ACS Partner Journal

Journal of the American Society for Mass Spectrometry

pubs.acs.org/jasms

Research Article

Resolution of Identity in Gas-Phase Dissociations of Mono- and Diprotonated DNA Trinucleotide Codons by ^{15}N-Labeling and Computational Structure Analysis

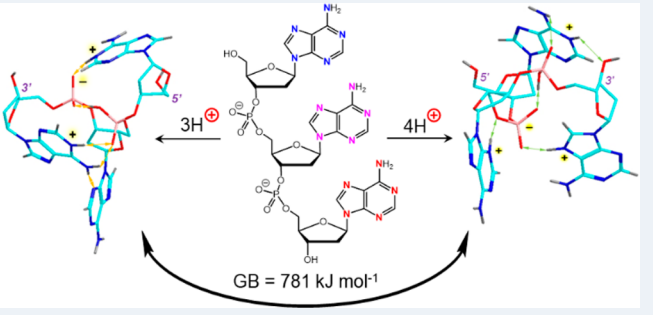
Jiahao Wan, Břetislav Brož, Yue Liu, Shu R. Huang, Aleš Marek,* and František Tureček*

Cite This: *J. Am. Soc. Mass Spectrom.* 2022, 33, 1936−1950

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Dissociations of DNA trinucleotide codons as gas-phase singly and doubly protonated ions were studied by tandem mass spectrometry using ^{15}N-labeling to resolve identity in the nucleobase loss and backbone cleavages. The monocations showed different distributions of nucleobase loss from the 5′-, middle, and 3′-positions depending on the nucleobase, favoring cytosine over guanine, adenine, and thymine in an ensemble-averaged 62:27:11:<1 ratio. The distribution for the loss of the 5′-, middle, and 3′-nucleobase was 49:18:33, favoring the 5′-nucleobase, but also depending on its nature. The formation of sequence w_2^+ ions was unambiguously established for all codon mono- and dications. Structures of low-Gibbs-energy protomers and conformers of dAAA$^+$, dGGG$^+$, dCCC$^+$, dTTT$^+$, dACA$^+$, and dATC$^+$ were established by Born−Oppenheimer molecular dynamics and density functional theory calculations. Monocations containing guanine favored classical structures protonated at guanine N7. Structures containing adenine and cytosine produced classical nucleobase-protonated isomers as well as zwitterions in which two protonated bases were combined with a phosphate anion. Protonation at thymine was disfavored. Low threshold energies for nucleobase loss allowed extensive proton migration to occur prior to dissociation. Loss of the nucleobase from monocations was assisted by neighboring group participation in nucleophilic addition or proton abstraction, as well as allosteric proton migrations remote from the reaction center. The optimized structures of diprotonated isomers for dAAA^{2+} and dACA^{2+} revealed combinations of classical and zwitterionic structures. The threshold and transition-state energies for nucleobase-ion loss from dications were low, resulting in facile dissociations involving cytosine, guanine, and adenine.



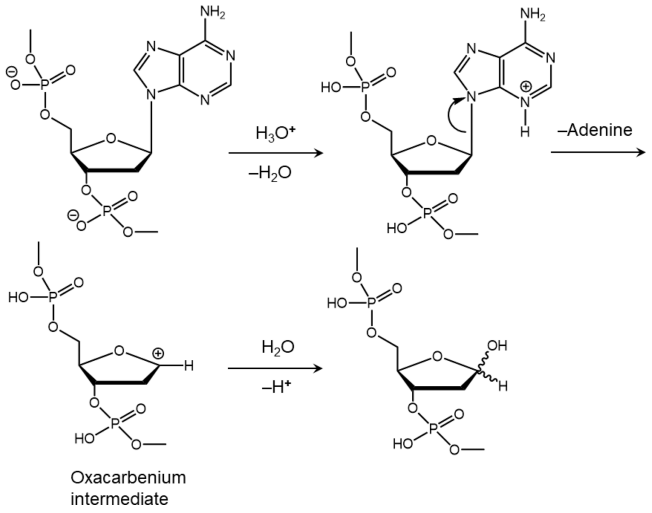
1. INTRODUCTION

Acid hydrolysis of DNA results in elimination of purine nucleobases in a process known as depurination.[1−3] Although nonenzymatic in nature, acid-catalyzed depurination is thought to play a role in nucleic acid metabolism in acidic gastric juice and in other low-pH environments such as in lysosomes.[1] The DNA depurinated sites are vulnerable to strand breaks[4,5] and, if unrepaired, can contribute to carcinogenesis.[6] The stability of DNA is important not only for biological materials but also in connection with the developing data storage technology where loss of a nucleobase would impair information retrieval.[7] Acid degradation of oligonucleotides in aqueous solution proceeds with first-order kinetics[1,8−11] with respect to the nucleotide and depends on pH. This has led to the suggested mechanism of depurination that presumes protonation of the nucleobase in the first step, followed by glycosidic bond cleavage, which forms a deoxyribose C1′ oxocarbenium cation as an intermediate (Scheme 1).[11]

However, the presumed oxacarbenium intermediate has not been characterized, as it is rapidly attacked by solvent, introducing a hydroxyl group at sugar C1′ and thereby finishing the hydrolysis. As in other numerous cases of reactive intermediates of chemical reactions,[12] gas-phase chemistry of isolated ions is ideally suited to tackle this problem by eliminating the solvent and allowing one to study the intrinsic properties of oligonucleotide-related oxocarbenium cations by tandem mass spectrometry. An additional benefit of studying the ion dissociations in the gas phase stems from the possibility of protonating any of the four DNA nucleobases, which is hindered in solution because of the low basicity of pyrimidine nucleobases, especially thymine. Dissociations of simple nucleosides and nucleotides have been addressed by Rodgers and co-workers in their energy-resolved guided ion beam studies of protonated adenine,[13] guanine,[14] and cytosine[15] nucleosides, as well as some fluorine-modified nucleosides.[16,17]

Received: July 10, 2022
Revised: August 22, 2022
Accepted: August 23, 2022
Published: August 30, 2022

Scheme 1. Previously Suggested Mechanism of Acid-Catalyzed Nucleobase Elimination



However, structures and dissociation energetics of more complex gas-phase oligonucleotide cations have not been addressed. Multiple protonation, in particular, can affect the relative stabilities of nucleobase protomers, as reported for singly protonated 2′-deoxyadenosine[18] and doubly protonated dinucleotide $(dAA+2H)^{2+}$.[19]

We now report a combined experimental and computational study of gas-phase dissociations of trinucleotide mono- and dications covering the DNA genetic code alphabet. A previous thorough study by O'Hair, Reid, and their co-workers has laid out the basic dissociations of DNA codon cations,[20] focusing on the loss of nucleobase and backbone cleavages in dependence on the nature and position of the nucleobases. However, in symmetrical trinucleotides of the dXXX and dXYX type, loss of identical neutral nucleobases X from the different positions was unresolved. The same applies to backbone cleavages occurring between the 5′-, middle, and 3′-nucleotide positions that can produce isobaric fragments of identical elemental composition. We achieve a resolution of this identity by analyzing dissociations of trinucleotides in which the nucleobases are distinguished by site-specific ^{15}N labeling, resulting in mass shifts by 5 Da for $[^{15}N_5]$adenine and $[^{15}N_5]$guanine, 3 Da for $[^{15}N_3]$cytosine, and 2 Da for $[^{15}N_2]$thymine. In addition, for several codons we report a thorough structural analysis of monocation and dication structures and dissociation energetics and kinetics using a combination of Born−Oppenheimer molecular dynamics (BOMD), density functional theory (DFT) calculations, and Rice−Ramsperger−Marcus−Kassel (RRKM) theory. By combining experimental and computational data, we wish to show that intramolecular interactions and neighboring-group participation play an important role in the proton-driven degradation of the codons in the gas phase.

2. EXPERIMENTAL SECTION

2.1. Materials. DNA codon trinucleotides (95% pure) were purchased from Integrated DNA Technologies, Inc. (Coralville, IA), and their quality and sequences were checked by tandem mass spectrometry. ^{15}N-labeled trinucleotides of the dXXX type, d[$^{15}N_5$-A]AA, dA[$^{15}N_5$-A]A, dAA[$^{15}N_5$-A], d[$^{15}N_5$-G]GG, dG[$^{15}N_5$-G]G, d[$^{15}N_3$-C]CC, dC[$^{15}N_3$-C]C, d[$^{15}N_2$-T]TT, dT[$^{15}N_2$-T]T, as well as trinucleotides of the dXYX type, d[$^{15}N_5$-A]CA, dAC[$^{15}N_5$-A], d[$^{15}N_5$-A]GA, d-[$^{15}N_5$-A]TA, d[$^{15}N_3$-C]AC, d[$^{15}N_3$-C]GC, d[$^{15}N_3$-C]TC, d-[$^{15}N_5$-G]AG, d[$^{15}N_5$-G]CG, d[$^{15}N_5$-G]TG, d[$^{15}N_2$-T]AT, d[$^{15}N_2$-T]CT, d[$^{15}N_2$-T]GT, and d[$^{15}N_2$-T]TG, were manually synthesized on a Controlled Pore Glass (CPG) synthesis column (500 Å) loaded with a desired nucleoside (200 nmol) according to standard procedures.[21−23] ^{15}N-labeled phosphoramidites were purchased from Cambridge Isotope Laboratories (Andover, MA); unlabeled phosphoramidites and synthesis columns were obtained from Bioautomation Corporation (Irving, TX); trichloroacetic acid solution (TCA) was purchased from Glen Research Corporation (Sterling, VA). The other chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Acetonitrile (HPLC gradient grade, >99.9%, 10 ppm of H_2O) was dried for several days over molecular sieves (4 Å, 500 mL) and filtered (0.22 μm PVDF syringe filter) immediately before use. Detailed synthetic, deprotection, and isolation procedures for the ^{15}N-labeled trinucleotides will be described in a separate publication.

2.2. Methods. Mass spectra were measured on a Bruker amaZon 3D-ion trap tandem mass spectrometer (Bruker Daltonik, Gmbh, Bremen, Germany). Samples of unlabeled trinucleotides were dissolved in 50:50:1 acetonitrile−water−acetic acid to achieve 10−20 μM concentrations and ionized by electrospray to form singly and doubly protonated ions. Samples of ^{15}N-labeled trinucleotides were dissolved in 1.0 mL of the solvent mixture and electrosprayed. The charge state of interest was selected by mass and subjected to collision induced dissociation (CID) at excitation amplitudes that were varied in 5−6 steps within 0.40−0.60 V and 0.15−0.30 V and for mono and dications, respectively. Spectra obtained at excitation amplitudes causing the precursor ion relative intensity to decrease to 50 ± 5% were used for comparisons among the different trinucleotides. CID-MS^3 spectra of selected ions were obtained analogously by mass-selecting and dissociating the fragment ions from CID-MS^2.

2.3. Calculations. Conformation analysis of monocations and dications was performed using Born−Oppenheimer molecular dynamics (BOMD) calculations as described previously.[24] Briefly, initial structures were generated for multiple protomers of adenine (N1, N3, N7), guanine (N7), cytosine (O2, N3), and thymine (O2 and O4) that were placed in the 5′-, middle, and 3′-nucleobase positions of the trinucleotide ions. BOMD trajectories for each protomer were run for 20 ps in 1 fs steps with Berendsen thermostat[25] set at 510 K. The hybrid semiempirical PM6-D3H4 method[26] was used that includes corrections for hydrogen bonds and dispersion interactions. PM6-D3H4 was run with MOPAC[27] under the high-level Cuby4 platform.[28] Since BOMD includes both nuclear and valence-electron motion, the initial proton positions are not fixed and protons can move among accessible positions in the same or another nucleobase, phosphate oxygen, or deoxyribose hydroxyl groups. Two hundred snapshots were extracted in regular intervals from each 20000-step trajectory and their geometries were gradient optimized with PM6-D3H4. The resulting structures were sorted out to remove duplicates, and 15−20 structures were fully gradient optimized with B3LYP[29] and the 6-31G(d,p) basis set including harmonic frequency analysis. Several low-energy B3LYP structures were further reoptimized with M06-2X[30] and the 6-31+G(d,p) basis set, and the reoptimized geometries were used for single-point energy M06-2X calculations with the 6-311++G(2d,p) basis set. This level of

https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936−1950

theory has been shown to give very similar relative nucleoside ion energies compared to those obtained by coupled-clusters calculations with single, double, and disconnected triple excitations (CCSD(T)) that were expanded to the complete basis set.[31] In addition, our previous studies indicated that dinucleotide[18,32] and tetranucleotide ion geometries[33,34] that were obtained by M06-2X/6-31+G(d,p) were consistent with ion structures determined by action spectroscopy measurements. Solvation energies were addressed with self-consistent reaction field calculations using the polarizable continuum model (PCM)[35] in water dielectric. Reference nucleoside proton affinities and gas-phase basicities were calculated using the DFT and CCSD(T) data reported previously[31] and are summarized in Table S1 (Supporting Information). Rice–Ramsperger–Kassel–Marcus calculations of unimolecular rate constants were performed with the QCPE 644 program of Zhu and Hase[36] that has been modified by expanding the limit for the number of internal degrees of freedom to 1000 and recompiled to run under Windows 7.[37] Vibrational state densities were obtained by a direct count of quantum states, and the rotational states were treated adiabatically. The microscopic rate constants [k(E,J,K)] were Boltzmann-averaged over the thermal distribution of rotational J and K states at 298 and 473 K.

3. RESULTS AND DISCUSSION

3.1. Survey of Monocation Spectra. Electrospray ionization produced mono- and diprotonated ions from most trinucleotides. Exceptions were the high-thymine combinations that either did not form dications (dTTT) or whose double protonation was very inefficient, as in dTTA, dTTC, dTAT, dTGT, and dTCT. The general pattern of CID-MS^2 monocation dissociations is illustrated with the spectra of isomeric trinucleotide ions $dACG^+$, $dCAG^+$, and $dAGC^+$ (Figure 1a–c), in which the adenine, guanine, and cytosine positions were permutated, and the different nucleobase masses allowed us to unambiguously assign the fragment ions. The main dissociations consisted of losses of the nucleobases that occurred unevenly from the 5′, middle, and 3′ positions, and also depended on the nucleobase.

Loss of the 3′-nucleobase was accompanied by elimination of water which was highly specific for this position for all nucleobases. Backbone dissociations were chiefly represented by cleavages between the 5′ and middle nucleotides, yielding w_2^+ ions with a minor formation of w_1^+ fragment ions. For oligonucleotide sequence ions nomenclature, see refs 38–40. The w_2^+ ions further lost the 3′-nucleobase, again followed by loss of water. These general features of monocation dissociations were very similar to those reported by O'Hair and co-workers.[20] Evaluation of the fragment ion relative intensities for loss of nucleobase and backbone cleavage from codon monocations is summarized in Table 1. The relative intensities of nucleobase loss from different positions were normalized to the total ion intensity for nucleobase loss. The percentages for the loss of the 3′ nucleobases were expressed as sum of the $(MH - 3'\text{-base})^+$ and $(MH - 3'\text{-base} - H_2O)^+$ relative intensities. The latter term is given separately in parentheses. The w_2^+ ion relative intensities were expressed relative to the sum of the loss-of-base and w_2^+ fragment ion intensities. Minor fragment ion intensities were not included in the normalization. To ascertain compatibility of the relative intensities reported in Table 1, they were measured at excitation amplitudes resulting in 50 ± 5% relative intensity

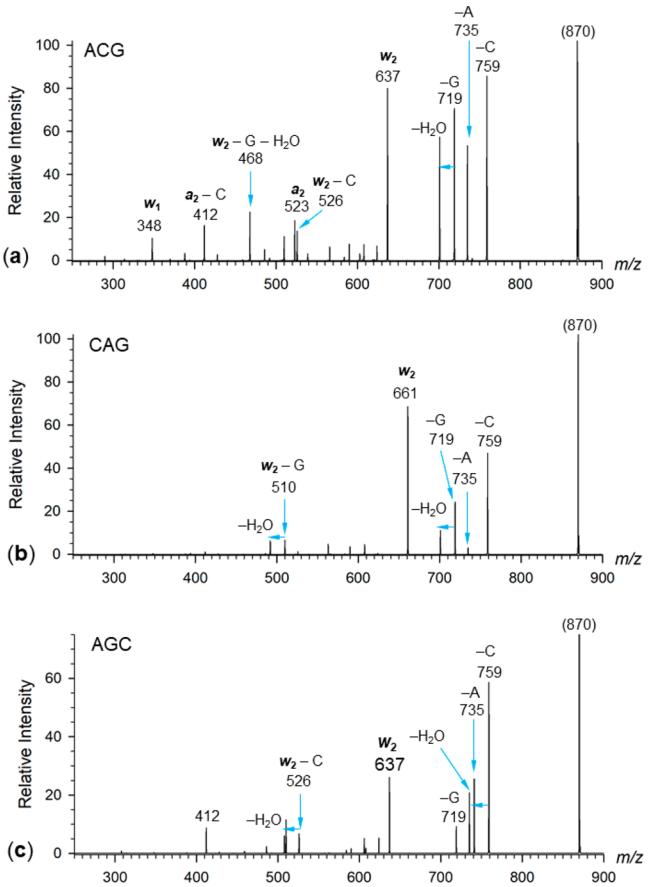


Figure 1. CID-MS^2 spectra of singly protonated (a) ACG, (b) CAG, and (c) AGC.

of the precursor ion. The relative intensities for loss of nucleobase for individual bases and their positions in the codons widely varied and depended on the trinucleotide sequence and presence of other bases. On average, loss of 5′-nucleobases was most common amounting to 49%, followed by loss of 3′-bases at 33%. The middle bases were lost less frequently at 18%. Cytosine was on average the most vulnerable base to be eliminated (62%), followed by guanine (27%), and adenine (11%). In contrast, loss of thymine was very rare, occurring at <1%. The formation of the w_2^+ ions represented on average 34% of the fragment ions, again showing a 0.4–98% variability among the codon monocations. A plot of the w_2^+ ion intensities versus those for the 5′-base loss was scattered, showing a weak correlation coefficient at r^2 = 0.38 (Figure S1, Supporting Information). There were some conspicuous outliers, such as dTTA, dTTT, and dATT, that gave high w_2^+ ion relative intensities in spite of showing a weak loss of the 5′ base. This indicated that the loss of the 5′-nucleobase may not trigger spontaneous elimination of the 5′-deoxyribose moiety to form the w_2^+ ion. This appeared to be in contrast to the mechanism suggested by Rodgers et al. in their study of dinucleotide cation dissociations.[41]

To further address this point, we investigated the dependence of the fragment ion relative intensities on the ion trap excitation amplitude, noting a previous study of oligonucleotide anions that has reported strong effects on product ion spectra.[42] The spectra of trinucleotide monocations obtained within the 0.46–0.58 V range showed only a weak dependence of the relative intensities of the nucleobase loss fragment ions

https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936–1950

Table 1. Fragment Ion Relative Intensities from Monocations

relative fragment ion intensity[a]					relative fragment ion intensity[a]				
codon[b]	5′	middle	3′ + water[c]	w_2[d]	codon[b]	5′	middle	3′ + water[c]	w_2[d]
AAA	50	13	37 (17)	54	**CTC**	48		52 (31)	37
GGG	93	4	3 (0.6)	28	TCC			100 (34)	<1
CCC	18	21	61 (23)	11	*CCG*	29	33	38 (21)	24
TTT	27	16	57 (35)	87	**CGC**	73	7	20 (5)	37
TTG	4	1	95 (25)	33	GCC	68		32 (7)	28
TGT	4	94	2	3	*CCA*	44	48	8 (3)	12
GTT[e]	97		3 (1.2)	95	**CAC**	27	4	69 (30)	29
TTA	12	7	81	75	ACC	9		91 (19)	3
TAT	4	84	12 (8)	36	AGC	18	8	74 (23)	17
ATT	59		41 (20)	98	ACG	20	32	48 (22)	23
TTC	1		99 (27)	6	GCA	73	21	6 (2)	41
TCT	1	97	2	4	GAC	74		26 (5)	38
CTT	97		3 (1)	94	CAG	55	4	41 (13)	44
AAT	99		1	73	CGA	92	6	3 (1)	46
ATA	61		39 (14)	32	AGT	67	32	1	42
TAA			100 (60)	5	ATG	24		76 (38)	52
AAG	45		55 (24)	58	GTA	93		7 (4)	58
AGA	55	22	23 (8)	46	GAT	100			38
GAA	94		6 (2)	59	CGT	96	4		32
AAC	14		86 (27)	14	CTG	64		36 (22)	61
ACA	33	30	38 (19)	29	GTC	46		54 (16)	41
CAA	67		33 (13)	59	GCT	88		12	56
TGG			100 (9)	0.6	TCG		33	66 (51)	0.3
GTG	88		12 (4)	46	TGC		20	80 (34)	0.4
GGT	93	4		25	ACT	26	73	2	42
GGC	84	2	14 (2)	32	ATC	9		91 (29)	8
GCG	82	8	10	38	CAT	97		3	44
CGG	58		42 (4)	46	CTA	93		7 (5)	56
GGA	98		2	39	TCA		65	34 (27)	1
GAG	92	1	7 (1)	37	TAC		4	96 (22)	<0.1
AGG	18		82 (12)	26					
CCT	99		1	43					

[a]Percent of summed base-loss ion intensities. [b]^{15}N-labeled nucleosides shown as bold characters; [c]Combined relative intensities for loss of 3′-base and water. Relative intensities for loss of water in parentheses. [d]Percent of combined base loss and w_2^+ ion intensities. [e]Nucleobase loss from positions shown in italics was not resolved, and relative intensities were summed for the unresolved positions.

which changed only marginally with the excitation amplitude. Within this interval, the relative intensity of the precursor ions dropped from 95% to <1% for most trinucleotide monocations (Figure S2a–f). Consecutive dissociations, such as the loss of water following the loss of the 3′-nucleobase and formation of w_2^+ ions were enhanced by <20% for most trinucleotide ions over the entire energy interval (Figure S2a-d). However, dissociations of dGTA$^+$ and dGGG$^+$ showed a larger increase from the very low to highest excitations (Figure S2e,f). This indicated that the trinucleotide ion composition affected the energetics and kinetics of the dissociations.

Further, we selected (MH − base)$^+$ fragment ions from CID-MS^2 spectra of selected codons for which loss of both the 5′- and 3′-nucleobase was observed, and we obtained their CID-MS^3 spectra. The spectra of the (MH − 5′-nucleobase)$^+$ ions from dGTC$^+$ and dCAG$^+$ and dCTG$^+$ (Figure 2a–c) showed major w_2^+ fragment ions by elimination of 2-hydroxymethylfuran ($C_5H_6O_2$) from the 5′-deoxyribose residue. Hence, when given sufficient internal excitation, the (MH − 5′-nucleobase)$^+$ ions did dissociate to form the sequence fragment ions. The consecutive dissociations of the w_2^+ ions depended on the remaining bases. For example, the w_2(TC)$^+$ and w_2(AG)$^+$ ions showed a facile loss of the 3′-nucleobase (C and G, respectively), followed by elimination of water from the 3′-sugar. In contrast, the w_2(TG)$^+$ ion was substantially more resistant to follow-up dissociation (Figure 2c). These differences indicated that the dissociations may be assisted by interactions between the sugar moiety and the bases remaining in the fragment ions that may exhibit chemical specificity. The specific nature of these interactions is addressed later in sections 3.3–3.8.

The CID-MS^3 spectra of the (MH − 3′-nucleobase)$^+$ ions from dGTC$^+$ and dCAG$^+$ and dCTG$^+$ (Figure S3a–c, Supporting Information) showed a major loss of water, which was consistent with the analogous fragmentations accompanying the loss of the 3′-nucleobase in the CID-MS^2 spectra of protonated trinucleotides. However, the competitive and consecutive dissociations of these (MH − 3′-nucleobase)$^+$ ions differed depending on the remaining nucleobases. For example, with dGTC$^+$ and dCTG$^+$ these other dissociations resulted in the elimination of the 3′-sugar moiety, forming the corresponding d_2^+ fragment ions at m/z 652 and 612, respectively. The d_2^+ ions further dissociated by loss of the 5′-nucleobase which was accompanied by a facile loss of water, forming the dominant fragment ions at m/z 483 from both dGTC$^+$ and dCTG$^+$ (Figure S3a,b). In contrast, the (MH −

https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936−1950

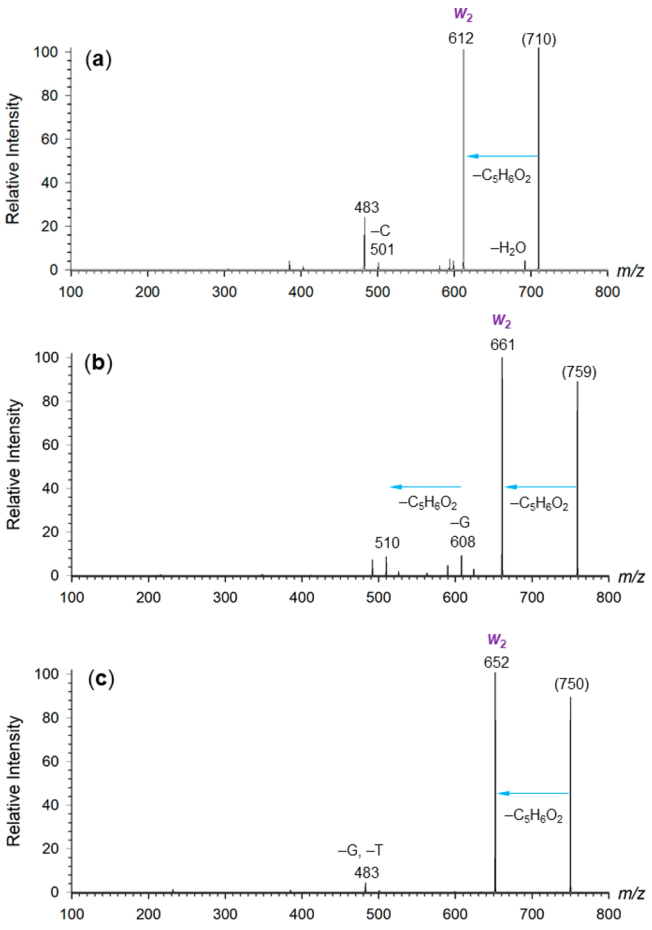


Figure 2. CID-MS^3 spectra of (MH − 5′-base)$^+$ ions from (a) dGTC, (b) dCAG, and (c) dCTG.

G)$^+$ ion from dCAG$^+$ underwent competitive loss of water and cytosine (m/z 701 and 608, respectively) while no formation of the $\boldsymbol{d}_2^+$ ion, expected at m/z 621, was observed (Figure S3c).

The middle-position nucleobase also had an effect on the spontaneous loss of water from the (MH − 3′-nucleobase)$^+$ ions. Table 1 data indicate that loss of water was particularly facile from trinucleotide ions having a pyrimidine base in the middle position. For example, over 50% of (MH − 3′-nucleobase)$^+$ ions from dTTT$^+$, dCTC$^+$, dCCG$^+$, dGTA$^+$, dTCG$^+$, dCTA$^+$, and dTCA$^+$ spontaneously dissociated by loss of water in CID-MS^2 of the monocations.

3.2. Resolution of Identity in Dissociations of Codon Monocations. The combined effects of the nucleobase nature and position in codons of the dXYZ type can lead to competitive enhancement or suppression of dissociations originating at any given position. This makes it difficult to assign specific reactivity to the nucleobase and position. To address this issue, we analyzed the dissociations of trinucleotide ions of the dXXX and dXYX type where the structurally identical nucleobases in different positions were distinguished by ^{15}N labeling. Starting with dAAA$^+$, loss of adenine occurred from the 5′-, middle, and 3′-positions in a 50:13:37 ratio, where the fraction for the loss of 3′-adenine included the ion intensity from subsequent specific loss of water. In contrast, dGGG$^+$ showed a dominant loss of 5′-guanine at 93% whereas the other positions were much less reactive. It is noteworthy that despite the highly abundant loss of 5′-G, the formation of $\boldsymbol{w}_2^+$ fragment ions from dGGG$^+$ was less abundant than from dAAA$^+$. Loss of cytosine from dCCC$^+$ proceeded from the 3′-nucleotide, with the 5′-, middle, and 3′-positions giving an 18:21:61 distribution. Similarly, loss of thymine from dTTT$^+$, albeit overall not very abundant, showed a 27:16:57 distribution for the 5′-, middle, and 3′-positions. The overall low propensity for loss of thymine was underscored by the spectra of dTGT$^+$, dTAT$^+$, and dTCT$^+$, that showed major loss of the middle nucleobases (Table 1). Insertion of thymine in the middle position, as in dATA$^+$, dGTG$^+$, and dCTC$^+$, did affect the distribution for the loss of the 5′- and 3′-nucleobases. With dATA$^+$, thymine slightly enhanced the loss of 5′-adenine (61%) compared to the same fraction in dAAA$^+$ (50%). In dGTG$^+$, it slightly decreased the loss of 5′-guanine (88%) compared to the same fraction in dGGG$^+$ (93%). In dCTC$^+$, it increased the loss of 5′-cytosine (48%) more visibly when compared to the same fraction in dCCC$^+$ (18%). The effect on the loss of nucleobases of a middle cytosine was chiefly competitive, as shown for dACA$^+$

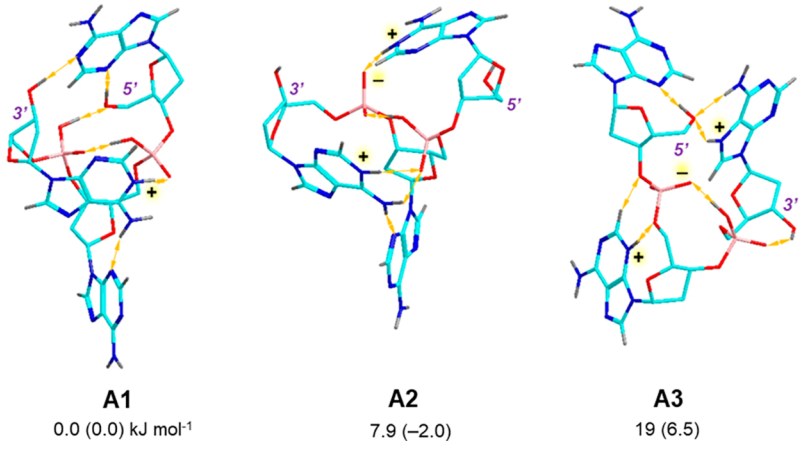
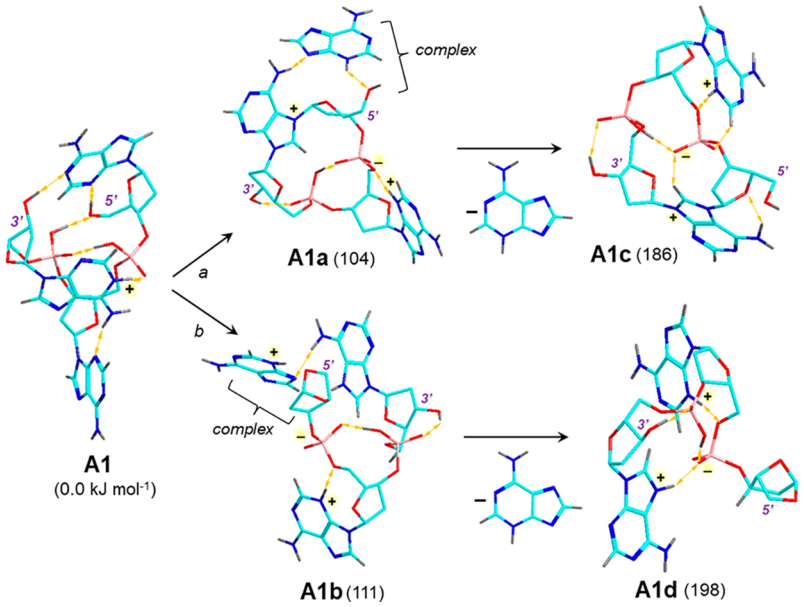


Figure 3. M06-2X/6-31+G(d,p)-optimized structures of low Gibbs energy dAAA$^+$ ions. Relative Gibbs energies are from single-point M06-2X/6-311++G(2d,p) calculations including B3LYP/6-31G(d,p) zero-point energies and 310 K enthalpies and entropies. Values in parentheses include PCM solvation energies in water dielectric. Atom color coding: cyan = C, blue = N, red = O, bronze = P, gray = H. Only the nucleobase and exchangeable hydrogen atoms are shown. Protonation and deprotonation sites are labeled with plus and minus signs. Hydrogen bonds are visualized by yellow arrows.

https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936−1950

Scheme 2. Reaction Pathways for the Loss of 5′-Adenine from dAAA$^+$[a]



[a]Relative energies in parentheses in kJ mol^{-1} are from M06-2X/6-311++G(2d,p) single-point energy calculations and include zero-point energies, referring to 0 K.

(33:30:38) and dGCG$^+$ (82:8:10). Combinations of dCXC$^+$ with adenine and guanine in the middle position affected the loss of the 5′- and 3′-cytosine to a different extent. In dCAC$^+$, the nucleobase-loss distribution (27:4:69) resembled that in dCCC$^+$ (16:21:61), showing an enhanced loss of cytosine from both the 5′- and 3′-positions. In contrast, in dCGC$^+$, the presence of the middle guanine resulted in a reversed preference for loss of the 5′-cytosine in a 73:7:20 distribution. The adenine-guanine combination in dAGA$^+$ resulted in an enhanced loss of guanine which was chiefly on account of a diminished loss of 3′-adenine. Finally, dGAG$^+$ showed a nucleobase loss distribution (92:1:7) that was very similar to that in dGGG$^+$ (93:4:3).

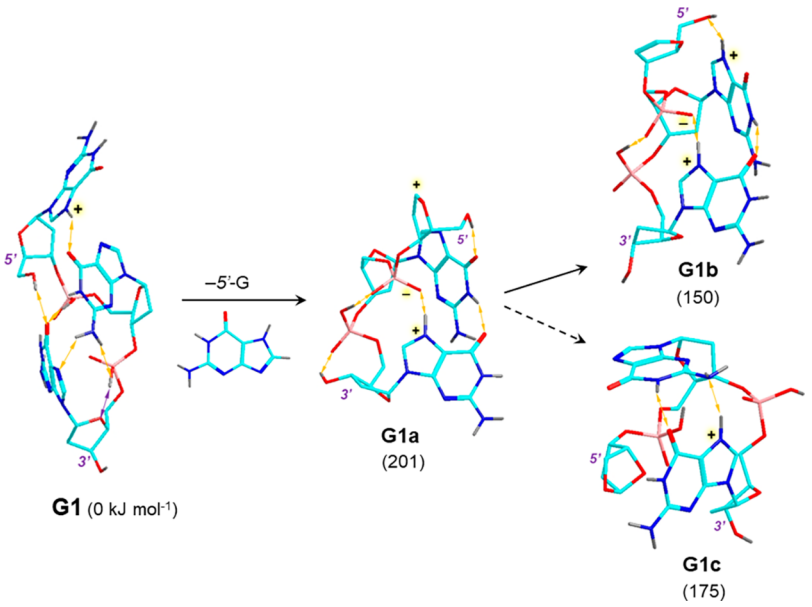
3.3. Codon Monocation Structures and Dissociation Mechanisms. The effects of the nucleobase nature and position that were revealed by resolution of identity in dissociations of trinucleotide ions of the dXXX$^+$ and dXYX$^+$ type pointed to specific interactions between the nucleobases in the codon monocations which could be of relevance to the reaction mechanisms of acid-induced DNA degradation. To address this point in structural detail at the single-atom resolution level, we obtained fully optimized geometries of multiple protomers and conformers of dAAA$^+$, dGGG$^+$, dCCC$^+$, dTTT$^+$, dACA$^+$, and dATC$^+$. Selected low-Gibbs energy isomers are presented and discussed here.

3.4. dAAA$^+$ Monocations. BOMD conformational analysis followed by DFT geometry optimizations yielded structures **A1**−**A3** as the lowest Gibbs energy isomers (Figure 3). The structures differed in both their protonation pattern and conformation but were closely spaced by energy. Structure **A1** was protonated at N1 of 3′-adenine and displayed strong hydrogen bonds involving the phosphoester and sugar hydroxy groups. Structures **A2** and **A3** were zwitterions in which a proton from the phosphate hydroxyl was transferred to a basic adenine position. This included N1 in both the 5′- and 3′-adenines in **A2**, N3 in the middle adenine, and N7 in the 3′-adenine in **A3**. The phosphate anions in the zwitterions were massively internally solvated by the sugar hydroxyl and nucleobase protons to create elaborate networks of hydrogen bonds. Interestingly, N3-protonated tautomers, which are favored in adenosine ions,[18,43] were not represented among the lowest-energy dAAA$^+$ structures. Nonspecific solvation by water, as estimated by polarizable continuum model calculations including full gradient geometry optimization did not appreciably change the relative Gibbs energies of **A1**−**A3** which remained within 7 kJ mol^{-1} (Figure 3). The very similar relative energies of **A1**−**A3** are arguably within the accuracy of the DFT calculations used here, indicating no preference for the classical or zwitterionic structures. In addition, the hydrogen-bond networks in both the classical and zwitterionic structures, along with the energy similarity among them, were prone to facilitate rapid proton migration within the ions in the course of dissociation. Accordingly, we found several dAAA$^+$ protomers with different positions of the charging proton that were within 40−80 kJ mol^{-1} above the global energy minimum and may be accessed via proton migration as reactive intermediates upon vibrational excitation.

Proton migrations are shown to be involved in the loss of 5′-adenine (Scheme 2). Dissociation of the N9−C1′ glycosidic bond to 5′-adenine in **A1** was triggered by a relay phosphate → 5′-OH → N3 proton transfer that charged the adenine and increased its leaving group ability. It is worth noting that this dissociation was not predicated on 5′-adenine protonation in the local energy minimum structure of the reactant. The glycosidic bond cleavage was assisted by neighboring group participation by N7 of the 3′-adenine (path a, Scheme 2), or 5′-OH (path b), forming new covalent bonds at C1′ and leading to complexes **A1a** and **A1b** at 104 and 111 kJ mol^{-1}, respectively, relative to **A1**. Loss of neutral adenine as the less stable N3-H tautomer from these complexes forming isomeric products ions **A1c** and **A1d**, required 186 and 198 kJ mol^{-1}, respectively, although it is possible that the departing adenine molecule can undergo an exergonic isomerization to the canonical N9-H tautomer (ΔH_{isom} = −37 kJ mol^{-1})[44] in the

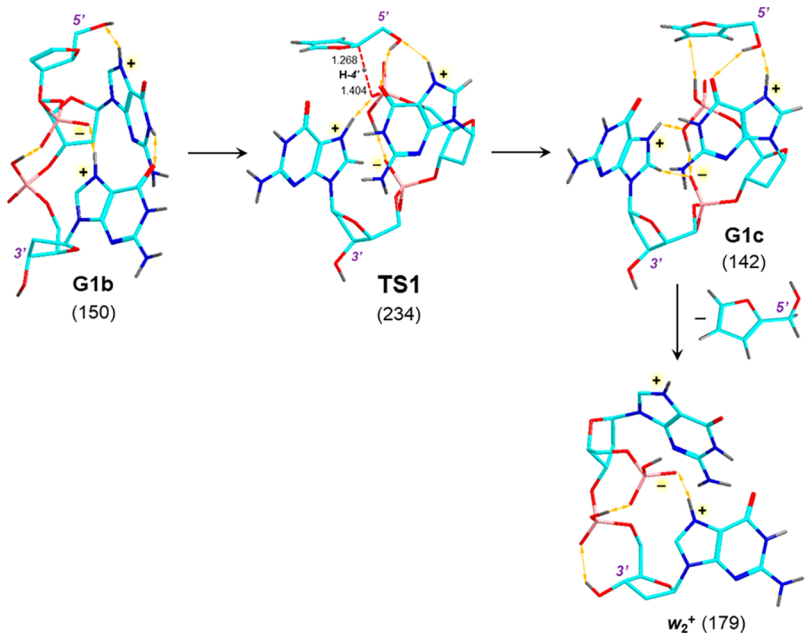
https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936−1950

Scheme 3. Pathways for Loss of 5′-Guanine from G1[a]



[a]Description as in Scheme 2.

Scheme 4. Formation of the w_2^+ Ion from (GGG-5′-G)$^+$[a]



[a]Description as in Scheme 2.

complex and lower the threshold energy. It is interesting to note that the calculated threshold energy for the formation of **A1c** was very similar to that measured for the loss of adenine from protonated 2′-deoxyadenosine (193 kJ mol^{-1}).[13]

3.5. dGGG$^+$ Monocations. In contrast to dAAA$^+$, the dGGG$^+$ ions unambiguously favored classical structures protonated at N7 of 5′-guanine, as illustrated by the global energy minimum **G1** (Scheme 3). Other structures that were protonated at the middle and 3′-guanine were >30 kJ mol^{-1} less stable than **G1**. Structures that were initiated with a zwitterionic proton distribution among the guanine residues and phosphate groups collapsed to N7-protonated protomers upon BOMD and further DFT gradient optimization. This can be related to the exceptional gas-phase basicity of N7 in guanosine (GB = 1036 kJ mol^{-1}, Table S1, Supporting Information) and its orientation with respect to the phosphate groups that disfavored proton transfer and zwitterion formation. Loss of 5′-guanine from **G1** can initially form a C1′ oxocarbenium cation (**G1a**, Scheme 3) which would be analogous to structures suggested for DNA depurination in water (Scheme 1). Interestingly, the loss of 5′-guanine was associated with internal proton migration from the proximate phosphate to 3′-guanine, forming the intermediate ion **G1a** as a zwitterion. **G1a** can get stabilized by deprotonation of the 5′-oxocarbenium ion by the middle guanine residue, forming zwitterionic product **G1b** at 150 kJ mol^{-1} relative to **G1**.

https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936−1950

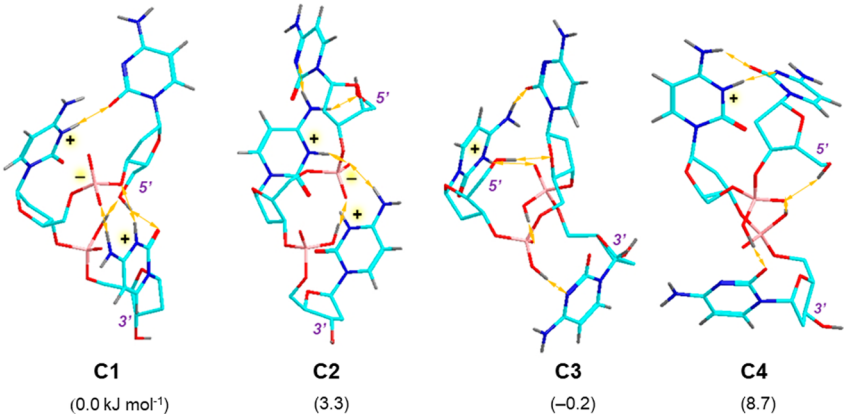


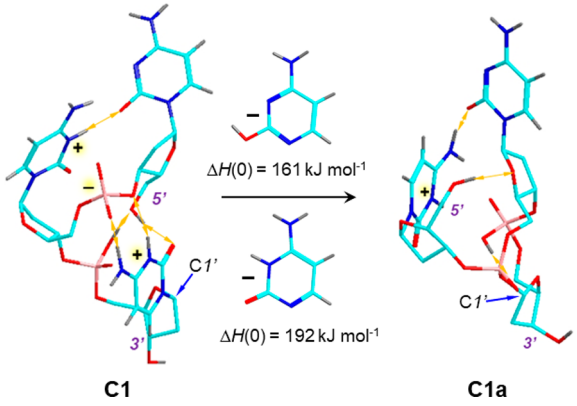
Figure 4. Low Gibbs energy dCCC$^+$ ions. Description as in Figure 3.

Another, albeit somewhat less favorable pathway, can include direct nucleophilic attack at C1′ by the adjacent 5′-OH, protonating the phosphate anion and closing a dioxabicyclo[2.2.1]heptane ring in **G1c** at 175 kJ mol^{-1} relative to **G1** (Scheme 3). The calculated dissociation energies in **G1** were similar to the threshold energy for loss of guanine from protonated 2′-deoxyguanosine (166 kJ mol^{-1}).[14]

The consecutive dissociation of the (MH − 5′-G)$^+$ ion **G1b** forming the w_2^+ fragment ion was investigated by finding the transition state for a proton transfer from C4′ of the 5′-sugar moiety (**TS1**, Scheme 4). The H4′ proton abstraction was facilitated by a relay proton transfer from the proximate phosphoester group to the other phosphate zwitteranion that increased the proximate phosphate basicity. Deprotonation of the 5′-deoxyribose ring followed the cleavage of the O−C3′ bond to reach **TS1** and proceeded by forming a noncovalent complex of the departing hydroxymethylfuran molecule with the w_2^+ ion (**G1c**, 142 kJ mol^{-1} relative to **G1**). The overall formation of w_2^+ from **G1** was calculated to be 179 kJ mol^{-1} endergonic, not including basis set superposition error. The calculated TS energy was then used for calculations of RRKM rate constants for crossing **TS1**, which, being the highest-energy point at 84 kJ mol^{-1} above **G1b**, was the rate-determining step. The rate constant (Figure S4, top panel, Supporting Information), along with the calculated vibrational energy distribution in **G1b** (Figure S4, bottom panel) clearly indicated that **G1b** was kinetically stable at the ion trap temperature (310 K). To proceed on the ion trap time scale (50 ms), the dissociation of **G1b** to the w_2^+ ion required internal excitation of 250−300 kJ mol^{-1}, or an equivalent temperature of 480 K (Figure S4). This was qualitatively consistent with both the low extent of dGGG$^+$ dissociation to w_2^+ following 5′-guanine loss, and facile dissociation of (MH − 5′-G)$^+$ ions upon further collisional activation.

3.6. dCCC$^+$ Monocations. Analysis of BOMD trajectories starting with all theoretically possible combinations of cytosine N3- and O2-protonated tautomers and combined with DFT geometry optimizations yielded several low-energy structures for the dCCC$^+$ monocations that are represented by **C1**−**C4** (Figure 4). All these low-energy dCCC$^+$ ions favored protonation at cytosine positions N3 over O2. However, in contrast to both dAAA$^+$ and dGGG$^+$, the cytidine trinucleotides did not favor protonation at 5′-cytosine. Two low-energy structures, **C1** and **C2**, were zwitterions showing protonation at the middle and 3′-cytosines, in which the latter protonated base was stabilized by strong hydrogen bonds to the phosphate anion (Figure 4). The other low-energy dCCC$^+$ ions were classical structures with N3 protonation at the 5′ and middle cytosines in **C3** and **C4**, respectively. Structure **C3** showed a network of hydrogen bonds allowing a facile shuttle of a proton from 5′-cytosine to the 3′-base which may play a role in dissociations resulting in loss of cytosine. Loss of 3′-cytosine from **C1** was assisted by phosphate participation at C1′, forming the cyclic phosphate triester **C1a** (Scheme 5). The

Scheme 5. Loss of 3′-Cytosine from dCCC$^+$[a]



[a]Description as in Scheme 2.

calculated threshold energy for the loss of 3′-cytosine depended on the neutral cytosine tautomer. A direct loss of an N3-H tautomer, consistent with the proton distribution in **C1**, led to a dissociation energy threshold at 192 kJ mol^{-1}. A prototropic isomerization of the cytosine N3 to the more stable O2-H tautomer in an ion−molecule complex would lower the threshold energy to 161 kJ mol^{-1}. Overall, the preference for 3′-protonated structures was consistent with the favored loss of the 3′-base upon CID. We note that our calculated threshold energies for the loss of 3′-cytosine from C1 were higher than the value reported for the loss of cytosine from 2′-deoxycytidine (146 kJ mol^{-1}).[15]

3.7. dTTT$^+$ Monocations. Initial dTTT$^+$ structures in which combinations of all possible O2 and O4 protonated positions were considered collapsed upon BOMD and geometry optimization to a single lowest energy protomer **T1** (Figure 5). In stark contrast to dAAA$^+$, dGGG$^+$, and dCCC$^+$, structure **T1** was protonated at the 3′-phosphoester

https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936−1950

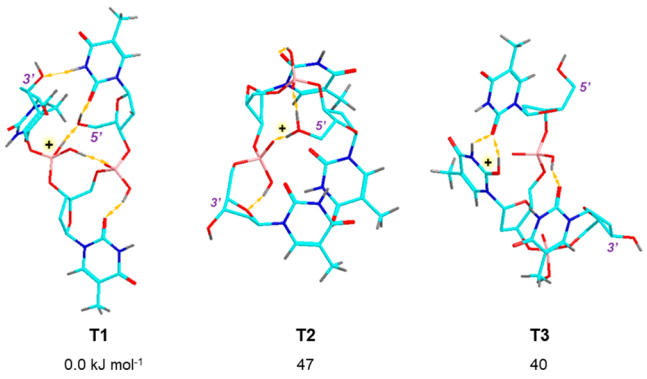


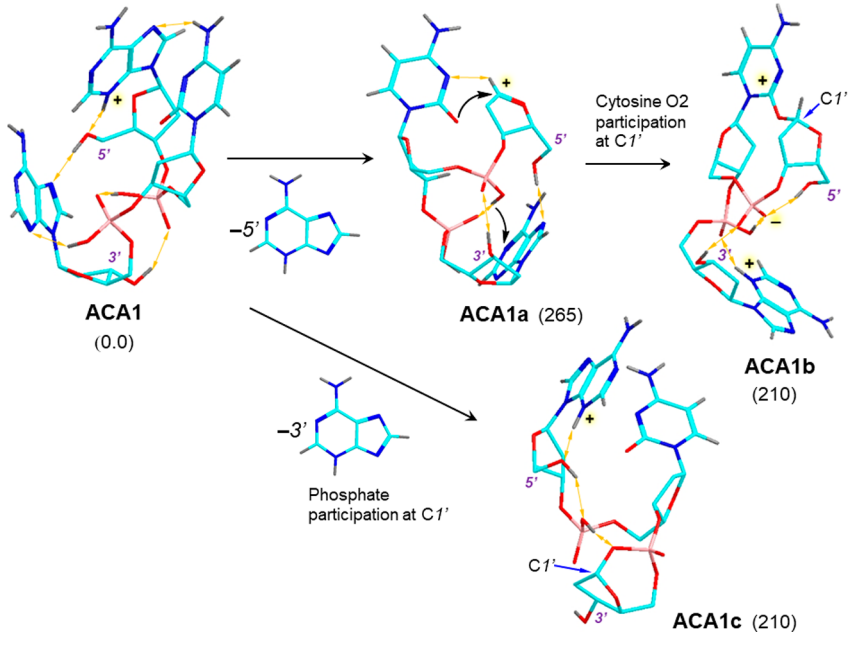
Figure 5. Low Gibbs energy $dTTT^+$ ions. Description as in Figure 3.

group. Thymine protonated structures, as represented by **T3**, were >40 kJ mol^{-1} less stable than **T1**. Comparing these energy differences with the basicity of the components, one finds the gas-phase basicity of triethyl phosphate (879 kJ mol^{-1})[45] to be slightly lower that the GB of 2′-deoxythymidine (886 kJ mol^{-1}, Table S1). However, this small difference can be readily overturned by more favorable phosphate hydrogen bonding in $dTTT^+$ ions. Phosphate protonation in $dTTT^+$ was consistent with a previous study of 2′-deoxythymidine that also preferred phosphate protonated structures.[46] The salient feature of **T1** was the elaborate hydrogen-bond network involving the bases, phosphates, 3′-O, and 5′-O that can allow a proton to migrate among the 3′-, middle, and 5′-thymines and induce the base loss. The CID-MS^2 result, which showed a favored loss of the 3′-thymine, may be due to a favorable dissociation thermodynamics. At the same time, loss of thymine from $dTTT^+$ at 13% combined relative intensity was less efficient in competition with the backbone dissociation forming w_2^+ ions (Table 1).

3.8. $dACA^+$ and $dATC^+$ Monocations. The prevalence of low energy zwitterionic structures for $dAAA^+$ and $dCCC^+$ made it of interest to examine the proton distribution in $dACA^+$ and $dATC^+$ to establish the most stable protomers for these codon sequences and relate the structures to the dissociations by loss of adenine, cytosine, and backbone cleavage. $dACA^+$ ions favored 5′-N3-protonated adenines **ACA1** and **ACA2** as the lowest energy structures, whereas a doubly adenine- protonated zwitterion **ACA3** was 11 kJ mol^{-1} higher by energy than **ACA1** (Figure S5, Supporting Information). Loss of the 3′- and 5′-adenine was found to occur at similar rates, 33% and 38%, respectively (Table 1). The calculated threshold energies for the adenine loss from **ACA1** indicated different reaction pathways. Loss of the 3′-base from **ACA1** can be assisted by participation of the proximate phosphate, forming a cyclic phosphoester **ACA1c** at 210 kJ mol^{-1} threshold energy (Scheme 6). In contrast, loss of 5′-adenine proceeded via a C1′oxacarbenium ion (**ACA1a**) at 265 kJ mol^{-1}, and the product was stabilized in the final product **ACA1b** (at 210 kJ mol^{-1}) by neighboring group participation by 5′-cytosine O2. Interestingly, this stabilization was in part due to the formation of an internally hydrogen bonded zwitterionic structure in **ACA1b** (Scheme 6). Thus, multiple reaction pathways for loss of adenine can occur competitively in $dACA^+$.

Low-energy structures of $dATC^+$ were obtained as zwitterion conformers **ATC1** and **ATC2** and the less stable classical structure **ATC3** (Figure S6, Supporting Information). Both **ATC1** and **ATC2** showed a relay of hydrogen bonds interconnecting adenine and cytosine via the deprotonated phosphate and 5′-hydroxyl groups. Loss of adenine from **ATC1** can proceed with participation by the phosphate anion, forming again a cyclic phosphoester at the 5′-sugar (**ATC1a**, Scheme 7). This dissociation had a low threshold energy (138 kJ mol^{-1}) but likely involved an activation barrier in the transition state. The subsequent loss of the 5′-sugar residue was only mildly endergonic (ΔH_0 = 40 kJ mol^{-1}) to produce the $w_2(TC)^+$ fragment ion. The competitive loss of cytosine

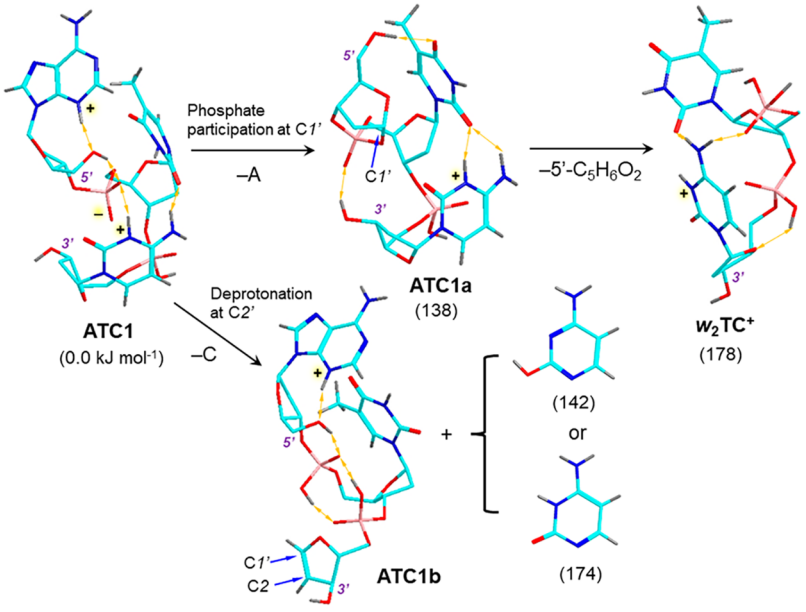
Scheme 6. Loss of 5′- and 3′-Adenine from $dACA^+$[a]



[a]Description as in Scheme 2.

https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936−1950

Scheme 7. Dissociations of the dATC$^+$ ions[a]



[a]Description as in Scheme 2.

Table 2. Fragment Ion Relative Intensities from Dications

	relative fragment ion intensity[a]					relative fragment ion intensity[a]			
codon[b]	5′	middle	3′ + water[c]	w_2[d]	codon[b]	5′	middle	3′ + water[c]	w_2[d]
AAA	96	1	3	16	*ATT*			100 (7)	1
GGG	59	3	38 (2)	18	*GTT*	10		90 (9)	15
CCC	50	1	48 (1)	5	*CTT*	24		76 (15)	38
ATA	47		53 (16)	18	*CAA*	98		2	6
ACA	69	7	24 (3)	17	*GAA*	96		4	7
AGA	26	13	61 (5)	21	*TAA*	5		95 (59)	13
GAG	89	0.4	10	9	ATC	3		97 (3)	0.5
GCG	65	1	34 (3)	7	ACT	8	92		3
GTG	76		24 (7)	26	CAT	99		1	22
CAC	63	0.4	37 (3)	11	CTA	99		1	51
CGC	13	0.3	86 (2)	3	TCA	5	50	45 (25)	6
CTC	64		36 (8)	55	TAC			100 (2)	
AAG[e]	68		32	24	GTC	45		55	3
AAC	46		54 (4)	9	GCT	71	28	1	11
AAT	100			36	CGT	100			5
GGA	88		12	25	CTG	84		16 (3)	46
GGC	44		56	5	TCG	2	25	74 (23)	4
GGT	100			17	TGC		3	97 (6)	0.2
CCA	94		6	14	ACG	50	18	32 (3)	12
CCG	78		22 (1)	8	AGC	4		96 (2)	0.7
CCT	100			14	GAC	85	1	14 (2)	13
TTG	93	2	5	46	GCA	76	1	22 (2)	8
TTA	94		6	18	CGA	71	2	27 (3)	32
TTC	94		6	23	CAG	82	5	14 (2)	23
ACC	16		84 (3)	3	ATG	13		87 (17)	9
GCC	40		60	3	AGT	62	37	1	22
TCC	1		99	6	GAT	99	1		18
AGG	8		92 (7)	5	GTA	95		5 (1)	12
CGG	27		73 (3)	12					
TGG			100 (14)	1					

[a]Percent of summed base-loss ion intensities. [b]^{15}N-labeled nucleosides shown as bold characters; [c]Combined relative intensities for loss of 3′-base and water. Relative intensities for loss of water in parentheses. [d]Percent of combined base loss and w_2^+ ion intensities. [e]Nucleobase loss from positions shown in italics was not resolved, and relative intensities were summed for the unresolved positions.

https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936−1950

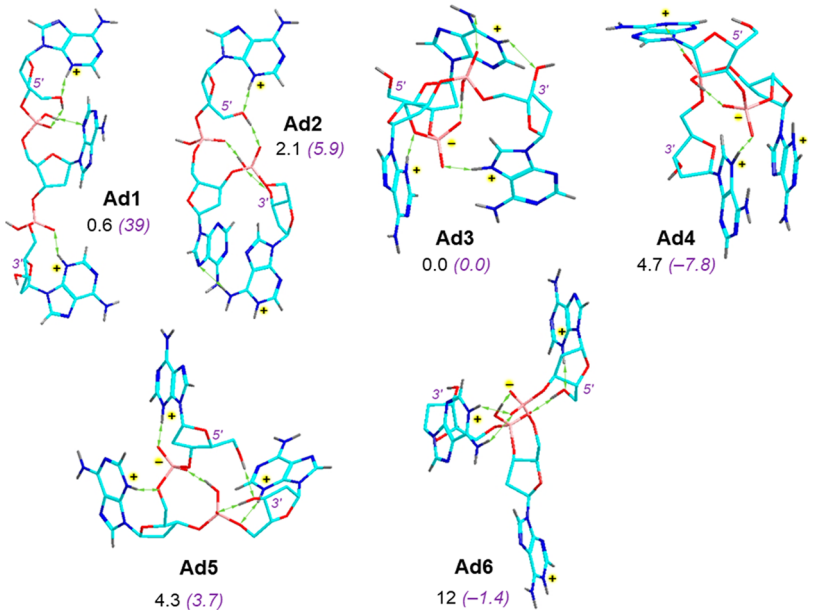


Figure 6. Structures of low Gibbs energy $dAAA^{2+}$ ions. Description as in Figure 3. Values in parentheses include PCM solvation energies (kJ mol^{-1}) in water dielectric.

was another low-energy dissociation of **ATC1** that proceeded with proton abstraction by the phosphate anion, forming **ATC1b** with a 1,2-double bond in the 3′-sugar residue at threshold energies ranging from 142 to 174 kJ mol^{-1}, depending on the neutral cytosine tautomer formed (Scheme 7).

In summary, the monocation structures revealed very different protonation patterns depending on the nucleobase type and position. Guanine-containing trinucleotides favored protonation at the most basic guanine position N7 but without extensive internal solvation by hydrogen bond networks, as illustrated with $dGGG^+$. Adenine- and cytosine-containing ions were more ambiguous, displaying low-energy classical structures with a single protonated nucleobase, as well as zwitterions in which another base was internally protonated by a proximate phosphate and stabilized by an extensive hydrogen bonding network. Thymine residues entirely avoided protonation in their lowest energy protomers which were protonated at the phosphoester groups instead. However, the dissociation energies for the loss of the nucleobase were all substantially above 100 kJ mol^{-1}, allowing access by proton migration to a large number of vibrationally excited monocation protomers and conformers as reactive intermediates within this energy range. Thus, given sufficient internal energy to dissociate, the monocation protomers can isomerize, making the charging proton mobile to access multiple bases in the excited ion. The proton migration can be further facilitated by hydrogen-bond networks in the ions. The dissociation outcome, as to the base loss specificity, appears to be chiefly affected by the stabilization of the intermediate oxacarbenium ions due to neighboring group participation. The stabilizing reactions were identified to include nucleophilic attack and proton abstraction by proximate phosphate groups or nucleobases that appear to be more favorable to occur at the 5′- and 3′-terminal nucleosides.

3.9. Resolution of Identity in Dissociations of Codon Dications. The survey of CID-MS^2 spectra of trinucleotide dications is compiled in Table 2. The survey was limited by the lack of double protonation in dTTT, dTAT, dTGT, and dTCT. Overall, the CID-MS^2 spectra were dominated by competitive losses of protonated nucleobases, AH^+, GH^+ and CH^+, that on average amounted to 84% of the singly charged fragment ion intensities. The formation of singly charged $\boldsymbol{w}_2^+$ and $\boldsymbol{w}_1^+$ ions was observed as the main backbone cleavages, as established by resolution of identity in $dXXX^{2+}$ and $dXYX^{2+}$ dications, as well as read directly from the CID-MS^2 spectra of the $dXYZ^{2+}$ trinucleotides. The average fragment ion relative intensity distribution showed 56%, 8%, and 36% of protonated base loss from the 5′-, middle, and 3′-positions, respectively. Loss of the 3′-protonated base was accompanied by elimination of water, which was less frequent for dications than it was for monocations. The average base-loss frequency favored cytosine (55%) over guanine (33%) and adenine (11%). Loss of protonated thymine was uniformly below 1%. The site specificity for the base loss was further clarified with the help of the ^{15}N-labeled codons. The CID-MS^2 spectrum of $dAAA^{2+}$ showed a predominant loss of 5′-AH^+ at 96%. This was different from the $dAAA^+$ monocation dissociations where loss of neutral 3′-adenine was competitive at 37% (Table 1). The CID-MS^2 spectrum of $dGGG^{2+}$ showed a more balanced loss of GH^+ from the 5′- and 3′-positions at 59% and 38%, respectively. Again, this contrasted the dissociations of $dGGG^+$ monocations where loss of 5′-guanine prevailed at 93% (Table 1). The $dAXA^{2+}$ dications (X = C, G, and T) showed loss of AH^+ from the 5′- and 3′-positions that was affected by the middle base, favoring 3′-adenine for X = T and G, while loss of 5′-AH^+ was favored for X = C. The $dGXG^{2+}$ dications showed loss of 5′-GH^+ which paralleled the loss of neutral 5′-guanine from $dGXG^+$ monocations. The loss of protonated cytosine from $dCXC^{2+}$ dications depended on the middle base. For $dCXC^{2+}$ dications with X = A and T, loss of 5′-CH^+ was favored at 63−64%. In contrast, with X = G, the major dissociation was loss of the 3′-CH^+. We note that complementary pairs of $(nucleobase + H)^+$ and loss-of-base ions were observed for AH^+ and GH^+ in the spectra where their m/z was above the ion trap low-mass cutoff.

https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936−1950

Loss of thymine from the dications showed a dichotomy. In dications having thymine in combination with two basic nucleobases (A, G, or C), loss of thymine was completely suppressed. In contrast, in dications having two thymines, loss of the 5′- or 3′-TH$^+$ was dominant. This may indicate that, in contrast to dTTT$^+$ monocations, thymine was protonated in dications. However, thymine protonation may occur transiently in high-energy structures involved in dissociation rather than in the lowest energy minima.

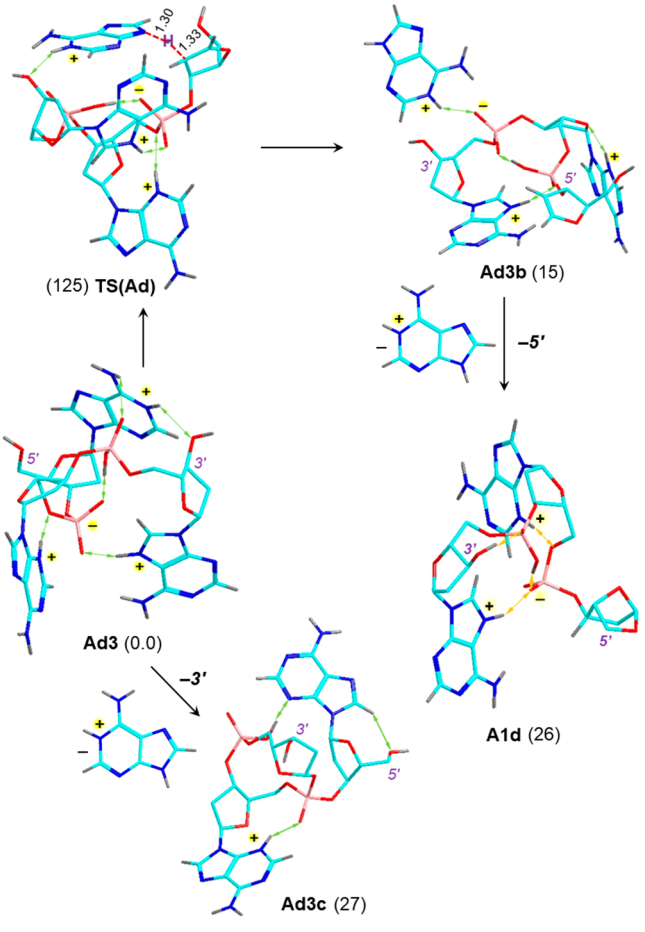
It is noteworthy that the middle nucleobase also had an effect on the competitive backbone dissociations forming w_2^+ ions. Data for dCTC^{2+}, dCTA^{2+}, and dCTG^{2+} showed the relative intensity of w_2^+ ions being close to 50% (Table 2). We attempted to correlate the formation of the w_2^+ ions with the loss of protonated 5′-nucleobases as shown in Figure S7 (Supporting Information). However, the plot showed a low correlation coefficient, $r^2 = 0.23$, indicating that the formation of w_2^+ ions depended on factors other than just the preceding loss of the 5′-nucleobase. The sequential nature of the w_2^+ ion formation was indirectly indicated by the absence of the complementary a_1^+ ions in the CID-MS^2 spectra of the dications. The sequential formation of w_2^+ ions was also consistent with the CID-MS^2 spectra of trinucleotide dications that were measured at several excitation amplitudes and showed a weak increase of the w_2^+ ion relative intensity at the highest excitation (Figure S8, Supporting Information). In contrast, the relative intensities of fragment ions by competitive loss of the nucleobases were insensitive to the excitation amplitude within the range causing 5 to >95% precursor ion dissociation.

3.10. Codon Dication Structures and Dissociation Mechanisms. The effects on trinucleotide dication dissociations of the nucleobase nature and position, as revealed by the CID-MS^2 spectra, were multivariate, showing no simple trends. We attempted to gain insight into the dissociation mechanisms from calculations of structures and relative energies of selected protomers and conformers. For dAAA^{2+}, all 18 theoretically possible combinations, resulting from placing two protons at the basic N1, N3, and N7 adenine positions, were considered in the initial structures submitted to BOMD. The calculated lowest Gibbs energy dAAA^{2+} ion structures showed a substantial variability as to the protonation sites and conformations. Among those, the lowest-energy classical isomers **Ad1** and **Ad2** were protonated at N3 and N1 in the 5′- and 3′-adenines. These ions showed extended conformations in which the neutral middle adenine was hydrogen bonded to the phosphate and 3′-adenine ion in **Ad1** and **Ad2**, respectively (Figure 6). Low energy zwitterionic isomers were obtained that underwent phosphate proton transfer to the third adenine residue. The lowest energy zwitterion (**Ad3**) showed protonation at N3, N1, and N7 in the 5′-, middle, and 3′-adenines, respectively. The phosphate groups in **Ad3** participated in a hydrogen-bonding network, which can facilitate proton transfer that is needed for the elimination of protonated adenine. Similar patterns, albeit with different proton distribution among the adenine basic positions were obtained for low energy zwitterionic protomers **Ad4**-**Ad6** (Figure 6). Solvation by water, as modeled by PCM calculations, in general favored the zwitterionic structures of which **Ad4** was the global Gibbs energy minimum.

The abundant formation of doubly charged dAAA^{2+} and other trinucleotide ions upon electrospray was remarkable when one considers that the trinucleotides were present at equilibrium as phosphate dianions in aqueous solution, and so four protons had to be attached in electrospray microdroplets to form the gas-phase dications. We examined by calculations the gas-phase basicity for proton abstraction from the lowest Gibbs energy dication (**Ad3**) to the lowest energy monocation (**A1**). The GB_{310} value, 781 kJ mol^{-1}, was substantially lower than GB_{310}(2′-deoxyadenosine) (953 kJ mol^{-1} at the same level of theory, Table S1 Supporting Information). This energy lowering may be in part caused by charge repulsion in the multiply charged ion. However, the diminished gas phase basicity of **A1** was sufficient to resist deprotonation of the doubly charged ions by solvent molecules in the electrospray plume.

The facile dissociation of dAAA^{2+} was examined by calculations of threshold energies for the loss of AH$^+$ from the 3′- and 5′-positions. Protonated adenine has two nearly isoenergetic protomers (N1-H, N9-H, and N3-H, N7-H),[47−49] so considering any of these as fragments gave practically identical threshold energies. The dissociations showed very low threshold energies for loss of both the 5′- and 3′-adenine ions, forming ions **A1d** and **Ad3c** at ΔH_0 = 26, and 27 kJ mol^{-1}, respectively (Scheme 8). These dissociation energies were substantially lower than those for the loss of neutral adenine from dAAA$^+$ monocations (186 kJ mol^{-1}, Scheme 2), indicating that the protonated adenine moiety was a substantially better leaving group than neutral adenine. We

Scheme 8. Dissociation Pathways for the Loss of 5′- and 3′-Adenine Ions from dAAA^{2+}[a]



[a]Description as in Scheme 2.

https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936−1950

note that the classical ion structure **Ad3c** in Scheme 8 was not the global energy minimum among the $(AAA^{2+} - AH^{+})^{+}$ ions. Prototropic isomerization to the lower-energy zwitterionic isomer **A1c** (Scheme 2) would lower the dissociation energy of **Ad3** to 17 kJ mol^{-1}, respectively, further increasing the energy gap between the monocation and dication dissociations. The dissociation thermochemistry was qualitatively consistent with the CID-MS^2 spectra where loss of the protonated nucleobase was the dominating process. However, the slightly lower energy threshold for the loss of the 3′-adenine ion contrasted the more favorable loss of the 5′-adenine ion in the spectrum. This indicated that the dissociation kinetics was determined by activation energies in the reaction steps preceding fragment ion separation.

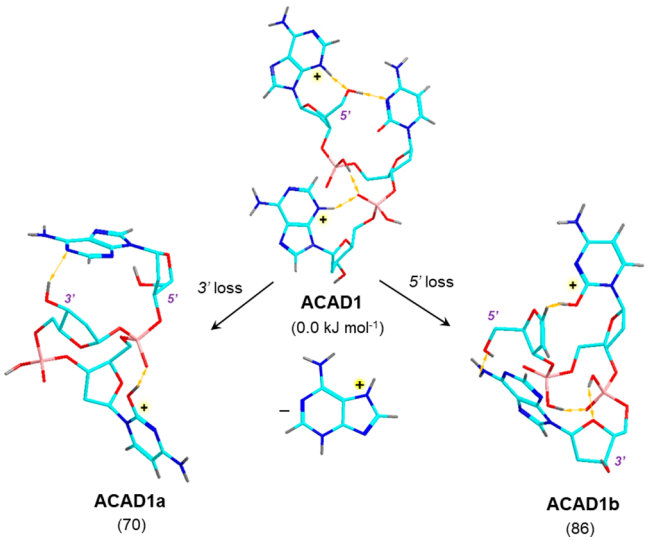
To address this issue, we investigated in detail the reaction coordinate for the 5′-adenine loss from **Ad3** as summarized in Figure S9 (Supporting Information). One-dimensional elongation of the 5′-adenine N9−C1′ bond was continuously endothermic (Figure S9, top panel), according to DFT calculations. Adding atomic thermal motion at 510 K, resulted in a reorientation of the departing adenine moiety, starting at an N9−C1′ bond length of 3.1 Å and leading to exothermic transfer of the β-H2′ to adenine N9 (Figure S9, bottom panel) within a few picoseconds, forming a 2,3-glycal at the 5′-deoxyribose. We were able to find a transition state ($\mathbf{TS_{Ad}}$) for the proton transfer at 125 kJ mol^{-1} relative to **Ad3** (Scheme 8). The $\mathbf{TS_{ad}}$ energy was significantly lower than the TS energies for adenine loss from $dAAA^{+}$ monocations, attesting to the enhanced leaving group ability of charged adenine, and consistent with the general features of the CID-MS^2 spectra. Interestingly, the newly formed adenine cation did not depart the $(MH - 5'\text{-adenine})^{+}$ fragment ion (**A1d**) but remained bound in a stable ion−ion complex (**Ad3b**) at 15 kJ mol^{-1} above **Ad3** that was weakly held by $\Delta H_0 = 22$ kJ mol^{-1} ($\Delta G_{310} = -7.5$ kJ mol^{-1}) against dissociation.

A bound cation−cation complex such as **Ad3b** was unexpected, indeed, in view of the long-range Coulomb repulsion between the ions. Inspection of the **Ad3b** structure revealed that the adenine cation was hydrogen bonded to the negative pole of the phosphate zwitteranion, which contributed to the complex overall stability. Both the product and intermediate complex energies in the dication dissociations were substantially lower than those calculated for the 5′-adenine loss from the singly charged ion (Scheme 2). This was also consistent with the CID-MS^2 spectra where loss of protonated bases from the dications were the dominant dissociations.

We made another comparison between the structures and dissociation energies for dACA monocations (Scheme 6) and dications. Among the 12 theoretically possible protomers combining double protonation at adenine N1, N3, and cytosine O2 and N3, the low-energy dication structures obtained by BOMD and DFT geometry optimization all had the 5′-adenine N3-H, 3′-adenine N3-H protonation pattern. The global energy minimum conformer **ACAD1** is shown in Scheme 9. The other protomeric structures that were protonated at cytosine N3 (**ACAD2**) or 3′-adenine N1-H were less stable (Figure S10, Supporting Information). No low energy zwitterionic structures were identified for dACA dications.

Deprotonation of **ACAD1** to the lowest energy monocation **ACA1** was endergonic by $\Delta G_{310} = 807$ kJ mol^{-1}, which represented the gas-phase basicity of the monocation. It is

Scheme 9. Dissociations of $dACA^{2+}$ Dications[a]



[a]Description as in Scheme 2.

noteworthy to compare the basicity of $dACA^{+}$ with that of $dAAA^{+}$ (781 kJ mol^{-1}, vide supra), indicating a lower stabilization of the additional protonated adenine residue in the $dAAA^{2+}$ dications. According to the CID-MS^2 data, protonated adenine accounted for 93% of the base loss from dACA dications, favoring 5′-adenine over 3′-adenine. Therefore, we investigated by calculations the adenine ion loss thermochemistry and structures of the fragment ions. Loss of the 3′-adenine ion was assisted by neighboring group participation by the 5′-phosphate, forming a cyclic phosphotriester **ACAD1a** at 70 kJ mol^{-1} relative to **ACAD1**. Loss of the 5′-adenine ion was accompanied by H2′ proton transfer from the 5′-deoxyribose oxacarbenium ion to the cytosine O2, forming a glycal structure for the product ion **ACAD1b** at 86 kJ mol^{-1} relative to **ACAD1**. These threshold energies were substantially lower than those for loss of neutral adenine from dACA monocations (210 kJ mol^{-1}, Scheme 4), indicating again that protonated adenine was a better leaving group than the neutral molecule. The relative dissociation energies for the 5′- and 3′-adenine ion loss from **ACAD1** did not conform to the distribution of the corresponding fragment ions in the identity-resolved CID-MS^2 spectrum, which favored the 5′-base loss over the 3′-base loss by a 69:24 ratio (Table 2). The structures of the fragment ions indicated complex pathways that included cleavage of the glycosidic C1′-N9 adenine bond, followed by proton transfer to the departing adenine moiety. Stabilization of the putative oxacarbenium ions by nucleophilic attack by the phosphate or proton abstraction by the cytosine O2 could occur in the intermediate complex or after protonated adenine departure.

4. CONCLUSIONS

The combined results of this experimental and computational study allow us to arrive at the following conclusions. Resolution of identity in dissociations of trinucleotide ions related to DNA codons revealed in detail the positional propensity of adenine, guanine, cytosine, and thymine to dissociate as neutral molecules or protonated nucleobases. Loss of the nucleobase from the terminal 5′- or 3′-positions was favored. According to combined BOMD and DFT

https://doi.org/10.1021/jasms.2c00194
J. Am. Soc. Mass Spectrom. 2022, 33, 1936−1950

calculations, low-energy protomers of trinucleotide mono- and dications were represented by classical nucleobase-protonated structures as well as zwitterions in which an additional nucleobase was internally protonated by a phosphate group. Dissociations of adenine and cytosine containing ions were accompanied by proton migrations that were induced by vibrational excitation and facilitated by extensive networks of intramolecular hydrogen bonds. These proton migrations occurred not only at the reaction center, but also were associated with the allosteric reorganization of the hydrogen bonding framework. The calculated dissociation and transition-state energies demonstrated that protonated adenine was a substantially better leaving group than neutral adenine, a conclusion that can be generalized for all other nucleobases.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jasms.2c00194.

Table of gas-phase basicities, supplementary CID-MS^2 spectra, optimized structures, kinetic plots (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Aleš Marek − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Prague 6, Czech Republic;* orcid.org/0000-0001-9031-8263; Email: ales.marek@uochb.cas.cz

František Tureček − *Department of Chemistry, University of Washington, Seattle, Washington 98195-1700, United States;* orcid.org/0000-0001-7321-7858; Phone: +1-206-685-2041l; Email: turecek@uw.edu

Authors

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

Břetislav Brož − *Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 16610 Prague 6, Czech Republic;* orcid.org/0000-0002-2142-1045

Yue Liu − *Department of Chemistry, University of Washington, Seattle, Washington 98195-1700, United States*

Shu R. Huang − *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.2c00194

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, Grant No. CHE-1951518. F.T. acknowledges support by the Klaus and Mary Ann Saegebarth Endowment. Research at the IOCB was supported by the Ministry of Education, Youth and Sport (MSMT INTER-EXCELLENCE LTAUSA19094).

■ REFERENCES

(1) An, R.; Jia, Y.; Wan, B.; Zhang, Y.; Dong, P.; Li, J.; Liang, X. Non-Enzymatic Depurination of Nucleic Acids: Factors and Mechanisms. *PLoS One* **2014**, *9*, e115950.

(2) Lindahl, T. Instability and Decay of the Primary Structure of DNA. *Nature* **1993**, *362*, 709−715.

(3) Lindahl, T.; Nyberg, B. Rate of Depurination of Native Deoxyribonucleic Acid. *Biochemistry* **1972**, *11*, 3610−3618.

(4) Lindahl, T.; Andersson, A. Rate of Chain Breakage at Apurinic Sites in Double-Stranded Deoxyribonucleic Acid. *Biochemistry* **1972**, *11*, 3618−3623.

(5) Lindahl, T. Instability and Decay of the Primary Structure of DNA. *Nature* **1993**, *362*, 709−715.

(6) Schaaper, R. M.; Glickman, B. W.; Loeb, L. A. Role of Depurination in Mutagenesis by Chemical Carcinogens. *Cancer Res.* **1982**, *42*, 3480−3485.

(7) Matange, K.; Tuck, J. M.; Keung, A. J. DNA Stability: A Central Design Consideration for DNA Data Storage Systems. *Nature Commun.* **2021**, *12*, 1358.

(8) Garrett, E. R.; Mehta, P. Solvolysis of Adenine Nucleosides. I. Effects of Sugars and Adenine Substituents on Acid Solvolyses. *J. Am. Chem. Soc.* **1972**, *94*, 8532−8541.

(9) Suzuki, T.; Ohsumi, S.; Makino, K. Mechanistic Studies on Depurination and Apurinic Site Chain Breakage in Oligodeoxyribonucleotides. *Nucleic Acids Res.* **1994**, *22*, 4997−5003.

(10) Garrett, E. R. Kinetics of the Hydrolytic Degradation of a Nucleoside, the Antibiotic Psicofuranine. *J. Am. Chem. Soc.* **1960**, *82*, 827−832.

(11) Hevesi, L.; Wolfson-Davidson, E.; Nagy, J.; Nagy, O.; Bruylants, A. Contribution to the Mechanism of the Acid-Catalyzed Hydrolysis of Purine Nucleosides. *J. Am. Chem. Soc.* **1972**, *94*, 4715−4720.

(12) Tureček, F. Transient Intermediates of Chemical Reactions by Neutralization-Reionization Mass Spectrometry. *Top. Curr. Chem.* **2003**, *225*, 77−129.

(13) Wu, R. R.; Rodgers, M. T. Mechanisms and Energetics for N-Glycosidic Bond Cleavage of Protonated Adenine Nucleosides: N3 Protonation Induces Base Rotation and Enhances N-Glycosidic Bond Stability. *Phys. Chem. Chem. Phys.* **2016**, *18*, 16021−16032.

(14) Wu, R. R.; Chen, Yu; Rodgers, M. T. Mechanisms and Energetics for N-Glycosidic Bond Cleavage of Protonated 2′-Deoxyguanosine and Guanosine. *Phys. Chem. Chem. Phys.* **2016**, *18*, 2968−2980.

(15) Wu, R. R.; Rodgers, M. T. O2 Protonation Controls Threshold Behavior for N-Glycosidic Bond Cleavage of Protonated Cytosine Nucleosides. *J. Phys. Chem. B* **2016**, *120*, 4803−4811.

(16) Devereaux, Z. J.; He, C. C.; Zhu, Y.; Roy, H. A.; Cunningham, N. A.; Hamlow, L. A.; Berden, G.; Oomens, J.; Rodgers, M. T. Structures and Relative Glycosidic Bond Stabilities of Protonated 2′-Fluoro-Substituted Purine Nucleosides. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1521−1536.

(17) Devereaux, Z. J.; Roy, H.A.; He, C.C.; Zhu, Y.; Cunningham, N.A.; Hamlow, L.A.; Berden, G.; Oomens, J.; Rodgers, M.T. Influence of 2′-Fluoro Modification on Glycosidic Bond Stabilities and Gas-Phase Ion Structures of Protonated Pyrimidine Nucleosides. *J. Fluorine Chem.* **2019**, *219*, 10−22.

(18) Wu, R. R.; Yang, Bo; Berden, G.; Oomens, J.; Rodgers, M. T. Gas-Phase Conformations and Energetics of Protonated 2′-Deoxyadenosine and Adenosine: IRMPD Action Spectroscopy and Theoretical Studies. *J. Phys. Chem. B* **2015**, *119*, 2795−2805.

(19) 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.

(20) Vrkic, A. K.; O'Hair, R. A. J.; Foote, S.; Reid, G. E. Fragmentation Reactions of all 64 protonated Trimer Oligonucleotides Via Tandem Mass spectrometry in an Ion Trap. *Int. J. Mass Spectrom.* **2000**, *194*, 145−164.

(21) White, H. A.: Manual of Oligonucleotide Synthesis Using the Phosphoramidite Method. In *Methods in Molecular Biology: New Nucleic Acid Techniques*; Walker, J. M., Ed.; Humana Press, 1988; Vol. 4, pp 193−213.

(22) Pradeepkumar, P. I.; Hobartner, C.; Baum, D. A.; Silverman, S. K. DNA-Catalyzed Formation of Nucleopeptide Linkages. *Angew. Chem., Int. Ed.* **2008**, *47*, 1753−1757.
(23) ATDBio Ltd.: Solid-Phase Oligonucleotide Synthesis, https://atdbio.com/nucleic-acids-book/Solid-phase-oligonucleotide-synthesis (accessed 2020-06-12).
(24) Liu, Y.; Turecek, F. Photodissociative Crosslinking of Diazirine-Tagged Peptides with DNA Dinucleotides in the Gas Phase. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1992−2006.
(25) 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.
(26) 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.
(27) Stewart, J. J. P. *MOPAC 16. Stewart Computational Chemistry*; Colorado Springs, CO, 2016.
(28) Řezáč, J. Cuby: An Integrative Framework for Computational Chemistry. *J. Comput. Chem.* **2016**, *37*, 1230−1237.
(29) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(30) 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.
(31) Huang, S. R.; Tureček, F. Noncanonical Isomers of Nucleoside Cation Radicals: An Ab Initio Study of the Dark Matter of DNA Ionization. *J. Phys. Chem. A* **2022**, *126*, 2480−2497.
(32) 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.
(33) 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.
(34) 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.
(35) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999−3093.
(36) Zhu, L., Hase, W. L.: *Quantum Chemistry Program Exchange*. Indiana University, Bloomington, 1994; QCPE 644.
(37) Gregersen, J. A.; Tureček, F. Mass-Spectrometric and Computational Study of Tryptophan Radicals (Trp + H)● Produced by Collisional Electron Transfer to Protonated Tryptophan in the Gas Phase. *Phys. Chem. Chem. Phys.* **2010**, *12*, 13434−13447.
(38) McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L. Tandem Mass Spectrometry of Small, Multiply Charged Oligonucleotides. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60−70.
(39) Murray, K. K. DNA Sequencing by Mass Spectrometry. *J. Mass Spectrom.* **1996**, *31*, 1203−1215.
(40) 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.
(41) Rodgers, M. T.; Campbell, S.; Marzluff, E. M.; Beauchamp, J. L. Site-Specific Protonation Directs Low-Energy Dissociation Pathways of Dinucleotides in the Gas Phase. *Int. J. Mass Spectrom. Ion Processes* **1995**, *148*, 1−23.
(42) Huang, T.; Kharlamova, A.; Liu, J.; McLuckey, S. A. Ion Trap Collision-Induced Dissociation of Multiply Deprotonated RNA: c/y Ions versus (a-B)/w-Ions. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1832−1840.
(43) Zima, V.; Liu, Y.; Tureček, F. Radical Cascade Dissociation Pathways to Unusual Nucleobase Cation Radicals. *J. Am. Soc. Mass Spectrom.* **2022**, *33*, 1038−1047.
(44) Huan g, S. R.; Dang, A.; Tureček, F. Ground and Excited States of Gas-Phase DNA Nucleobase Cation Radicals. A UV-vis Photodissociation Action Spectroscopy and Computational Study of Adenine and 9-Methyladenine. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 1271−1281.
(45) *NIST Chemistry Webbook, Standard Reference Database Number 69*; https://webbook.nist.gov/chemistry (accessed June 2022).
(46) Wu, R. R.; Hamlow, L. A.; He, C. C.; Nei, Y.-w.; Berden, G.; Oomens, J.; Rodgers, M. T. The Intrinsic Basicity of the Phosphate Backbone Exceeds that of Uracil and Thymine Residues: Protonation of the Phosphate Moiety is Preferred over the Nucleobase for pdThd and pUrd. *Phys. Chem. Chem. Phys.* **2017**, *19*, 30351−30361.
(47) Tureček, F.; Chen, X. Protonated Adenine: Tautomers, Solvated Clusters, and Dissociation Mechanisms. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1713−1726.
(48) Marian, C.; Nolting, D.; Weinkauf, R. The Electronic Spectrum of Protonated Adenine: Theory and Experiment. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3306−3316.
(49) Hud, N. V.; Morton, T. H. DFT Energy Surfaces for Aminopurine Homodimers and Their Conjugate Acid Ions. *J. Phys. Chem. A* **2007**, *111*, 3369−3377.