.937453	1.073623	C	-4.323299	-0.568913	-1.074949	C	0.090031	-2.933396	-0.239725
C	3.887163	0.518661	-0.843151	C	-4.872539	0.569829	1.410354	C	0.696344	-2.670882	0.999161
C	3.572361	1.226990	1.457198	C	-5.327156	0.394351	-0.967396	O	1.995436	-2.980132	1.262314
C	4.305930	0.514446	0.494909	C	-5.604166	0.968373	0.280096	H	-1.995880	0.589793	1.094491
O	5.425568	-0.202993	0.790602	O	-6.596038	1.898423	0.325167	H	-2.728325	-0.082847	-1.390245
H	-0.196583	0.470497	-1.730958	H	-0.775493	-0.688993	-2.094238	H	-3.953505	-1.283715	1.160078
H	-0.589721	1.839531	0.856767	H	-0.880603	-0.761993	0.819863	H	-3.907184	-2.057896	-0.431798
H	0.710508	2.843840	-1.740397	H	-2.568358	-2.566602	-1.012780	H	-1.946679	-1.258682	2.603536
H	0.683965	3.672395	-0.180562	H	-2.424644	-2.677574	0.745906	H	-1.680246	-2.738646	-1.439020
H	2.443901	1.241631	-2.255991	H	-3.317647	-0.697525	2.175500	H	0.416917	-1.914413	2.993249
H	1.884256	2.500449	1.830065	H	-4.144049	-1.034685	-2.046428	H	0.652524	-3.448758	-1.018187
H	4.476420	-0.020899	-1.582664	H	-5.095855	1.003646	2.386327	H	2.458633	-3.220627	0.444623
H	3.904638	1.243330	2.496457	H	-5.916183	0.698059	-1.831267	O	-5.262718	-0.149591	-1.771888
H	5.682819	-0.078155	1.713306	H	-6.736174	2.210468	1.228430	H	-6.164101	0.188655	-1.912556
O	-1.999381	3.757570	-0.138881	O	0.307085	-3.032280	0.620368	N	-2.631480	1.015072	0.406644
H	-2.815237	4.158563	-0.484770	H	1.150596	-3.522060	0.463407	H	-3.348093	1.563451	0.897906
N	-0.499965	0.518660	-0.756229	N	-1.029140	-0.232214	-1.209031	N	2.138987	2.630258	1.772897
H	-1.434268	0.092398	-0.730966	H	-0.279281	0.490229	-1.021116	H	2.557482	3.346279	2.365700

2.4 Other low-energy structures

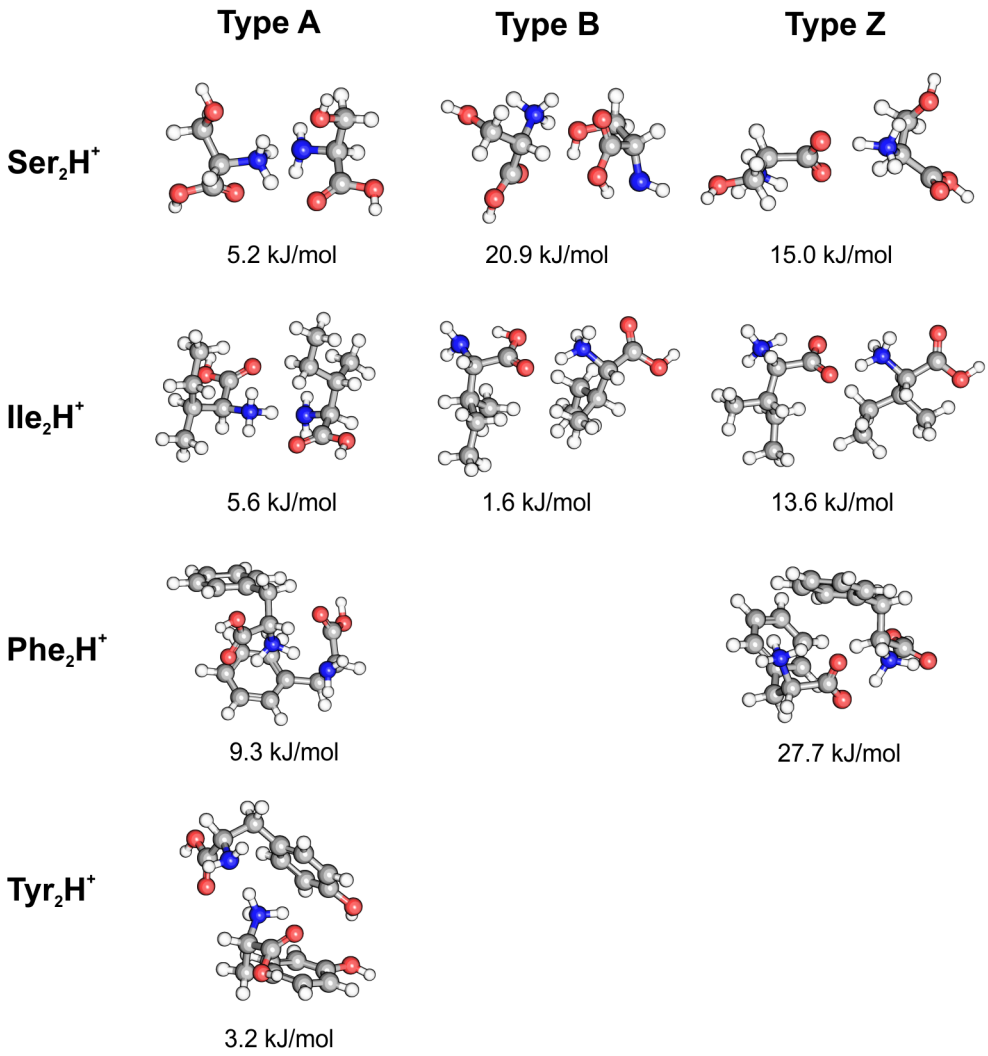

Figure S4. Several more low-energy structures of protonated dimers. The second stable structure per each of type is shown for each protonated amino acid dimer. The ΔG_{298K} values relative to the lowest-energy structure are given together. In case of Phe_2H^+ type B, Tyr_2H^+ type B, and Tyr_2H^+ type Z, only one structure which is shown in Figure 3 was found from conformational search with 30 kJ/mol energy window and thus, no second stable structure is shown here.

2.5 Vibrational Mode Assignment

Table S5. Vibrational modes assigned to the infrared bands of Ser_2H^+

IR bands (in cm^{-1})		assigned vibrational mode
observed	calculated (**SA**)	
1775	1769	ν(C=O)
1710	1734	ν(C=O)
1592	1605	$\delta_{scis}(NH_3^+)$
	1589	$\delta_{scis}(NH_3^+)$
	1578	$\delta_{scis}(NH_2)$
1479	1522	$\delta_{umb}(NH_3^+)$
1400	1407	δ(C–O–H) in -COOH + δ(C–O–H) in side chain
	1389	δ(C–O–H) in -COOH + δ(C–O–H) in side chain
1153	1156	ν(C–OH) + δ(C–O–H) in -COOH
	1143	ν(C–OH) + δ(C–O–H) in -COOH
1050	1087	ν(C–OH) in -COOH + δ(C–O–H) in side chain

Table S6. Vibrational modes assigned to the infrared bands of Ile_2H^+

IR bands (in cm^{-1})		assigned vibrational mode
observed	calculated (**IB**)	
1770	1740	ν(C=O)
1706	1690	ν(C=O)
1597	1610	$\delta_{scis}(NH_3^+)$
	1592	$\delta_{scis}(NH_2)$
	1586	$\delta_{scis}(NH_3^+)$
1479	1508	$\delta_{umb}(NH_3^+)$
1400	1401	δ(C–O–H)
1153	1155	ν(C–OH) + δ(C–O–H)

Table S7. Vibrational modes assigned to the infrared bands of Phe_2H^+

IR bands (in cm^{-1})		assigned vibrational mode
observed	calculated (**FA**)	
1772	1748	ν(C=O)
1749	1723	ν(C=O)
1710		
1588	1599	$\delta_{scis}(NH_3^+)$
	1572	$\delta_{scis}(NH_2)$
	1538	$\delta_{scis}(NH_3^+)$
1481	1475	ν(C=C) + δ(C–H) in -C_6H_5 ring
	1429	$\delta_{scis}(CH_2)$
1400	1384	δ(C–O–H)
	1300	δ(C–O–H)
1150	1143	ν(C–OH) + δ(C–O–H)
	1125	ν(C–OH) + δ(C–O–H)

Table S8. Vibrational modes assigned to the infrared bands of Tyr_2H^+

IR bands (in cm^{-1})		assigned vibrational mode
observed	calculated (**YA**)	
1769	1752	ν(C=O)
1743	1732	ν(C=O)
1705		
1600	1607	$\delta_{scis}(NH_3^+)$
	1582	$\delta_{scis}(NH_2)$
1505	1498	ν(C=C) + ν(C–OH) + δ(C–H) in $-C_6H_5OH$ ring
1398	1383	δ(C–O–H) in COOH
1325	1327	δ(C–O–H) in $-C_6H_4OH$ ring
	1316	δ(C–O–H) in $-C_6H_4OH$ ring
1274	1258	ν(C–OH) in $-C_6H_4OH$ ring
1170	1161	δ(C–O–H) in $-C_6H_4OH$ ring
1150	1143	ν(C–OH) + δ(C–O–H) in -COOH
1100	1103	δ(C–H) in $-C_6H_4OH$ ring