ACS Partner Journal

Journal of the American Society for Mass Spectrometry

pubs.acs.org/jasms

Article

Do d(GCGAAGC) Cations Retain the Hairpin Structure in the Gas Phase? A Cyclic Ion Mobility Mass Spectrometry and Density Functional Theory Computational Study

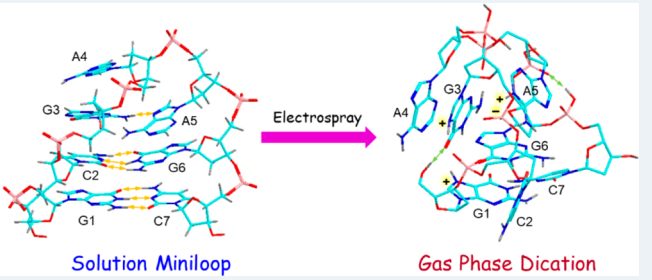
Jiahao Wan, Marianna Nytka, Haocheng Qian, Karel Lemr,* and František Tureček*

Cite This: *J. Am. Soc. Mass Spectrom.* 2023, 34, 2323−2340

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: d(GCGAAGC) is the smallest oligonucleotide with a well-defined hairpin structure in solution. We report a study of multiply protonated d(GCGAAGC) and its sequence-scrambled isomers, d(CGAAGCG), d(GCGAACG), and d(CGGAAGC), that were produced by electrospray ionization with the goal of investigating their gas-phase structures and dissociations. Cyclic ion mobility measurements revealed that dications of d(GCGAAGC) as well as the scrambled-sequence ions were mixtures of protomers and/or conformers that had collision cross sections (CCS) within a 439−481 Å^2 range. Multiple ion conformers were obtained by electrospray under native conditions as well as from aqueous methanol. Arrival time distribution profiles were characteristic of individual isomeric heptanucleotides. Extensive Born−Oppenheimer molecular dynamics (BOMD) and density functional theory (DFT) calculations of d(GCGAAGC)$^{2+}$ isomers indicated that hairpin structures were high-energy isomers of more compact distorted conformers. Protonation caused a break up of the C2···G6 pair that was associated with the formation of strong hydrogen bonds in zwitterionic phosphate anion-nucleobase cation motifs that predominated in low energy ions. Multiple components were also obtained for d(GCGAAGC)$^{3+}$ trications under native and denaturing electrospray conditions. The calculated trication structures showed disruption of the G···C pairs in low energy zwitterions. A hairpin trication was calculated to be a high energy isomer. d(GCGAAGC)$^{4+}$ tetracations were produced and separated by c-IMS as two major isomers. All low energy d(GCGAAGC)$^{4+}$ ions obtained by DFT geometry optimizations were zwitterions in which all five purine bases were protonated, and the ion charge was balanced by a phosphate anion. Tetracations of the scrambled sequences were each formed as one dominant isomer. The CCS calculated with the MobCal-MPI method were found to closely match experimental values. Collision-induced dissociation (CID) spectra of multiply charged heptanucleotides showed nucleobase loss and backbone cleavages occurring chiefly at the terminal nucleosides. Electron-transfer-CID tandem mass spectra were used to investigate dissociations of different charge and spin states of charge-reduced heptanucleotide cation radicals.



INTRODUCTION

The question of biomolecular ion structure upon transitioning from solution to the gas phase has been of interest to the research community ever since it became possible to obtain large gas-phase ions by electrospray ionization.[1,2] Multiply charged protein cations, in particular, have been a frequent subject of structure-related studies using H/D exchange[2] and ion mobility.[3−8] The major factors affecting protein ion stability in the gas phase have been identified as being due to Coulomb repulsion and the internal solvation of polar and charged groups. It has been recognized that the experimental methods used to tease out gas-phase protein ion structures have limited resolution and are often complemented by molecular dynamics modeling.[9] Compared with peptides and proteins, much less is known about the gas-phase structures of oligonucleotide ions. Due to their acid−base nature, oligonucleotides have been mostly studied as gas-phase polyanions using methods such as temperature induced melting,[10] collisional activation,[11−13] and ion mobility.[14−18] These and related studies have concluded that some features of the oligonucleotide solution structure may be retained in the gas-phase anions.[14,19−21] Even less is known about gas-phase structures of multiply charged oligonucleotide cations, despite that multiple protonation of oligonucleotides by electrospray is remarkably effective, involving both the phosphate groups and nucleobases.[22−27] An ion mobility study of GAA·TCC triplex ions concluded on the basis of force-field molecular dynamics calculations that the oligonucleo-

Received: June 26, 2023
Revised: August 31, 2023
Accepted: September 4, 2023
Published: September 11, 2023

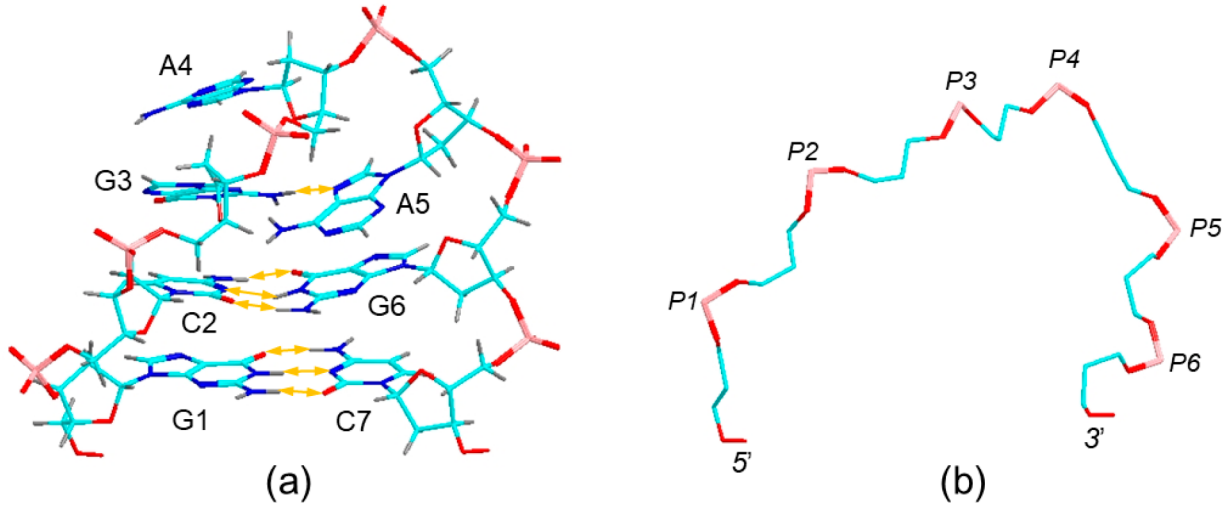


Figure 1. (a) d(GCGAAGC) solution structure from refined NMR analysis.[32] Atom color coding is as follows: cyan = C, blue = N, red = O, pink = P, gray = H. Hydrogen bonds between nucleobases are shown as double-headed ochre arrows. (b) Extracted d(GCGAAGC) backbone consisting of O5′-C5′-C4′-C3′-O3′-P segments starting from the 5′ terminus.

tide cations retained their solution structure.[28] Remarkably, these authors considered only cytosine-protonated tautomers in their structure assignment.[28] Protonation affects the electrostatic interactions in the gas-phase ions in several ways. Negatively charged phosphates are converted to neutral phosphoester groups that can provide hydroxyl groups to participate in hydrogen bonding networks.[29] Protonation of nucleobases, with the exception of the less basic thymine, affects their proton donor and acceptor sites that can disrupt standard hydrogen bonding patterns, such as the Watson−Crick (WC) pairwise G-C and A-T interactions. In addressing these questions at the fundamental level, it is desirable to investigate oligonucleotide models that have well-established structures in solution and can be transferred into the gas phase in the form of multiply charged cations. At the same time, a suitable model should have a limited size to be amenable to a rigorous computational investigation of energy and structure that can be utilized to obtain experimentally relevant parameters, such as collision cross sections (CCS). The d(GCGAAGC) miniloop, also called the d(GCGAAGC) hairpin, appeared to be such a suitable model. The miniloop conformation in the condensed phase has been established by X-ray crystallography[30] and NMR spectroscopy,[31] and the structure details have been refined using NMR with residual dipolar couplings.[32] The structure is maintained by WC pairs between terminal G1-C7 and next inner C2-G6, which are supported by π-stacking with G3 and A4 (Figure 1a).

The deoxyribose-phosphate backbone makes a smooth loop that allows the WC pairs to assume the energetically most favorable positions (Figure 1b). Despite its thermodynamic stability in solution, the hairpin can be disrupted when attached to the 3′-terminus of a larger oligonucleotide and hybridized with a complementary strand.[33] Size-wise, the doubly through quadruply protonated forms of d(GCGAAGC) consist of 230−232 atoms, which make them amenable to a computational analysis by Born−Oppenheimer molecular dynamics (BOMD) and gradient geometry optimization using quantum chemical methods. Here, we report a combined experimental and computational study of di-, tri-, and tetracations produced by electrospray protonation of d(GCGAAGC) under native conditions of 5 mM ammonium acetate and from aqueous methanol. As reference models, we include an investigation of polycations produced from modified heptanucleotide sequences. The reverse sequence, d(CGAAGCG), may theoretically include two WC G-C pairs, although its condensed-phase structure has not been established. Two other sequences that we investigated, d(GCGAACG) and d(CGGAAGC), have the terminal G and C nucleobases in "wrong" mutual positions and thus do not allow WC base pairing in a loop structure. We show that these gas-phase cations can develop a number of canonical and zwitterionic structures within a broad range of relative Gibbs energies that can be characterized by comparing experimental and theoretical collision cross sections. We also report collision-induced dissociation (CID) and tandem electron-transfer-CID (ET-CID-MS^3) spectra of multiply charged heptanucleotide cations to elucidate the effects of sequence and charge on their gas-phase ion chemistry.

EXPERIMENTAL SECTION

Materials. Heptanucleotides d(GCGAAGC), d(CGAAGCG), d(GCGAACG), and d(CGGAAGC) were custom synthesized by Integrated DNA Technologies, Inc. (Coralville, IA), and their quality and sequences were checked by tandem mass spectrometry.

Methods. Tandem mass spectra were measured on a Bruker amaZon 3D ion trap mass spectrometer furnished with an auxiliary chemical ionization source for electron and proton transfer studies.[34] Multiply charged ions were generated by electrospray ionization of 10−20 μM solutions in 50:50:1 methanol−water−acetic acid and selected by mass. Collision-induced dissociations were performed with r.f. excitation amplitudes that were adjusted to achieve >50% dissociation of the precursor ions. Electron transfer dissociation was carried out with fluoranthene anion radicals ($C_{16}H_{10}^{-\bullet}$) at 100−150 ms ion−ion reaction times. Proton transfer reactions were performed with fluoranthene anions ($C_{16}H_9^-$) at 100−150 ms reaction times. Ion mobility measurements were made on a Waters Select Series Cyclic IMS Q-TOF instrument equipped with a cyclic traveling wave device (Waters Corp., Wilmslow, U.K.).[35,36] Sample solutions were directly infused into the electrospray ion source at a rate of 5 μL/min. The instrumental parameters were set as follows: capillary voltage 2 kV, cone

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

voltage 40 V, source offset 10 V, source temperature 100 °C, desolvation temperature 280 °C, cone gas 30 L/h, desolvation gas 600 L/h, nebulizer gas 6.0 bar, step wave body gradient 20 V, head gradient 10 V, trap DC bias 35 V, transfer CE 4 V, helium flow rate 120 mL/min, nitrogen flow rate 40 mL/min, racetrack bias 70 V (for other parameters see Table S1−S8). The precursor ions were selected by a quadrupole (highmass 15, lowmass 4.9). The IMS experiment sequence consisted of injection, separation, and ejection steps. The separation time was kept at 2 ms. Five TOF pushes per data bin were set for all measurements. Ions for IMS measurements were generated from both 50:50:1 methanol/water/acetic acid and 5 mM ammonium acetate solutions. The latter were used to mimic "native" ionization conditions that are believed to preserve solution conformations of biomolecules.[9] The instrument control was carried out by Masslynx 4.2 with modified DriftScope 2.9 (all Waters Corp., Wilmslow, U.K.) to analyze data. CCS values were determined as averages of six single-pass ion mobility measurements with manual calibration using a standard Waters procedure based on a logarithmic fit[37−39] (Table S2, Figure S1, Supporting Information). Arrival time profiles obtained by single-pass c-IMS measurements were least-squares fitted with multiple Gaussian peaks using the Excel solver. The fitting parameters were constrained to accommodate the increasing peak broadening due to diffusion as a function of arrival time. The CCS for the maxima of the fitted peaks was calculated by using the experimental calibration curve. The fitted data were compared with arrival time distributions from multipass measurements, increasing the ion path in c-IMS from 98 cm for 1 pass up to 490 cm for 5 passes. The maximum number of passes was chosen to avoid wrap-around overlaps.

Calculations. Conformational analysis was carried out by Born−Oppenheimer molecular dynamics calculations (BOMD) as described previously.[29] Briefly, the semiempirical PM6 method[40] that was augmented by dispersion and hydrogen bonding interactions (D3H4)[41] was used to treat electron energy. The total energy stability was maintained by the Berendsen thermostat.[42] Initial structures of d(GCGAAGC) cations were derived from the reported Cartesian coordinates of the DNA miniloop that were obtained from the NMR spectra.[32] Protons were placed on the phosphate groups and guanine and adenine nucleobases to generate +2 through +4 charge states. On the basis of previous energy analysis of gas-phase di-,[24] tri-,[29] and tetranucleotide cations,[26,27] we considered protonation at N7 of guanine and N3 of adenine as the most energetically favorable sites. This created 10 initial combinations of doubly charged protomers that are denoted as G1G3, G1G6, G3G6, G1A4, G1A5, G3A4, G3A5, A4G6, A5G6, and A4A5. Trication protomers were considered for the initial G1G3G6, G1G3A4, G3A4G6, and G1A4A5 combinations. Tetracation calculations started from the G1G3A4G6 protomer. BOMD trajectories were run for all of these combinations for 20 ps at 300, 500, 600, and 800 K using the high-end Cuby4 platform.[43] Since BOMD included both electron and nuclear motion, the initial proton positions were not fixed and the ions were allowed to isomerize by proton migration in the course of the BOMD trajectory, as indeed observed and discussed later in the paper. It should be noted, however, that in none of the low-energy structures did we observe proton migration onto a cytosine ring. The 20000 snapshots obtained by BOMD for each initial structure and temperature were used to extract 200 snapshots at regular intervals that were then fully gradient optimized with PM6-D3H4 run under MOPAC.[44] Structures optimized in this cycle were sorted out to remove duplicates and high-energy ions, and the selected candidates were considered for geometry optimization with density functional theory (DFT) calculations. In contrast to our previous experience with DFT calculations of di-, tri-, and tetranucleotide mono and dications[24−27,29] using hybrid density functionals such as B3LYP,[45] M06-2X,[46] and ωB97-XD[47,48] with the 6-31+G(d,p) basis set, these calculations were failing for all charge states of the heptanucleotide ions. The failure typically consisted of a poor wave function and energy gradient guess when starting the next optimization cycle, which resulted in a sudden jump of energy and a distorted structure. Due to the size of these systems (230−232 atoms) and the 6-31+G(d,p) basis set (3176-3187 basis functions), the calculations became intractable. Fortunately, we found out that including in B3LYP calculations empirical corrections of dispersion energy with Becke−Johnson damping (GD3-BJ)[49] resulted in smooth convergence of geometry and energy. Geometry optimizations were therefore run with B3LYP/6-31G(d,p) + GD3-BJ that also provided harmonic frequencies for selected ions. Single-point energies were calculated with B3LYP/6-31++G(d,p) + GD3-BJ that also provided atomic charge densities for collision cross section calculations. The reliability of the B3LYP/6-31G(d,p) + GD3-BJ data was checked by benchmarking against M06-2X/6-31+G(d,p) calculations that were run for a series of trinucleotide cations where relative energies obtained by both methods showed the same trend.[29] The reasons for the failure of standard DFT geometry optimizations of heptanucleotide cations were not clear but may warrant further analysis. These calculations were carried out with Gaussian 16 (Revision B.01) that was licensed from Gaussian Inc. (Wallingford, CT). The fully optimized geometries were used in calculations of Mulliken and dipole restricted B3LYP-GD3-BJ/6-31++G(d,p) atomic charge densities according to the Mertz−Singh−Kollman (MK) scheme.[50] The ion trajectory method[51] was used to calculate CCS in nitrogen. We used two different sets of parameters; in one we employed the parameters reported by Campuzano et al.[52] and Kim et al.[53] using Mulliken charge densities. The other was based on MK charge densities using the MobCal-MPI package of Ieritano, Hopkins et al.[54,55] In general, we found the CCS calculated by MobCal-MPI to give closer fits than those from standard MobCal.

RESULTS AND DISCUSSION

Ion Formation and Characterization. Electrospray ionization of aqueous methanol solutions of all heptanucleotide sequences under study produced doubly, triply, and quadruply protonated ions at m/z 1065.7, 710.8, and 533.4, respectively. The relative intensities of these multiply charged ions were tuning-dependent and optimized for measurements involving the given charge state. We note that the extent of multiple protonation by electrospray was remarkable, considering that in aqueous solution the heptanucleotides exist as polyanions, so the formation of +4 states in gas-phase ions may require a take up of up to 10 protons. Increased acidity in electrospray microdroplets, as first reported in 1994,[56] may in part account for the facile multiple protonation of the heptanucleotides. Electrospray ionization of d(GCGAAGC) and d(CGAAGCG) solutions in 5 mM ammonium acetate under "native" conditions[7] gave doubly and triply charged ions along with quadruply charged dimers. Quadruply charged monomers were formed at very low intensities under native conditions. The cations were selected

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

by mass and examined by cyclic ion mobility measurements, CID-MS^2 and ET-CID-MS^3 spectra of pertinent charge states.

Ion Mobility Separations and CCS. Heptanucleotide ions in the +2, +3, and +4 charge states were investigated by ion mobility measurements of arrival times using the TWAVE-cyclic IMS. The four sequences showed different arrival time profiles, as discussed for different charge states. The +2 state of the miniloop sequence, d(GCGAAGC)$^{2+}$, showed a composite peak with arrival times between 36 and 42 ms for the single-pass measurement (Figure 2a). The composite arrival time profile

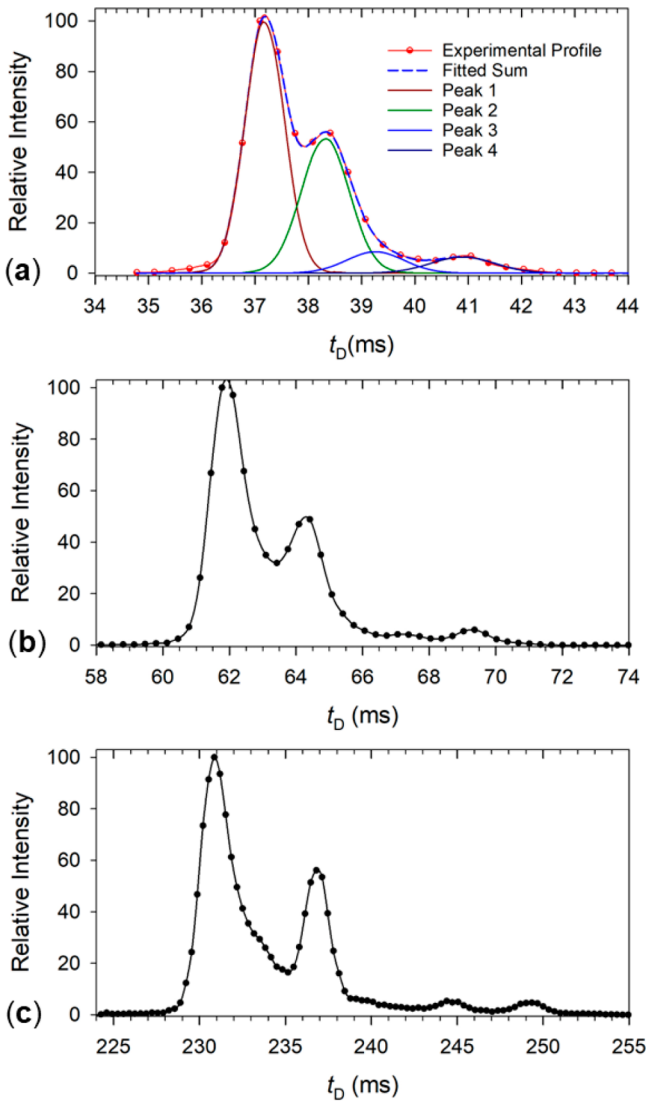


Figure 2. Arrival time profiles from c-IMS of d(GCGAAGC)$^{2+}$ dications after (a) one, (b) two, and (c) five passes. The ion population in (c) was stored after two passes and after slicing reinjected for an additional five c-IMS passes. Gaussian fits of partially resolved and unresolved peaks are shown in (a).

was deconvoluted by fitting with Gaussian peaks having maxima at t_D = 37.2, 38.3, and 40.9 ms and constituting 53, 35, and 7% of the total integrated dication intensity. An improved fit to the experimental profile reaching rmsd = 0.5% was obtained by including another Gaussian at t_D = 39.3 ms (7% intensity), resulting in a near-perfect overlap. The accuracy of the Gaussian fit in Figure 2a was probed by c-IMS measurements using 2 and 5 passes (Figure 2b and 2c, respectively). These showed improved separation of the four peaks from the single-pass measurement (Figure 2a) that appeared with similar relative intensities. A possible fifth component was indicated in Figure 2c as a shoulder at 233.5 ms. d(GCGAAGC)$^{2+}$ ions produced under "native" conditions by electrospray from 5 mM aqueous ammonium acetate displayed arrival time profiles that were virtually identical with those obtained for ions from aqueous-methanol solutions. The c-IMS data of d(GCGAAGC)$^{2+}$ that were obtained over a period of several months differed in the relative intensity of the band for dimeric tetracations that were readily distinguished by their isotope pattern and did not interfere with the dications (Figure S2a,b, Supporting Information). In contrast, the composite peak of the dication arrival time distribution was unchanged when the ions were produced by electrospray from either aqueous methanol or ammonium acetate solution. This indicated that the electrospray solvent, either "denaturing" or "native", had a negligible effect on the formation of different ion isomers. We further used a composite c-IMS sequence consisting of injecting and separating the ions in one or two cycles, which was followed by storing an ion population in a prearray store (slicing)[35] within a narrow arrival-time range. The slicing, i.e., ion mobility isolation of selected ions) eliminated the interference of quadruply charged dimers in multipass measurements. These ions were reinjected into c-IMS and separated in more cycles. The sequence is visualized in Figure S3 (Supporting Information). Comparison of Figure 2b and 2c arrival time distributions indicated that there were no significant changes in the peak relative positions and intensities upon ion manipulation.

The inverted dication, d(CGAAGCG)$^{2+}$, also showed a composite peak consisting of several overlapping components (Figure 3a). In this case, it was more difficult to attain unambiguous deconvolution into individual Gaussian peaks, because of the substantial overlap. Our best fit with rmsd = 0.4% showed three peaks with t_D = 35.8, 36.8, 38.0 ms that gave 8, 16, and 50% of total ion intensity. This fit was improved by including another Gaussian at t_D = 38.2 ms, which amounted to 25% of total dication intensity. Analogous results were obtained for c-IMS of d(CGAAGCG)$^{2+}$ that were produced under "native" conditions that also showed composite peak profiles (Figure S4, Supporting Information). Increasing the number of c-IMS passes revealed the presence of two minor components at longer arrival times (Figure 3b, c). For example, after five passes, the data showed at least four components (Figure 3c). Interestingly, the unresolved shoulder at the lowest t_D in Figure 3a and 3b appeared to merge into the major peak after the slicing and five-pass separation (Figure 3c). This could possibly indicate that the d(CGAAGCG)$^{2+}$ isomers interconvert during the passage through the c-IMS. This possibility was investigated by collisionally activating the population of stored (CGAAGCG)$^{2+}$ ions separated by passage through one cycle of IMS. However, subsequent c-IMS by three cycles of separation of the collisionally activated ions revealed no dependence on the excitation voltage in the range 0−70 V (Figure S5, Supporting Information).

The arrival time profile from c-IMS of d(GCGAACG)^2 dications showed a major component that was accompanied by multiple peaks that were partially resolved in the single-pass scan (Figure 4a). After deconvolution, the arrival time profile of d(GCGAACG)$^{2+}$ was composed of a major peak with a maximum at t_D = 38.6 ms and side peaks at 36.5, 37.4, 39.8, and 41.3 ms (Figure 4a), giving a tight fit with rmsd = 0.2%. The

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

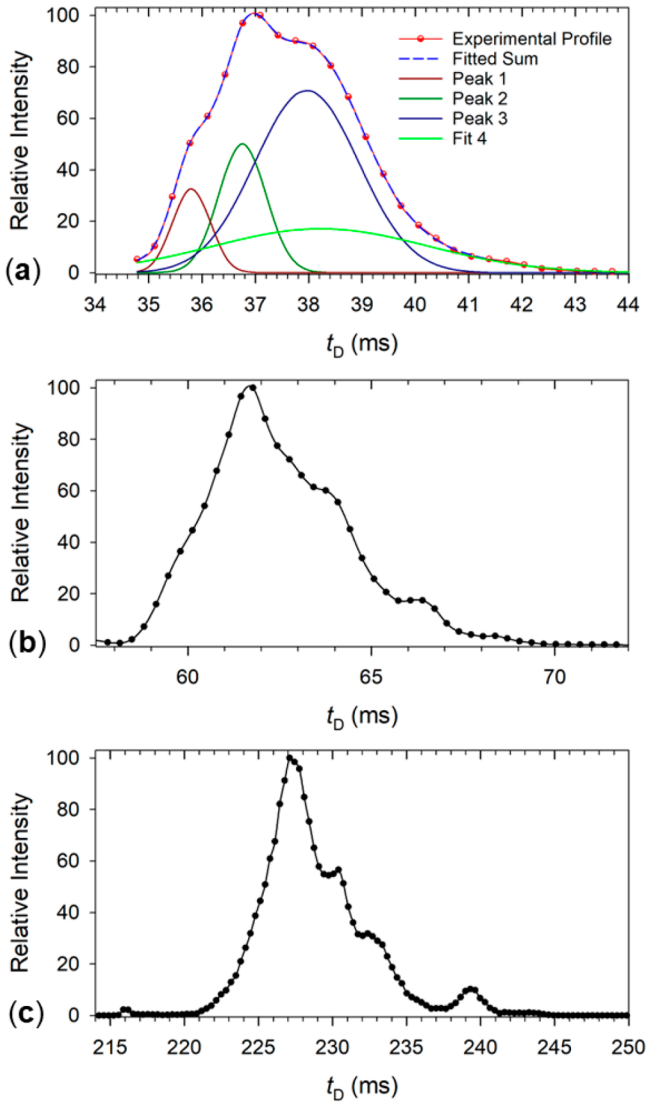


Figure 3. Arrival time profiles from c-IMS of $d(CGAAGCG)^{2+}$ dications after (a) one, (b) two, and (c) five passes. The ion population in (c) was stored (slicing) after two passes and reinjected for an additional five c-IMS passes. Gaussian fits of partially resolved and unresolved peaks are shown in (a).

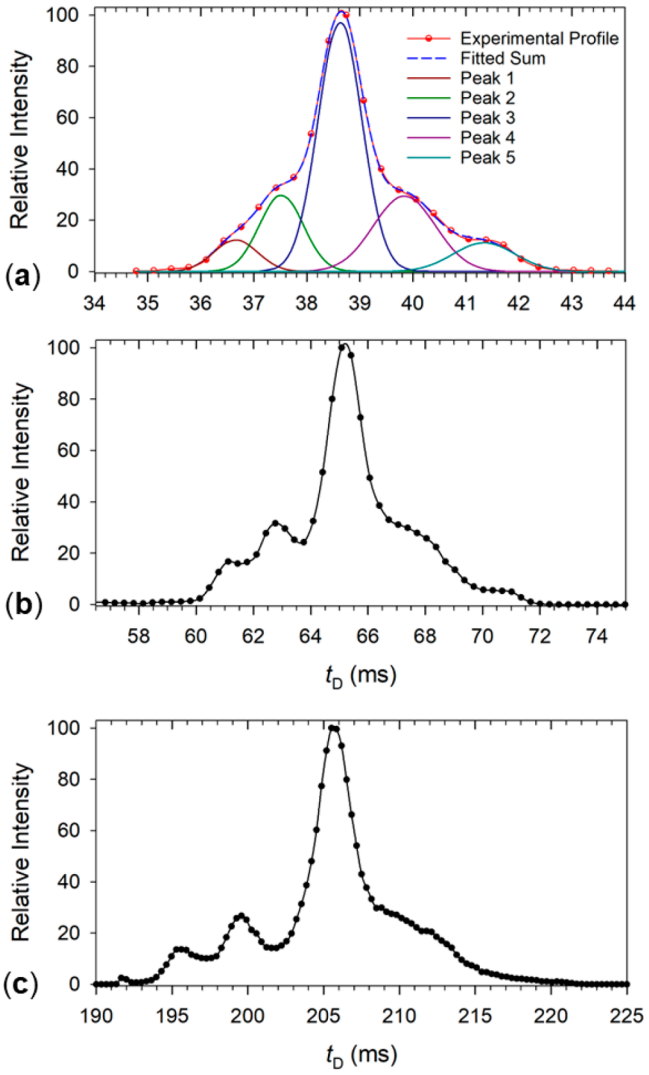


Figure 4. Arrival time profiles from c-IMS of $d(GCGAACG)^{2+}$ dications after (a) one, (b) two, and (c) five passes. The ion population in (c) was stored after one pass (slicing) and reinjected for an additional five c-IMS passes. Gaussian fits of partially resolved peaks are shown in (a).

Gaussian decomposition of the single-pass profile was corroborated by multiple-pass scans shown in Figure 4b, c. After two passes, the first two components were resolved to show two distinct peak maxima, which was further documented by the five-pass scan, where these peak maxima were further separated by 4 ms (Figure 4c). In contrast, the two minor peaks appearing at longer arrival times remained incompletely resolved even after five passes despite their stretched time separation window.

The dication with the other scrambled sequence, $d(CGGAAGC)^{2+}$, also showed a composite profile, consisting of a major component with a maximum at t_D = 38.4 ms and minor components at t_D = 36.9 and 41.0 ms (Figure 5a). Fitting in two more Gaussian peaks with t_D = 38.7 and 41.3 ms allowed us to achieve a rmsd = 0.4%. Arrival time profiles obtained at longer path lengths corroborated the Gaussian fit-assigned peaks for the major components. Extending the ion path length to two and five passes (196 and 490 cm, respectively) allowed us to completely resolve all five components (Figure 5b, c). The only deviation between the Gaussian fit and the resolved scans was in the width of the minor fourth component, which was fitted as a single broad peak. The arrival times of the Gaussian-fitted peaks were converted to CCS that are listed in Table 1. The major components of $d(GCGAAGC)^{2+}$ and $d(CGAAGCG)^{2+}$ showed very similar CCS values that were within the 448−470 Å^2 range. $d(GCGAAGC)^{2+}$ showed two major components with CCS = 450 and 459 Å^2 that when combined, amounted to 88% of the total ion intensity. These were accompanied by two minor components of CCS = 466 and 478 Å^2, at 7% and 5%, respectively. $d(CGAAGCG)^{2+}$ was composed of three major components with CCS = 448, 459, and 470 Å^2 that when combined, accounted for 91% of ion intensity. A minor peak with CCS = 440 Å^2 was distinguished by arrival-time profile deconvolution. Dications of the scrambled sequences each consisted of one major component that had CCS = 461 and 459 Å^2 at 49% and 54% for $d(GCGAACG)^{2+}$ and $d(CGGAAGC)^{2+}$, respectively (Table 1). In both cases, the major isomers were accompanied by several minor components within the 445−479 Å^2 range (Table 1).

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

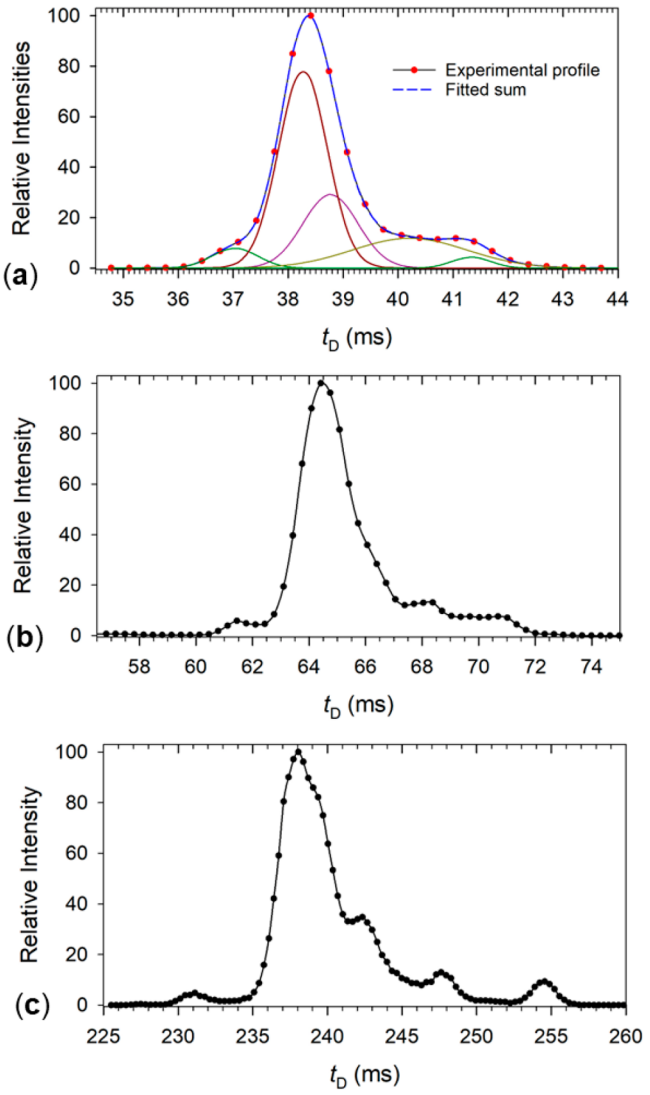


Figure 5. Arrival time profiles from c-IMS of d(CGGAAGC)$^{2+}$ dications after (a) one, (b) two, and (c) five passes. The ion population in (c) was stored after two passes (slicing) and reinjected for an additional five c-IMS passes. Gaussian fits of partially resolved peaks are shown in (a).

The data indicated that the miniloop and inverted-sequence dications had gas-phase structures of quite similar degrees of compactness that arose from their solution conformations or resulted from conformational changes due to protonation in the gas phase. This was irrespective of the electrospray solvent, suggesting that the ion structures were governed by their gas-phase properties or the nature of the charging events in electrospray rather than their solution properties. The other aspect that could be gleaned from the c-IMS data was the heterogeneity of the ion populations for all four heptanucleotide sequences, which was particularly salient for d(GCGAAGC)$^{2+}$ that originated from a single stable miniloop solution structure.

Composite peaks were also obtained by c-IMS of the trications (Figure 6a–d). The miniloop trication gave several components after deconvolution, with a clearly separated major peak at t_D = 25.3 ms and two other peaks at t_D = 26.0 and 26.8 ms. This group was flanked by two minor peaks that were distinguished by Gaussian fitting and had t_D = 24.4 and 27.8 ms. Hence the miniloop trications also showed structure heterogeneity in the gas phase. The inverted trication showed one major peak at 26.9 ms followed by a smaller peak at t_D = 27.8 ms.

Table 1. Collision Cross Sections of Heptanucleotide Ions

	collision cross section[a]	
ion	CCS_{exp}	CCS_{fit}
d(GCGAAGC)$^{2+}$	450, 459, 479	450 (53)[b], 459 (35), 466 (7)[c], 478 (5)
d(CGAAGCG)$^{2+}$	448, 467	440 (8), 448 (16), 459 (50), 470 (25)
d(GCGAACG)$^{2+}$	461, 482	445 (6), 452 (15), 461, (49), 470 (21), 479 (8)
d(CGGAAGC)$^{2+}$	459, 480	448 (5), 459 (54), 462 (24)[c], 468 (8)[c], 479 (8)
d(GCGAAGC)$^{3+}$	510	494 (1), 509 (48), 521 (25), 533 (21), 549 (5)
d(CGAAGCG)$^{3+}$	510, 534, 555	510 (5), 522 (9), 531 (15), 535 (59), 550 (12)
d(GCGAACG)$^{3+}$	523, 565	495 (1), 510 (15), 524 (58), 538 (15), 552 (3), 565 (7)
d(CGGAAGC)$^{3+}$	517, 537, 553	489 (1), 513 (16), 523 (13), 532 (19), 542 (23), 554 (28)
d(GCGAAGC)$^{4+}$	617	584 (2)[b], 608 (46), 624 (51)
d(CGAAGCG)$^{4+}$	587, 624	596 (7), 625 (88), 636 (5)[b]
d(GCGAACG)$^{4+}$	624	593 (0.2)[b], 601 (6)[b], 624 (94)
d(CGGAAGC)$^{4+}$	607, 637	608 (83), 627 (17)

[a]In Å^2. [b]Integrated relative intensities of deconvoluted ion mobility peaks in parentheses. [c]Broad peaks of unresolved components.

The scrambled sequence trication, d(GCGAACG)$^{3+}$, gave a composite peak that was fitted by six Gaussian components at t_D = 24.5, 25.4, and 26.2 (major), 27.1, 28.0, and 28.8 ms. Finally, d(CGGAAGC)$^{3+}$ gave a composite peak consisting of at least three components that were fitted by six Gaussians with t_D = 24.2, 25.5, 26.1, 26.7, 27.3, and 28.1 ms. Multipass measurements gave arrival time profiles that were consistent with the Gaussian fits in the single-pass spectra. For d(CGAAGCG)$^{3+}$, we obtained three well-separated peaks after five passes (Figure S6a, Supporting Information). Three passes were sufficient to resolve the five peaks of d(GCGAACG)$^{3+}$ (Figure S6b) as predicted by Gaussian fitting in Figure 6c. The largely unresolved multiplet of d(CGGAAGC)$^{3+}$ (Figure 6d) showed partial peak separation, although the previously resolved peaks at t_D = 27.3 and 28.1 ms coalesced after four passes (Figure S6c, Supporting Information).

The trication CCS showed a broader range of values compared to those of the dications. The major components from d(GCGAAGC)$^{3+}$ and d(CGAAGCG)$^{3+}$ were assigned CCS at 509 Å^2 and 535 Å^2, respectively (Table 1). The scrambled sequences had a larger CCS at 524 Å^2 and 513–554 Å^2 for the major components of d(GCGAACG)$^{3+}$ and d(CGGAAGC)$^{3+}$, respectively.

In contrast to the di- and trications, all four tetracations displayed simpler arrival time profiles (Figure 7a–d). Thus, d(GCGAAGC)$^{4+}$ gave two nearly resolved peaks of similar relative intensity at t_D = 22.4 and 23.0 ms (Figure 7a), whereas the inverted tetracation, d(CGAAGCG)$^{4+}$, showed one major peak at t_D = 23.0 ms and only minor peaks at t_D = 21.9 and 23.5 ms (Figure 7b). The d(GCGAACG)$^{4+}$ ion gave one dominant peak at t_D = 23.0 ms, which, judged by its width, could consist of two unresolved components of very similar arrival times. Finally, d(CGGAAGC)$^{4+}$ gave one major peak at t_D = 22.4 ms, which was accompanied by a minor component at t_D = 23.1 ms. The major components from d(GCGAAGC)$^{4+}$ had CCS = 608 and 624 Å^2 which was similar to the major peak from d-(CGAAGCG)$^{4+}$ that had CCS = 625 Å^2. The scrambled

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323–2340

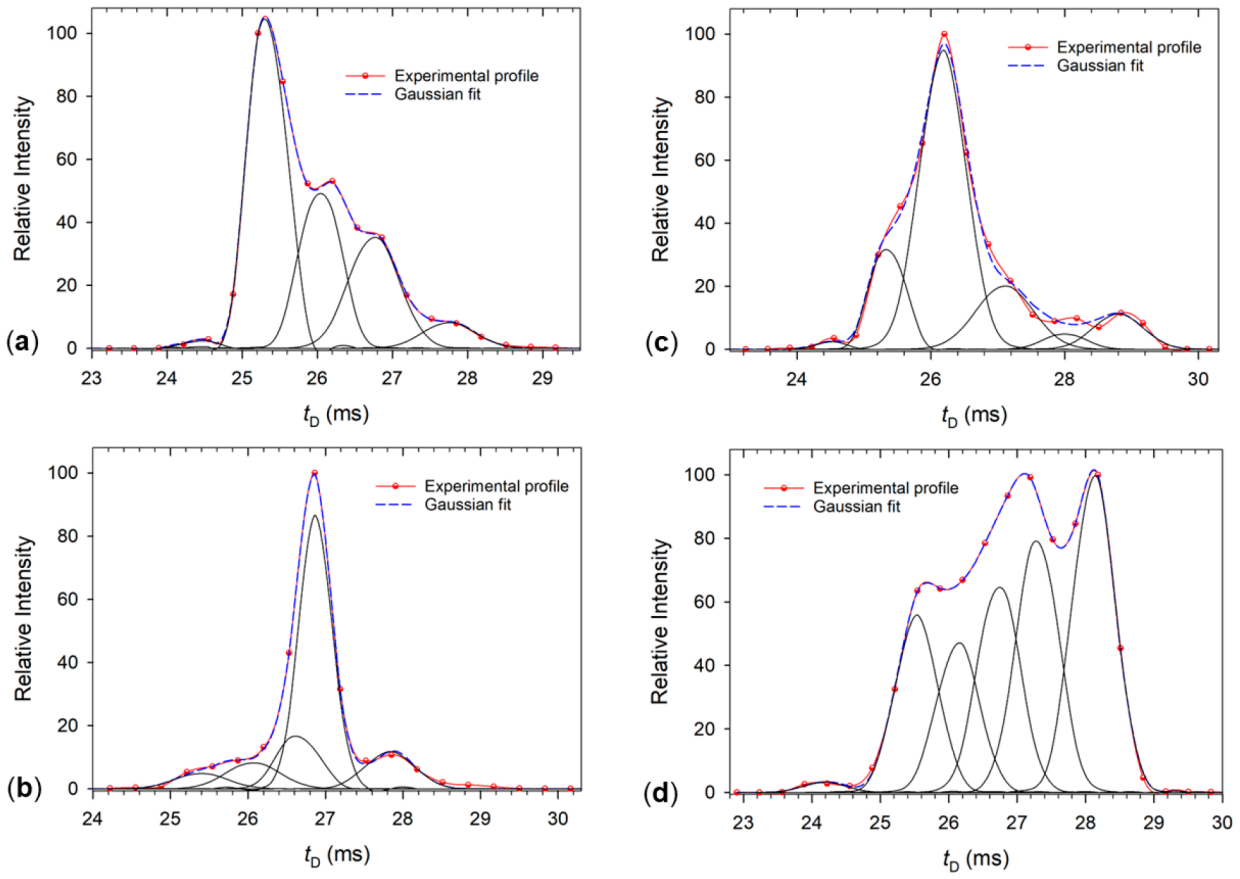


Figure 6. Arrival time profiles from single-pass c-IMS of trications. (a) $d(GCGAAGC)^{3+}$, (b) $d(CGAAGCG)^{3+}$, (c) $d(GCGAACG)^{3+}$, and (d) $d(CGGAAGC)^{3+}$ with best Gaussian fits.

sequences showed chiefly one major peak, which was at CCS = 624 and 608 Å^2 for $d(GCGAACG)^{4+}$ and $d(CGGAAGC)^{4+}$, respectively.

The results of these measurements revealed two facts. First, all charge states of all four sequences consisted of mixtures of isomers, which could be represented by different protomers, conformers, or combinations thereof. Second, the CCS of the miniloop ion components was not significantly different from those of the inverted and scrambled sequences. This raised the question of whether or not the gas-phase ions produced by electrospray retained the main features of the miniloop structure as defined by the two WC G-C pairs.[30−32] It is worth noting that arrival time profiles were characteristic of individual heptanucleotides (cf. Figures 2−7), as earlier observed for oligosaccharides.[57] The differences in their profiles were evident for all followed charge states and were well stable over time. ATD profiles may be useful in distinguishing isomeric oligonucleotides in general.

Ion Structures and Theoretical CCS. Because of the aforementioned challenges to obtaining fully optimized structures for the heptanucleotide ions, we focused on calculations of the d(GCGAAGC) miniloop and d(CGAAGCG) reversed miniloop sequence ions. Optimized structures were obtained for dications, trications, and tetracations that along with atomic charge densities were used to calculate theoretical CCS in nitrogen. The relative energies and theoretical CCS are compiled in Table 2. As mentioned above, the Mobcal-MPI-calculated CCS was smaller than those from the MobCal calculations and, in general, provided better fits with the experimental data. Therefore, we limit the discussion to the MPI theoretical data.

Dications. Out of over two hundred fully optimized $d(GCGAAGC)^{2+}$ dication structures that we obtained, we selected those that had theoretical CCS within 3% of the components that were identified by c-IMS measurements. At the same time, the DFT calculations provided ion relative energies that were used to assess the possibility of protomers and conformers coexisting under equilibrium conditions. The structures are labeled with the initial protonation sites (e.g., G1G3) and temperature that were used in the BOMD trajectory calculations. All the presented optimized

d(GCGAAGC) ion structures are oriented with G1 in the lower left corner and the loop going clockwise. Among the several G1G3 dication structures that were optimized from the 300 K BOMD trajectories, the lowest-energy structure ($\mathbf{G1G3a_{300}}$) preserving the essential features of the miniloop is shown in Figure 8. $\mathbf{G1G3a_{300}}$ showed WC pairing of G1 with C7 and C2 with G6 which was enforced by multiple hydrogen bonds that are visualized by ochre double headed arrows in Figure 8a. Hydrogen bonds involving the backbone are shown as light green arrows (Figure 8a). The deoxyribose-phosphoester backbone formed a flat loop, as shown by the extracted chain of the O5′-C5′-C4′-C3′-O3′-P segments (Figure 8b,c). However, the calculated properties of $\mathbf{G1G3a_{300}}$ strongly indicated that it was not the structure belonging to any of the major observed $d(GCGAAGC)^{2+}$ dications. In particular, $\mathbf{G1G3a_{300}}$ was 243 kJ mol^{-1} less stable than the isomer corresponding to the global energy minimum ($\mathbf{G1G3b_{500}}$), and its CCS_{MPI} (471 Å^2) was larger than the experimental values for the major components

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

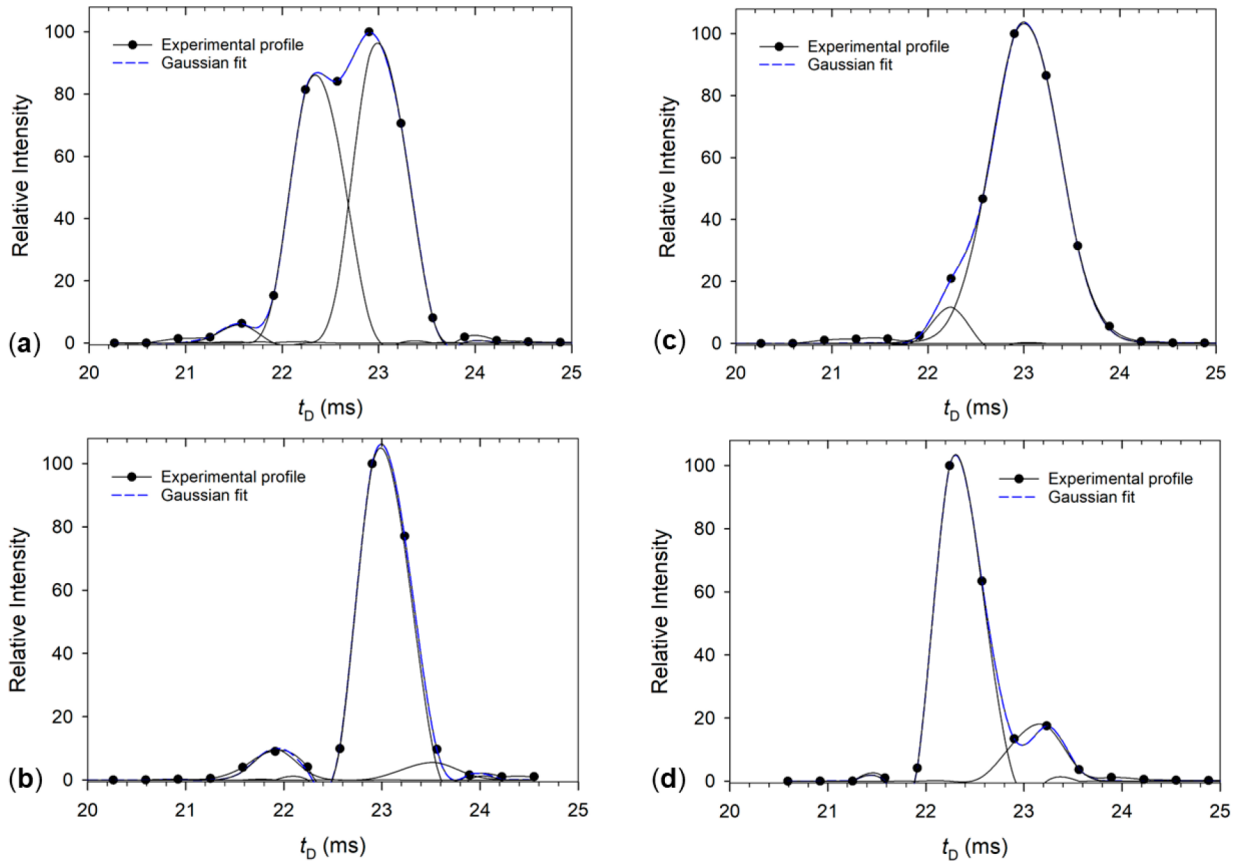


Figure 7. Arrival time profiles from single-pass c-IMS of tetracations. (a) d(GCGAAGC)$^{4+}$, (b) d(CGAAGCG)$^{4+}$, (c) d(GCGAACG)$^{4+}$, and (d) d(CGGAAGC)$^{4+}$.

Table 2. d(GCGAAGC) Dication Relative Energies and Calculated CCS

dication	ΔH_0 [a,b]	ΔG_{310} [a,c]	CCS_{TM} [d,e]	CCS_{MPI} [d,f]
G1A4a$_{500}$	23	19	444	428
G1G6$_{500}$	166	153	446	431
G1A4b$_{500}$	98	95	449	433
G1G3b$_{500}$	0	0	455	445
G1G3c$_{500}$	57	49	469	458
G1A5$_{300}$	115	101	472	458
A4G6$_{600}$	181	172	471	458
G1G3a$_{300}$	255	243	484	471

[a]In kJ mol^{-1}. [b]Including zero-point vibrational energies and referring to 0 K. [c]Including 310 K enthalpies and entropies. [d]In Å^2. [e]From standard ion trajectory calculations using Mulliken charge densities. [f]From MobCal-MPI calculations using MK charge densities.

(Table 1 and 2). The similarity between the calculated CCS of **G1G3a$_{500}$** and that of the minor components in the c-IMS data (CCS_{fit} = 466 Å^2, Δ_{rel} = 1.1%, and CCS_{fit} = 478 Å^2, Δ_{rel} = −1.5%) may indicate a small fraction of kinetically trapped miniloop ions, although we do not have other supporting evidence for such a conclusion.

The lowest-energy isomer (**G1G3b$_{500}$**) had a CCS_{MPI} = 445 Å^2 that closely matched that of the most abundant peak of d(GCGAAGC)$^{2+}$ at a CCS = 450 Å^2 from the ion mobility measurements (Table 1). Upon BOMD, ion **G1G3b$_{500}$** underwent a proton transfer from P2-OH onto N1 in A5, forming a zwitterion structure (Figure 9). The P2 anion was stabilized by hydrogen bonds to protonated G3 and A5. A major distortion in **G1G3b$_{500}$** from the miniloop conformation can be seen in (*i*) the outward rotation of C2 that rearranged its hydrogen bonding to G6, (*ii*) π-stacking of A4 with charged G3, and (*iii*) the absence of π-stacking of A5 with G6. Notably, hydrogen bonds involving the phosphoester and P2-phosphate groups contributed to both a more compact ion structure and very significant energy stabilization compared to the miniloop structure.

Distortion of hydrogen bonding between C2 and G6 was a major feature of the structure of **G1A4a$_{500}$** which was the second lowest-energy isomer at ΔG_{310} = 19 kJ mol^{-1} relative to **G1G3b$_{500}$** (Figure 10). Ion **G1A4a$_{500}$** retained the initial protonation sites at N7 of G1 and at N3 of A4. The ion shape can be seen best in a rotated projection (Figure 10b) that shows C2 that was rotated by almost 90° with respect to G6 and connected to it by a single O2···H−N1 hydrogen bond. Another distinct feature of **G1A4a$_{500}$** was the strong hydrogen bond at 1.354 Å of a Zundel type[58] between the P1-OH and O6 at G6, which was similar to that in **G1G3b$_{500}$**. The CCS_{MPI} of **G1A4a$_{500}$** (428 Å^2) indicated it was the most compact dication structure of those studied; however, this CCS_{MPI} was outside the experimental CCS range from the ion mobility measurements (Table 1). Fits to the second IMS peak of CCS = 459 Å^2 were realized by structures **G1G3c$_{500}$**, **G1A5$_{300}$**, and **A4G6$_{600}$** that each had CCS_{MPI} = 458 Å^2. Ion **G1G3c$_{500}$** which was 49 kJ mol^{-1} above **G1G3b$_{500}$**, was a zwitterion in which the acidic proton from P2-OH migrated to N1 of A5. This isomer showed a distortion of the C2-G6 pair, in which C2 was perpendicular to the G1-C7 pair. In addition, A4, G3, and A5 were arranged in a π-stacking pattern (Figure 11). Ion **G1A5$_{300}$**, had the best preserved

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

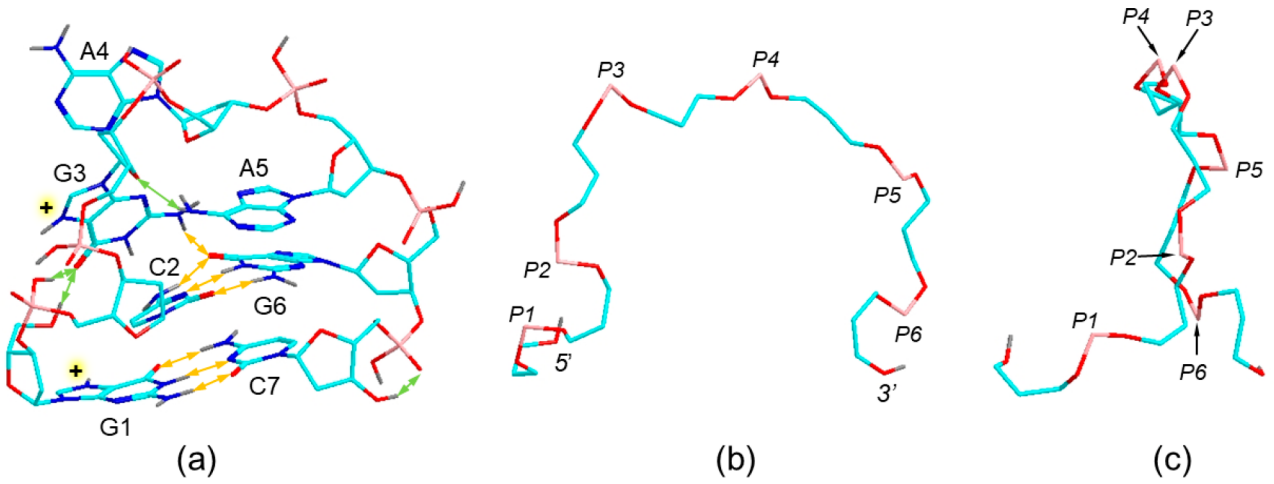


Figure 8. (a) B3LYP/6-31G(d,p) + GD3BJ optimized structure of the d(GCGAAGC)$^{2+}$ miniloop dication isomer **G1G3a**$_{300}$. Atom color coding is as follows: cyan = C, blue = N, red = O, pink = P, gray = H. Only exchangeable hydrogen atoms are shown to reduce clutter. Hydrogen bonds between nucleobases are shown as double-headed ochre arrows. Hydrogen bonds involving the backbone are shown as light green arrows. (b) View of the **G1G3a**$_{300}$ backbone consisting of the **G1G3a**$_{300}$ backbone consisting of the O5−C5−C4−C3−O3−P segments starting from the 5′ terminus. (c) **G1G3a**$_{300}$ backbone rotated by 90°.

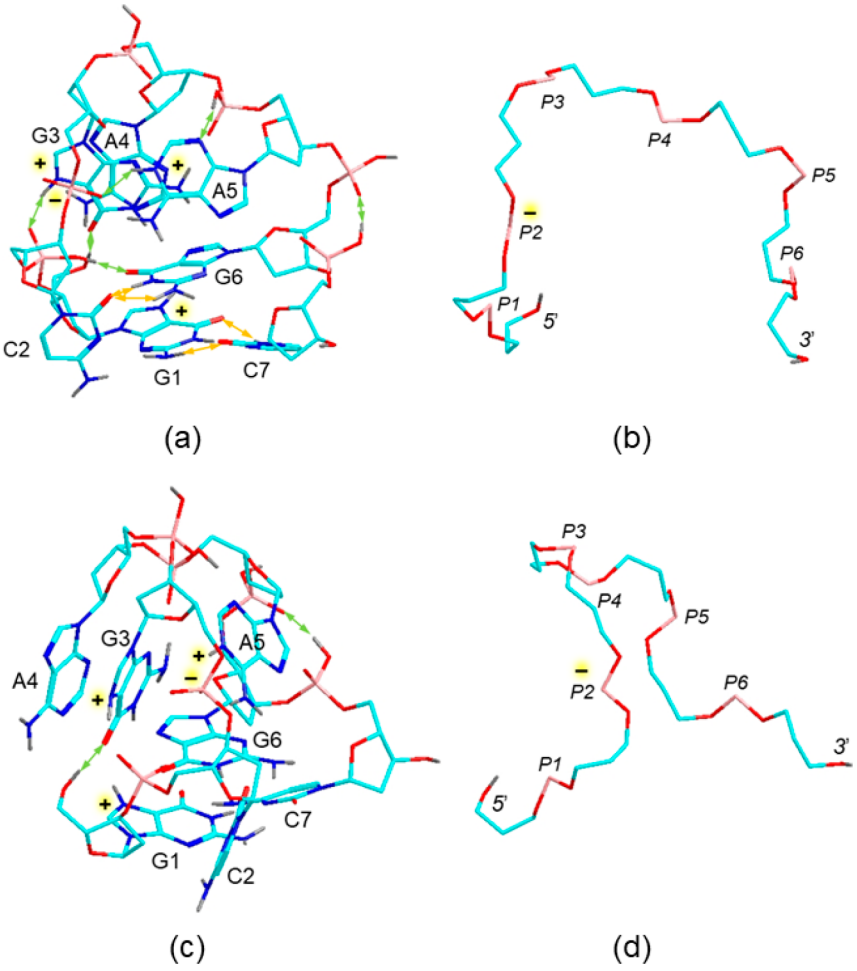


Figure 9. (a, c) Different views of the lowest-energy dication **G1G3b**$_{500}$. (b, d) Views of the backbone skeleton. Structure description is as in Figure 8. The phosphate anion is marked with a minus sign.

miniloop conformation, which showed WC pairs of G1-C7 and C2-G6. A noteworthy feature of this dication was that its protonation structure isomerized to a zwitterion, in which the P1-OH proton migrated to N7 at G3 (Figure 12).

However, its high Gibbs energy, ΔG_{310} = 101 kJ mol^{-1} relative to **G1G3b**$_{500}$ (Table 2), clearly indicated that this miniloop dication could not coexist with the low-energy isomers under equilibrium conditions. Its potential presence in the ion mixture

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

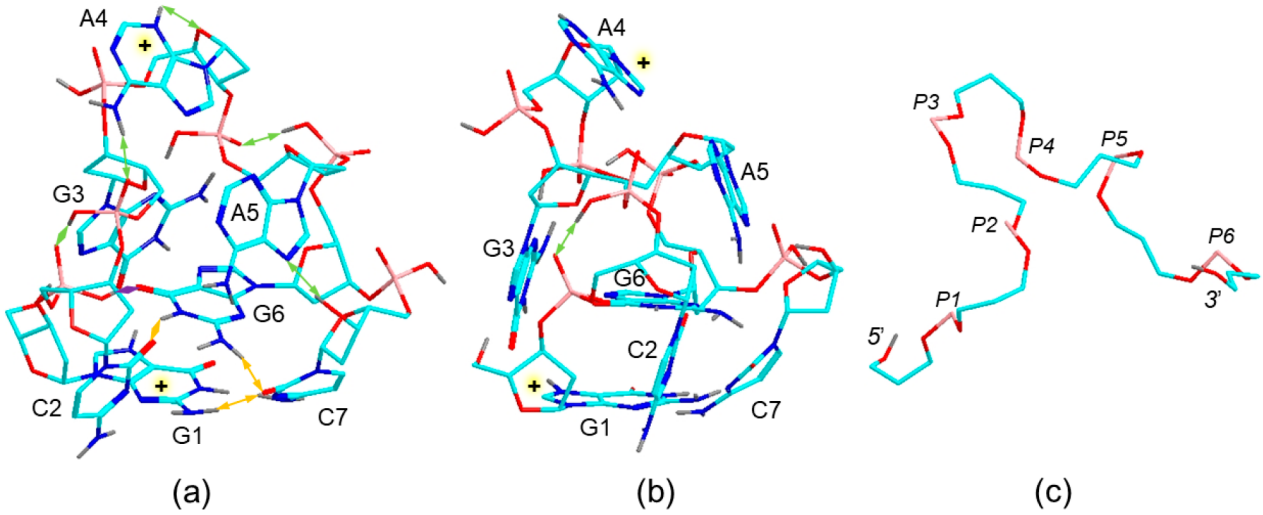


Figure 10. (a, b) Different views of d(GCGAAGC)$^{2+}$ dication **G1A4a**$_{500}$. (c) View of the **G1A4a**$_{500}$ backbone skeleton in the (b) orientation. Structure description as in Figure 8.

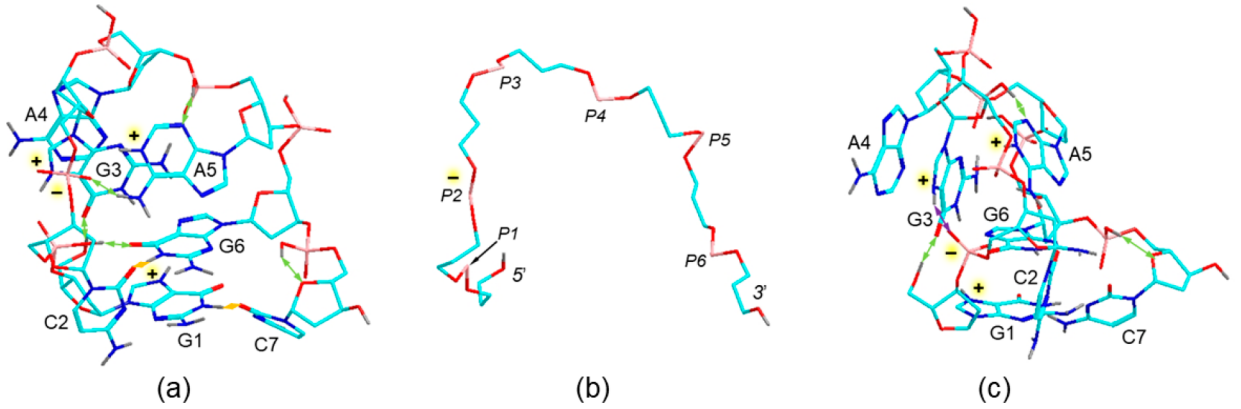


Figure 11. (a, c) Different views of the d(GCGAAGC)$^{2+}$ dication **G1G3c**$_{500}$. (b) View of the **G1G3c**$_{500}$ backbone skeleton in the (a) orientation. Structure description as in Figure 8.

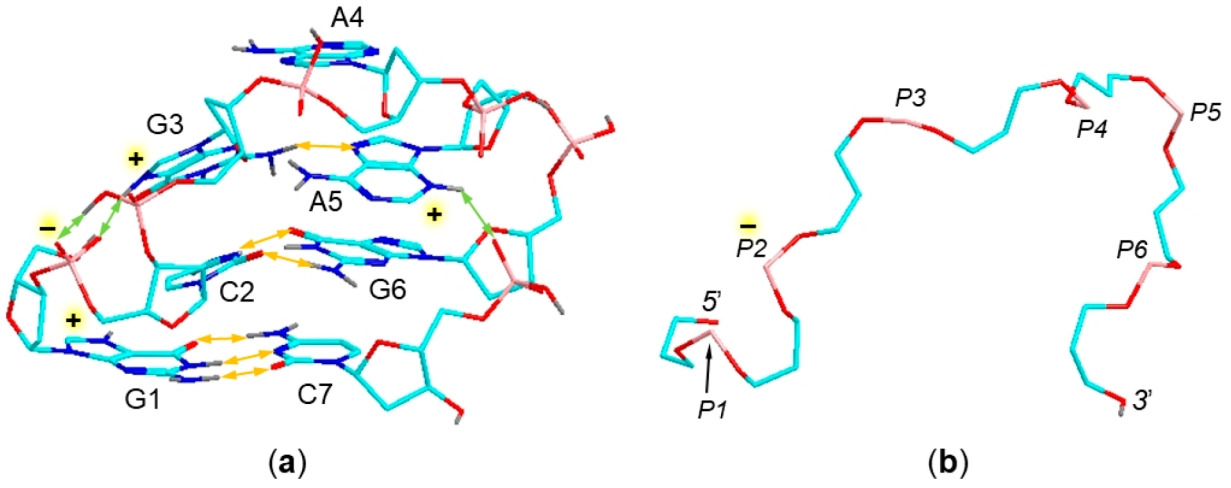


Figure 12. (a) Optimized structure of d(GCGAAGC)$^{2+}$ dication **G1A5**$_{300}$. (b) View of the **G1A5**$_{300}$ backbone skeleton. Structure description as in Figure 8.

could be due to kinetic factors that prevented rearrangement of the miniloop during ion formation and storage. It is noteworthy that in contrast to the other miniloop structure (**G1G3a**$_{300}$), **G1A5**$_{300}$ had the P1 anion on the periphery where it was stabilized by hydrogen bonds with P2-OH and N7−H at G3. This prevented protonation at G6 and possibly contributed to the retention of the G1-C7 and C2-G6 WC pairs.

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

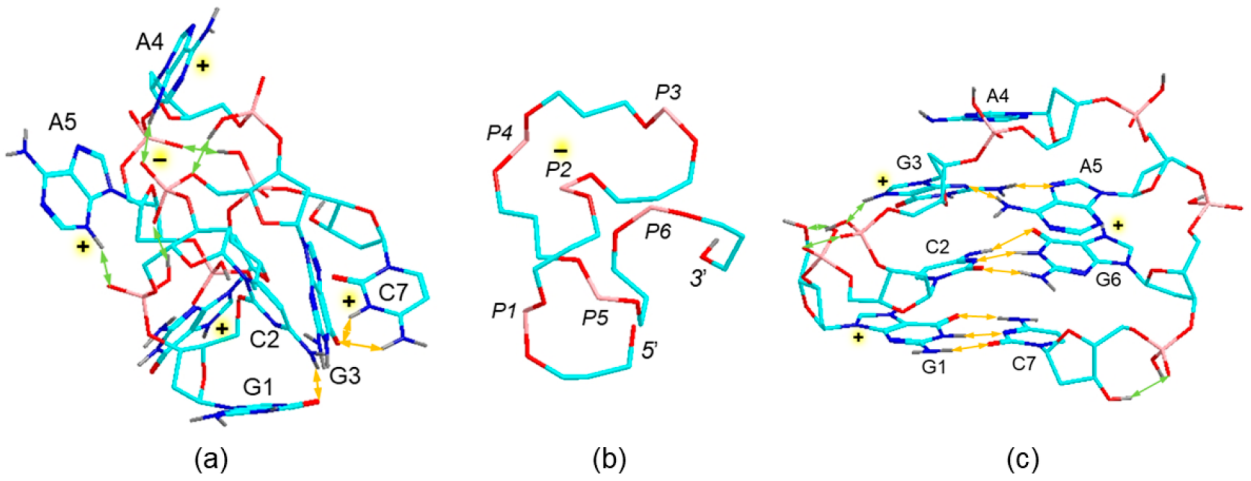


Figure 13. Optimized structures of d(GCGAAGC)$^{3+}$ trications (a) **GAG$_{500}$**$^{3+}$ and (c) **GGGa$_{300}$**$^{3+}$. (b) View of the **GAG$_{500}$**$^{3+}$ backbone skeleton. Structure description as in Figure 8.

Ion **A4G6$_{600}$** (CCS_{MPI} = 458 Å^2) was another zwitterion that was generated by BOMD at 600 K followed by DFT geometry optimization, which resulted in a substantial structure rearrangement. The optimized structure of **A4G6$_{600}$** indicated that the P3-OH proton migrated to C7, resulting in a severely distorted structure in which C2 was remote from G1 and became hydrogen bonded to A5 (Figure S7, Supporting Information). The structure was dominated by multiple internal solvations of the P3 anion by hydrogen bonds from the protonated nucleobases A4, G6, and C7. The latter was realized as an O2-protonated tautomer in **A4G6$_{600}$**. Overall, ion **A4G6$_{600}$** was 172 kJ mol^{-1} less stable than the global energy minimum and represented another high-energy isomer.

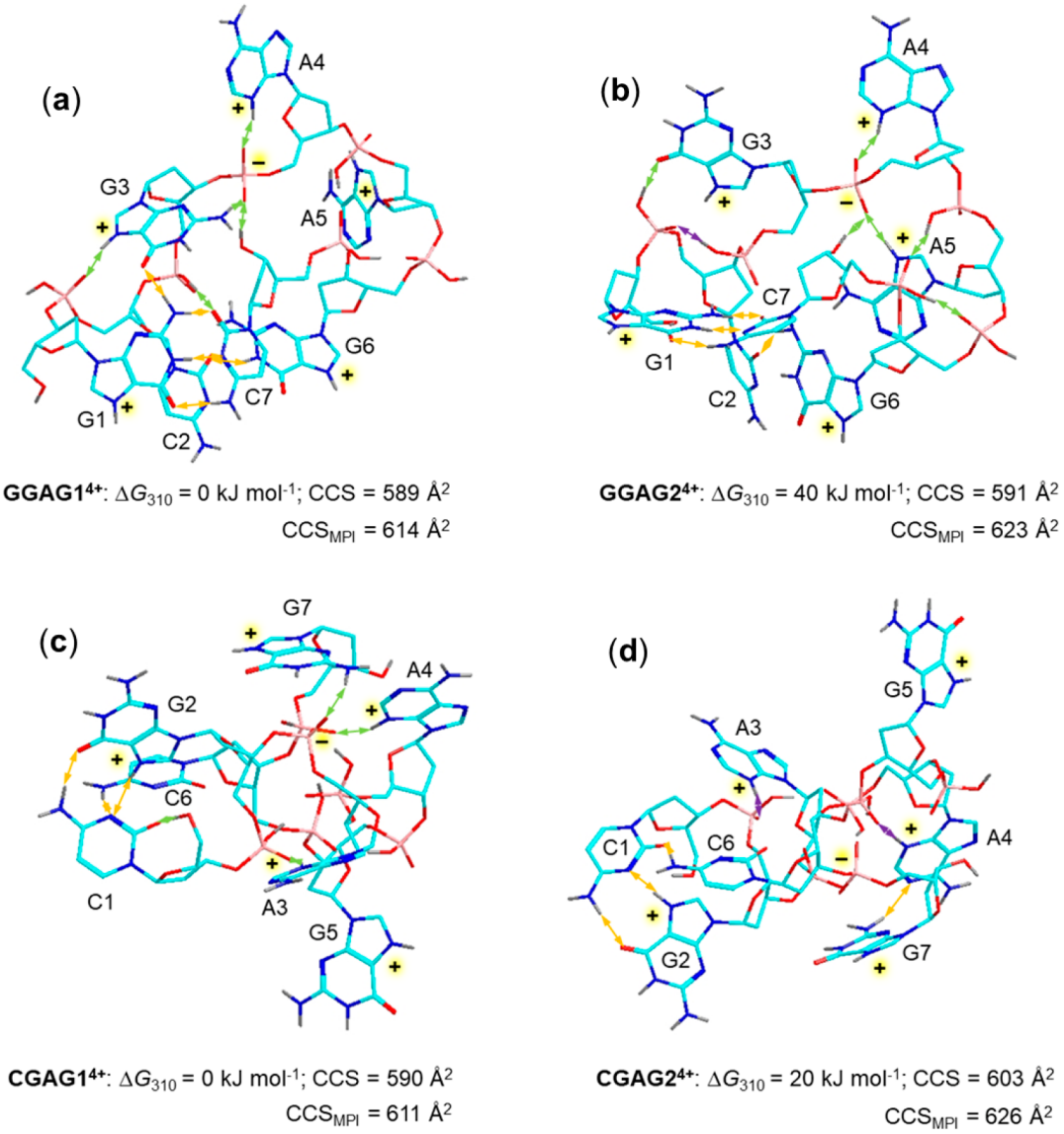
Two other compact dication structures, **G1A4b$_{500}$** (CCS_{MPI} = 433 Å^2) and **G1G6$_{500}$** (CCS_{MPI} = 431 Å^2) were less stable than **G1G3b$_{500}$** by 95 and 153 kJ mol^{-1}, respectively. Both of these structures showed a complete disruption of the A5-G6 π-stacking (Figure S8, Supporting Information). In addition, partial or complete disruption of G-C hydrogen bonds was apparent for **G1A4b$_{500}$** and **G1G6$_{500}$**, respectively, which probably contributed to their higher Gibbs energy. Note that **G1G6$_{500}$** was a zwitterion formed by the transfer of the acidic proton from P1−H to O6 of G3.

To summarize the aforementioned dication structures, all low-energy d(GCGAAGC)$^{2+}$ isomers showed major deviations from the original miniloop conformation. When preserved, miniloop features with two WC G-C pairs were represented by high-energy isomers in the gas-phase dications. The overall abundance of low-energy zwitterionic structures was remarkable, albeit in line with previous structure studies of tri- and tetranucleotide dications.[24−27,29] This appears to support the general conclusion that the formation of oligonucleotide zwitterions by the transfer of a proton from phosphate ester groups onto nucleobases is energetically favorable in gas-phase ions. Both the zwitterions and canonical dications developed a number of strong hydrogen bonds that involved the deoxyribose and phosphate groups. This contrasted the solution miniloop structure[31,32] in which the phosphates were located on the heptanucleotide perimeter where they were exposed to stabilizing solvation by water. In the absence of solvent, internal solvation in gas-phase dications involving the phosphoester OH groups became the structure determining factor. Furthermore, the calculated CCS indicated that the collapse of the miniloop structure in gas-phase ions resulted in more compact structures when compared to those in which the miniloop was substantially or partially preserved.

In contrast to d(GCGAAGC), there is no established solution or crystal structure for the reversed d(CGAAGCG) sequence. Our BOMD conformational search of the d(CGAAGCG)$^{2+}$ dications started from random conformations in which the protonation patterns followed the G2G5, G2G7, G5G7, G2A3, G2A4, A3G5, A4G5, A3G7, and A4G7 combinations in which protons were placed at guanine N7 and adenine N3. Several DFT-optimized structures were obtained, which all showed disordered conformations. Selected low-energy structures are displayed in Figure S9 (Supporting Information). The MobCal calculated CCS of these low-energy d(CGAAGCG)$^{2+}$ isomers were in the 476−544 Å^2 range, which significantly exceeded the IMS data (440−470 Å^2, Table 1). A closer agreement was achieved by implementing the MobCal-MPI scheme that gave CCS_{MPI} in the 460−482 Å^2 range (Figure S9). In particular, the low Gibbs-energy ions **G5G7−68** and **G2G7−156** had calculated CCS_{MPI} = 460 and 466 Å^2, respectively, which were very close to those of the most abundant peak in the c-IMS spectrum at CCS = 459 Å^2. With the lowest-energy ion **A3G5−185**, the MobCal-MPI-calculated CCS_{MPI} = 474 Å^2 was close to that of the second most abundant peak from c-IMS with a CCS = 470 Å^2. Although the structure assignment for d(CGAAGCG)$^{2+}$ ions may not be definite when based on the CCS alone, the presence of multiple isomers as revealed by c-IMS, as well as the calculated structures, indicated that gas-phase dications preferred disordered geometries with no tendency for multiple WC pairs.

Trications. d(GCGAAGC)$^{3+}$ initial ion structures were constructed by placing the charging protons on the G and A bases. The optimized structures to be discussed were selected according to the match of their calculated CCS_{MPI} with those from IMS. The lowest-energy structure **GAG$_{500}$**$^{3+}$ had CCS_{MPI} = 508 Å^2 which provided a close match (Δ = −0.2%) to the most abundant peak in the ion mobility spectrum (509 Å^2). **GAG$_{500}$**$^{3+}$ was a zwitterion in which the P2-OH proton moved to N3 at A5 (Figure 13a). In addition, the hydrogen-bonding pattern of multiple phosphoester groups favored a proton shift from N7 at G3 to N3 at C7. Multiple hydrogen bonds including the phosphoester OH resulted in a highly coiled conformation, as shown by the backbone projection in Figure 13b.

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340



GGAG1^{4+}: ΔG_{310} = 0 kJ mol^{-1}; CCS = 589 Å^2
CCS_{MPI} = 614 Å^2

GGAG2^{4+}: ΔG_{310} = 40 kJ mol^{-1}; CCS = 591 Å^2
CCS_{MPI} = 623 Å^2

CGAG1^{4+}: ΔG_{310} = 0 kJ mol^{-1}; CCS = 590 Å^2
CCS_{MPI} = 611 Å^2

CGAG2^{4+}: ΔG_{310} = 20 kJ mol^{-1}; CCS = 603 Å^2
CCS_{MPI} = 626 Å^2

Figure 14. Optimized structures of tetracations (a) **GGAG1**$^{4+}$, (b) **GGAG2**$^{4+}$, (c) **CGAG1**$^{4+}$, and (d) **CGAG2**$^{4+}$, with relative Gibbs energies and calculated CCS. Structure description is in Figure 8.

A stacked trication, **GGGa$_{300}$**$^{3+}$ was obtained whose CCS_{MPI} = 509 Å^2 also was a close match to the major peak in the ion mobility spectrum (CCS = 509 Å^2). Ion **GGGa$_{300}$**$^{3+}$ was a canonical structure with protonation at the N7 positions in G1, G3 and G6 that retained the main features of the miniloop structure, including the G1-C7 and C3-G6 Watson−Crick pairs and A5-G6 ring π-stacking (Figure 13c). However, **GGGa$_{300}$**$^{3+}$ was a high-energy isomer at ΔG_{310} = 180 kJ mol^{-1} relative to **GAG$_{500}$**$^{3+}$ and, thus, represented a thermodynamically unstable conformation in the gas phase. A related miniloop structure, **GGGb$_{300}$**$^{3+}$ (Figure S10a,b, Supporting Information), which had ΔG_{310} = 126 kJ mol^{-1} relative to **GAG$_{500}$**$^{3+}$, was thermodynamically more favorable than **GGGa$_{300}$**$^{3+}$ and had a matching CCS_{MPI} = 511 Å^2 that fitted the second peak in the ion mobility data. Overall, there were several trication structures that gave very close CCS_{MPI} values that would be indistinguishable by c-IMS. Regarding their relative Gibbs energies, the global energy minimum **GAG$_{500}$**$^{3+}$ appeared to be the most likely candidate for the gas-phase ion.

Among the d(CGAAGCG)$^{3+}$ ion structures, we found ion **GA3G$_{500}$**$^{3+}$ that was the global energy minimum at ΔG_{310} = −57 kJ mol^{-1} relative to **GAG$_{500}$**$^{3+}$. **GA3G$_{500}$**$^{3+}$ was a canonical trication with protonation sites at G2, G5, and G7 (Figure S10c,d). It showed that A3, A4, and G5 clustered at a backbone loop, whereas C1-G2 and C6-G7 formed local hydrogen-bonded pairs. However, this hydrogen bonding was mediated by the guanine N7 protons rather than the usual WC pattern.

Tetracations. In contrast to dications and trications, the structures of the d(GCGAAGC) and d(CGAAGCG) tetracations did not preserve the miniloop conformations. Two low-energy d(GCGAAGC)$^{4+}$ ions, **GGAG1**$^{4+}$ and **GGAG2**$^{4+}$, both showed a G1-C7 WC pair, while the C2-G6 moiety was disordered (Figure 14).

Remarkably, and despite the 4-fold protonation, these low energy structures were zwitterions in which all guanines and adenines were protonated while the P3 phosphate was deprotonated. The calculated CCS_{MPI} of **GGAG1**$^{4+}$ (614 Å^2) was close to that of the first abundant component in c-IMS (608 Å^2; Figure 7a). The CCS_{MPI} of **GGAG2**$^{4+}$ (623 Å^2) was a close

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

match with the second abundant c-IMS peak at 624 Å^2. These CCS-matching structures indicated a substantial distortion of the ion conformation upon 4-fold protonation. The low-energy reversed-sequence tetracations, **CGAG1**$^{4+}$ and **CGAG2**$^{4+}$, were also zwitterions (Figure 14). **CGAG2**$^{4+}$ showed a C1-G2 WC pair, while the other protonated nucleobases were on the ion periphery. The calculated CCS_{MPI} for **CGAG2**$^{4+}$ (626 Å^2) gave a close match with the CCS of the most abundant c-IMS peak at 625 Å^2 (Table 1), while that for the lowest-energy tetracation **CGAG1**$^{4+}$ (CCS_{MPI} = 611 Å^2) was 2.2% underestimated relative to that for the major d(CGAAGCG)$^{4+}$ component.

Ion Dissociations in CID-MS^2 Spectra. Although the ion mobility data indicated isomer mixtures for the heptanucleotide cations, it appeared to be of interest to investigate their dissociations upon collisional activation. The spectra are briefly described here, and the major dissociations are summarized in Table 3. CID of the cations resulted in the combined loss of nucleobase and backbone cleavages, analogous to those reported earlier for di-, tri-, and tetranucleotide cations.[27,29] The dications showed a loss of neutral nucleobases that preferentially occurred from the terminal positions but also depended on the sequence. Nucleobase loss from the 3′-position can be recognized by the associated loss of water, which is specific for this terminal position.[29] CID-MS^2 of d(GCGAAGC)$^{2+}$ showed a major loss of 5′-guanine whereas losses of cytosine and adenine were minor (Figure 15a). CID-MS^2 of the reverse-sequence dication d(CGAAGCG)$^{2+}$ showed a major loss of 3′-guanine that was accompanied by loss of cytosine and adenine (Figure S11a, Supporting Information). Similarly, the scrambled-sequence dication d(CGGAAGC)$^{2+}$ showed a loss of cytosine that originated from both the 3′- and 5′-positions (Figure S13a, Supporting Information). In this case the lower relative intensity of the combined loss of cytosine and water indicated that the major fraction of cytosine loss occurred from the 5′-position. In contrast, the other scrambled-sequence dication d-(GCGAACG)$^{2+}$ with terminal guanines showed a major loss of adenine from one of the internal positions that was accompanied by the loss of guanine (Figure S11a, Supporting Information). It should be noted that loss of the nucleobase requires that there be an available proton either placed on the base during ionization or transferred from another protonated base upon collisional excitation. These scenarios are not distinguished in the CID spectra, because proton transfer in oligonucleotide cations is a low energy process that can proceed prior to dissociation.[29] This was further investigated with the current heptanucleotide ion set by generating the lower charge states by proton-transfer reactions[34] of higher-charge states with the fluoranthene anion, $C_{16}H_9^-$. d(GCGAAGC)$^{2+}$ was generated in the ion trap by proton transfer from the mass-selected +3 and +4 states and its CID-MS^3 spectra were obtained. These were found to be nearly identical with the CID-MS^2 spectrum of d(GCGAAGC)$^{2+}$ produced by electrospray, indicating that the proton locations in the dication do not have a major effect on the dissociations after collisional excitation.

CID of trications showed mainly losses of protonated nucleobases forming doubly charged fragment ions. d-(GCGAAGC)$^{3+}$ showed a competitive loss of GH^+ and 3-CH^+, the latter being accompanied by loss of water (Figure 15b). The inverted trication, d(CGAAGCG)$^{3+}$, showed chiefly loss of GH^+ that was accompanied by consecutive loss of water and backbone cleavage eliminating the a_1 neutral fragment and forming the abundant fragment ion at m/z 876.5 (Figure S12b and S13b, Supporting Information). The scrambled trications d(GCGAACG)$^{3+}$ and d(CGGAAGC)$^{3+}$ showed chiefly losses of the terminal nucleobases, GH^+ and CH^+, respectively, that most likely originated from both 3′- and 5′-positions, as judged by the consecutive loss of water. Interestingly, the loss of AH^+ from d(GCGAACG)$^{3+}$ was only minor (Figure S11b). The loss of a protonated nucleobase requires transfer of two protons. This can be viewed as a two step process proceeding via a complex of the neutral nucleobase with the complementary multiply charged fragment ion, in which the departing nucleobase picks up the charging proton.

CID of the tetracations gave rise to much simpler spectra compared with those of the other charge states. d-(GCGAAGC)$^{4+}$ showed loss of GH^+ and 3′-CH^+ that were analogous to the dissociations of the trication (Figure 15c). The inverted tetracation, d(CGAAGCG)$^{4+}$, showed a very weak nucleobase ion loss, with the main dissociations occurring in the backbone (Figure S11c). In contrast, d(GCGAACG)$^{4+}$ and d(CGGAAGC)$^{4+}$ showed major losses of the terminal nucleobases, GH^+ and CH^+, respectively (Figure S12c and S13c), which were analogous to the dissociations of the respective trications. Interestingly, the loss of GH^+ and CH^+ from the tetracations was accompanied by much less prominent loss of water (Figure S12c, S13c), suggesting that these dissociations chiefly proceeded from the 5′-positions.

Common features among the dication sequences were also observed for backbone dissociations. Most backbone cleavages occurred at the 5′-terminus, forming $\boldsymbol{w}_6^{2+}$ ions.[29] For dications, this was illustrated by d(GCGAAGC)$^{2+}$ and d(CGAAGCG)$^{2+}$ where the 5′- and 3′-termini were discerned by mass (Figure 15a–c and S11a–c). With d(GCGAACG)$^{2+}$ and d-(CGGAAGC)$^{2+}$ the dominant backbone fragment ions could arise from either terminus and therefore were labeled as $\boldsymbol{w}_6/\boldsymbol{d}_6^{2+}$ (Figure S12a, S13a). In addition to $\boldsymbol{w}_6^{2+}$, backbone dissociations

Table 3. Summary of the Ion Dissociations

Loss of Base	
d(GCGAAGC)$^{2+}$	5′-G ≫ 3′-C, A
d(GCGAAGC)$^{3+}$	5′-GH^+ > 3′-CH^+, very weak loss of AH^+
d(GCGAAGC)$^{4+}$	5′-GH^+ > 3′-CH^+, no loss of AH^+
d(CGAAGCG)$^{2+}$	3′-G (-H_2O) > C > A
d(CGAAGCG)$^{3+}$	3′-G (-H_2O) > C > A
d(CGAAGCG)$^{4+}$	5′-C
d(GCGAACG)$^{2+}$	A > 5′-G ≫C
d(GCGAACG)$^{3+}$	GH^+ ≫ CH^+ > AH^+
d(GCGAACG)$^{4+}$	GH^+ ≫ CH^+, no loss of AH^+
d(CGGAAGC)$^{2+}$	C > G > A
d(CGGAAGC)$^{3+}$	3′-CH^+ ≫ GH^+, no loss of AH^+
d(CGGAAGC)$^{4+}$	CH^+ dominant
Backbone Cleavages	
d(GCGAAGC)$^{2+}$	$\boldsymbol{w}_6^{2+}$; consecutive loss of G, C, and A; $\boldsymbol{w}_5/\boldsymbol{d}_5^{2+}$; $\boldsymbol{w}_5/\boldsymbol{d}_5^{+}$; $\boldsymbol{w}_4^{+}$
d(GCGAAGC)$^{3+}$	$\boldsymbol{w}_6^{2+}$, consecutive loss of C; $\boldsymbol{w}_4^{+}$
d(GCGAAGC)$^{4+}$	$\boldsymbol{d}_6^{2+}$, consecutive loss of G, $\boldsymbol{a}_1$; $\boldsymbol{w}_6^{3+}$
d(CGAAGCG)$^{2+}$	$\boldsymbol{w}_6^{2+}$, consecutive loss of G + H_2O; $\boldsymbol{w}_5/\boldsymbol{d}_5^{+}$; $\boldsymbol{w}_4^{+}$
d(CGAAGCG)$^{3+}$	$\boldsymbol{w}_6^{2+}$, consecutive loss of G + H_2O; $\boldsymbol{w}_5/\boldsymbol{d}_5^{2+}$; $\boldsymbol{w}_5/\boldsymbol{d}_5^{+}$; $\boldsymbol{w}_4^{+}$
d(CGAAGCG)$^{4+}$	$\boldsymbol{w}_6^{3+}$; $\boldsymbol{d}_6^{2+}$, consecutive loss of C + H_2O
d(GCGAACG)$^{2+}$	$\boldsymbol{w}_6/\boldsymbol{d}_6^{2+}$; $\boldsymbol{w}_5/\boldsymbol{d}_5^{+}$
d(GCGAACG)$^{3+}$	$\boldsymbol{w}_6/\boldsymbol{d}_6^{2+}$; $\boldsymbol{w}_5/\boldsymbol{d}_5^{2+}$
d(GCGAACG)$^{4+}$	$\boldsymbol{w}_6/\boldsymbol{d}_6^{3+}$; $\boldsymbol{w}_6/\boldsymbol{d}_6^{2+}$, consecutive loss of C, G; $\boldsymbol{w}_5/\boldsymbol{d}_5^{2+}$
d(CGGAAGC)$^{2+}$	$\boldsymbol{w}_6/\boldsymbol{d}_6^{2+}$, consecutive loss of G; $\boldsymbol{w}_5/\boldsymbol{d}_5^{2+}$
d(CGGAAGC)$^{3+}$	$\boldsymbol{w}_6/\boldsymbol{d}_6^{2+}$, consecutive loss of C; $\boldsymbol{w}_5/\boldsymbol{d}_5^{2+}$
d(CGGAAGC)$^{4+}$	$\boldsymbol{w}_6/\boldsymbol{d}_6^{2+}$; $\boldsymbol{w}_6/\boldsymbol{d}_6^{3+}$

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

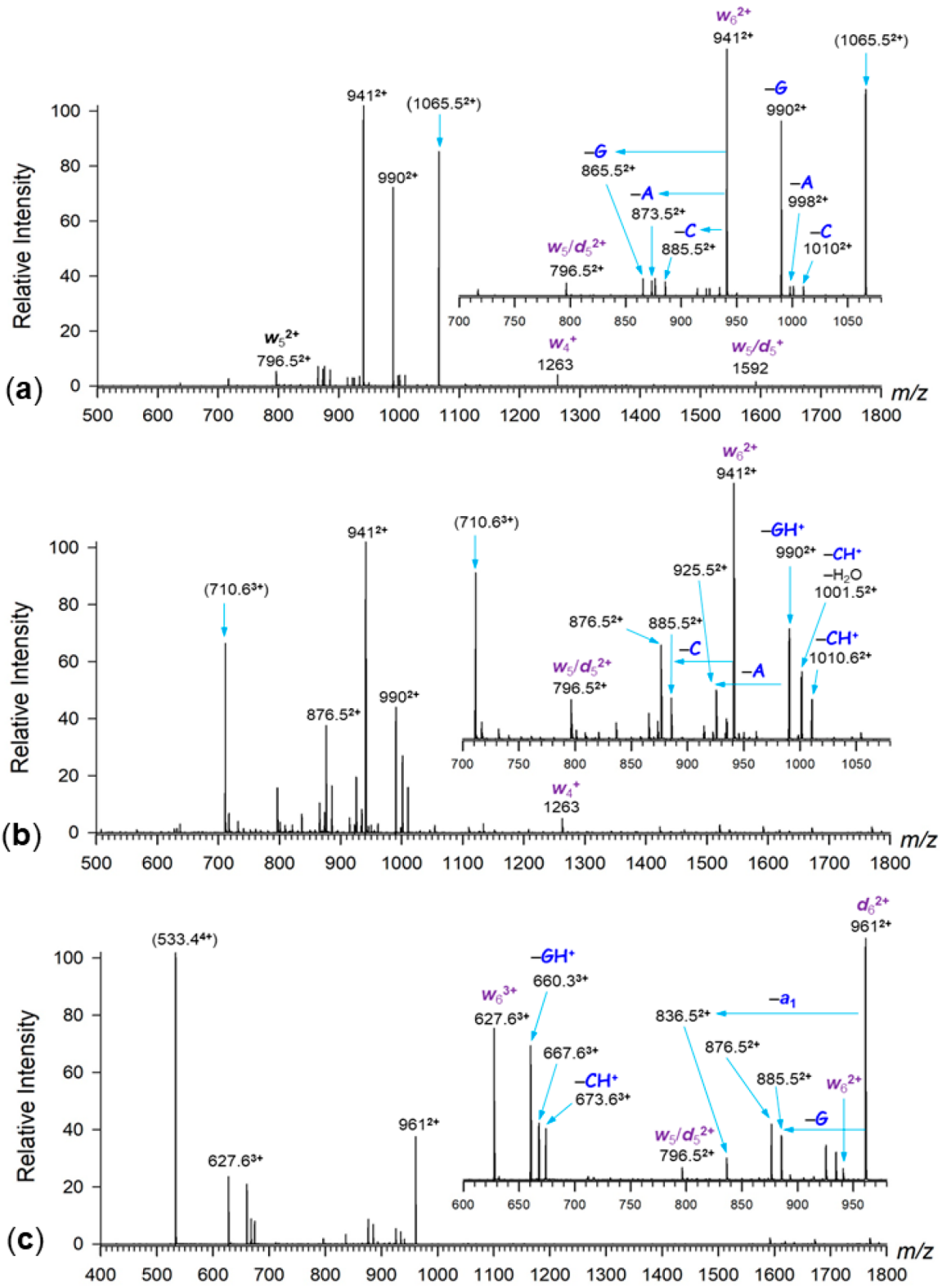


Figure 15. CID-MS^2 spectra of (a) GCGAAGC^{2+} at m/z 1065.5, (b) GCGAAGC^{3+} at m/z 710.6, and (c) GCGAAGC^{4+} at m/z 533.4. Insets show the expanded regions of doubly and triply charged fragment ions.

leading to w_5^{2+} and w_5/d_5^{2+} ions were also observed for all dication sequences. Singly charged fragment ions, e.g., w_5/d_5^+ and w_4^+ represented less abundant dissociation products, as shown for d(GCGAAGC)$^{2+}$, d(CGAAGCG)$^{2+}$, and d-(GCGAACG)$^{2+}$. The triply charged ions showed a very similar pattern, in which w_6/d_6^{2+} fragment ions dominated the CID-MS^2 spectra. These ions underwent consecutive base loss, e.g., loss of cytosine from w_6(CGAAGC)$^{2+}$, and loss of guanine and water from w_6(GAAGCG)$^{2+}$.

The quadruply charged ions showed more evident differences depending on the sequence. CID of d(GCGAAGC)$^{4+}$ formed the d_6^{2+} ion as an abundant sequence fragment by cleavage

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

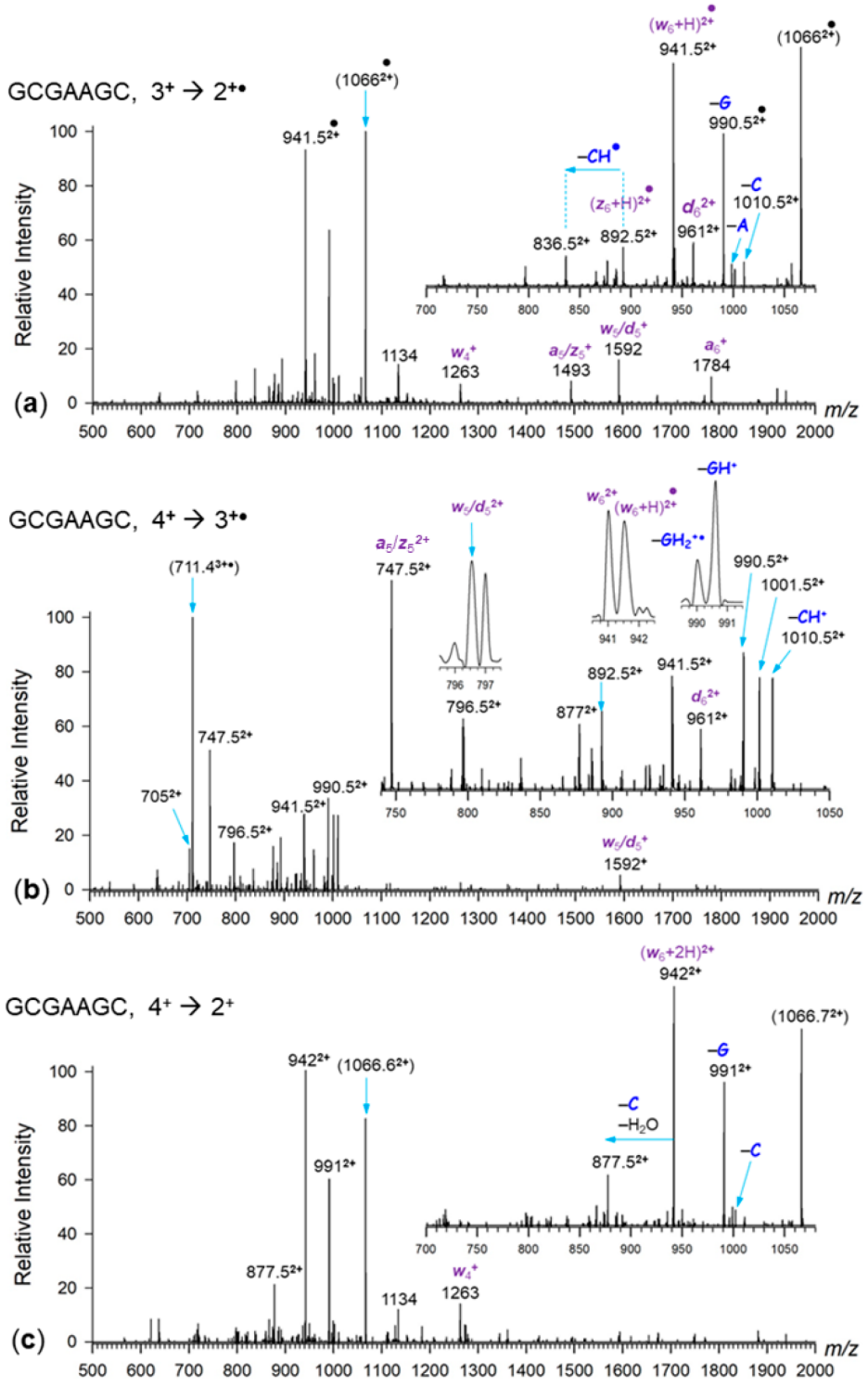


Figure 16. ET-CID-MS^3 spectra of d(GCGAAGC) cations: (a) d(GCGAAGC+3H)$^{2+\bullet}$ at m/z 1066; (b) d(GCGAAGC+4H)$^{3+\bullet}$ at m/z 711.4; (c) d(GCGAAGC+4H)$^{2+}$ at m/z 1066.6.

between G6 and C7, in addition to the major $\boldsymbol{w}_6^{3+}$ ion (Figure 15c). The tetracation of the inverted d(CGAAGCG)$^{4+}$ sequence also produced a $\boldsymbol{d}_6^{2+}$ fragment ion which in this case underwent further consecutive loss of 5′-cytosine and water (m/z 876.5, Figure S11c). Backbone dissociations of d(GCGAACG)$^{4+}$ and d(CGGAAGC)$^{4+}$ also proceeded near the termini, forming the $\boldsymbol{w}_6/\boldsymbol{d}_6^{3+}$ and $\boldsymbol{w}_6/\boldsymbol{d}_6^{2+}$ ions which were unresolved for these sequences (Figure S12c, S13c). The formation of $\boldsymbol{w}_6/\boldsymbol{d}_6^{2+}$ ions from the tetracations was interesting because it involved the loss of two charges within the formally complementary mononucleoside fragment. This can be accommodated by consecutive elimination of a protonated nucleobase, forming a glycal cation intermediate that can undergo standard phosphate ester elimination of the deoxyribose $C_5H_7O_2^+$ ion fragment. The loss-of-base and backbone dissociations are summarized in Table 3.

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

Cation-Radical Dissociations in ET-CID-MS^3 Spectra. Electron transfer to oligonucleotide cations has been reported to cause dissociation at various levels, which were very low to practically none for hexanucleotides,[23] weak for tetranucleotides,[26,27] and moderate to abundant for di- and trinucleotides.[25,59] In an effort to characterize the present heptanucleotide cations, we obtained ET-CID-MS^3 spectra of various charge states that were generated by nondissociative electron transfer, e.g., one-electron $3^+ \rightarrow 2^{+\bullet}$, $4^+ \rightarrow 3^{+\bullet}$, and two-electron $4^+ \rightarrow 2^+$ charge reducing reactions. The former two one-electron reductions generated open shell systems that were potentially susceptible to radical-induced dissociations, which are briefly described here. CID of d(GCGAAGC+3H)$^{2+\bullet}$ showed two major dissociations which were loss of G and backbone cleavage forming the $(\boldsymbol{w}_6+\mathrm{H})^{2+\bullet}$ ions (Figure 16a). Both dissociations generated product ions that retained the radical sites and were analogous to dissociations observed for even-electron ions d(GCGAAGC)$^{2+}$. ET-CID differed from CID for the formation of $\boldsymbol{a}_6^+$, $\boldsymbol{a}_5/\boldsymbol{z}_5^+$, and $\boldsymbol{d}_6^{2+}$ fragment ions from d(GCGAAGC +3H)$^{2+\bullet}$, which involved loss of neutral radical fragments. CID of d(GCGAAGC+4H)$^{3+\bullet}$ showed losses of 3′-CH$^+$ and GH$^+$ that were also observed for CID of the even-electron trication d(GCGAAGC)$^{3+}$ (Figure 15b). The major differences pertinent to the radical trication included the formation of $\boldsymbol{d}_6^{2+}$ ions which resulted from the loss of $\boldsymbol{z}_1^{+\bullet}$ radical cations containing (cytosine +2H) as a proton and radical carrier. Hydrogen transfer upon dissociation was also indicated by the loss of (guanine+2H)$^{+\bullet}$ accompanying the loss of GH$^+$, and the formation of $\boldsymbol{w}_6^{2+}$ and $\boldsymbol{w}_5/\boldsymbol{d}_5^{2+}$ ions that were accompanied by $(\boldsymbol{w}_6+\mathrm{H})^{2+\bullet}$, and $(\boldsymbol{w}_5+\mathrm{H})/(\boldsymbol{d}_5+\mathrm{H})^{2+\bullet}$ ions by loss of doubly charged fragments. The major distinctive feature of the CID spectrum of d(GCGAAGC+4H)$^{3+\bullet}$ was the formation of $\boldsymbol{a}_5/\boldsymbol{z}_5^{2+}$ ions by backbone cleavage. It is worth noting that the $\boldsymbol{a}_5/\boldsymbol{z}_5^{2+}$ ions were not accompanied by hydrogen atom adduct radicals in the spectrum. The ET-CID-MS^3 spectrum of the inverted sequence dication radical, d(CGAAGCG+3H)$^{2+\bullet}$ showed prominent fragment dication radicals by loss of neutral nucleobases 5′-G > A ≈ C (Figure S14a, Supporting Information). Loss of 5′-G was accompanied by a consecutive elimination of a neutral $\boldsymbol{a}_1$ fragment, forming the prominent m/z 877 dication radical. Backbone cleavages producing the $(\boldsymbol{w}_6+\mathrm{H})^{2+\bullet}$ and $(\boldsymbol{w}_5+\mathrm{H})/(\boldsymbol{d}_5+\mathrm{H})^{2+\bullet}$ ions proceeded without the involvement of radical sites that remained in the product ions. The trication radical of the inverted sequence, d(CGAAGCG+4H)$^{3+\bullet}$, showed major $\boldsymbol{w}_5/\boldsymbol{d}_5^{2+}$ ions at m/z 796.5 by loss of $(\boldsymbol{a}_2+\mathrm{H})/\boldsymbol{z}_2+\mathrm{H})^{+\bullet}$ cation radicals that appeared at m/z 540 (Figure S14b). The ET-CID-MS^3 spectra of cation radicals with scrambled sequences, d(GCGAACG)$^{2+\bullet}$, d(GCGAACG)$^{3+\bullet}$, d(CGGAAGC)$^{2+\bullet}$, and d(CGGAAGC)$^{3+\bullet}$ also showed chiefly $\boldsymbol{w}_6$ and $\boldsymbol{w}^5$ type fragment dications and their hydrogen atom adducts, as shown in Figure S15a,b and S16a,b (Supporting Information).

The products of consecutive two-electron reduction of the tetracations, $4^+ \rightarrow 3^{+\bullet} \rightarrow 2^+$, chiefly showed $(\boldsymbol{w}_6+2\mathrm{H})^{2+}$ backbone fragment ions (Figure 14c) that were analogous to those formed from the corresponding dication radicals, allowing for the additional hydrogen atom. The main difference was that the CID-MS^3 dissociations of these +2 ions proceeded with much less hydrogen atom transfer, and all four hydrogen atoms introduced by protonation were retained in the $(\boldsymbol{w}_6+2\mathrm{H})^{2+}$ ions. Similarly, the loss of nucleobases from the 2+ ions was devoid of hydrogen atom migrations, and no loss of GH$^\bullet$, CH$^\bullet$, or AH$^\bullet$ was observed (Figure S14c and S15c). This may indicate that the transfer of the second electron in the $3^{+\bullet} \rightarrow 2^+$ reduction step resulted in the formation of a closed-shell dication and not a dication diradical. Multistep electron transfer to high charge states of proteins has been claimed to produce diradicals on the basis of their reaction with residual oxygen.[60] In contrast, the dissociations of doubly reduced heptanucleotide ions reported here indicated that no diradicals were formed.

In summary, the analysis of the CID-MS^2 and ET-CID-MS^3 spectra revealed the general features of ion backbone dissociations that chiefly occurred at the 5′-terminus regardless of the nature of the nucleobase (G or C) and the heptanucleotide sequence. Similarly, preferential loss of the terminal bases (C and G) generally occurred with the exception of doubly charged d(GCGAACG) that lost one of the internal adenines from d(GCGAACG+2H)$^{2+}$, d(GCGAACG+3H)$^{2+\bullet}$, and d(GCGAACG+4H)$^{2+}$, regardless of the ion formation mode and presence or absence of nucleobase radicals. Loss of the nucleobase is known to be facilitated by base protonation,[29] indicating protonation at adenine in d(GCGAACG) dications. Nevertheless, the proton distribution in these multiply charged ions and the ion conformation were not possible to rationally link to the dissociation patterns that reflected the transition states rather than the initial ion structures.

CONCLUSIONS

The combined experimental data and calculations allowed us to arrive at the following conclusions. The miniloop structure of GCGAAGC underwent distortions in gas-phase cations generated by electrospray under both native and denaturing conditions. Protonation in di-, tri-, and tetracations resulted in the formation of zwitterionic species in which stabilizing strong hydrogen bonds in phosphate anion-nucleobase cation motifs disrupted Watson−Crick pairing between C2 and G6. The collapse of the miniloop in GCGAAGC^{2+} manifested itself by compact structures that were indicated by collision cross sections measured by c-IMS. Cations derived from sequence-scrambled heptanucleotides also showed multiple isomers with CCS comparable to those of the GCGAAGC ions.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jasms.3c00228.

Additional figures with arrival time profiles, optimized ion structures, CID-MS^2 and ET-CID-MS^3 spectra (PDF)

AUTHOR INFORMATION

Corresponding Authors

Karel Lemr − *Department of Analytical Chemistry, Faculty of Science, Palacký University, 779 00 Olomouc, Czech Republic; Institute of Microbiology of the Czech Academy of Sciences, 142 20 Prague, Czech Republic;* orcid.org/0000-0003-3158-0637; Phone: +420-585 634 415; Email: karel.lemr@upol.cz

František Tureček − *Department of Chemistry, University of Washington, Seattle, Washington 98195-1700, United States;* orcid.org/0000-0001-7321-7858; Email: turecek@uw.edu

Authors

Jiahao Wan − *Department of Chemistry, University of Washington, Seattle, Washington 98195-1700, United States*

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340

Marianna Nytka − *Department of Analytical Chemistry, Faculty of Science, Palacký University, 779 00 Olomouc, Czech Republic; Institute of Microbiology of the Czech Academy of Sciences, 142 20 Prague, Czech Republic*

Haocheng Qian − *Department of Chemistry, University of Washington, Seattle, Washington 98195-1700, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/jasms.3c00228

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Research at the University of Washington was supported by the Chemistry Division of the U.S. National Science Foundation, Grants CHE-1951518. F.T. acknowledges support by the Klaus and Mary Ann Saegebarth Endowment. Research at Palacký University was supported by the Czech Science Foundation, Grant 23-07254S.

■ REFERENCES

(1) Henry, K. D.; Williams, E. R.; Wang, B. H.; McLafferty, F. W.; Shabanowitz, J.; Hunt, D. F. Fourier-Transform Mass Spectrometry of Large Molecules by Electrospray Ionization. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 9075−9078.

(2) Suckau, D.; Shi, Y.; Beu, S. C.; Senko, M. W.; Quinn, J. P.; Wampler, F. M., III; McLafferty, F. W. Coexisting Stable Conformations of Gaseous Protein Ions. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 790−793.

(3) Wyttenbach, T.; Bowers, M. T. Structural Stability from Solution to the Gas Phase: Native Solution Structure of Ubiquitin Survives Analysis in a Solvent-Free Ion Mobility-Mass Spectrometry Environment. *J. Phys. Chem. B* **2011**, *115*, 12266−12275.

(4) Jurneczko, E.; Barran, P. E. How Useful Is Ion Mobility Mass Spectrometry for Structural Biology? The Relationship between Protein Crystal Structures and Their Collision Cross Sections in the Gas Phase. *Analyst* **2011**, *136*, 20−28.

(5) Eldrid, C.; Ujma, J.; Kalfas, S.; Tomczyk, N.; Giles, K.; Morris, M.; Thalassinos, K. Gas Phase Stability of Protein Ions in a Cyclic Ion Mobility Spectrometry Traveling Wave Device. *Anal. Chem.* **2019**, *91*, 7554−7561.

(6) Seo, J.; Hoffmann, W.; Warnke, S.; Bowers, M. T.; Pagel, K.; von Helden, G. Retention of Native Protein Structures in the Absence of Solvent: A Coupled Ion Mobility and Spectroscopic Study. *Angew. Chem., Int. Ed.* **2016**, *55*, 14173−14176.

(7) Susa, A. C.; Xia, Z.; Williams, E. R. Native Mass Spectrometry from Common Buffers with Salts that Mimic the Extracellular Environment. *Angew. Chem., Int. Ed.* **2017**, *56*, 7912−7915.

(8) Clemmer, D. E.; Russell, D. H.; Williams, E. R. Characterizing the Conformationome: Toward a Structural Understanding of the Proteome. *Acc. Chem. Res.* **2017**, *50*, 556−560.

(9) Bakhtiari, M.; Konermann, L. Protein Ions Generated by Native Electrospray Ionization: Comparison of Gas Phase, Solution, and Crystal Structures. *J. Phys. Chem. B* **2019**, *123*, 1784−1796.

(10) Schnier, P. D.; Klassen, J. S.; Strittmatter, E. F.; Williams, E. R. Activation Energies for Dissociation of Double Strand Oligonucleotide Anions: Evidence for Watson-Crick Base Pairing in Vacuo. *J. Am. Chem. Soc.* **1998**, *120*, 9605−9613.

(11) Griffey, R. H.; Greig, M. J.; An, H.; Sasmor, H.; Manalili, S. Targeted Site-Specific Gas-Phase Cleavage of Oligoribonucleotides. Application in Mass Spectrometry-Based Identification of Ligand Binding Sites. *J. Am. Chem. Soc.* **1999**, *121*, 474−475.

(12) Gabelica, V.; Rosu, F.; Houssier, C.; DePauw, E. Gas Phase Thermal Denaturation of an Oligonucleotide Duplex and Its Complexes with Minor Groove Binders. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 464−467.

(13) Gabelica, V.; DePauw, E. Collision-Induced Dissociation of 16-mer DNA Duplexes with Various Sequences: Evidence for Conservation of the Double Helix Conformation in the Gas Phase. *Int. J. Mass Spectrom.* **2002**, *219*, 151−159.

(14) Gidden, J.; Ferzoco, A.; Baker, E. S.; Bowers, M. T. Duplex Formation and the Onset of Helicity in Poly d(CG)n Oligonucleotides in a Solvent-Free Environment. *J. Am. Chem. Soc.* **2004**, *126*, 15132−15140.

(15) Butcher, D.; Chapagain, P.; Leng, F.; Fernandez-Lima, F. Differentiating Parallel and Antiparallel DNA Duplexes in the Gas Phase Using Trapped Ion Mobility Spectrometry. *J. Phys. Chem. B* **2018**, *122*, 6855−6861.

(16) Khristenko, N.; Amato, J.; Livet, S.; Pagano, B.; Randazzo, A.; Gabelica, V. Native Ion Mobility Mass Spectrometry: When Gas-Phase Ion Structures Depend on the Electrospray Charging Process. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1069−1081.

(17) Saintmont, F.; Hoyas, S.; Rosu, F.; Gabelica, V.; Brocorens, P.; Gerbaux, P. Structural Characterization of Dendriplexes In Vacuo: A Joint Ion Mobility/Molecular Dynamics Investigation. *J. Am. Soc. Mass Spectrom.* **2022**, *33*, 1555−1568.

(18) Sipe, S. N.; Fouque, K. J. D.; Garabedian, A.; Leng, F.; Fernandez-Lima, F.; Brodbelt, J. S. Exploring the Conformations and Binding Location of HMGA2·DNA Complexes Using Ion Mobility Spectrometry and 193 nm Ultraviolet Photodissociation Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2022**, *33*, 1092−1102.

(19) Vairamani, M.; Gross, M. L. G-Quadruplex Formation of Thrombin-Binding Aptamer Detected by Electrospray Ionization Mass Spectrometry. *J. Am. Chem. Soc.* **2003**, *125*, 42−43.

(20) Arcella, A.; Portella, G.; Ruiz, M. L.; Eritja, R.; Vilaseca, M.; Gabelica, V.; Orozco, M. Structure of Triplex DNA in the Gas Phase. *J. Am. Chem. Soc.* **2012**, *134*, 6596−6606.

(21) Ickert, S.; Hofmann, J.; Riedel, J.; Beck, S.; Pagel, K.; Linscheid, M. W. Charge-Induced Geometrical Reorganization of DNA Oligonucleotides Studied by Tandem Mass Spectrometry and Ion Mobility. *Eur. J. Mass Spectrom.* **2018**, *24*, 225−230.

(22) Schürch, S. Characterization of Nucleic Acids by Tandem Mass Spectrometry - The Second Decade (2004−2013): From DNA to RNA and Modified Sequences. *Mass Spectrom. Rev.* **2016**, *35*, 483−523.

(23) Hari, Y. I.; Leumann, C.; Schürch, S. What Hinders Electron Transfer Dissociation (ETD) of DNA Cations? *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 2677−2685.

(24) Korn, J. A.; Urban, J.; Dang, A.; Nguyen, H. T. H.; Turecek, F. UV-Vis Action Spectroscopy Reveals a Conformational Collapse in Hydrogen-Rich Dinucleotide Cation Radicals. *J. Phys. Chem. Lett.* **2017**, *8*, 4100−4107.

(25) Liu, Y.; Korn, J. A.; Dang, A.; Turecek, F. Hydrogen-Rich Cation Radicals of DNA Dinucleotides. Generation and Structure Elucidation by UV-Vis Action Spectroscopy. *J. Phys. Chem. B* **2018**, *122*, 9665−9680.

(26) Liu, Y.; Huang, S. R.; Tureček, F. Guanine-Adenine Interactions in DNA Tetranucleotide Cation Radicals Revealed by UV/Vis Photodissociation Action Spectroscopy and Theory. *Phys. Chem. Chem. Phys.* **2020**, *22*, 16831−16842.

(27) Huang, S. R.; Liu, Y.; Tureček, F. UV-Vis Photodissociation Action Spectroscopy Reveals Cytosine-Guanine Hydrogen Transfer in DNA Tetranucleotide Cation Radicals upon One-Electron Reduction. *J. Phys. Chem. B* **2020**, *124*, 3505−3517.

(28) Li, J.; Begbie, A.; Boehm, B. J.; Button, A.; Whidborne, C.; Pouferis, Y.; Huang, D. M.; Pukala, T. L. Ion Mobility-Mass Spectrometry Reveals Details of Formation and Structure for GAA·TCC DNA and RNA Triplexes. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 103−112.

(29) Wan, J.; Brož, B.; Liu, Y.; Huang, S. R.; Marek, A.; Tureček, F. Resolution of Identity in Gas-Phase Dissociations of Mono and Diprotonated DNA Codons by ^{15}N-Labeling and Computational Structure Analysis. *J. Am. Soc. Mass Spectrom.* **2022**, *33*, 1936−1950.

(30) Sunami, T.; Kondo, J.; Tsunoda, M.; Sekiguchi, T.; Hirao, I.; Watanabe, K.; Miura, K.; Takenaka, A. X-ray Structure of d-

(GCGAAGC); Switching of Partner for G:A Pair in Duplex Form. *Nucleic Acids Res.* **2002**, *2*, 181−182.
(31) Hirao, I.; Kawai, G.; Yoshizawa, S.; Nishimura, Y.; Ishido, Y.; Watanabe, K.; Miura, K. Most Compact Hairpin-Turn Structure Exerted by a Short DNA Fragment, d(GCGAAGC) in Solution: An Extraordinarily Stable Structure Resistant to Nucleases and Heat. *Nucleic Acids Res.* **1994**, *22*, 576−582.
(32) Padrta, P.; Štefl, R.; Králík, L.; Žídek, L.; Sklenář, V. Refinement of d(GCGAAGC) Hairpin Structure Using One- and Two-Bond Residual Dipolar Couplings. *J. Biomol. NMR* **2002**, *24*, 1−14.
(33) Jolles, B.; Refregiers, M.; Laigle, A. Opening of the Extraordinarily Stable Mini-Hairpin d(GCGAAGC). *Nucleic Acids Res.* **1997**, *25*, 4608−4613.
(34) Hartmer, R.; Kaplan, D. A.; Gebhardt, C. R.; Ledertheil, T.; Brekenfeld, A. Multiple Ion/Ion Reactions in the 3D Ion Trap: Selective Reagent Ion Production for ET and PTR from a Single Compound. *Int. J. Mass Spectrom.* **2008**, *276*, 82−90.
(35) Giles, K.; Ujma, J.; Wildgoose, J.; Pringle, S.; Richardson, K.; Langridge, D.; Green, M. A Cyclic Ion Mobility-Mass Spectrometry System. *Anal. Chem.* **2019**, *91*, 8564−8573.
(36) Ujma, J.; Ropartz, D.; Giles, K.; Richardson, K.; Langridge, D.; Wildgoose, J.; Green, M.; Pringle, S. Cyclic Ion Mobility Mass Spectrometry Distinguishes Anomers and Open-Ring Forms of Pentasaccharides. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1028−1037.
(37) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.; Robinson, C. V. Ion Mobility Mass Spectrometry Analysis of Large Protein Complexes. *Nat. Protoc.* **2008**, *3*, 1139−1152.
(38) Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Anal. Chem.* **2010**, *82*, 9557−9565.
(39) Picache, J. A.; Rose, B. S.; Balinski, A.; Leaptrot, K. L.; Sherrod, S. D.; May, J. C.; McLean, J. A. Collision Cross Section Compendium to Annotate and Predict Multi-Omic Compound Identities. *Chem. Sci.* **2019**, *10*, 983−993.
(40) Stewart, J. J. P. Optimization of Parameters for Semi-Empirical Methods. V. Modification of NDDO Approximations and Application to 70 Elements. *J. Mol. Model.* **2007**, *13*, 1173−1213.
(41) Řezáč, J.; Fanfrlík, J.; Salahub, D.; Hobza, P. Semi-empirical Quantum Chemical PM6Method Augmented by Dispersion and H Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes. *J. Chem. Theory Comput.* **2009**, *5*, 1749−1760.
(42) Berendsen, H. J.; Postma, J. V.; van Gunsteren, W. F.; DiNola, A. R. H. J.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(43) Řezáč, J. Cuby: An Integrative Framework for Computational Chemistry. *J. Comput. Chem.* **2016**, *37*, 1230−1237.
(44) Stewart, J. J. P. *MOPAC 16*; Stewart Computational Chemistry: Colorado Springs, CO, 2016.
(45) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098−3100.
(46) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(47) Chai, J. D.; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128*, 084106−084106.
(48) Chai, J. D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(49) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456−65.
(50) Singh, U. C.; Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* **1984**, *5*, 129−145.
(51) Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. Structural Information from Ion Mobility Measurements: Effects of the Long Range Potential. *J. Phys. Chem.* **1996**, *100*, 16082−16086.
(52) Campuzano, I. D. G.; Bush, M. F.; Robinson, C. V.; Beaumont, C.; Richardson, K.; Kim, H.; Kim, H. I. Structural Characterization of Drug-like Compounds by Ion Mobility Mass Spectrometry: Comparison of Theoretical and Experimentally Derived Nitrogen Collision Cross-sections. *Anal. Chem.* **2012**, *84*, 1026−1033.
(53) Kim, H.; Kim, H. I.; Johnson, P. V.; Beegle, L. W.; Beauchamp, J. L.; Goddard, W. A.; Kanik, I. Experimental and Theoretical Investigation into the Correlation between Mass and Ion Mobility for Choline and Other Ammonium Cations in N2. *Anal. Chem.* **2008**, *80*, 1928−1936.
(54) Ieritano, C.; Crouse, J.; Campbell, J. L.; Hopkins, W. S. A Parallelized Molecular Collision Cross Section Package with Optimized Accuracy and Efficiency. *Analyst* **2019**, *144*, 1660−1670.
(55) Ieritano, C.; Hopkins, W. S. Assessing Collision Cross Section Calculations Using MobCal-MPI with a Variety of Commonly Used Computational Methods. *Mater. Today Commun.* **2021**, *27*, 102226.
(56) Gatlin, C. L.; Tureček, F. Acidity Determination in Droplets Formed by Electrospraying Methanol-Water Solutions. *Anal. Chem.* **1994**, *66*, 712−718.
(57) Borovcová, L.; Hermannová, M.; Pauk, V.; Šimek, M.; Havlíček, V.; Lemr, K. Simple Area Determination of Strongly Overlapping Ion Mobility Peaks. *Anal. Chim. Acta* **2017**, *981*, 71−79.
(58) Zündel, G.; Metzger, H. Energiebänder der Tunnelnden Uberschuß-Protonen in Flüssigen Säuren. Eine IR-Spektroskopische Untersuchung der Natur der Gruppierungen H5O2+. *Z. Phys. Chem.* **1968**, *58*, 225−245.
(59) Wan, J.; Brož, B.; Liu, Y.; Huang, S. R.; Marek, A.; Tureček, F. The DNA Radical Code. Resolution of Identity in Dissociations of Trinucleotide Cation Radicals in the Gas Phase. *J. Am. Soc. Mass Spectrom.* **2023**, *34*, 304−319.
(60) Baba, T.; Campbell, J. L. Capturing Polyradical Protein Cations after an Electron Capture Event: Evidence for their Stable Distonic Structures in the Gas Phase. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1695−1701.

https://doi.org/10.1021/jasms.3c00228
J. Am. Soc. Mass Spectrom. 2023, 34, 2323−2340