PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2019, **21**, 12625

Infrared multiple photon dissociation action spectroscopy of protonated glycine, histidine, lysine, and arginine complexed with 18-crown-6 ether†

Christopher P. McNary,[a] Y.-W. Nei,[b] Philippe Maitre,[c] M. T. Rodgers *[b] and P. B. Armentrout *[a]

Received 22nd April 2019,
Accepted 17th May 2019

DOI: 10.1039/c9cp02265a

rsc.li/pccp

Complexes of 18-crown-6 ether (18C6) with four protonated amino acids (AAs) are examined using infrared multiple photon dissociation (IRMPD) action spectroscopy utilizing light generated by the infrared free electron laser at the Centre Laser Infrarouge d'Orsay (CLIO). The AAs examined in this work include glycine (Gly) and the three basic AAs: histidine (His), lysine (Lys), and arginine (Arg). To identify the $(AA)H^+(18C6)$ conformations present in the experimental studies, the measured IRMPD spectra are compared to spectra calculated at the B3LYP/6-311+G(d,p) level of theory. Relative energies of various conformers and isomers are provided by single point energy calculations carried out at the B3LYP, B3P86, M06, and MP2(full) levels using the 6-311+G(2p,2d) basis set. The comparisons between the IRMPD and theoretical IR spectra indicate that 18C6 binds to Gly and His *via* the protonated backbone amino group, whereas protonated Lys prefers binding *via* the protonated side-chain amino group. Results for Arg are less definitive with strong evidence for binding to the protonated guanidino side chain (the calculated ground conformer at most levels of theory), but contributions from backbone binding to a zwitterionic structure are likely.

Introduction

Crown ethers are macrocyclic oligomers, commonly with a repeating $-CH_2CH_2O-$ unit. They have high binding affinities for cationic species on the basis of favorable electrostatic interactions of the electron-donor oxygen sites with cationic electron-acceptor sites. The binding selectivity of crown ethers can be modified by the size of the crown cavity, which affects the coordination shell that oxygen atoms can form with cations.[1] This selectivity has been exploited in the technique called selective noncovalent adduct protein probing (SNAPP), which has become a useful method for exploring protein structure and folding states in the liquid phase.[2–10] SNAPP relies on the selective binding of crown ethers to basic amino acid (AA) residues, in particular lysine (Lys), to facilitate identification and characterization of protein sequence, structure, and conformational changes using mass spectrometry (MS). The number of 18-crown-6 (18C6) ligands that bind to the protein is directly correlated to the protein structure and can be easily determined by the mass shift. Therefore, SNAPP can be used to provide information that is useful in understanding functional behavior in biological systems at the molecular level. 18-Crown-6 is most commonly employed as a protein side-chain tag because of its enzyme-like specificity in its interactions with the protonated lysine side chain, which can form three strong hydrogen bonds with alternate oxygens of 18C6. The extent of 18C6 attachment to a protein is generally determined by the degree of lysine side-chain accessibility. Intramolecular interactions within the protein, such as hydrogen bonds or a salt bridge, of a Lys side chain generally prevent the attachment of 18C6 from occurring. Julian and coworkers applied a site-directed mutagenesis approach, where the Lys residues of a series of ubiquitin mutants were exchanged for asparagine one at a time, as a means to investigate the mechanism of the SNAPP method.[9] They observed that non-interacting Lys residues are more likely to bind 18C6 than those

[a] *Department of Chemistry, University of Utah, 315 S. 1400 E. Room 2020, Salt Lake City, Utah 84112, USA. E-mail: armentrout@chem.utah.edu*
[b] *Department of Chemistry, Wayne State University, Detroit, Michigan 48202, USA. E-mail: mrodgers@chem.wayne.edu*
[c] *Université Paris Sud, Laboratoire de Chimie Physique, UMR8000 CNRS, Faculté des Sciences, Bâtiment 350, 91405 Orsay Cedex, France*

† Electronic supplementary information (ESI) available: Tables S1–S4 list the relative energies at four levels of theory for all conformations located up to ~26, 27, 23, and 15 kJ mol^{-1} for $(Gly)H^+(18C6)$, $(His)H^+(18C6)$, $(Lys)H^+(18C6)$, and $(Arg)H^+(18C6)$, respectively, at 0 K and the B3LYP level of theory. Fig. S1 and S2 show comparisons between IRMPD spectra for $(Gly)H^+(18C6)$ and $(His)H^+(18C6)$ with additional conformers. The *XYZ* coordinates for all of the conformations listed in Tables S1–S4 are also provided. See DOI: 10.1039/c9cp02265a

engaged in hydrogen bonding, and both are more likely to bind 18C6 than Lys residues participating in salt bridges. Interestingly, they also observed complexation of up to six 18C6 ligands although the number of Lys residues was only five in the ubiquitin mutant, indicating that the protonated N-terminus or residues other than Lys must also contribute to the SNAPP distribution.

The use of molecular recognition of crown ethers by various protein sequences and conformations has also been pursued by Schalley and coworkers.[11] They applied molecular recognition between 18C6 and oligolysine peptides to investigate molecular mobility, which has attracted considerable attention in supramolecular chemistry and biochemistry. They utilized H/D exchange methods to investigate whether 18C6 moves along an oligolysine scaffold by hopping from one Lys side chain to another. They observed dynamic motion of 18C6 along the oligolysine chain and suggested that many biologically relevant noncovalently bound complexes may exhibit dynamic behavior that has yet to be recognized. They proposed a mechanism for the motion that proceeds by simultaneous transfer of 18C6 from its ammonium ion binding site to a nearby Lys amino group together with an excess proton. Brodbelt and co-workers[12] have reported the use of an 18C6 derivative chromophore to study fragmentation patterns of peptides. The chromophore facilitates peptide fragmentation by absorbing UV irradiation and transferring it to the peptide by intramolecular vibrational redistribution (IVR) in the gas phase.

Gas-phase threshold collision-induced dissociation (TCID) investigations of metal cation–crown ether complexes have probed the inherent energetics of interactions between the cation and crown ether that stabilize such complexes.[13–18] Smaller alkali cations bind more strongly to the crown ether compared to larger alkali cations, largely an electrostatic effect. For a particular metal cation, larger crown ethers exhibit higher binding energies as a result of the greater number of oxygen atom binding sites. In addition, TCID has been used to obtain thermochemical information regarding the binding between 18C6 and a series of protonated peptidomimetic bases that serve as mimics of the N-terminal amino group and the side chains of the basic AAs in peptides and proteins.[19] This work included isopropylamine (IPA) as a mimic of the N-terminal amino group, *n*-butylamine (NBA) and other primary amines as mimics for the side chain of Lys, imidazole (IMID) and 4-methylimidazole (4MeIMID) as mimics for the side chain of histidine (His), and 1-methylguanidine (MGD) as a mimic for the side chain of arginine (Arg). The measured 18C6 binding affinities for those protonated complexes follow the order: IPA > NBA > IMID > MGD > 4MeIMID, suggesting that binding to the N-terminal amino group may be most favorable, followed by the Lys side chain. The relative binding affinities of the His and Arg side-chain mimics make it unclear whether His or Arg will bind to 18C6 more effectively.

Expanding on the TCID of protonated peptidomimetic bases, another TCID study determined binding affinities between 18C6 and protonated AAs directly.[20] Here the measured 18C6 binding affinities for the protonated AAs followed a similar trend to that of the protonated peptidomimetic bases, glycine (Gly) > alanine (Ala) > Lys > His > Arg. Interestingly, the theoretical ground conformations for Gly, Ala, Arg, and His binding to 18C6 were predicted to follow the motif of backbone amino group binding, whereas Lys prefers binding to 18C6 *via* the protonated side-chain heteroatom. (Although technically, these single amino acids do not have the backbone of a peptide, we use the term here as a succinct designation of non-side-chain functional groups.) In all cases, the binding occurred *via* three nearly equal NH···O hydrogen bonds. There the measured BDEs were explained by the steric interactions between 18C6 and the AA side chains, where Gly and Ala bind the strongest because they possess the smallest side-chain substituents, H and CH_3, and thus experience the least steric repulsion with 18C6. According to theory, binding to the protonated backbone amino group was favored over binding to the protonated side chain of His and Arg, likely a result of the additional steric repulsive interactions from the side chain as well as less than optimal hydrogen bond orientations. However, the preferences for side-chain or backbone binding to His and Arg could not be determined experimentally.[20] In an attempt to further understand this relative preference, acetylated versions (which block that binding site) of those protonated AA–18C6 complexes were examined using TCID.[21] There, the 18C6 binding affinities of the protonated acetylated AAs were found to be N_α-AcLys > N_ε-AcLys > N_α-AcArg > N_α-AcHis, where N_α is the backbone amine and N_ε is the side-chain amine. These results indicate that the Lys side chain is still the preferred binding site for 18C6 among the basic AAs in proteins and peptides. Interestingly, all of these studies concluded that the protonated primary amines are favored binding sites for 18C6 because they form three strong NH···O hydrogen-bonding interactions.[2,20,22,23] Nevertheless, preferred binding preferences for some protonated peptidomimetic bases and AAs remain undetermined.

In the present work, we employ infrared multiple photon dissociation (IRMPD) action spectroscopy coupled with electronic structure theory calculations to potentially determine the binding preferences of 18C6 to the four protonated AAs: H^+(Gly), H^+(His), H^+(Lys), and H^+(Arg).[20] As noted above, Lys, His, and Arg offer the best targets for molecular recognition of specific side chains in peptides and proteins as they are the sites most commonly protonated, whereas Gly is an ideal model for molecular recognition of the protonated N-terminus. The comparison between the calculated IR spectra for several low-lying conformations and the experimental IRMPD spectrum of each (AA)H^+(18C6) complex should provide further insight into the molecular recognition of protonated AAs, and by inference, peptides and proteins by 18C6.

Experimental and theoretical methods

Centre Laser Infrarouge d'Orsay (CLIO)

The protonated AA–18C6 complexes, (AA)H^+(18C6), were prepared by adding 50 μL of the AA and 18C6 stock solutions (1 mM each), and 5 μL of acetic acid to 5 mL of 1 : 1 water/methanol solvent.

The ions were generated by electrospray ionization (ESI) using flow rates of 1–2 μL min^{-1}, spray voltages of 2000–4000 V, drying gas flow of 2–5 L min^{-1}, nebulizer pressure of 1.5 bar, and a drying gas temperature of 200 °C.

Experiments were performed on a modified Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS; Bruker APEX-Qc system).[24] The FT-ICR MS was equipped with an Apollo II ESI ion source, a quadrupole mass filter, a collision/thermalizing cell (hexapole), and a 7 Tesla magnet. The (AA)H^+(18C6) complexes were first mass selected using the quadrupole mass filter and then accumulated in the hexapole filled with Ar for approximately 200–4000 ms. Ions were subsequently accelerated along the axis of the magnetic field, decelerated, and trapped in the ICR cell with a background pressure of $\sim 1.5 \times 10^{-9}$ mbar. IRMPD action spectroscopy was then performed in the ICR cell by focusing tunable IR laser radiation from the infrared free electron laser (IR-FEL) at the Centre Laser Infrarouge d'Orsay (CLIO)[25] with a 2 meter focal mirror.[24]

The CLIO IR-FEL is based on a 10 to 50 MW electron linear accelerator[24] and provides 8 μs long macropulses fired at a repetition rate of 25 Hz. Each macropulse was composed of 500 micropulses, each a few ps long and separated by 16 ns. For a typical IR average power of 500 mW, the corresponding micropulse and macropulse energies are 20 μJ and 40 mJ, respectively. The electron beam energy was set to 44.4 MeV. While scanning the photon energy over the 800–2000 cm^{-1} range, the power would change linearly by a factor of two at the most. The IR-FEL spectral width was adjusted through a tuning of the optical cavity length and found to have a full width at half maximum (fwhm) less than 0.5% of the central wavelength. The irradiation times and IR-FEL beam intensity in the ICR cell were adjusted to assure that no more than 40% depletion of the precursor ion was obtained. For the IR-FEL regions where the precursor ion was depleted more than 40%, shorter irradiation times and/or 3–9 db attenuators were used. All of the experimental scans for each (AA)H^+(18C6) complex were averaged with a common baseline normalization to zero to produce the experimental IRMPD spectra measured in this work. The IR FEL wavelength was monitored online while recording the IRMPD spectrum. For this purpose, a small fraction of the IR beam was used to record the IR absorption spectrum of a polystyrene film. As a result, at each wavelength during the IR FEL scan, polystyrene absorption and a MS^2 mass spectrum were simultaneously recorded. Wavelength corrections can thus be made during the data treatment.

Theoretical calculations

To obtain stable geometries, vibrational frequencies, and energies for the (AA)H^+(18C6) complexes, theoretical calculations were performed using UCSF Chimera,[26] Amber suite,[27] NWChem suite,[28] and the Gaussian 09 Rev. D[29] suite of programs. The (AA)H^+(18C6) complexes exhibit many stable low-energy structural conformations. Therefore, potential low-energy conformations were obtained *via* a 20 000 cycle simulated annealing procedure employing the Amber14SB[27] force field. Briefly, a three-phase molecular dynamic distance-restrained simulated annealing process was used with each cycle beginning and ending at 100 K. Each cycle lasted for 1.25 ps and achieved a maximum simulation temperature of 1600 K. Heating and cooling times for each cycle were 0.85 ps each, allowing 0.4 ps for the ions to sample conformational space at the simulation temperature. The conformations accessed at the end of each annealing cycle were subjected to a geometry optimization minimization and relative energies were computed using molecular mechanics methods every 0.001 ps followed by a quantum mechanics geometry optimization calculation at the HF/6-31G level of theory. Conformations above a relative energy of ~120 kJ mol^{-1} were not included in further calculations. Further optimizations of the low-energy conformations were then performed at the B3LYP/6-31G(d) level[30–32] utilizing the opt = loose (maximum step size of 0.01 a.u. and an RMS force of 0.0017 a.u.)[29] criterion. Of those conformations, only those at low energy (<35 kJ mol^{-1}) were selected for final geometry optimizations and vibrational frequency calculations, which were performed at the B3LYP/6-311+G(d,p) level of theory with the vibrations used in zero-point energy (ZPE) and thermal (298 K) corrections scaled by a factor of 0.989.[33]

For comparison to experimental IRMPD spectra, the calculated vibrational frequencies were scaled by a factor of 0.975 and broadened using a 20 cm^{-1} fwhm Gaussian line shape. This scaling factor and broadening account for the finite laser bandwidth, unresolved rotational structure of the ions (which should be near room temperature), anharmonicity of the vibrational modes, and broadening as a result of the multiple-photon absorption process.[34] The 0.975 scaling factor leads to good agreement between calculated and experimental vibrational peaks, as also shown in previous studies.[35–40] Single-point energy calculations were performed at the B3LYP, B3P86,[30,41] M06,[42] and MP2(full)[43–46] (where full indicates correlation of all electrons) levels of theory using the 6-311+G(2d,2p) basis set with the optimized geometries calculated at the B3LYP/6-311+G(d,p) level of theory. Optimized geometries with empirical dispersion included were also computed for the lowest-lying conformations of the (AA)H^+(18C6) complexes, but this approach was abandoned because the agreement between these theoretical IR spectra and the measured IRMPD spectra degraded.

Results and discussion

IRMPD fragmentation

The primary dissociation pathway for the IRMPD of (Gly)H^+(18C6) is the loss of Gly, where the primary H^+(18C6) product further dissociates *via* elimination of two to four C_2H_4O units from the 18C6 molecule. These results coincide with previous TCID results fairly well,[20] except at high energies, TCID induces competitive loss of 18C6 (forming H^+(Gly)) and the $H^+(C_2H_4O)$ (n = 1) product was also observed. Likewise, the IRMPD and TCID mass spectra of (His)H^+(18C6) and (Lys)H^+(18C6) also agree well, with H^+(His) and H^+(Lys) being the dominant fragments, followed by sequential loss of CO + H_2O from H^+(His) and NH_3 and NH_3 + CO + H_2O

from H^+(Lys). For (Arg)H^+(18C6), the primary dissociation pathway for IRMPD and TCID was loss of 18C6, whereas the sequential losses differ slightly. The loss of water was observed with IRMPD, whereas the loss of ammonia was observed with TCID at high energies, with additional fragments arising from the guanidinium side chain at higher energies. For comparison, Harrison and coworkers examined the CID fragmentation of protonated AAs and found the same fragments for H^+(His) and H^+(Lys), whereas H^+(Arg) fragmented with losses of both NH_3 and H_2O in a 2 : 1 ratio at the lowest energies along with the guanidine fragments.[47]

Nomenclature: 18-crown-6

The various conformations of 18C6 can be identified according to the nomenclature of Hill and Feller,[48] then adopted by El-Azhary and coworkers (although they appear to shift the sequence of angles),[49] by the ∠OCCO, ∠CCOC, and ∠COCC dihedral angles going in the counterclockwise direction, where "+" indicates angles between 0° and 120°, "0" indicates angles from 120° to 240°, and "−" indicates angles between 240° and 360°. For 18C6, Feller and coworkers[50–52] and El-Azhary and coworkers eliminated enantiomers by subjecting each 18C6 structure to the twelve possible perturbations for each conformation namely (−00/+00/−−**0**/−00/+00/−−0) equals (−−**0**/−00/+00/−−0/−00/+00), *etc.* and then inversion of each of these angles (− → + and + → −). In the present work, the conformations of 18C6 in the (AA)H^+(18C6) complexes were re-examined by removing the protonated AA and re-optimizing the free 18C6 conformer to test for stability of the conformation. These optimized conformers of 18C6 were then identified by the naming scheme mentioned above. Comparison of the free 18C6 conformers to those of (AA)H^+(18C6) showed that several conformers of 18C6 are stabilized by intermolecular interactions with the H^+(AA). For comparison to the previous work on free 18C6, we reproduced three higher symmetry conformations: D_{3d}, C_i, and S_6, see Table 1 and Fig. 1. Glendening, Feller, and Thompson find that C_i is 22.6 kJ mol^{-1} more stable than D_{3d} at the MP2/6-31+G(d) level of theory (they did not consider the S_6 conformer),[50] whereas El-Azhary and coworkers find S_6 is their ground conformer, lying 7.7 kJ mol^{-1} below C_i, with D_{3d} another 25.4 kJ mol^{-1} higher at the MP2/6-31+G(d) level. Our own B3LYP/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) calculations (including zero point energy corrections) find the C_i conformer is lowest in energy, with S_6 and D_{3d} lying 6.8 and 3.4 kJ mol^{-1} higher, respectively, at 0 K. The latter results appear to agree better with experimental studies of crystalline 18C6 in which the C_i conformer is observed.[53]

Table 1 Sequences and theoretical relative energies at 0 (298) K of neutral 18C6 conformations

Conformer	Sequence of dihedral angles[a]	Energy (kJ mol^{-1}) This work[b]	Literature
D_{3d}-1	−00/+00/−00/+00/−00/+00	3.4 (3.2)	22.6,[c] 25.4[d]
B-1	−−0/−00/+00/−−0/−00/+00	16.0 (16.6)	
C-1	+00/−00/+0+/+00/−−0/−00	16.4 (16.2)	
D-1	+0+/+00/−00/+00/−−0/−00	16.8 (14.2)	
E-1	−00/+00/+0+/+00/−00/+00	19.0 (17.4)	
C_i	+−0/000/+00/−+0/000/−00	0.0 (0.0)	0.0,[c] 0.0[d]
S_6	+0+/−0−/+0+/−0−/+0+/−0−	6.8 (11.7)	−7.7[d]

[a] The sequence of dihedral angles is designated according to the nomenclature of Hill and Feller[48] by the ∠OCCO, ∠CCOC, and ∠COCC dihedral angles, where + indicates angles between 0° and 120°, 0 indicates angles from 120° to 240°, and − indicates angles between 240° and 360°. [b] B3LYP/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) including zero point energy corrections. [c] MP2/6-31+G(d) calculations of Glendening, Feller, and Thompson.[50] [d] MP2/6-31+G(d) calculations of El-Azhary and coworkers.[49]

Table 1 and Fig. 1 include a select number of free 18C6 conformations that are directly correlated to the bound 18C6 conformations found in the (AA)H^+(18C6) complexes, and include the D_{3d} conformation. These are labeled as D_{3d} (equivalent to A) to E in order of their relative energies at 0 K, with their dihedral angle sequences listed in Table 1. It can be seen that the A–E conformers are all approximately circular, relatively planar, and have no hydrogen atoms pointing inside the crown, thereby allowing a strong interaction of multiple oxygen atoms with the protonated amino acids. These structures also differ from one another in how the oxygen atoms are oriented, with the symmetric D_{3d} conformer having three pointing up (u) and three down (d) in an alternating ududud sequence. Conformers B–E all have four up and two down oxygens; B and D have uuduud sequences, whereas C and E have uuudud sequences. In contrast, the C_i conformer is oblong and planar and the S_6 conformer is circular and nonplanar, and both structures have multiple hydrogen atoms pointing inside the crown; geometries that do not allow particularly strong H^+(AA) binding. Except for the high symmetry species, we were unable to match these conformers with the 47 conformers listed by El-Azhary and coworkers, even though both sets have similar energies relative to D_{3d}. This is presumably because different methodologies were used to construct both sets of free 18C6 conformations. The free 18C6 conformations presented in Table 1 were optimized directly from the (AA)H^+(18C6) complexes, whereas El-Azhary and coworkers[49] located their conformers by using the CONFLEX method and allowing optimization of unconstrained 18C6. Apparently, the intramolecular bonds formed during binding of H^+(AA) to 18C6 constrain the number of conformations that can be found within our low-energy criteria, <35 kJ mol^{-1}, and potentially biases the conformations to those not easily found using the CONFLEX method.

For naming the (AA)H^+(18C6) complexes, the starting sequence of dihedral angles in 18C6 begins at the shortest NH···O hydrogen bond and is indicated by one through six, *e.g.*, B-1 equals −−0/−00/+00/−−0/−00/+00 whereas B-3 equals +00/−−0/−00/+00/−−0/−00. Note, enantiomers for conformers B and E were also observed in this study and have four sequences of inverted ∠**OCCO** dihedral angles and two sequences of inverted ∠OCCO angles coupled with ∠CCOC and ∠COCC dihedral angles that are inverted and swapped, *e.g.*, +**00**/−**00**/+0+/+**00**/−**00**/+0+ equals −**00**/+**00**/−−0/−**00**/+**00**/−−0 (B-2).

Protonated AAs complexed to 18-crown-6

To identify the various conformations of (AA)H^+(18C6), we use bracketed nomenclature specifying the site of protonation on

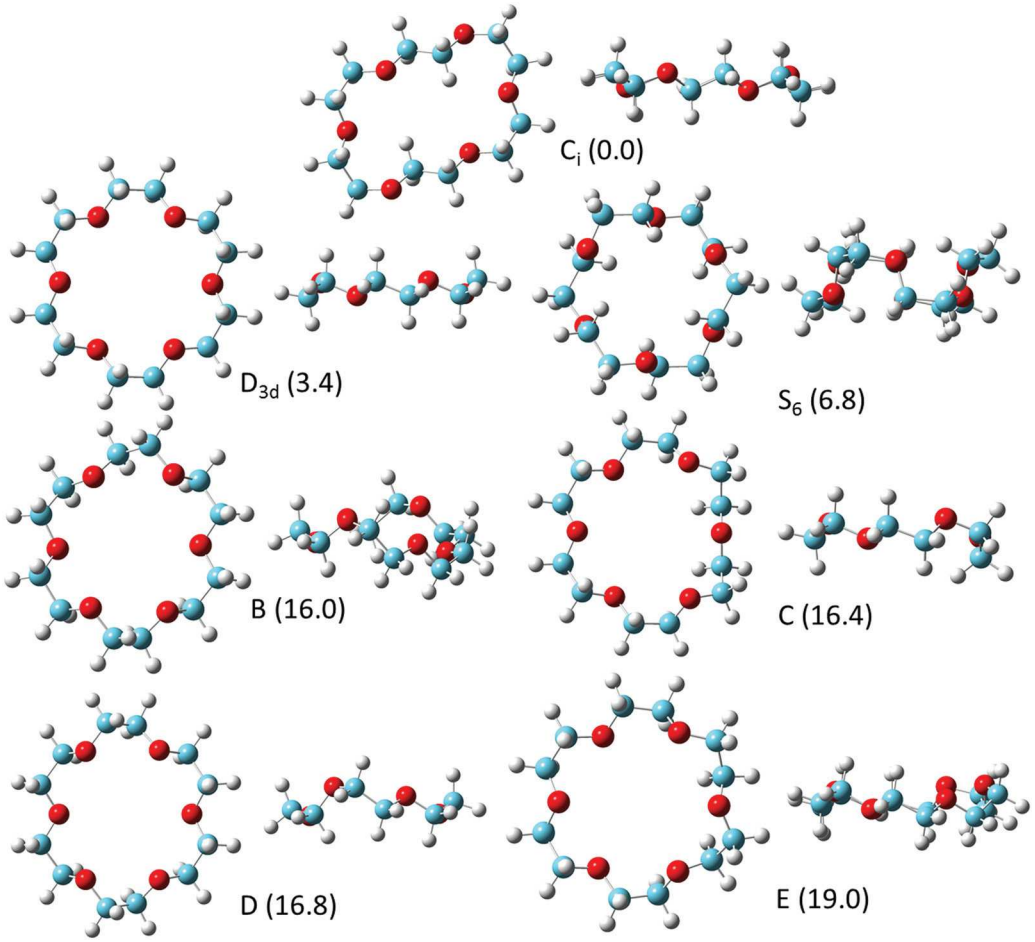

Fig. 1 Structures of the 18C6 conformers listed in Table 1 of the main text. Values in parenthesis are the relative energies (kJ mol^{-1}) at 0 K calculated at the B3LYP/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) level of theory.

the AA. Here, Gly (G) is protonated on the backbone amino group, $[N_\alpha]$, as is possible for all of the other amino acids. In addition, His (H) can be protonated on either nitrogen of the imidazole side chain, $[N_\pi]$ or $[N_\tau]$, where the nitrogens are denoted by pros ("near", abbreviated π and also referred to as N_1) and tele ("far", abbreviated τ or N_3). Lys (K) can be protonated on the amino side-chain group, $[N_\varepsilon]$. Arg (R) can be protonated on the guanidino side-chain group, $[N_\omega]$. In each case, the protonation site is followed by a series of dihedral angles unique to each H^+(AA) species and starting at the carboxylic acid hydrogen. H^+(Gly) uses two dihedral angles that proceed to the N-terminus (*i.e.*, ∠HOCC and ∠OCCN). H^+(His) uses four dihedral angles ending at the imidazole side-chain nitrogen (N_π) (∠HOCC, ∠OCCC, ∠CCCC, and ∠CCCN_π). H^+(Lys) uses six dihedral angles ending at the ε-amino side-chain group (∠HOCC_α, ∠OCC_αC, ∠CC_αCC, ∠C_αCCC, ∠CCCC_ε, and ∠CC$C_\varepsilon N_\varepsilon$). H^+(Arg) uses six dihedral angles ending at the guanidino side-chain carbon (∠HOCC_α, ∠OCC_αC, ∠CC_αCC ∠C_αC$C_\delta N_\delta$, ∠CC$C_\delta N_\delta$C, and ∠C$C_\delta N_\delta$C). These dihedrals are described using *c* (*cis*, for angles between 0–45°), *g* (*gauche*, 45–135°), or *t* (*trans*, 135–180°), and + or − for the *gauche* angles when needed, to distinguish similar conformers. The conformation of the H^+(AA) are then followed by the structure of the 18C6 conformer, *i.e.*, (D-1), (B-2), *etc.*

The most stable 0 and 298 K conformers of the (AA)H^+(18C6) complexes were calculated at the B3LYP, B3P86, M06, and MP2(full) levels with the 6-311+G(2d,2p) basis set using optimized geometries calculated at the B3LYP/6-311+G(d,p) level of theory. As detailed further below, these calculations indicate that 18C6 prefers to bind to the protonated backbone amino group in the Gly and His complexes, whereas the protonated side-chain substituent is preferred for binding for Lys and Arg (except at 298 K, MP2(full) calculations find a backbone binding ground conformer). In most cases, the binding occurs *via* three nearly ideal NH···O hydrogen bonds. Exceptions include Arg, where binding to the side chain forms several NH···O hydrogen bonds with 18C6, and special cases, in which the three NH···O hydrogen bonds are augmented by additional binding between 18C6 and the carboxylic acid backbone group of the AA.

We also considered the temperature of the ions, calculating relative energies at both 0 and 298 K. In general, the 0 K enthalpies and 298 K free energies follow similar orderings although it will be seen below that some changes occur with

temperature. As there are many low-lying conformations, it is certainly the case that mixtures of conformations may be formed under the source conditions used, as discussed more thoroughly in each case below.

Conformations of (Gly)H$^+$(18C6). The four lowest-lying conformations of the (Gly)H$^+$(18C6) complex all have H$^+$(Gly) in a *tt* conformation that orients the carbonyl group on the same side of the backbone as the protonated amino group. These are shown in Fig. 2 and listed in Table 2, along with the lowest-energy complex having the *tc* conformation. The *tt* conformation is also the ground conformer (GC) of isolated H$^+$(Gly) at all levels of theory. B3LYP, M06, and MP2(full) levels predict the H$^+$G[N]-*tt*(D_{3d}-1) conformer to be the GC at 0 and 298 K, whereas B3P86 predicts this conformer to be the GC only at 298 K with the H$^+$G[N]-*tt*(D-1) conformer being the GC at 0 K, but only by 0.01 kJ mol^{-1}. We located a total of 31 H$^+$G[N]-*tt*(18C6) conformations within 25–32 kJ mol^{-1} of the two predicted GCs at the levels of theory explored, all of which are included in Table S1 (ESI†). In addition, we also found two excited conformers involving H$^+$G[N]-*tc* conformers in which the carboxylic acid group has rotated 180°. H$^+$G[N]-*tc*(D_{3d}-1) and H$^+$G[N]-*tc*(D-1) lie 16–18 and 17–24 kJ mol^{-1}, respectively, above the calculated 0 K GC. These excitation energies are comparable to those of excited H$^+$G[N]-*tc* without the crown ether, 18–22 kJ mol^{-1}.[54] In all of the 33 structures, the protonated amino group interacts with 18C6 *via* three nearly ideal NH$\cdots$O hydrogen bonds (*e.g.*, 1.86, 1.86, and 2.02 Å in H$^+$G[N]-*tt*(D_{3d}-1) and 1.86, 1.88, and 1.88 Å in H$^+$G[N]-*tt*(D-1)) with one or two longer range CO$\cdots$HC$_{18C6}$ hydrogen bonds as well ($\sim$2.5 Å).

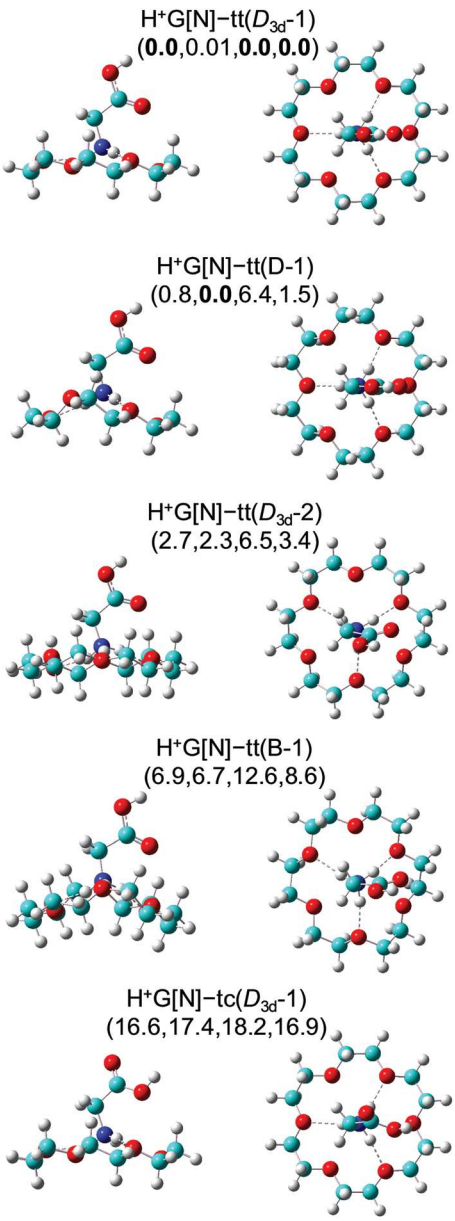

Fig. 2 Optimized conformations calculated at the B3LYP/6-311+G(d,p) level of theory for the (Gly)H$^+$(18C6) complex. Relative energies at 0 K in kJ mol^{-1} are given at the B3LYP, B3P86, M06, and MP2(full) levels.

Table 2 Relative energies in kJ mol^{-1} at 0 K (Gibbs energies at 298 K) of (Gly)H$^+$(18C6) and (His)H$^+$(18C6) complexes[a]

(AA)H$^+$(18C6)	B3LYP		B3P86		M06		MP2(full)	
(Gly)H$^+$(18C6)								
H$^+$G[N]-*tt*(D_{3d}-1)	**0.0**	(**0.0**)	0.01	(**0.0**)	**0.0**	(**0.0**)	**0.0**	(**0.0**)
H$^+$G[N]-*tt*(D-1)	0.8	(3.2)	**0.0**	(2.4)	6.4	(8.9)	1.5	(4.0)
H$^+$G[N]-*tt*(D_{3d}-2)	2.7	(3.2)	2.3	(2.8)	6.5	(7.0)	3.4	(3.9)
H$^+$G[N]-*tt*(B-1)	6.9	(6.7)	6.7	(7.5)	12.6	(12.4)	8.6	(8.4)
H$^+$G[N]-*tc*(D_{3d}-1)	16.6	(14.2)	17.4	(15.0)	18.2	(15.8)	16.9	(14.5)
(His)H$^+$(18C6)								
H$^+$H[N$_\alpha$]-*cggc*(B-2)	**0.0**	(**0.0**)	**0.0**	(**0.0**)	0.4	(2.4)	**0.0**	(0.6)
H$^+$H[N$_\alpha$]-*cggc*(D_{3d}-1)	4.1	(2.0)	4.0	(1.9)	**0.0**	(**0.0**)	1.5	(**0.0**)
H$^+$H[N$_\alpha$]-*cggc*(D-5)	2.5	(4.3)	2.2	(3.9)	2.9	(6.7)	2.8	(5.1)
H$^+$H[N$_\alpha$]-*cggc*(D-2)	5.7	(1.9)	5.8	(2.0)	5.5	(3.7)	9.0	(5.8)
H$^+$H[N$_\pi$]-*tgg_g_*(B-2)	18.6	(11.0)	26.0	(18.4)	27.5	(22.0)	36.2	(29.1)

[a] Ground conformations in bold. All values calculated at the level of theory indicated using the 6-311+G(2d,2p) basis set with geometries, zero-point energies, and thermal energy corrections calculated at the B3LYP/6-311+G(d,p) level of theory.

Conformations of (His)H$^+$(18C6). For (His)H$^+$(18C6), the proton is predicted to preferentially bind to the backbone amino group of His, forming H$^+$H[N$_\alpha$], which can then bind to 18C6 *via* three NH$\cdots$O hydrogen bonds. The low-lying structures are shown in Fig. 3 and listed in Table 2. The four lowest-energy structures all have a H$^+$H[N$_\alpha$]-*cggc* moiety that is stabilized by an intramolecular OH$\cdots$N$_\pi$ hydrogen bond between the backbone hydroxyl hydrogen and side-chain imine nitrogen. In contrast, the isolated H$^+$(His) molecule prefers to be protonated on the side chain, a position stabilized by hydrogen bonds to either the N$_\alpha$ amine (preferred) or carbonyl oxygen.[55] B3LYP and B3P86 levels predict H$^+$H[N$_\alpha$]-*cggc*(B-2) to be the GC at 0 and 298 K, whereas M06 predicts H$^+$H[N$_\alpha$]-*cggc*(D_{3d}-1) is the GC at both temperatures. In contrast, the MP2(full) level predicts H$^+$H[N$_\alpha$]-*cggc*(B-2) to be the GC at 0 K and H$^+$H[N$_\alpha$]-*cggc*(D_{3d}-1) to be the GC at 298 K. The differences in relative energies between these two conformers are 0.6–2.4 kJ mol^{-1} at 298 K, which indicates that population of both conformations is possible at this temperature. Overall, 33 conformations of His protonated on the backbone amine complexed to 18C6 were located within about 18 kJ mol^{-1} of the GC, Table S2 (ESI†).

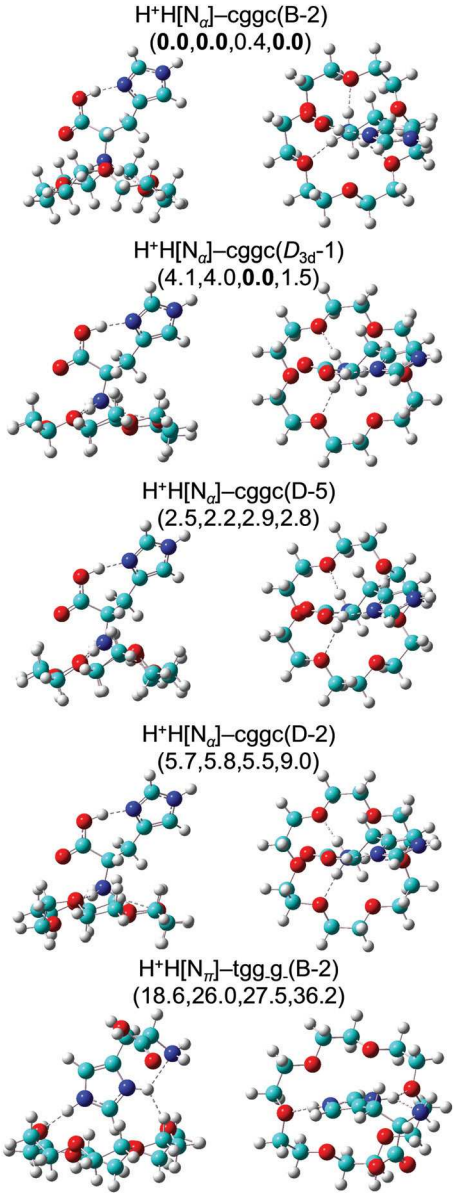

Fig. 3 Optimized conformations calculated at the B3LYP/6-311+G(d,p) level of theory for the (His)H$^+$(18C6) complex. Relative energies at 0 K in kJ mol^{-1} are given at the B3LYP, B3P86, M06, and MP2(full) levels.

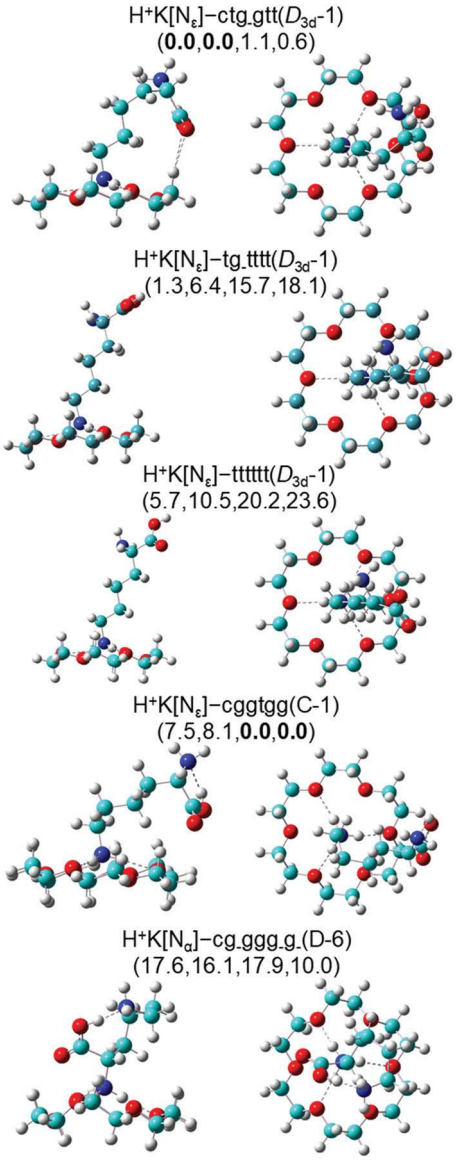

Fig. 4 Optimized conformations calculated at the B3LYP/6-311+G(d,p) level of theory for the (Lys)H$^+$(18C6) complex. Relative energies at 0 K in kJ mol^{-1} are given at the B3LYP, B3P86, M06, and MP2(full) levels.

Six stable conformations containing His protonated on the side chain were also found (Table S2, ESI†) where H$^+$H[N$_\pi$] binds to 18C6 *via* N$_\pi$H$\cdots$O and N$_\tau$H$\cdots$O hydrogen bonds. The lowest energy protonated side-chain conformer, H$^+$H[N$_\pi$]-*tgg_g_*(B-2), is calculated to be 18–36 kJ mol^{-1} less stable at 0 K (11–29 kJ mol^{-1} at 298 K) than the GC at the four levels of theory explored, Table 2. Attempts were made to optimize a salt-bridge conformation, in which the imidazole side chain and amino group are both protonated and the carboxylic acid group is deprotonated; however, these conformations would always optimize to one of the charge-solvated (non-zwitterionic) conformations.

Conformations of (Lys)H$^+$(18C6). Low-energy conformers of (Lys)H$^+$(18C6) are shown in Fig. 4 with relative energies listed in Table 3. Three of the lowest-energy structures involve binding Lys protonated on its side-chain amino group (N$_\varepsilon$) to the D_{3d}-1 conformer of 18C6 and differ only in the side-chain conformation. The B3P86 level of theory predicts H$^+$K[N$_\varepsilon$]-*ctg_gtt*(D_{3d}-1) to be the GC at 0 and 298 K, whereas B3LYP predicts this conformer to be the GC at 0 K, and M06 and MP2(full) theory find this conformer is low-lying ($\leq$1.1 kJ mol^{-1}). B3LYP yields H$^+$K[N$_\varepsilon$]-*tg_tttt*(D_{3d}-1) as the GC at 298 K, with B3P86 also finding that this structure is low in energy, but M06 and MP2(full) find that it is relatively high in energy (above 10 kJ mol^{-1}). The alternative H$^+$K[N$_\varepsilon$]-*tttttt*(D_{3d}-1) structure is also potentially low-lying, 0.6–1.7 kJ mol^{-1} at 298 K above the GC at the B3LYP and B3P86 levels, but 11–14 kJ mol^{-1} at the M06 and MP2(full) levels. These three H$^+$K[N$_\varepsilon$](D_{3d}-1) conformers lie within 1.7–14 kJ mol^{-1} of one another at 298 K, which could indicate

population of all three species. Meanwhile, M06 and MP2(full) predict $H^+K[N_\varepsilon]$-*cggtgg*(C-1) to be the GC at 0 and 298 K, with B3LYP and B3P86 indicating this lies over 8 kJ mol^{-1} above their respective GCs at 298 K. The differences in relative energy between these conformers are such that all levels indicate that $H^+K[N_\varepsilon]$-*ctg_gtt*(D_{3d}-1) should be populated at 298 K, with $H^+K[N_\varepsilon]$-*tg_tttt*(D_{3d}-1) and $H^+K[N_\varepsilon]$-*tttttt*(D_{3d}-1) also populated according to B3LYP and B3P86 and $H^+K[N_\varepsilon]$-*cggtgg*(C-1) populated according to M06 and MP2(full) levels.

In all of these conformers, $H^+K[N_\varepsilon]$ binds to 18C6 *via* three NH$\cdots$O hydrogen bonds, Fig. 4. $H^+K[N_\varepsilon]$-*tg_tttt*(D_{3d}-1) is further stabilized by an intramolecular $N_\alpha H\cdots OC$ hydrogen bond between a backbone amino hydrogen and the carbonyl oxygen, and exhibits an extended lysine conformation, resulting in the protonated amino group of the side chain interacting with 18C6 *via* three nearly ideal NH$\cdots$O hydrogen bonds (1.91–1.94 Å). The $H^+K[N_\varepsilon]$-*tttttt*(D_{3d}-1) is very similar but has a $N_\alpha H\cdots OH$ intramolecular H-bond, leading to its slightly higher energy. The other predicted GCs, $H^+K[N_\varepsilon]$-*ctg_gtt*(D_{3d}-1) and $H^+K[N_\varepsilon]$-*cggtgg*(C-1), have the H^+(Lys) moiety stabilized by an intramolecular $OH\cdots N_\alpha$ hydrogen bond between the backbone hydroxyl and amino groups, with the former also exhibiting intermolecular $CO\cdots HC_{18C6}$ and $C(H)O\cdots HC_{18C6}$ hydrogen bonds (2.50 and 2.66 Å, respectively) to the oxygens of the carboxylic acid. These conformations exhibit a bent conformation of the AA, resulting in the protonated side-chain amino group interacting with 18C6 *via* three NH$\cdots$O hydrogen bonds showing slightly more variability (1.88–1.97 Å). In addition to these conformers, another 34 excited conformations with 18C6 bound to the protonated side chain of Lys within 28–38 kJ mol^{-1} of the GC were located at the levels of theory explored, Table S3 (ESI†).

Three excited conformations where 18C6 binds to the protonated backbone amino group of lysine were also found, Table S3 (ESI†). The most stable of these according to B3LYP, $H^+K[N_\alpha]$-*cg_ggg_g_*(D-6), is located 10–18 kJ mol^{-1} higher in energy at 0 K (15–26 kJ mol^{-1} at 298 K) than the side-chain-bound GC, Table 3 and Fig. 4. These relative energies match those previously calculated,[20] where 18C6 preferred binding to the side chain of H^+(Lys) over the backbone by 17 kJ mol^{-1} at the B3LYP level of theory. In isolated H^+K, the lowest energy conformer has a protonated side-chain amine that is stabilized by hydrogen bonds to both the carbonyl oxygen and amino nitrogen groups of the backbone.[56] Clearly, these two stabilizing hydrogen bonds are much less favorable than the three formed with 18C6.

It can also be noted that the excitation energies for the isolated $H^+K[N_\varepsilon]$-*tg_tttt* conformer are 65–79 kJ mol^{-1} above the GC, $H^+K[N_\varepsilon]$-*tg_gg_gg_*, at the four levels of theory explored.[20,57–59] The unbound $H^+K[N_\varepsilon]$-*ctg_gtt* and $H^+K[N_\varepsilon]$-*cggtgg* conformers are not stable and would always optimize to a $H^+K[N_\varepsilon]$-*ctgtgg* conformation, indicating that binding to 18C6 is needed to stabilize these two conformations.

In addition, several attempts were made to calculate zwitterionic salt-bridge structures for (Lys)H^+(18C6), where both amino groups of the side chain and backbone are protonated, while the carboxylic acid group is deprotonated. The zwitterion salt bridge structure would enable 18C6 to bind to either the protonated side-chain or backbone amino groups. These calculations would always optimize to one of the low-lying charge-solvated conformations located. This seems reasonable, given that the isolated H^+(Lys) zwitterionic complexes are $\sim$39 kJ mol^{-1} less stable than the protonated side-chain GC.[57]

Conformations of (Arg)H^+(18C6). For (Arg)H^+(18C6), the B3LYP, B3P86 and MP2(full) levels all predict $H^+R[N_\omega]$-*tcg_gg_t*(B-3) to be the GC at 0 K, as shown in Fig. 5 and listed in Table 3. M06 predicts $H^+R[N_\omega]$-*tcg_g_gg*(E-2) to be the GC at 0 K. In both structures, there are three $N_\omega H\cdots O$ hydrogen bonds between the guanidinium side chain and 18C6, two additional $O\cdots HC_{18C6}$ hydrogen bonds between the carboxylic acid group and 18C6, and an intramolecular $N_\delta H\cdots N_\alpha$ hydrogen bond. The B3LYP and B3P86 levels predict the $H^+R[N_\omega]$-*cggg_tg_*(B-2) conformer is the GC at 298 K, whereas M06 predicts $H^+R[N_\omega]$-*cgggtg*(E-1) is the GC at 298 K. These structures have three $N_\omega H\cdots O$ and one $N_\delta H\cdots O$ hydrogen bonds between the guanidinium side chain and 18C6, and intramolecular $N_\omega H\cdots OC$ and $OH\cdots N_\alpha$ hydrogen bonds. Uniquely, MP2(full) predicts that a zwitterionic salt-bridge structure, $H^+R[N_\alpha]$-*gg_gg_t*(D_{3d}-1), is the GC at 298 K. Here, the backbone amino group and side chain are both protonated, leaving the carboxylic acid group deprotonated. The protonated backbone amino group in H^+(Arg) binds to a slightly distorted D_{3d} conformer of 18C6 *via* three $N_\alpha H\cdots O$ hydrogen bonds (1.9–2.1 Å) and longer range $CO\cdots HC_{18C6}$ hydrogen bonds (2.43 and 2.51 Å). This H^+(Arg) moiety is stabilized by intramolecular $N_\delta H\cdots O$ and $N_\omega H\cdots O$ hydrogen bonds with both of the backbone carboxylate oxygen atoms. In addition to these protonated GCs, we found 49 additional excited conformations to be within 14–31 kJ mol^{-1} of the GC at 0 K (12–23 kJ mol^{-1} at 298 K), as detailed in Table S4 (ESI†).

Interestingly, the isolated $H^+R[N_\omega]$-*tcg_gg_t* conformer optimized to the $H^+R[N_\omega]$-*tg_g_gg_t* conformer when 18C6 was removed, lying 8–9 kJ mol^{-1} above the isolated $H^+R[N_\omega]$-*tg_g_g_gg* GC at these levels of theory.[57,59,60] The free

Table 3 Relative energies in kJ mol^{-1} at 0 K (Gibbs energies at 298 K) for (Lys)H^+(18C6) and (Arg)H^+(18C6) complexes[a]

(AA)H^+(18C6)	B3LYP		B3P86		M06		MP2(full)	
(Lys)H^+(18C6)								
$H^+K[N_\varepsilon]$-*ctg_gtt*(D_{3d}-1)	**0.0**	(3.7)	**0.0**	(**0.0**)	1.1	(0.7)	0.6	(0.2)
$H^+K[N_\varepsilon]$-*tg_tttt*(D_{3d}-1)	1.3	(**0.0**)	6.4	(1.4)	15.7	(10.3)	18.1	(12.7)
$H^+K[N_\varepsilon]$-*tttttt*(D_{3d}-1)	5.7	(0.6)	10.5	(1.7)	20.2	(10.9)	23.6	(14.3)
$H^+K[N_\varepsilon]$-*cggtgg*(C-1)	7.5	(11.6)	8.1	(8.5)	**0.0**	(**0.0**)	**0.0**	(**0.0**)
$H^+K[N_\alpha]$-*cg_ggg_g_*(D-6)	17.6	(26.4)	16.1	(21.2)	17.9	(22.6)	10.0	(14.7)
(Arg)H^+(18C6)								
$H^+R[N_\omega]$-*tcg_gg_t*(B-3)	**0.0**	(12.8)	**0.0**	(11.2)	1.4	(1.4)	**0.0**	(1.9)
$H^+R[N_\omega]$-*cggg_tg_*(B-2)	1.1	(**0.0**)	2.6	(**0.0**)	15.3	(1.5)	24.4	(12.4)
$H^+R[N_\omega]$-*cgggtg*(E-1)	2.7	(5.9)	3.1	(4.8)	9.6	(**0.0**)	17.1	(9.4)
$H^+R[N_\omega]$-*tcg_g_gg*(E-2)	9.6	(24.2)	8.1	(21.1)	**0.0**	(1.8)	0.8	(4.4)
$H^+R[N_\omega]$-*tg_g_g_gg*(B-2)	3.4	(6.2)	7.1	(8.4)	18.2	(8.2)	24.9	(16.8)
$H^+R[N_\alpha]$-*gg_gg_t*(D_{3d}-1)	13.9	(23.2)	10.2	(18.0)	5.5	(2.0)	1.6	(**0.0**)

[a] Ground conformations in bold. All values calculated at the level of theory indicated using the 6-311+G(2d,2p) basis set with geometries, zero-point energies, and thermal energy corrections calculated at the B3LYP/6-311+G(d,p) level of theory.

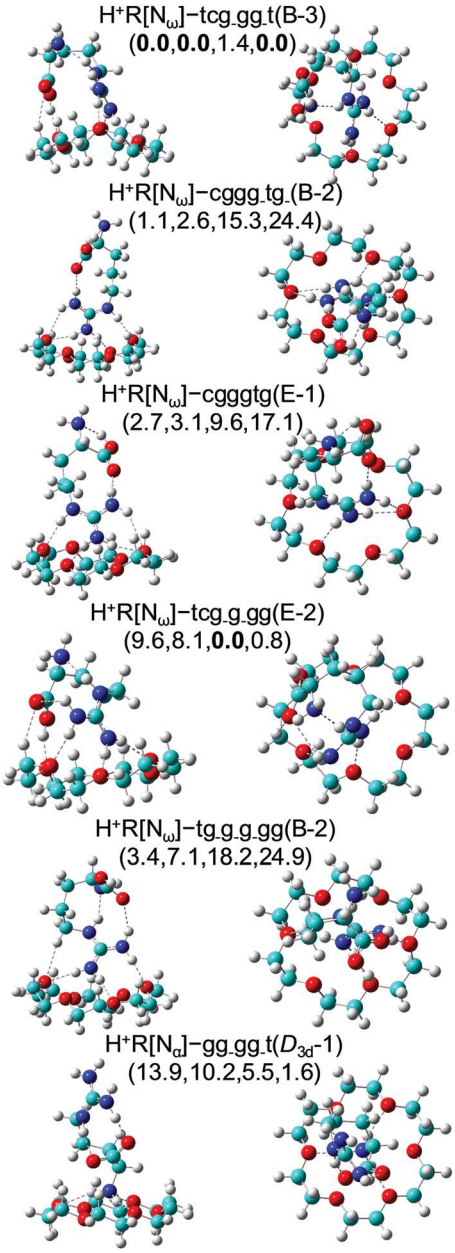

Fig. 5 Optimized conformations calculated at the B3LYP/6-311+G(d,p) level of theory for the (Arg)H^+(18C6) complex. Relative energies at 0 K in kJ mol^{-1} are given at the B3LYP, B3P86, M06, and MP2(full) levels.

$H^+R[N_\omega]$-*tcg_g_gg* conformer optimized to the ground isomer, $H^+R[N_\omega]$-*tg_g_g_gg*, whereas the free $H^+R[N_\omega]$-*cggg_tg_*, $H^+R[N_\omega]$-*cgggtg*, $H^+R[N_\omega]$-*tg_g_g_gg* (GC), and $H^+R[N_\alpha]$-*gg_gg_t* conformations did not change within these designations. Clearly, binding to 18C6 helps stabilize some of the less stable free isomers.

Comparison of experimental IRMPD and theoretical IR spectra

A general comparison between the IRMPD and theoretical IR spectra of the (AA)H^+(18C6) complexes shows similar vibrational bands in the finger print region, which are associated with the carboxylic acid C=O stretch near 1750–1800 cm^{-1}, methylene wagging of the 18C6 backbone near 1340–1350 cm^{-1},

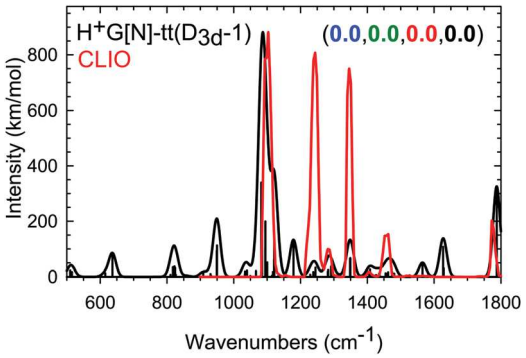

Fig. 6 Comparison of the measured IRMPD action spectrum of (Gly)H^+(18C6) (red) with the theoretical linear IR spectra for the ground conformation of (Gly)H^+(18C6) at 298 K (black). Relative energies at 298 K in kJ mol^{-1} are given at the B3LYP, B3P86, M06, and MP2(full) levels.

methylene twisting of the 18C6 backbone at 1240–1280 cm^{-1}, C–O stretch of the 18C6 backbone near 1050–1090 cm^{-1}, and methylene rocking of the 18C6 backbone near 925–950 cm^{-1}. Additional IRMPD bands are observed for the (AA)H^+(18C6) complexes and are discussed below. Saturation was observed in the C–O stretching region for the four systems. As a result, the corresponding region of the IRMPD spectra was recorded with an attenuated (−6 or −9 dB) IR FEL beam.

(Gly)H^+(18C6). Fig. 6 shows the spectrum for the (Gly)H^+(18C6) complex, which has three intense bands at 1103, 1244, and 1346 cm^{-1}, with weaker bands at 1283, 1404, 1463, and 1773 cm^{-1}. Below 1075 cm^{-1}, there was insufficient FEL power to dissociate the (Gly)H^+(18C6) complex. All of the low-energy conformers of (Gly)H^+(18C6) predict very similar spectra (even *tt* and *tc* isomers of Gly), consistent with the similar binding between the protonated amino group of Gly and 18C6. These are all shown in Fig. S1 of the ESI.† Fig. 6 uses the 0 K GC, $H^+G[N]$-*tt*(D_{3d}-1), as exemplary, in order to identify the character of the observed bands. The major bands are associated with C–O stretches of the 18C6 backbone (1082 and 1085 cm^{-1}) with a shoulder at 1121 cm^{-1} corresponding to a NH_3 rock/COH bend, methylene twisting of the 18C6 backbone (1229 cm^{-1}), and methylene wagging of the 18C6 backbone (1348, 1349 cm^{-1}). The weaker bands correspond to methylene twisting of the 18C6 backbone (1282, 1283 cm^{-1}), glycine CH_2 wag coupled with 18C6 methylene wags (1408 cm^{-1}), 18C6 methylene scissors (1455, 1456, 1458, 1462 cm^{-1}), and glycine carboxylic acid C=O stretch (1788 cm^{-1}). Although the predicted spectra match the experimental spectrum reasonably well, relative intensities are not accurately predicted, presumably a consequence of the multiple photon character of the IRMPD process. Further, reasonably intense bands are also predicted at 949 cm^{-1} (CH_2 wags of 18C6), 1179 cm^{-1} (another NH_3 rock/COH bend), 1565 cm^{-1} (NH_3 umbrella mode), and 1627 cm^{-1} (NH_3 bends), but not observed in the experimental spectrum. The failure to observed predicted bands near 1600 cm^{-1} is not unique to the Gly system, but is also found for His and Lys complexes with 18C6 (see below). The origins of this failure are not apparent to us. It is possible that the stronger binding of

H^+Gly to 18C6 leads to more inefficient photodissociation, which might explain the lack of the low-frequency 949 cm^{-1} and more minor 1179 cm^{-1} bands.

Because of the similarity of the predicted spectra, a definitive assignment of the conformation formed experimentally is not possible, although the IR spectra for the lowest-energy $H^+G[N]$-*tt*(D_{3d}) and $H^+G[N]$-*tt*(D-1) conformers exhibit slightly better agreement with the experimental spectrum. Certainly, these conformers alone are sufficient to explain the experimental spectrum observed.

(His)H^+(18C6). The experimental IRMPD spectrum of (His)H^+(18C6), shown in Fig. 7, is substantially more complicated than that for glycine. Here, bands are found at 572, 857, 941, 1063, 1082, 1104, 1158, 1250, 1352, 1460–1500, and 1773 cm^{-1}. Again the low-energy conformers predict similar spectra because the conformation of the His moiety is the same and they all bind to 18C6 *via* the protonated-backbone amino group. Fig. 7 compares the predicted spectra of select conformers with the IRMPD spectrum, with other conformers included in Fig. S2 of the ESI.† Using the $H^+H[N_\alpha]$-*cggc*(B-2) conformer as exemplary, the intense triplet of bands at 1063, 1082, and 1104 cm^{-1} has the same character as the 1103 cm^{-1} band observed in the (Gly)H^+(18C6) system, and the bands at 1352 (methylene wag) and 1773 (C–O stretch) cm^{-1} parallel the similar bands for glycine. The weak band at 572 cm^{-1} is an out-of-plane bend in the imidazole ring (568 cm^{-1}). The band at 857 cm^{-1} is associated with several bands (822–852 cm^{-1}) corresponding to methylene twists in 18C6 (with one imidazole out-of-plane bend at 823 cm^{-1}). The band at 941 cm^{-1} is a methylene twist/CC stretch of 18C6 (941, 945 cm^{-1}). Unlike the (Gly)H^+(18C6) complex, the band at 1250 cm^{-1} is predicted to correspond to a C–OH stretch/C_βH bend in the His moiety (1239 cm^{-1}). The broad band at 1460–1500 cm^{-1} is associated with a strong predicted band at 1464 cm^{-1} (COH bend in His) coupled with several 18C6 methylene scissors modes (as for Gly). As for the (Gly)H^+(18C6) complex, bands at 1564 (CC imidazole stretch/NH_3 umbrella), 1573 (NH_3 umbrella), 1628 (NH_3 bends), and 1643 (NH_3 bends) cm^{-1} are predicted but not observed. The broad band observed at 1158 cm^{-1} is not predicted well by any of the conformations, although there are weak predicted bands in this area (1169 and 1187 cm^{-1}) corresponding to NH_3 rock/CH twist in His and CH_2/CH twist in His, respectively.

All of the low-lying $H^+H[N_\alpha]$ conformers of the (His)H^+(18C6) complex are similar to one another and agree fairly well with the IRMPD spectrum (exceptions noted above). Although the $H^+H[N_\pi]$-*tgg_g_*(B-2) conformer protonated on the side chain (Fig. 7) also has reasonable agreement with most of the IRMPD spectrum, the main peak at 1100 cm^{-1} is blue shifted from experiment and predicted bands between 600–750 and at 895 and 1400 cm^{-1} are missing from the experiment. These additional discrepancies with experiment suggest that H^+(His) is complexed to 18C6 *via* the protonated backbone rather than the side chain, in agreement with the lowest-energy conformers predicted by theory, Table 2.

(Lys)H^+(18C6). The experimental IRMPD spectrum exhibits strong bands at 941, 1084, and 1347 with weaker features at 1246, 1288, 1458, 1589, and 1760 cm^{-1}. This spectrum is compared with theoretical IR spectra of representative low-lying conformers of (Lys)H^+(18C6) in Fig. 8, with relative energies in Table 3. Here, it can be seen that while most

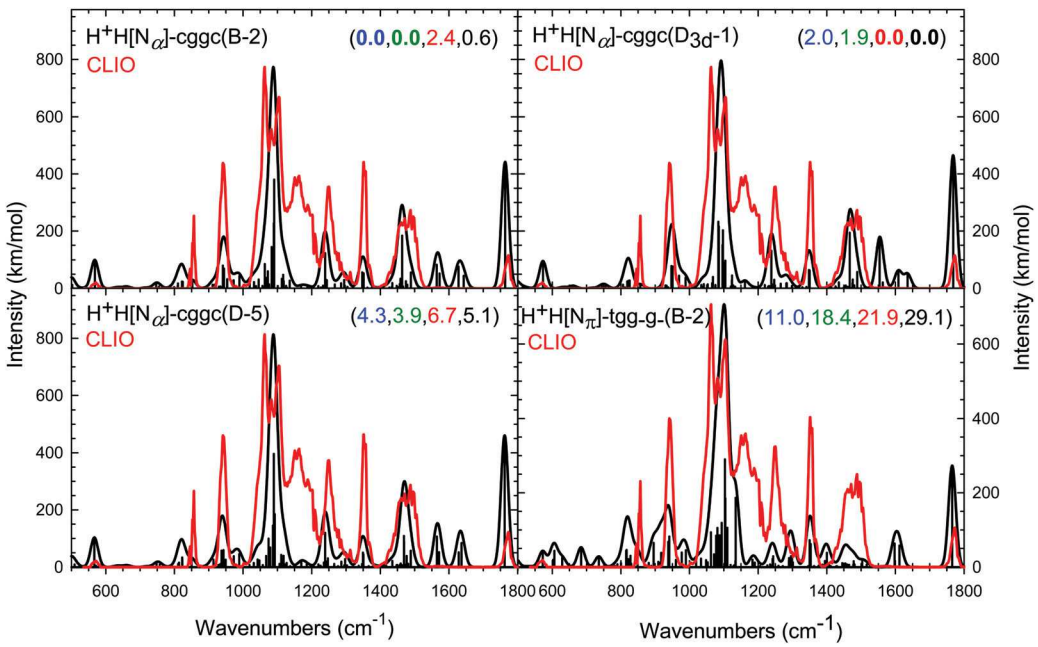

Fig. 7 Comparison of the measured IRMPD action spectrum of (His)H^+(18C6) (red) with the theoretical linear IR spectra for the ground and selected stable low-energy conformations of (His)H^+(18C6) at 298 K (black). Relative energies at 298 K in kJ mol^{-1} calculated at the B3LYP, B3P86, M06, and MP2(full) levels of theory are given in parentheses in each panel. To facilitate comparison of the measured and computed spectra, the IRMPD spectrum is overlaid (in red) with each computed spectrum and normalized to match the intensity of the most intense feature in each spectrum.

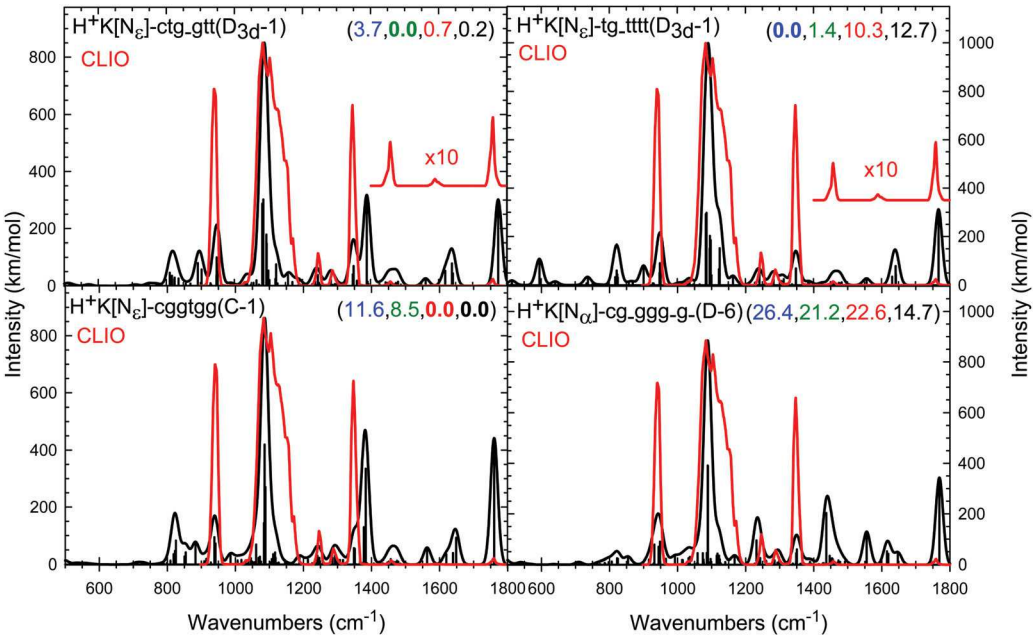

Fig. 8 Comparison of the measured IRMPD action spectrum of (Lys)H^+(18C6) with the theoretical linear IR spectra for the ground and selected stable low-energy conformations of (Lys)H^+(18C6) at 298 K. Relative energies at 298 K in kJ mol^{-1} calculated at the B3LYP, B3P86, M06, and MP2(full) levels of theory are given in parentheses in each panel. To facilitate comparison of the measured and computed spectra, the IRMPD spectrum is overlaid (in red) with each computed spectrum and normalized to match the intensity of the most intense feature in each spectrum.

theoretical spectra are similar, the $H^+K[N_\varepsilon]$-*ctg_gtt*(D_{3d}-1) and $H^+K[N_\varepsilon]$-*cggtgg*(C-1) conformers predict a fairly intense band at ~1390 cm^{-1} (COH bend) that is not found in the experimental spectrum, whereas the $H^+K[N_\varepsilon]$-*tg_tttt*(D_{3d}-1) and $H^+K[N_\varepsilon]$-*tttttt*(D_{3d}-1) (not shown but very similar to *tg_tttt*) conformers do not have this band. This is because in the former two conformers, there is an OH···N_α hydrogen bond (as indicated by the first dihedral angle being *cis*) that shifts the COH bend frequency to higher energy (up from 1280–1300 cm^{-1}).

We can assign the observed bands using the $H^+K[N_\varepsilon]$-*tg_tttt*(D_{3d}-1) spectrum as exemplary. The bands at 1084 (C–O stretches of 18C6), 1246 (methylene twisting of 18C6), 1288 (methylene twisting of the 18C6), 1347 (methylene wagging of the 18C6), 1458 (18C6 methylene scissors), and 1760 (carboxylic acid C=O stretch) cm^{-1} match those seen in the (Gly)H^+(18C6) complex. None of the spectra predict a band near 1589 cm^{-1}. The remaining band observed at 941 cm^{-1} is associated with CC stretch/CCH bends in 18C6 and Lys (948, 949, and 951 cm^{-1}). The *tg_tttt* and *tttttt* conformers have slightly different C=O stretching frequencies (1768 and 1756 cm^{-1}, respectively) but both agree reasonably well with the observed peak at 1760 cm^{-1}. The backbone-protonated conformer $H^+K[N_\alpha]$-*cg_ggg_g_*(D-6) (Fig. 8) is predicted to have a strong band at 1420 cm^{-1} (again the COH bend, now shifted to higher energies by a OH···N_ε hydrogen bond), that is not observed experimentally. On this basis and the relative energetics, the observed spectrum is fully consistent with complexation of 18C6 to Lys protonated on the side-chain amino group (N_ε) and agrees best with the low-energy $H^+K[N_\varepsilon]$-*txtttt*(D_{3d}-1) conformers, which are calculated to be very low in energy at the B3LYP and B3P86 levels. The IRMPD spectrum appears inconsistent with the $H^+K[N_\varepsilon]$-*cg_ggtt*(D_{3d}-1) and $H^+K[N_\varepsilon]$-*cggtgg*(C-1) isomers, which are low-lying at the M06 and MP2 levels.

(Arg)H^+(18C6). The experimental IRMPD spectrum of (Arg)H^+(18C6) has intense peaks at 941, 1114, 1350, 1612, and 1663 cm^{-1}, with minor peaks at 620, 722, 845, 1250, 1292, 1450, and 1746 cm^{-1}. This spectrum is compared to theoretical IR spectra of representative low-lying conformers of (Arg)H^+(18C6) in Fig. 9 with relative energies in Table 3.

The bands that are observed can be assigned by referring to the $H^+R[N_\omega]$-*cggg_tg_*(B-2) spectrum (the 298 K GC for B3LYP and B3P86). The three high-frequency bands correspond to 1751 (carboxylic acid C=O stretch of Arg), 1657/1672 (NH and NH_2 bends), 1617 (NH and NH_2 bends) cm^{-1}, whereas a smaller intensity band at 1566 cm^{-1} (more NH and NH_2 bends) is not observed experimentally. A series of bands near 1450 (CH_2 scissors, mainly of 18C6 but also Arg), 1300 and 1350 (CH_2 wagging of the 18C6 and Arg), and 1240 (CH_2 rock of 18C6) cm^{-1} match those seen in the experimental spectrum. The broad band observed near 1100 cm^{-1} is associated with a series of 18C6 C–O stretches and is observed in all theoretical spectra although the intensity of the shoulder to the blue is not reproduced particularly well. The strong band observed at 941 cm^{-1} is assigned to rocking modes of the CH_2 groups of 18C6 (940 cm^{-1}). The broad and weak bands at 845, 722, and 620 cm^{-1} are assigned to wagging modes of the $N_\alpha H_2$ group (846 cm^{-1}) with a side band at ~820 cm^{-1} corresponding to rocking modes of the CH_2 groups of 18C6, out-of-plane $N_\delta H$ bends at 708 and 714 cm^{-1}, and out-of-plane wags of the $N_\omega H_2$ group at 615 and 619 cm^{-1}.

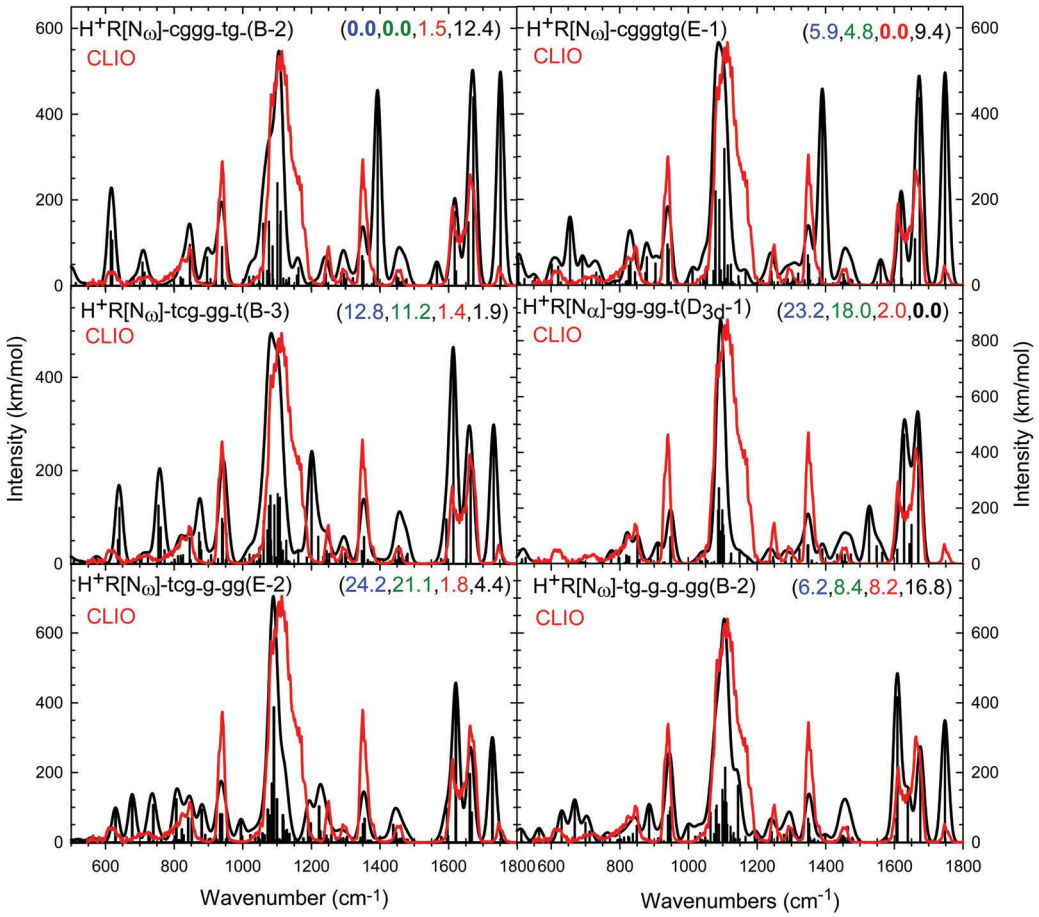

Fig. 9 Comparison of the measured IRMPD action spectrum of (Arg)H^+(18C6) with the theoretical linear IR spectra for the ground and selected stable low-energy conformations of (Arg)H^+(18C6) at 298 K. Relative energies at 298 K in kJ mol^{-1} calculated at the B3LYP, B3P86, M06, and MP2(full) levels of theory are given in parentheses in each panel. To facilitate comparison of the measured and computed spectra, the IRMPD spectrum is overlaid (in red) with each computed spectrum and normalized to match the intensity of the most intense feature in each spectrum.

Despite this agreement, the $H^+R[N_\omega]$-*cggg_tg_*(B-2) spectrum predicts bands of appreciable intensity at 896 and 1393 cm^{-1} that are not experimentally observed. Both bands correspond to motions of the carboxylic acid COH group: motion of the hydrogen out of the plane established by the COH$\cdots$$N_\alpha$ hydrogen bond and the COH bend, respectively. The position of these bands is determined by the intramolecular COH$\cdots$$N_\alpha$ hydrogen bond (as indicated by the first dihedral angle being *cis*), and is red shifted and loses intensity in conformations that have an intramolecular COH$\cdots$OC hydrogen bond instead (as indicated by a *trans* first dihedral angle). Thus, the spectrum of $H^+R[N_\omega]$-*cggg_tg_*(B-2) and $H^+R[N_\omega]$-*cgggtg*(E-1) (GC at 298 K for M06) are very similar above 800 cm^{-1}. They differ in the lower frequency range, with $H^+R[N_\omega]$-*cggg_tg_*(B-2) providing a better match to experiment. Among other low-lying conformers found, the $H^+R[N_\omega]$-*tcg_gg_t*(B-3) and $H^+R[N_\omega]$-*tcg_g_gg*(E-2) carbonyl stretches (1732 and 1728 cm^{-1}) are red shifted compared to experiment because the CO is hydrogen bonded to a CH group of the crown. Likewise, the COH is hydrogen bonded to a crown oxygen, such that the intense band associated with the COH bend has shifted to 1201 and 1191/1222 cm^{-1}, respectively, which again is not observed experimentally. Neither of the spectra of these conformers agree with experiment very well below 800 cm^{-1} either. The other side-chain-bound conformer included in Table 3 and Fig. 9, $H^+R[N_\omega]$-*tg_g_g_gg*(B-2), reproduces the experimental spectrum reasonably well except for extra bands of modest intensity at 669 and 885 cm^{-1} (out-of-plane $N_\omega H$ and $N_\delta H$ bends shifted by $N_\delta H\cdots N_\alpha$ and $N_\omega H\cdots O_{18C6}$ hydrogen bonds). In particular, the C–O stretch (1747 cm^{-1}) and the other two high frequency bands are reproduced well, as is the shape of the broad band centered at 1114 cm^{-1}. For this conformer, the intramolecular $N_\delta H\cdots N_\alpha$ and $N_\omega H\cdots$OC hydrogen bonds lead to the COH hydrogen binding solely to the carbonyl oxygen, such that the COH bend has shifted to 1144 cm^{-1}, thereby providing the shoulder to the intense central band.

We also considered the backbone-bonded conformer, $H^+R[N_\alpha]$-*gg_gg_t*(D_{3d}-1), which is the MP2 GC at 298 K. As can be seen in Fig. 9, this conformer reproduces many of the observed bands but predicts a moderately intense band at 1527 cm^{-1} (the NH_3^+ umbrella motion) that is not observed, and conversely does not reproduce the minor bands at 620, 722, and 1746 cm^{-1}. In all the $H^+R[N_\alpha]$ conformers found, the arginine is zwitterionic with both N_α and N_ω protonated and

the carboxylic acid deprotonated (with resulting $N_\delta H\cdots OC$ and $N_\omega H\cdots OC$ intramolecular hydrogen bonds), such that the CO stretch red shifts below 1700 cm^{-1}. Therefore, the presence of the 1746 cm^{-1} band is unequivocal evidence for the population of side-chain-bonded conformers, although its relatively weak intensity compared to the 1663 cm^{-1} band suggests its population may be small. Given this observation, contributions from the backbone-bonded conformer seem likely (especially because bands similar to the 1527 cm^{-1} predicted here were not observed in the $(AA)H^+(18C)$ systems for glycine, histidine, and lysine). Here, the B3LYP and B3P86 GCs seem unlikely to be populated and no level of theory predicts that $H^+R[N_\omega]$-*tg_g_g_gg*(B-2) is the GC, whereas MP2(full) and M06 would indicate population of the backbone-bonded $H^+R[N_\alpha]$-*gg_gg_t*(D_{3d}-1) conformer.

Conclusion

In this study, four protonated amino acids (AA = Gly, His, Lys, and Arg) were complexed to 18-crown-6 ether in an attempt to determine their gas-phase binding configurations by use of infrared multiple photodissociation (IRMPD) action spectroscopy utilizing light generated by the Centre Laser Infrarouge d'Orsay (CLIO). The IRMPD spectrum obtained for each $(AA)H^+18C6$ complex was compared to theoretical IR spectra calculated at the B3LYP/6-311+G(d,p) level of theory after a comprehensive conformational search. IRMPD leads to the primary loss of 18C6 for the $(His)H^+(18C6)$, $(Lys)H^+(18C6)$, and $(Arg)H^+(18C6)$ complexes, whereas the loss of Gly was observed for the $(Gly)H^+(18C6)$ complex. The comparison between the IRMPD and theoretical IR spectra confirms that 18C6 binds to Gly and His *via* the protonated-backbone amino group, whereas protonated Lys clearly prefers binding *via* the protonated side-chain amino group. Protonated Arg clearly binds *via* the guanidinium side chain, but contributions from conformers involving binding to the protonated backbone in a salt-bridge structure also seem likely.

As noted above, a previous TCID study of these same complexes measured the 18C6 binding affinities for the protonated AAs and found the trend that Gly > Lys > His > Arg.[20] Here, theory indicated that Gly, Arg, and His bind *via* the backbone amino group, whereas Lys binds *via* the protonated side-chain amine. This preferred binding site for Lys was confirmed by TCID studies of acetylated (Ac) versions that block particular binding sites.[21] Further, this study indicated that the 18C6 binding affinities to the protonated side chains fell in the order Lys > Arg > His. The present IRMPD results largely support these previous results, with Gly and His clearly binding at the protonated backbone amino group and Lys at the protonated side-chain amino group. For Arg, the present theoretical results (which were conducted at a higher level than those previously[20]) find that the GC at most levels of theory binds 18C6 at the side-chain guanidinium, with binding at the protonated backbone amino group competitive at only the M06 and MP2 levels. IRMPD results clearly indicate the side-chain binding site is operative, but also suggests this may be a minor contributor. Overall, all levels of theory provide accurate predictions of the GC for 18C6 complexes of protonated Gly and His, whereas B3LYP and B3P86 provide accurate predictions of the Lys complex. For Arg complexes, the likely presence of the side-chain bound conformer and absence of the B3LYP and B3P86 GCs suggest that MP2(full) and M06 are yielding more accurate energetics.

Combining the present structural results with the previous TCID thermochemistry has implications for molecular recognition and SNAPP. In particular, our results indicate that these methods should be sensitive to protonated exposed Lys side chains and the N-terminus (especially for residues with small side-chains) and that protonated His side chains are much less likely to bind, whereas protonated Arg side chains seem possible targets for binding 18C6, although still less strongly than Lys or the N-terminus.

Conflicts of interest

The authors declare no competing financial interest.

Acknowledgements

This work is supported by the National Science Foundation, Grants CHE-1709789 (MTR), CHE-1664618 (PBA), OISE-0730072 and OISE-1357787. The authors also thank the Center for High Performance Computing at The University of Utah, the Extreme Science and Engineering Discovery Environment (XSEDE), Grant TG-CHE140143, for computer time and support. The research leading to this result has been supported by the project CALIPSOplus under the Grant Agreement 730872 from the EU Framework Programme for Research and Innovation HORIZON 2020. Financial Support from the National FT-ICR network (FR3624 CNRS) for conducting the research is gratefully acknowledged.

References

1 C. N. Stedwell, J. F. Galindo, K. Gulyuz, A. E. Roitberg and N. C. Polfer, *J. Phys. Chem. A*, 2013, **117**, 1181–1188.
2 R. R. Julian and J. L. Beauchamp, *Int. J. Mass Spectrom.*, 2001, **210–211**, 613–623.
3 R. R. Julian, M. Akin, J. A. May, B. M. Stoltz and J. L. Beauchamp, *Int. J. Mass Spectrom.*, 2002, **220**, 87–96.
4 R. R. Julian and J. L. Beauchamp, *J. Am. Soc. Mass Spectrom.*, 2002, **13**, 493–498.
5 R. R. Julian, J. A. May, B. M. Stoltz and J. L. Beauchamp, *Int. J. Mass Spectrom.*, 2003, **228**, 851–864.
6 R. R. Julian and J. L. Beauchamp, *J. Am. Soc. Mass Spectrom.*, 2004, **15**, 616–624.
7 T. Ly and R. R. Julian, *J. Am. Soc. Mass Spectrom.*, 2006, **17**, 1209–1215.
8 T. Ly and R. R. Julian, *J. Am. Soc. Mass Spectrom.*, 2008, **19**, 1663–1672.

9 T. Ly, Z. Liu, B. G. Pujanauski, R. Sarpong and R. R. Julian, *Anal. Chem.*, 2008, **80**, 5059–5064.
10 G. K. Yeh, Q. Sun, C. Meneses and R. R. Julian, *J. Am. Soc. Mass Spectrom.*, 2009, **20**, 385–393.
11 D. P. Weimann, H. D. F. Winkler, J. A. Falenski, B. Koksch and C. A. Schalley, *Nat. Chem.*, 2009, **1**, 573.
12 J. J. Wilson, G. J. Kirkovits, J. L. Sessler and J. S. Brodbelt, *J. Am. Soc. Mass Spectrom*, 2008, **19**, 257–260.
13 D. Ray, D. Feller, M. B. More, E. D. Glendening and P. B. Armentrout, *J. Phys. Chem.*, 1996, **100**, 16116–16125.
14 M. B. More, D. Ray and P. B. Armentrout, *J. Phys. Chem. A*, 1997, **101**, 831–839.
15 M. B. More, D. Ray and P. B. Armentrout, *J. Phys. Chem. A*, 1997, **101**, 4254–4262.
16 M. B. More, D. Ray and P. B. Armentrout, *J. Phys. Chem. A*, 1997, **101**, 7007–7017.
17 M. B. More, D. Ray and P. B. Armentrout, *J. Am. Chem. Soc.*, 1999, **121**, 417–423.
18 P. B. Armentrout, *Int. J. Mass Spectrom.*, 1999, **193**, 227–240.
19 Y. Chen and M. T. Rodgers, *J. Am. Chem. Soc.*, 2012, **134**, 2313–2324.
20 Y. Chen and M. T. Rodgers, *J. Am. Chem. Soc.*, 2012, **134**, 5863–5875.
21 Y. Chen and M. T. Rodgers, *J. Am. Soc. Mass Spectrom.*, 2012, **23**, 2020–2030.
22 P. Hurtado, F. Gámez, S. Hamad, B. Martínez-Haya, J. D. Steill and J. Oomens, *J. Phys. Chem. A*, 2011, **115**, 7275–7282.
23 M. Meot-Ner, *J. Am. Chem. Soc.*, 1983, **105**, 4912–4915.
24 J. M. Bakker, T. Besson, J. Lemaire, D. Scuderi and P. Maître, *J. Phys. Chem. A*, 2007, **111**, 13415–13424.
25 F. Glotin, J. M. Ortega, R. Prazeres and C. Rippon, *Nucl. Instrum. Methods Phys. Res., Sect. B*, 1998, **144**, 8–17.
26 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.
27 D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu and P. A. Kollman, *AMBER 14*, University of California, San Francisco, 2014.
28 M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha and E. Apra, *et al.*, *Comput. Phys. Commun.*, 2010, **181**, 1477–1489.
29 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Revision D.01*, 2009.
30 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
31 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.
32 B. Miehlich, A. Savin, H. Stoll and H. Preuss, *Chem. Phys. Lett.*, 1989, **157**, 200–206.
33 C. W. Bauschlicher and H. Partridge, *J. Chem. Phys.*, 1995, **103**, 1788–1791.
34 N. C. Polfer, *Chem. Soc. Rev.*, 2011, **40**, 2211–2221.
35 R. A. Coates, C. P. McNary, G. C. Boles, G. Berden, J. Oomens and P. B. Armentrout, *Phys. Chem. Chem. Phys.*, 2015, **17**, 25799–25808.
36 G. C. Boles, R. A. Coates, G. Berden, J. Oomens and P. B. Armentrout, *J. Phys. Chem. B*, 2016, **120**, 12486–12500.
37 G. C. Boles, R. A. Coates, G. Berden, J. Oomens and P. B. Armentrout, *J. Phys. Chem. B*, 2015, **119**, 11607–11617.
38 R. A. Coates, G. C. Boles, C. P. McNary, G. Berden, J. Oomens and P. B. Armentrout, *Phys. Chem. Chem. Phys.*, 2016, **18**, 22434–22445.
39 G. C. Boles, R. L. Hightower, R. A. Coates, C. P. McNary, G. Berden, J. Oomens and P. B. Armentrout, *J. Phys. Chem. B*, 2018, **122**, 3836–3853.
40 A. M. Chalifoux, G. C. Boles, G. Berden, J. Oomens and P. B. Armentrout, *Phys. Chem. Chem. Phys.*, 2018, **20**, 20712–20725.
41 J. P. Perdew, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1986, **33**, 8822–8824.
42 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241.
43 M. Head-Gordon, J. A. Pople and M. J. Frisch, *Chem. Phys. Lett.*, 1988, **153**, 503–506.
44 S. Sæbø and J. Almlöf, *Chem. Phys. Lett.*, 1989, **154**, 83–89.
45 M. J. Frisch, M. Head-Gordon and J. A. Pople, *Chem. Phys. Lett.*, 1990, **166**, 275–280.
46 M. Head-Gordon and T. Head-Gordon, *Chem. Phys. Lett.*, 1994, **220**, 122–128.
47 N. N. Dookeran, T. Yalcin and A. G. Harrison, *J. Mass Spectrom.*, 1996, **31**, 500–508.
48 S. E. Hill and D. Feller, *Int. J. Mass Spectrom.*, 2000, **201**, 41–58.
49 N. A. Al-Jallal, A. A. Al-Kahtani and A. A. El-Azhary, *J. Phys. Chem. A*, 2005, **109**, 3694–3703.
50 E. D. Glendening, D. Feller and M. A. Thompson, *J. Am. Chem. Soc.*, 1994, **116**, 10657–10669.
51 E. D. Glendening and D. Feller, *J. Am. Chem. Soc.*, 1996, **118**, 6052–6059.
52 D. Feller, *J. Phys. Chem. A*, 1997, **101**, 2723–2731.

53 J. D. Dunitz, M. Dobler, P. Seiler and R. P. Phizackerley, *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.*, 1974, **30**, 2733–2738.
54 P. B. Armentrout, A. L. Heaton and S. J. Ye, *J. Phys. Chem. A*, 2011, **115**, 11144–11155.
55 M. Citir, C. S. Hinton, J. Oomens, J. D. Steill and P. B. Armentrout, *Int. J. Mass Spectrom.*, 2012, **330–332**, 6–15.
56 R. Wu and T. B. McMahon, *ChemPhysChem*, 2008, **9**, 2826–2835.
57 B. Gao, T. Wyttenbach and M. T. Bowers, *J. Phys. Chem. B*, 2009, **113**, 9995–10000.
58 A. S. Lemoff, M. F. Bush, J. T. O'Brien and E. R. Williams, *J. Phys. Chem. A*, 2006, **110**, 8433–8442.
59 M. F. Bush, J. Oomens and E. R. Williams, *J. Phys. Chem. A*, 2009, **113**, 431–438.
60 W. D. Price, R. A. Jockusch and E. R. Williams, *J. Am. Chem. Soc.*, 1997, **119**, 11988–11989.