Separating and Probing Tautomers of Protonated Nucleobases using Differential Mobility Spectrometry

Ahdia Anwar,[1] Jarrod Psutka,[1] Stephen W.C. Walker,[1] Thorsten Dieckmann,[1] John S. Janizewski,[2] J. Larry Campbell,[1,3,*] W. Scott Hopkins[1,**]

[1] Department of Chemistry, University of Waterloo, Waterloo, ON, N2L 3G1 Canada
[2] Pfizer Global Research and Development, Eastern Point Road, Groton, CT, 06340 USA
[3] SCIEX, 71 Four Valley Drive, Concord, ON, L4K 4V8 Canada

Corresponding Author(s):

*   Larry.Campbell@sciex.com
**   Scott.Hopkins@uwaterloo.ca

ABSTRACT:

The protonated nucleobases $(C + H)^+$, $(T + H)^+$, $(U + H)^+$, $(A + H)^+$, and $(G + H)^+$ are investigated in a combined experimental and computational study using differential mobility spectrometry (DMS), mass spectrometry, and electronic structure calculations. DMS is used to isolate individual tautomeric forms for each protonated nucleobase prior to characterization with HDX or CID. The population distributions of each protonated nucleobase formed by electrospray ionization (ESI) are dominated by a single tautomeric form, as is predicted by our calculations. However, all nucleobases present additional tautomers upon ESI, with these minor contributions to the ensemble populations attributed to additional higher energy metastable species. In addition to the tautomer-derived species, additional ion signals in the DMS data are attributed to larger nucleobase-containing clusters, which fragment post-DMS to yield bare ion and fragment ion signals that are consistent with those expected for the bare protonated nucleobases. Contributions from larger clustered species are instead distinguished by monitoring DMS ion signal as declustering potential voltages are ramped.

1. INTRODUCTION

Owing to the importance of DNA and RNA in encoding and expressing genetic information, and the central role that nucleobases play in establishing the structure and functionality of nucleic acid sequences, a great deal of experimental [1-4] and theoretical [5-8] effort has gone into determining the structures and properties of cytosine (C), guanine (G), adenine (A), thymine (T), and uracil (U). Of importance are the sites of protonation and the tautomeric forms that the nucleobases exhibit, since these variations are thought to impact mutagenic processes (*e.g.*, point mutation during nucleic acid replication) [9] and the stabilization of triplex structures.[10] Mass spectrometry and quantum chemical calculations have been employed to great success in determining nucleobase properties such as gas-phase acidity and basicity.[11-17] However, it has been shown previously that several different tautomers are likely to exist simultaneously in a given nucleobase ensemble.[18, 19] For example, Salpin *et al*. used infrared multiple photon dissociation (IRMPD) spectroscopy to demonstrate the presence of at least two tautomeric forms of $(C + H)^+$, $(T + H)^+$, and $(U + H)^+$ in ion populations generated by ESI.[20] Comparison of the experimental IRMPD spectra with IR spectra that were calculated at the B3LYP/6-31++G(d,p) level of theory indicated that the three protonated nucleobases existed predominantly as enolic tautomers, with a small sub-population of oxo tautomers. Subsequent work by Bakker *et al*. showed that the vibrational spectra of monohydrated protonated uracil, $(U + H)^+ \cdot OH_2$, and cytosine, $(C + H)^+ \cdot OH_2$, were also consistent with the presence of two tautomeric species arising from the production of two protonated forms of the associated nucleobases via ESI.[21, 22] This suggests that the molecular properties of protonated nucleobases as determined by mass spectrometry are likely to correspond to an ensemble average for the various tautomeric structures that are present under the experimental conditions employed during measurement. It is therefore desirable to separate the tautomeric species prior to mass spectrometric or spectroscopic interrogation.

Various forms of ion mobility spectrometry have been employed to separate tautomers prior to MS analysis.[23-25] For example, the Attygalle laboratory recently reported on the characterization of tautomer populations of deprotonated hydroxybenzoic acid with travelling wave ion mobility spectrometry.[26] This work challenges the notion that ESI-MS results reflect solution phase population distributions, and demonstrates that tautomer populations can be tuned by varying ESI source conditions. We have also recently reported on the use of ion mobility to characterize tautomer populations generated via ESI by using differential mobility spectrometry

(DMS)[27-30] to separate and probe the nitrogen- and oxygen-protonated tautomers of *para*-aminobenzoic acid.[31, 32] By taking advantage of the different DMS behaviors of the two tautomers, we could examine the MS/MS fragmentation patterns and HDX behaviors of each species individually and demonstrate that each structure did, indeed, exhibit its own characteristic physicochemical properties. We also demonstrated that a great deal of care had to be taken in HDX experiments since high vapor pressures of HDX reagent can drive *in situ* tautomerization via proton-transfer relay networks established upon ion-solvent clustering.[32] Studies like these show that ESI solvent effects are not necessarily the primary criteria that determine relative tautomer (or, by extension, isomer or conformer) population distributions. Instead, a variety of post-ESI instrument conditions could potentially contribute to the observed sub-populations within a gas phase ensemble.

Here, we utilize the DMS technique to separate and study the individual tautomeric forms of protonated adenine, $(A + H)^+$, guanine, $(G + H)^+$, cytosine, $(C + H)^+$, thymine, $(T + H)^+$, and uracil, $(U + H)^+$ that are generated via ESI. The various tautomers of these molecules are studied individually by HDX and CID, and we show that the relative tautomer populations can be manipulated post-ESI and post-DMS by using the instrument declustering potential to selectively fragment high-energy, kinetically trapped tautomers prior to MS characterization.

2. METHODS

2.2 Experimental Details

A SelexION differential mobility spectrometer was used in conjunction with a QTRAP 5500 system (SCIEX; Concord, ON).[29-31, 33] Instrument parameters included a ESI probe voltage of 5500 V, a source temperature of 32 °C, nebulizing gas pressure of 20 psi, and auxiliary gas pressure of 0 psi. The DMS was set to a temperature of 150 °C, and nitrogen was used as both the curtain gas (20 psi) and collisionally activated dissociation gas (~9 mTorr) for all experiments. Nucleobase solids were purchased from Sigma-Aldrich and subsequently dissolved in a 50:50 mixture of ultrapure water and methanol with 0.1% formic acid to yield solutions of 10 ng/mL. Analyte solutions were pumped into the ESI source at 7 μL/min. HPLC-grade methanol, isopropanol, and deuterium oxide were also purchased from Sigma-Aldrich and used without further purification or dilution.

DMS experiments involved the stepping of the separation voltage (SV) from 0 to 4000 V in 500 V increments. At each SV, the compensation voltage (CV) was scanned from −80 V to 15 V in increments of 0.1 V to produce an ionogram. A dispersion plot,[34, 35] which plots optimal conditions for ion transmission as a function of SV and CV, was then generated. Dispersion plots enable the identification of the DMS behavior of particular ions according to known patterns.[33-35] These data were acquired for each nucleobase in a pure N_2 DMS environment, as well as with DMS environments that had been seeded with 1.5% (mole ratio) methanol (MeOH) and isopropanol (IPA) chemical modifiers. Hydrogen-deuterium exchange (HDX) experiments were conducted through the infusion of deuterium oxide into the throttle gas. These experiments were conducted under two different HDX conditions. In the first implementation, the throttle gas was bubbled through D_2O to saturate the N_2 with HDX reagent. This yields maximum rates of HDX in the junction chamber between the DMS cell and the orifice of the mass spectrometer, as described in reference [32]. In the second implementation, the throttle gas sampled only the headspace above the D_2O HDX reagent vessel, resulting in a lower D_2O partial pressure and slower rates of HDX.[32] In this way, the DMS cell was used to select a specific tautomer prior to HDX, which was monitored by recording a full scan mass spectrum (Q1).

Enhanced product ion (EPI) scans were also conducted for each of the separated nucleobase tautomers. Following DMS isolation of a given tautomer, the collision energy (CE) of the Q2 ion trap was ramped from 0 V to 60 V in 0.25 V increments, while recording the complete mass spectrum at each interval. By plotting the fraction of the parent and each fragment ion present as a function of collision energy, breakdown curves were produced.[36] A schematic diagram of the DMS region is shown in Figure 1.
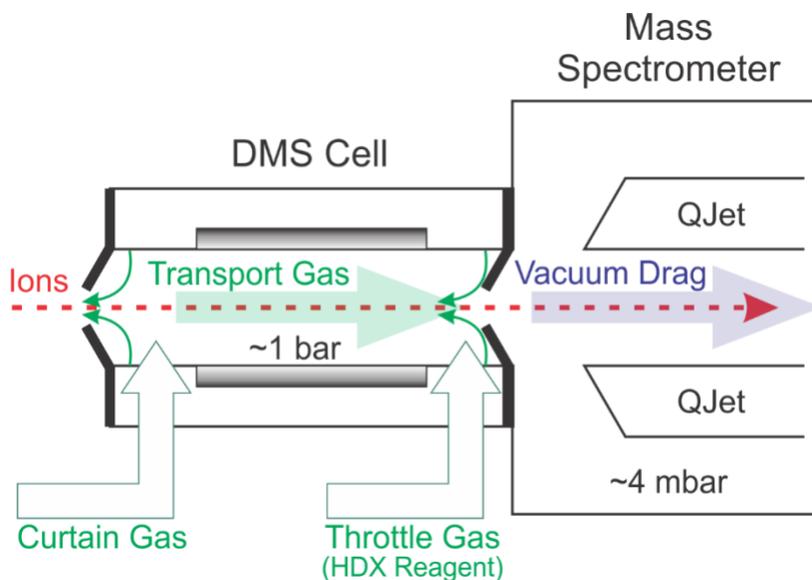
Figure 1. A schematic diagram of the DMS cell and the QJet region of the mass spectrometer. Adapted from reference [32].

2.1 Computational Details

All possible tautomeric forms of the protonated nucleobases (C, G, A, T, and U) were considered. Optimization and frequency calculations (T = 298.15 K, P = 1 atm) were performed at the B3LYP level of theory using a 6-311++G(d,p) basis set as implemented in Gaussian 09.[37] Harmonic frequency calculations were conducted for all tautomers to estimate thermochemical corrections to the DFT electronic energies. These calculations also generated harmonic vibrational spectra for the tautomers of $(C + H)^+$, $(T + H)^+$, and $(U + H)^+$ for comparison with the experimental IRMPD spectra reported in reference [20] as a means of validating our computational methodology. Using the calculated standard Gibbs' energies, the various tautomers were sorted energetically to determine the species most likely to be present in the probed ensembles. The four lowest energies tautomers of each protonated nucleobase were then carried forward for treatment with the coupled cluster single, double, and perturbative triple excitations method (*i.e.*, CCSD(T)/6-311++G(d,p) level of theory). These improved electronic energies were combined with the DFT thermochemical corrections to produce the standard Gibbs' energies that we report in this manuscript. Calculated structures and thermodynamic data are provided in the supporting information.

3. RESULTS

3.1 Protonated Cytosine, $(C + H)^+$

The dispersion plots recorded for $C \cdot H^+$ (m/z 112) in a pure N_2 environment, and in environments seeded with 1.5 % (mole ratio) MeOH or IPA vapor are shown in Figures 2A–C, respectively. Three major peaks are observed for $(C + H)^+$, thus suggesting the presence of at least three different tautomers in the sample. Upon seeding the DMS cell with MeOH vapor, the observed Type B ion trajectories are deflected to lower values of CV compared to those observed in the pure N_2 environment.[38-40] This is expected due to the ion-solvent clustering interactions experienced by the $(C + H)^+$ tautomers in the modified environment.[41] When the stronger clustering IPA modifier is introduced to the DMS cell, the tautomers all adopt Type A behavior. Interestingly, the two weakest features diminish in intensity and disappear by SV ≈ 2500 V in the IPA-modified environment. This is likely an indication of *in situ* tautomerization induced by the IPA clustering,[39] or of proton scavenging by the IPA at higher values of SV (*i.e.*, IPA has a higher gas phase basicity than those two tautomers at high field). The errors (2σ) given on the dispersion plots are determined from Gaussian fits of the peaks observed in the associated ionograms. An example is provided in Figure 2D, which plots the ionogram recorded when monitoring the m/z 112 peak with the separation voltage set to SV = 3500 V.

The three lowest energy tautomers of $(C + H)^+$ as identified by CCSD(T)//B3LYP calculations are shown in the inset of Figure 2. These are the same lowest energy species reported by Salpin *et al*. in reference [20]. It is worth highlighting the fact that Salpin *et al*. observed only two tautomers via IRMPD, whereas our DMS results show three peaks in the ionogram for m/z 112. There are several possible explanations for this discord, which include the fact that different ion sources might produce different population distributions and that similar structures (*e.g.*, $(C + H)^+$ isomers 1 and 3) might exhibit similar spectra in the region studied by Salpin *et al*. Indeed, the calculated vibrational spectra for isomers 1 and 3 exhibit very similar spectra in the $1000 – 2000 \text{ cm}^{-1}$ region. However, to explore further the properties of the separated species, we examined the CID behavior of the $(C + H)^+$ species associated with each peak in the ionogram. Figure 3 plots the breakdown curves obtained for each of the three peaks shown in the ionogram in Figure 2D. As expected, $(C + H)^+$ exhibits fragmentation channels associated with loss of NH_3, H_2O, and HNCO.[42-44] While all three DMS-separated species displayed the same fragmentation channels in roughly the same distributions, peak I showed an onset of fragmentation 4–5 V lower in energy than peaks II

and III. Although we expect that these measurements are somewhat coarse, they are reproducible and provide some evidence that the structure associated with peak I in the ionogram is, indeed, a unique tautomeric species.
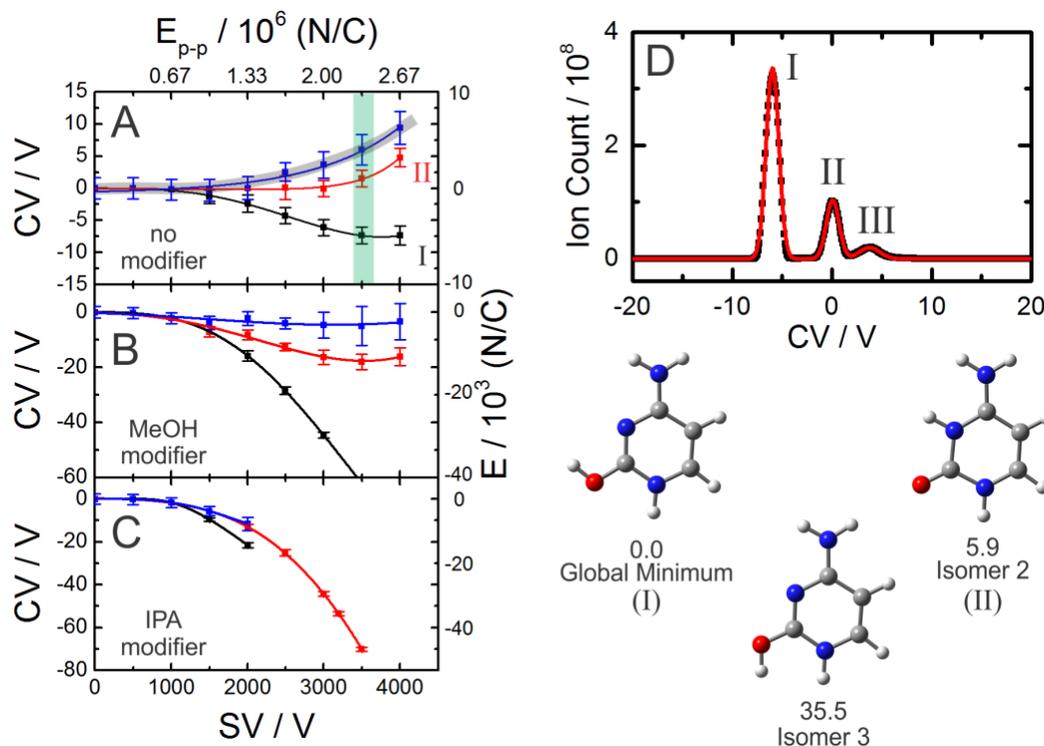
Figure 2. The dispersion plot obtained for $(C + H)^+$ (m/z 112) with a DMS cell containing (**A**) a pure N_2 environment, and a N_2 environment seeded with 1.5% (mole ratio) (**B**) methanol vapor, and (**C**) isopropyl alcohol vapor. Error bars are 2σ obtained from Gaussian fits to the ionogram peaks. (**D**) The ionogram recorded for the m/z 112 peak in a pure N_2 environment with SV = 3500 V (highlighted green in **A**). (**Inset**) The three lowest energy tautomers of $(C + H)^+$ as calculated at the CCSD(T)/6-311++G(d,p)//B3LYP/6-311++G(d,p) level of theory. Energies are reported as standard Gibbs' energies in kJ mol^{-1}.
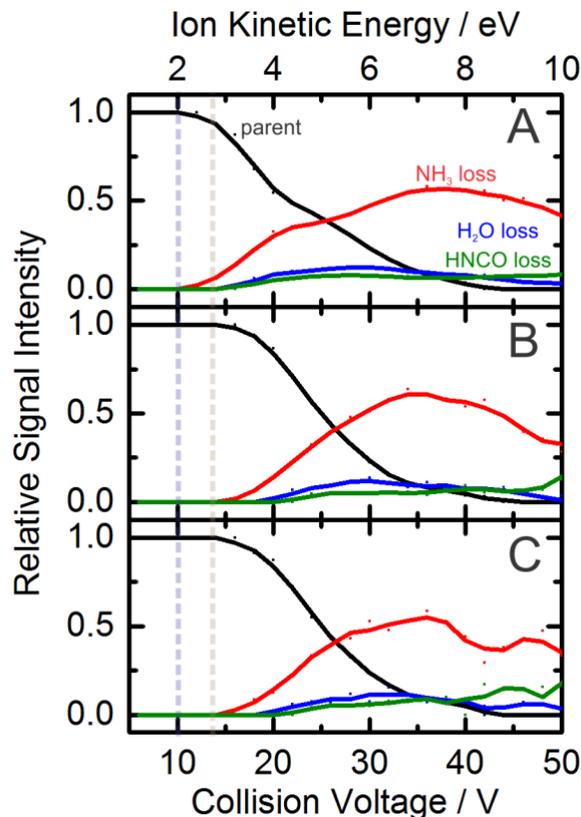
Figure 3. The breakdown curves obtained for $(C + H)^+$ (m/z 112) when isolating on the ionogram peaks plotted in Figure 2D. **(A)** peak I (CV = −6 V), **(B)** peak II (CV = 1 V), and **(C)** peak III (CV = 4 V). Collision energy was increased in 2.5 V increments from 0 to 50 V. Ion kinetic energy is calculated with respect to the center of mass frame for the collision partners.

To further probe the separated species, the relative rates of HDX for each $(C + H)^+$ structure were assessed.[32, 45] Figure 4 plots the results of the HDX experiments for $(C + H)^+$. It is clear from Figures 4B-D that the $(C + H)^+$ species associated with the three peaks in the ionogram all exhibit different HDX behavior. Although $(C + H)^+$ has four exchangeable protons, in all three cases we see little evidence of HDX of the third and fourth nuclei. When gating the DMS on peak I, rapid exchange of one H atom is observed under both the high and low D_2O partial pressure conditions. This accords with the expected acidity of the protonated carbonyl of the global minimum structure. The structure associated with peak II shows a relatively low rate of exchange for the first H atom, going to completion under high D_2O partial pressure, but remaining relatively unaffected in the low pressure D_2O environment. This might be evidence for saturation of the HDX environment under high partial pressure, or evidence of HDX-induced tautomerization.[32] The

lower rate of HDX for peak II accords with the lower relative acidity expected for the protonated nitrogen atoms of isomer 2.
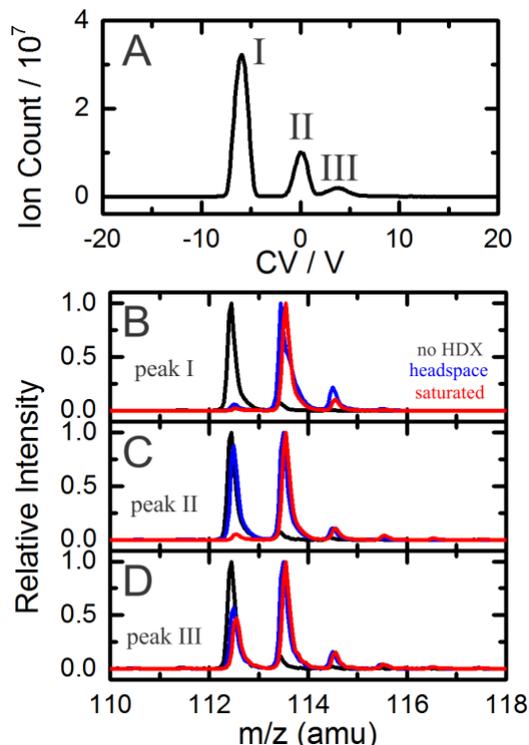
Figure 4. (**A**) The ionogram recorded when gating on $(C + H)^+$ (m/z 112). The results of HDX experiments when isolating on (**B**) peak I, (**C**) peak II, and (**D**) peak III. The black traces show the observed mass distributions in the absence of HDX reagent. The blue traces show the effect of introducing a low vapor pressure of D_2O, and the red traces are observed following HDX in N_2 at 18 °C seeded with a saturated partial pressure of D_2O.

The HDX profile of peak III is somewhat difficult to interpret in the context of peaks I and II. Unlike the species associated with peaks I and II, the structure responsible for peak III does not undergo complete exchange of the first proton in the saturated D_2O environment. This suggests that the rate of HDX for peak III is lower than that of peaks I and II. However, in the low pressure D_2O environment, the rate of HDX for peak III is higher than that of peak II. Moreover, the HDX profile for peak III is very nearly identical in both the high and low D_2O partial pressure environments. This intriguing behavior led us to rethink the DMS sampling procedure, and specifically, the identity of the molecules giving rise to the $(C + H)^+$ (m/z 112) peak in the mass spectrum. We began by investigating ESI source conditions, but ultimately found that there was little variation in the observed ionogram across the parameter range that we explored (*e.g.*, solution flow rate, ESI voltage, etc.). However, significant variation of the ionogram was observed upon

variation of the instrument declustering potential (DP). Figure 5 plots the ionogram recorded for C•H$^+$ (m/z 112) as DP is stepped from 0–300 V in 50 V increments. At DP = 0 V, four peaks are observed in the ionogram – a new, weak feature is observed at CV = −3 V. As DP is increased, peaks I and II, and the new feature at CV = −3 V deplete, indicating that the (C + H)$^+$ parent ion is fragmenting at the higher DP voltages. This behavior is expected since this voltage-ramp experiment can be viewed as a coarse version of CID; following selection by the DMS, the ions are accelerated/heated by the DP en route to the mass analyzer. The use of high DP voltages has previously been employed to activate ions prior to MS analysis in a form of ersatz "in-source MS/MS" fragmentation.[46] Interestingly, peak III, which showed the anomalous HDX behavior, *grows* in intensity with increasing DP voltage. We attribute this behavior to the fragmentation of larger clusters to produce (C + H)$^+$. This is supported by the mass spectra at low DP voltages where signals are observed at masses corresponding to (C + H)$^+$ clustered with, *e.g.*, water and formic acid. If a larger cluster fragments post-DMS to yield (C + H)$^+$, the trajectory of that larger cluster will appear in the (C + H)$^+$ ionogram and dispersion plot. Consequently, care must be taken to ensure that species separated by the DMS cell are attributed to the correct parent ion. This is an important consideration that has gone unaddressed in DMS-based studies to date. Ion behavior as a function of DP seems to be a satisfactory means of distinguishing target ions from larger clusters which contain the ions of interest. Note that there is a slight initial increase in the signals attributed to the bare (C + H)$^+$ ions at low DP voltages; we interpret this depletion at low DP voltages to clustering with trace amounts solvent vapor in the region between the DMS and the mass spectrometer (see Figure 1).
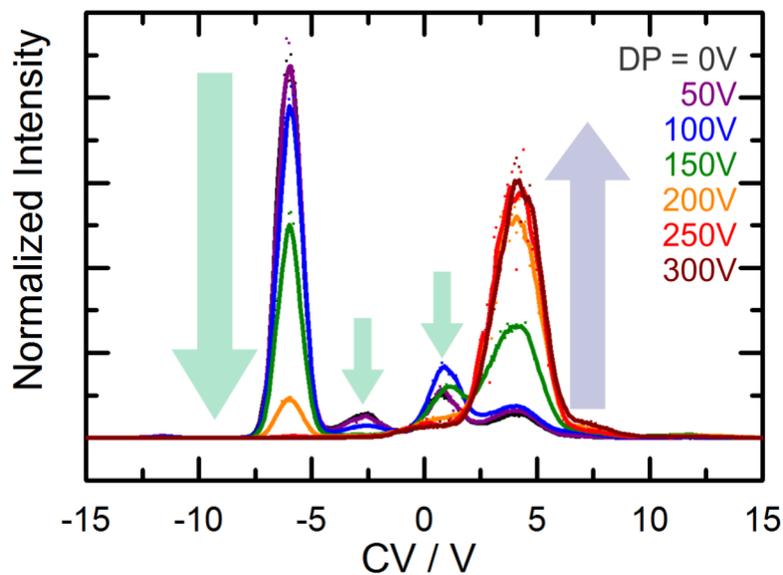
Figure 5. The ionogram recorded for $C \cdot H^+$ (m/z 112) in a pure N_2 environment at $SV = 3000$ V as the declustering potential, DP, is stepped from 0–300 V in 50 V increments.

Based on the analysis described above, we assign peaks I and II, and the weak feature at $CV = -3$ V (observed at low DP) to the three lowest energy isomers of $(C + H)^+$ (shown in Figure 2; highlighted with green downward arrows in Figure 5). These species are completely resolved by DMS, and they exhibit different breakdown curves and HDX profiles. Our calculations also suggest that the two lowest energy tautomers observed here are the same structures that were identified by Salpin *et al*. with IRMPD spectroscopy.[20] Note that these two species differ structurally by a simple proton transfer between adjacent proton acceptor sites. This proton transfer process is easily facilitated by a relay mechanism via an intermolecular hydrogen-bonding network with a protic solvent molecule.[32, 47-49] For example, the two lowest energy proton-bound cytosine-methanol complexes, $(C + H)^+ \cdot (MeOH)_n$ ($n = 1,2$), are shown in Figure 6. Given that these structures are likely to be present in a methanol-modified DMS environment, and that similar structures are likely to be present in the high partial pressure HDX environment, interconversion of isomers 1 and 2 is likely to occur prior to MS characterization under these conditions. This interpretation is supported by the loss in signal intensity for the weaker features in the alcohol-modified environments. Similar observations were previously reported in our study of protonated 4-aminobenzoic acid.[32]
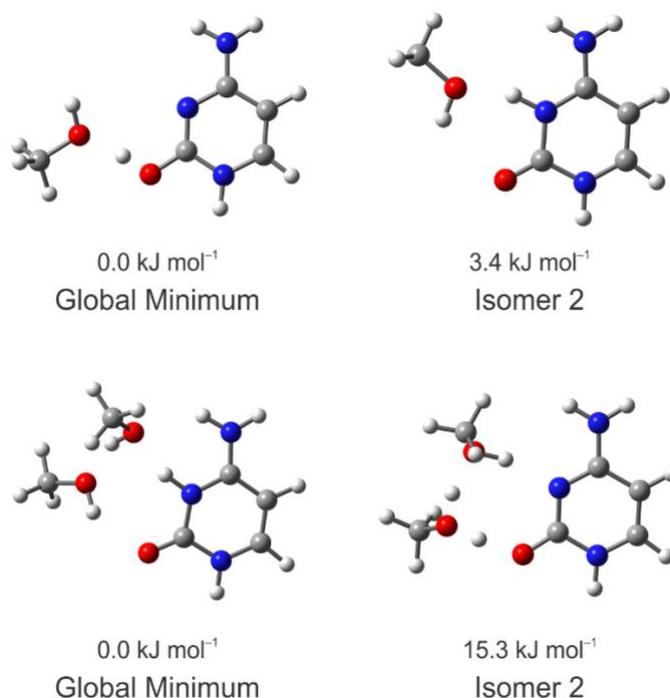
Figure 6. The lowest energy proton-bound cytosine-methanol clusters, $(C + H)^+ \cdot MeOH$ and $(C + H)^+ \cdot (MeOH)_2$. Proton transfer along the intermolecular hydrogen-bond network could facilitate interconversion of the two lowest energy isomers of $(C + H)^+$. Standard Gibbs' energies were calculated at the B3LYP/6-311++G(d,p) level of theory.

3.2 $(A + H)^+$, $(T + H)^+$, $(U + H)^+$, and $(G + H)^+$

Having established a method to clearly identify bare- and clustered-ion signal in the DMS data, we proceeded to conduct analogous studies for $(A + H)^+$, $(T + H)^+$, $(U + H)^+$, and $(G + H)^+$. The dispersion plots that were recorded for these species in an unmodified N_2 environment are plotted in Figure 7. The dispersion plots for the MeOH- and IPA-modified environments, the DP-scans, the HDX data, and the breakdown curves for these species are all available in the supporting information that accompanies this manuscript. The greyed-out traces in Figure 7 are associated with larger clusters, which produce the ion of interest upon fragmentation post-DMS cell. When running with a declustering potential of DP = 150 V (the standard instrument setting), we can clearly resolve two tautomers for $(A + H)^+$, $(T + H)^+$, and $(U + H)^+$, and three tautomers for $(G + H)^+$ (see supporting information for additional details). This accords with the work of Salpin *et al.*, who identified contributions from two tautomers in the IRMPD spectra of $(T + H)^+$, and $(U + H)^+$.[20] Moreover, as was the case with $(C + H)^+$, trace amounts of a third tautomeric species are observed for $(A + H)^+$, $(T + H)^+$, and $(U + H)^+$ under low declustering potential conditions (see DP

scans in supporting information). This suggests that $(C + H)^+$, $(A + H)^+$, $(T + H)^+$, and $(U + H)^+$ all have two tautomeric forms at relatively low energy, and a third at higher energy which fragments under low-to-moderate declustering potential voltages. In the case of $(G + H)^+$, a single tautomer accounts for most (*ca.* 80%) of the total ion signal. Two other weaker features persist to higher DP voltages (*ca.* 150-200 V), indicating the presence of two additional higher-energy tautomers in the ensemble (see Figure S26). At low DP, these two features exhibit intensities that are approximately equal, but the structure associated with the CV = −1 V ionogram peak (at SV = 3500 V) depletes at a substantially higher rate as DP is increased.
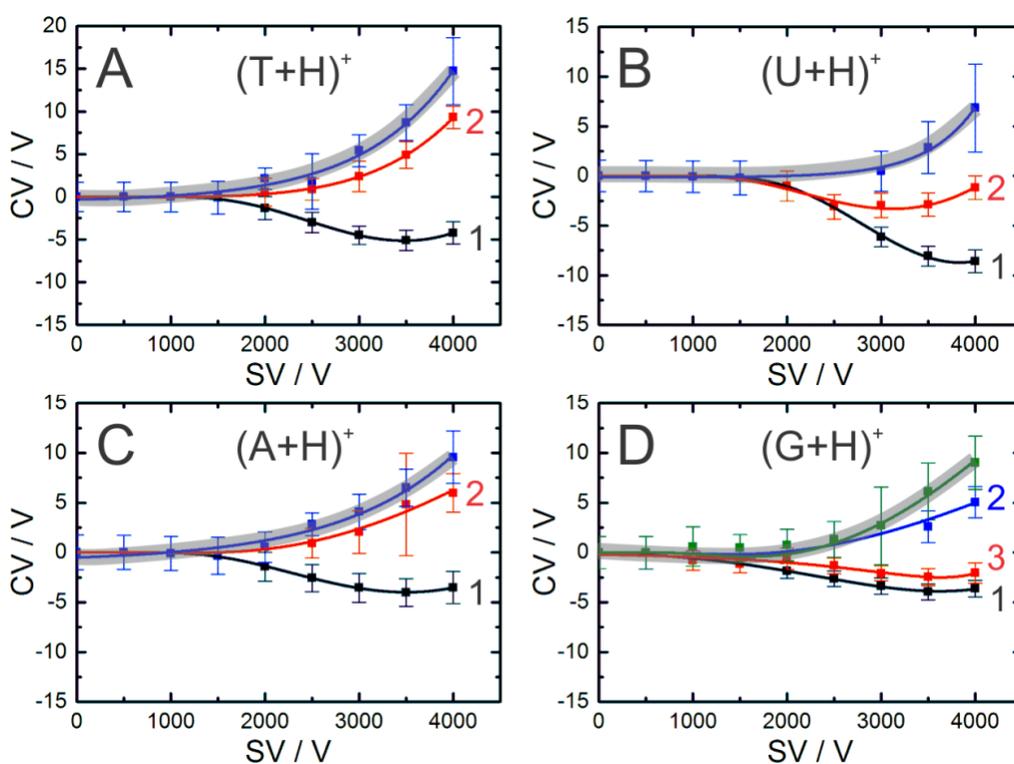
Figure 7. The dispersion plots obtained for (**A**) $(T + H)^+$ (m/z 117), (**B**) $(U + H)^+$ (m/z 113), (**C**) $(A + H)^+$ (m/z 136), and (**D**) $(G + H)^+$ (m/z 152) for a pure N_2 environment with DP set to 150 V. Curves that are greyed-out are associated with larger clusters which fragment to produce the ion of interest post-DMS. The numeric labels indicate the tautomer associated with a particular dispersion plot (see Figure 8).

To estimate the relative standard Gibbs' energies of the various protonated nucleobase tautomers, electronic structure calculations were undertaken at the CCSD(T)/6-311++G(d,p)//B3LYP/6-311++G(d,p) level of theory. The four lowest energy tautomers of the protonated nucleobases are shown in Figure 8. As expected based on the experimental results, (C

+ H)$^+$, (T + H)$^+$, and (U + H)$^+$ exhibit two tautomers at relatively low energy, which can interconvert via solvent-mediated proton-transfer between adjacent basic sites on the nucleobase. The third lowest energy tautomer for these species lies at least 20 kJ mol^{-1} above the global minimum structure. This suggests that the higher energy tautomers are kinetically trapped during the ESI process,[26, 48] since they are expected to have a negligible contribution to the ensemble population in a stochastic Boltzmann distribution at the experimental temperature and pressure (see Table 1). In the case of (A + H)$^+$, calculations suggest that three low-energy tautomers are likely to be observed in the probed ensemble. Our DMS experiments clearly separate two structures. At low declustering potential we observe an asymmetry/skewing to the low-CV side of the dominant ionogram peak (see Figure S25), which suggests the presence of a third, higher-energy tautomer, but attempts to fully resolve this signal have been unsuccessful. For (G + H)$^+$, our calculations indicate that the population distribution should be dominated by a single tautomer (as was observed experimentally). The second and third lowest energy tautomers of (G + H)$^+$ are calculated to lie 16.9 kJ mol^{-1} and 19.1 kJ mol^{-1} above the global minimum, respectively. This suggests that the two weak features observed in the (G + H)$^+$ ionogram are metastable species which were kinetically trapped during production. To compare the calculated fractional populations to those observed experimentally, the ionogram peaks were fit to Gaussian distributions and peak areas were extracted. The relative population percentages of the lowest energies tautomers for each protonated nucleobase are reported in Table 1. These values are provided for experiments where DP was set to 0 V and 150 V to illustrate how sensitive nucleobase tautomer populations are to the declustering potential. Although experiment and theory are in relatively good agreement when it comes to (U + H)$^+$ tautomer populations, for the most part there are significant differences between the observed and calculated tautomer populations of the protonated nucleobases. This, taken together with the tautomer population variability as a function of DP, indicates that the gas phase ensembles of the protonated nucleobases are generated and trapped in non-equilibrium conditions, and that these populations can be manipulated post-production via instrument conditions. It is also worth noting that, in the case of (T + H)$^+$, there is a significant difference between the relative tautomer populations at DP = 0 V and DP = 150 V (see Table 1). This is likely due to loss of the global minimum tautomer signal due to ion-solvent clustering at low DP since the expected relative populations are re-established at DP = 150 V.
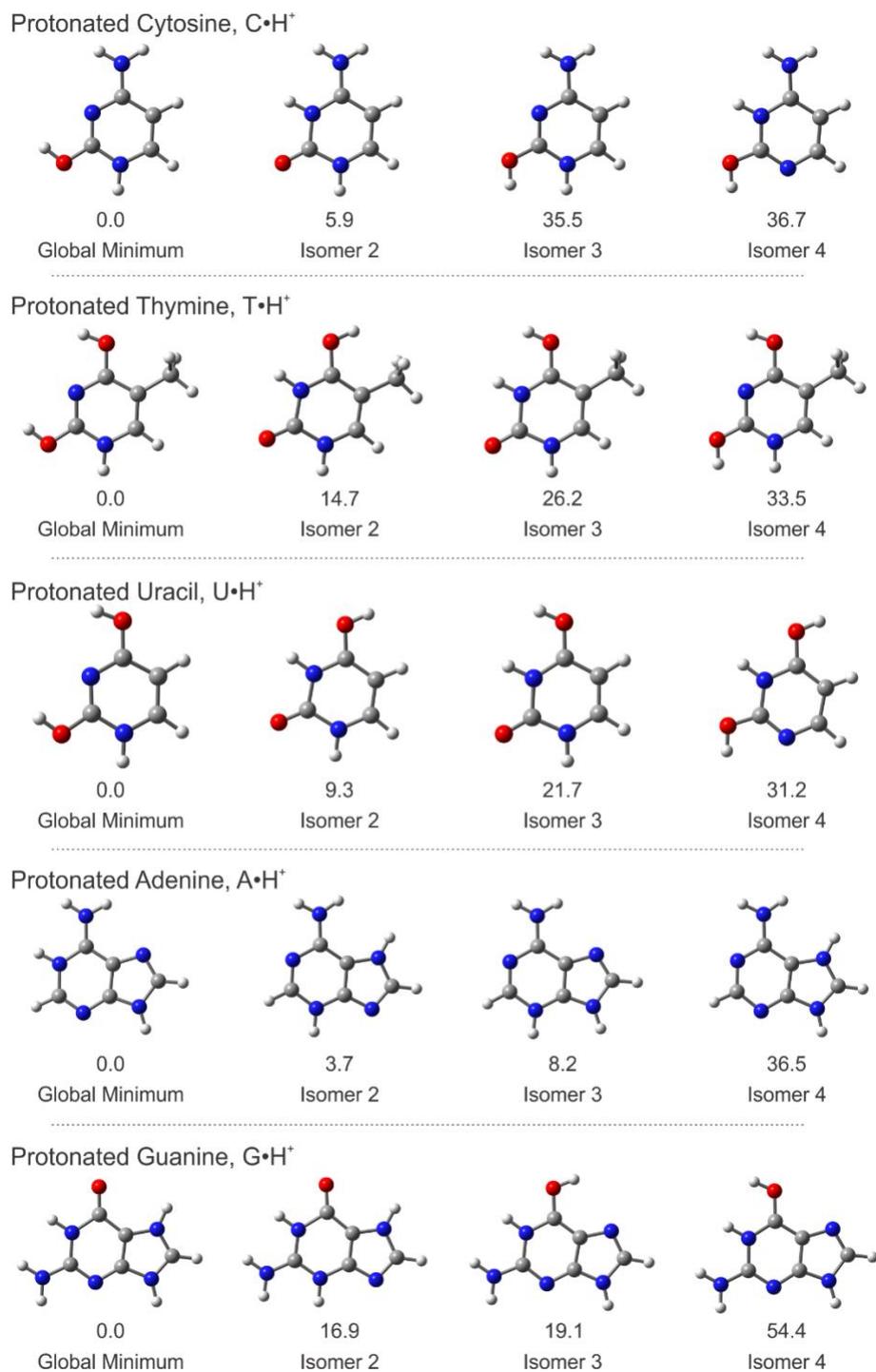
Figure 8. The four lowest energy tautomeric forms of $(C + H)^+$, $(T + H)^+$, $(U + H)^+$, $(A + H)^+$, and $(G + H)^+$. Electronic energies were calculated at the CCSD(T)/6-311++G(d,p) level of theory and thermochemical corrections were calculated at the B3LYP/6-311++G(d,p) level of theory. Standard Gibbs' energies are reported in kJ mol^{-1}.

Table 1. The percentage of the ensemble population of the three lowest energy tautomeric forms of $(C + H)^+$, $(T + H)^+$, $(U + H)^+$, $(A + H)^+$, and $(G + H)^+$. Experimental populations were determined by Gaussian fits of the ionogram distributions recorded at DP = 0 V and at DP = 150 V. Errors (1σ) are reported in parentheses. The calculated population percentages are based on relative standard Gibbs' energies which were calculated at the CCSD(T)//B3LYP level of theory employing a 6-311++G(d,p) basis set.

Species	Percentage of Population		
	DP = 0 V	DP = 150 V	Calculated
$(C + H)^+$			
Global Minimum	79.5 (0.4)	73.5 (0.5)	91.5
Tautomer 2	8.1 (0.5)	26.5 (0.9)	8.5
Tautomer 3	12.4 (0.5)	0.0	0.0
$(A + H)^+$			
Global Minimum	73.6 (4.2)	100.0	79.3
Tautomer 2	17.7 (1.2)	0.0	17.8
Tautomer 3	8.7 (4.1)	0.0	2.9
$(G + H)^+$			
Global Minimum	74.7 (0.4)	79.5 (0.4)	99.8
Tautomer 2	10.8 (4.5)	20.5 (1.3)	0.1
Tautomer 3	14.6 (1.8)	trace	0.0
$(T + H)^+$			
Global Minimum	30.9 (0.5)	86.6 (0.9)	99.7
Tautomer 2	39.4 (1.4)	13.4 (1.0)	0.3
Tautomer 3	29.7 (1.5)	0.0	0.0
$(U + H)^+$			
Global Minimum	91.1 (0.3)	94.6 (0.3)	97.7
Tautomer 2	6.2 (0.3)	5.4 (0.4)	2.3
Tautomer 3	2.7 (0.5)	0.0	0.0

4. CONCLUSIONS

Differential mobility spectrometry has been used to separate the various tautomers present in gas phase ensembles of $(C + H)^+$, $(T + H)^+$, $(U + H)^+$, $(A + H)^+$, and $(G + H)^+$. We find that these populations are dominated by contributions from a single tautomeric species for each protonated nucleobase, but the observation of weaker features in the DMS data indicates that higher-energy, metastable species are present in the sample. The unique identity of these tautomers has been confirmed via differences in the observed HDX profiles and breakdown curves, and structures are assigned based on CCSD(T)//B3LYP calculations. The observation of multiple protonated nucleobase tautomers is in accordance with an IRMPD spectroscopy study by Salpin *et al.*, who identified contributions from at least two tautomeric forms of $(C + H)^+$, $(T + H)^+$, and $(U + H)^+$.[20] Supporting electronic structure calculations yield results for relative tautomer energies which are in general agreement with the observed populations, and we find the same low energy structures for $(C + H)^+$, $(T + H)^+$, and $(U + H)^+$ as were determined spectroscopically.[20] This suggests that DMS could find use as a pre-filter to separate tautomeric species prior to laser interrogation, thereby facilitating deconvolution of spectra. Indeed, similar methodology has already been employed in studies of isomeric lipid and saccharide species.[50, 51]

It is, however, important to note that the DMS data for all five protonated nucleobases exhibited spurious ion signals that arose from fragmentation of larger nucleobase-containing clusters following DMS-separation and prior to MS detection. These species could not be distinguished from the bare protonated nucleobases simply by monitoring the ion signals corresponding to the bare protonated nucleobase or its product ions. Instead, we could identify contributions from larger clusters by monitoring ion signal as post-DMS DP voltages were ramped. The ion signals of the bare protonated nucleobases increased slightly at low DP voltages (DP < 50 V), and then decreased dramatically in intensity as the DP voltage was ramped up to 300 V. It should be noted that depletion of the various tautomer signals did not occur at the same rate, presumably due to relative differences in fragmentation energies. We are currently investigating this more detail. In contrast, the ion signals arising from fragmentation of larger clusters continued to increase in intensity well into the DP ramp, depleting only at very high voltages (DP > 200 V). This is an important consideration for future experiments wherein DMS is used to separate isomers, conformers, or tautomers prior to spectroscopic or mass spectrometric characterization, since ion populations that

are produced from declustering processes are likely to yield convoluted results owing to contributions from multiple structures.

ACKNOWLEDGEMENTS

APPENDIX A. SUPPORTING INFORMATION

The supporting information associated with this article can be found, in the online version, at [insert web link]

REFERENCES

[1] M. Kubota, T. Kobayashi, J. Electron Spectrosc. Relat Phenomena, 82 (1996) 61-70.

[2] J.S. Kwiatkowski, J. Leszczynski, J Mol. Struct., 376 (1996) 325-342.

[3] J. Smets, L. Adamowicz, G. Maes, J. Phys. Chem., 99 (1995) 6387-6400.

[4] L.C. Sowers, G.V. Fazakerley, R. Eritja, B.E. Kaplan, M.F. Goodman, Proc. Nat. Acad Sci., 83 (1986) 5434-5438.

[5] M.T. Baei, M.R. Taghartapeh, E.T. Lemeski, A. Soltani, Phys. B, 444 (2014) 6-13.

[6] C.X. Dong, J.C. Yang, H.M. Ning, C.P. Li, J. Mol. Struct. Theochem, 950 (2010) 64-71.

[7] T. Fornaro, M. Biczysko, S. Monti, V. Barone, Phys. Chem. Chem. Phys., 16 (2014) 10112-10128.

[8] S. Panigrahi, A. Bhattacharya, S. Banerjee, D. Bhattacharyya, J. Phys. Chem. C, 116 (2012) 4374-4379.

[9] R.R. Sinden, DNA Structure and Function, Academic Press Inc., San Diego, 1994.

[10] M.D. Frank-Kamenetskii, S.M. Mirkin, Annu. Rev. Biochem., 64 (1995) 65-95.

[11] A.K. Chandra, D. Michalska, R. Wysokinsky, T. Zeegers-Huyskens, J. Phys. Chem. A, 108 (2004) 9593-9600.

[12] L. Di Donna, A. Napoli, G. Sindona, C. Athanassopoulos, J. Am. Soc. Mass Spectrom., 15 (2004) 1080-1086.

[13] E.S. Kryachko, M.T. Nguyen, T. Zeegers-Huyskens, J. Phys. Chem. A, 105 (2001) 1288-1295.

[14] M.A. Kurinovich, L.M. Phillips, S. Sharma, J.K. Lee, Chem. Commun., (2002) 2354-2355.

[15] H. Moustafa, S. El-Taher, M.F. Shibl, R. Hilal, Int. J. Quantum Chem., 87 (2002) 378-388.

[16] M.T. Nguyen, A.K. Chandra, T. Zeegers-Huyskens, J. Chem. Soc. Faraday Trans., 94 (1998) 1277-1280.

[17] N. Russo, M. Toscano, A. Grand, F. Jolibois, J. Comput. Chem., 19 (1998) 989-1000.

[18] C. Colominas, F.J. Luque, M. Orozco, J. Am. Chem. Soc., 118 (1996) 6811-6821.

[19] Y. Podolyan, L. Gorb, J. Leszczynski, J. Phys. Chem. A, 104 (2000) 7346-7352.

[20] J.Y. Salpin, S. Guillaumont, J. Tortajada, L. MacAleese, J. Lemaire, P. Maitre, Chem. Phys. Chem., 8 (2007) 2235-2244.

[21] J.M. Bakker, J.Y. Salpin, P. Maitre, Int. J. Mass Spectrom., 283 (2009) 214-221.

[22] J.M. Bakker, R.K. Sinha, T. Besson, M. Brugnara, P. Tosi, J.Y. Salpin, P. Maitre, J. Phys. Chem. A, 112 (2008) 12393-12400.

[23] Z. Karpas, Z. Berant, R.M. Stimac, Struct. Chem., 1 (1990) 201-204.

[24] P.M. Lalli, B.A. Iglesias, H.E. Toma, G.F. de Sa, R.J. Daroda, J.C. Silva, J.E. Szulejko, K. Araki, M.N. Eberlin, J. Mass Spectrom., 47 (2012) 712-719.

[25] D. Schroder, M. Budesinsky, J. Roithova, J. Am. Chem. Soc., 134 (2012) 15897-15905.

[26] H.X. Xia, A.B. Attygalle, Anal. Chem., 88 (2016) 6035-6043.

[27] E.G. A., Z. Karpas, Ion Mobility Spectrometry, 3rd ed., CRC Press, Boca Raton, FL, 2014.

[28] S.A. A., Differential Ion Mobility: Non-linear Ion Transport and Fundamentals of FAIMS, CRC Press, Boca Raton, FL, 2009.

[29] J.L. Campbell, M. Zhu, W.S. Hopkins, J. Am. Soc. Mass Spectrom., 25 (2014) 1583-1591.

[30] E.V. Krylov, E.G. Nazarov, R.A. Miller, Int. J. Mass Spectrom., 266 (2007) 76-85.

[31] J.L. Campbell, J.C.Y. Le Blanc, B.B. Schneider, Anal. Chem., 84 (2012) 7857-7864.

[32] J.L. Campbell, A.M.C. Yang, L.R. Melo, W.S. Hopkins, J. Am. Soc. Mass Spectrom., 27 (2016) 1277-1284.

[33] R.W. Purves, R. Guevremont, Anal. Chem., 71 (1999) 2346-2357.

[34] W.S. Hopkins, Mol. Phys., 113 (2015) 3151-3158.

[35] D.S. Levin, P. Vouros, R.A. Miller, E.G. Nazarov, J.C. Morris, Anal. Chem., 78 (2006) 96-106.

[36] K. Vekey, J. Mass Spectrometry, 31 (1996) 445-463.

[37] 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, et al., Gaussian 09 Revision D.01, in: Gaussian, Inc. Wallingford CT 2009.

[38] B.B. Schneider, T.R. Covey, S.L. Coy, E.V. Krylov, E.G. Nazarov, Anal. Chem., 82 (2010) 1867-1880.

[39] B.B. Schneider, T.R. Covey, E.G. Nazarov, Int. J. Ion Mobil. Spectrom., 16 (2013) 207-216.

[40] B.B. Schneider, E.G. Nazarov, T.R. Covey, Int. J. Ion Mobil. Spectrom., 15 (2012) 141-150.

[41] B.B. Schneider, E.G. Nazarov, F. Londry, T.R. Covey, Int. J. Ion Mobil. Spectrom., 18 (2015) 159-170.

[42] P. Cheng, Y.N. Li, S.Q. Li, M.T. Zhang, Z. Zhou, Phys. Chem. Chem. Phys., 12 (2010) 4667-4677.

[43] C.C. Nelson, J.A. McCloskey, J. Am. Chem. Soc., 114 (1992) 3661-3668.

[44] C.C. Nelson, J.A. McCloskey, J. Am. Soc. Mass Spectrom., 5 (1994) 339-349.

[45] V. Blagojevic, D.K. Bohme, Int. J. Mass Spectrom., 378 (2015) 180-185.

[46] J.A. Buckley, J.B. French, N.M. Reid, Can. Aeronaut. Sp. J., 20 (1974) 231-233.

[47] G.C.J. T., Ann. Chim., 58 (1806) 54-73.

[48] W.S. Hopkins, R.A. Marta, V. Steinmetz, T.B. McMahon, Phys. Chem. Chem. Phys., 17 (2015) 28548-28555.

[49] V.R.I. Kaila, G. Hummer, Phys. Chem. Chem. Phys., 13 (2011) 13207-13215.

[50] D.S. Levin, P. Vouros, R.A. Miller, E.G. Nazarov, J. Am. Soc. Mass Spectrometry, 18 (2007) 502-511.

[51] T.P.I. Lintonen, P.R.S. Baker, M. Suoniemi, B.K. Ubhi, K.M. Koistinen, E. Duchoslav, J.L. Campbell, K. Ekroos, Anal. Chem., 86 (2014) 9662-9669.