Gas-Phase Conformations of Deprotonated and Protonated Mononucleotides Determined by Ion Mobility and Theoretical Modeling

Jennifer Gidden and Michael T. Bowers*

Department of Chemistry & Biochemistry, University of California, Santa Barbara, California 93106

Received: June 25, 2003; In Final Form: September 18, 2003

Collision cross-sections of protonated and deprotonated 2′-deoxy-5′-mononucleotide ions were measured in helium using ion mobility based methods. Various computational methods were then used to generate candidate structures of the ions and calculate their collision cross-sections for comparison to experiment. For the deprotonated systems, the ion mobility data indicate that only one family of conformers is present. Molecular mechanics calculations (AMBER) predict that the sugar is twisted into a $C_{3'}$-endo conformation with the phosphate and base above the plane of the sugar. For dAMP, dCMP, and dTMP, the base is in an "anti" conformation about the glycosidic bond but in dGMP, the guanine base is in a "syn" position with the amino group hydrogen-bonded to the deprotonated phosphate. DFT calculations yield similar low-energy structures as those predicted by AMBER. The calculated cross-sections of the $C_{3'}$-endo conformers agree very well with experimental values (1−2% difference). For the protonated mononucleotides, the ion mobility results also indicate that only one family of conformers is present. However, unlike the deprotonated systems, theory predicts that the sugars and the bases are in an opposite orientation ($C_{2'}$-endo vs $C_{3'}$-endo and syn vs anti) with respect to how they are positioned in the deprotonated series. DFT calculations predict that the lowest energy structures are protonated at N3 in dAMP, N7 in dGMP, N3 in dCMP, and the phosphate group in dTMP. AMBER calculations, with the bases protonated at the DFT predicted sites, yield structures that agree very well with experiment (1−2% difference).

Introduction

The original view of the structure of DNA as proposed by Watson and Crick is that of a right-handed (clockwise) double helix in which the two strands are held together by complimentary hydrogen-bonding between bases on each strand.[1] Later studies, however, showed that DNA is a somewhat flexible molecule and can actually adopt a number of different types of structures. It can form various double-helix structures[2,3] and even flip over to become a left-handed double helix.[4,5] DNA can bend due to specific sequences,[6,7] or protein binding,[8,9] or gene expression.[10] DNA has even been shown to form triplex and quadruplex structures.[11−15]

DNA is a relatively simple polymer consisting of repeating nucleotide units that are composed of a phosphate group, sugar, and one of four possible bases (A, C, G, or T). It is the composition and sequence of these nucleotides that allows DNA to encode the information necessary to develop living organisms. It is also the variations in the structures of the nucleotides (in particular, the sugar, base, and their relative positions to each other) that lead to the variations observed in the overall helical structures of DNA. For example, the switching of a sugar pucker from a $C_{2'}$-endo configuration to a $C_{3'}$-endo configuration and the rotation of a base from an "anti" conformation to a "syn" conformation accompanies the transition of a right-handed "B" helix into a left-handed "Z" helix.[5] Adenine (A) can protonate at low pH and form A^+·C and A^+·G base pairs instead of the "complimentary" A·T base pairs.[16] The protonation of cytosine (C) leads to C^+·G base pairs that help stabilize triplex formations.[17,18] Although plenty of data have supported the view of the "original" structure of DNA, relatively little is known about the alternate structures DNA can fold into. Thus, it is important to first understand the structural features of the nucleotides to better understand the overall structure of DNA and its various helices.

Most structural studies of DNA have been performed in the condensed phase, using X-ray crystallography or, more recently, 2-D and 3-D NMR, designed to mimic the actual environment in which DNA exists. Although the results yield valuable information about how DNA folds in solution, the solvent effects can hinder investigations focusing on the intrinsic structural properties of DNA. Recent advances in mass spectrometric methods have led to an increase in the number of structural studies on DNA performed in the gas phase. Base compositions and sequences of oligonucleotides have been determined by mass spectrometry[19−22] as have structural properties of DNA−protein complexes and DNA duplexes.[23−25] H/D exchange studies have been used to investigate the conformations and hydrogen-bonding interactions in oligonucleotides.[25−28] Collision-induced-dissociation (CID) and metastable decay studies have been used to investigate the structures and fragmentation mechanisms of oligonucleotides.[29−32]

Advances in computing methods have also led to an increase in the number of theoretical studies probing the structures of DNA. Tremendous effort has been put into developing the proper force field parameters that yield accurate predictions for DNA (as well as protein) folding. Higher level calculations are also being used to predict structures of DNA. The electronic properties of oligonucleotides have recently been investigated using photoelectron studies.[33,34] Interpretation of the spectra requires accurate information about the structures of the

* To whom correspondence should be addressed. E-mail address: bowers@chem.ucsb.edu.

10.1021/jp035894z CCC: $25.00
Published on Web 10/30/2003

oligonucleotides and the nature of the highest occupied molecular orbitals (HOMOs). Ab initio calculations are often used to provide the necessary information and predict vertical detachment energies for comparison with experiment.

Of course, the big question is how well these theoretical methods predict structures for DNA. Can DFT predict accurate structures for these relatively "large" molecules? How well parametrized are the molecular mechanics calculations? To answer some of these questions, the gas-phase conformations of protonated and deprotonated 2′-deoxy-5′-mononucleotides were examined by ion mobility experiments[35,36] and theoretical modeling and the results are presented here. The experiments consist of measuring the time it takes for a pulse of mass-selected ions to drift through a buffer gas under the influence of a weak electric field. This drift time is directly proportional to the shape of the ion. Molecular mechanics calculations (using the AMBER force field) and DFT calculations were then used to generate trial structures of the mononucleotide ions and compare them with the experimental results.

Experimental Methods

Detailed information about the instrument used for the ion mobility experiments has been published[37] so only a brief description is given here. Protonated and deprotonated mononucleotide ions were formed in a home-built MALDI source.[38] 2, 5-Dihydroxybenzoic acid (DHB) was used as the matrix and methanol and water were used as solvents. Approximately 100 μL of a 50:50 mixture of DHB (100 mg/mL in methanol) and the desired mononucleotide (1 mg/mL in 90% methanol/10% water) was applied to a stainless steel cylindrical target and dried. The target is then fitted onto a screw-threaded device in the MALDI source which, when turned by a motor, rotates and translates the target so that fresh areas of sample are exposed to the laser shots.

Ions exiting the MALDI source are accelerated to 5 kV and mass analyzed with a reverse-geometry sector mass spectrometer. For the ion mobility measurements, the mononucleotide ions of interest are mass selected, decelerated to ~10 eV, and injected into a 4 cm long drift cell filled with 3 Torr of helium. The ions drift through the gas at constant velocity under the influence of a weak electric field and are detected as a function of time, yielding an arrival time distribution or ATD (the timer is triggered by the laser pulse used to generate the ions in the MALDI source). These arrival times are directly proportional to the ions' collision cross-sections with helium and, therefore, give information about the geometric shape of the ion. Compact, folded ions with small collision cross-sections drift through the cell faster than more extended ions that have larger cross-sections and thus have shorter arrival times. If the ions have multiple conformations with significantly different cross-sections (>3%), each conformer will be separated in the drift cell and appear at different times in the ATD[39−42] (assuming the ions do not interconvert as they drift[42−45]). Cooled nitrogen flowing through passages surrounding the cell is used to measure mobilities of ions at different temperatures. The temperature can be controlled within ±1 K.

The mobilities (K_o) of the mononucleotide ions are determined from a series of ATD measurements, in which only the voltage across the drift cell is changed, using

$$t_A = t_d + t_o = \left(\frac{273l^2}{760T}\frac{p}{V}\frac{1}{K_o}\right) + t_o \qquad (1)$$

where t_A is the arrival time of the ions (determined from the center of the ATD peak), t_d is the amount of time the ions spend inside the drift cell (inversely proportional to K_o), t_o is the amount of time the ions spend outside the drift cell, l is the drift length, T is temperature, p is the pressure of the helium gas, and V is the voltage applied across the drift cell.[46] A plot of t_A vs p/V yields a straight line with a slope inversely proportional to K_o and an intercept of t_o. Once the mobilities of the ions are known, their corresponding collision cross-sections, σ, can be determined from kinetic theory using

$$\sigma = \frac{3q}{16N_o}\left(\frac{2\pi}{\mu kT}\right)^{1/2}\frac{1}{K_o} \qquad (2)$$

where q is the ion charge, N_o is the number density of helium at standard temperature and pressure, μ is the ion − He reduced mass, k is Boltzmann's constant, and T is temperature.[46]

Theoretical Methods

Conformational information about the ions is obtained by comparing calculated cross-sections of theoretical structures to the experimental values obtained from the ATDs using eqs 1 and 2. The AMBER 7.0 set of molecular mechanics/dynamics programs[47] was used to generate 100 candidate structures of each deprotonated and protonated mononucleotide via a simulated annealing procedure. In this process, an initial structure is energy minimized, run through a 30 ps MD simulation at 800 K, cooled to 0 K through another 10 ps MD run, and energy minimized again. This final structure and its energy are saved. This structure is then used as the starting point for another annealing cycle. The angle-averaged collision cross-section of each minimized structure was calculated using a temperature-dependent projection model[48,49] that has given reliable results for a number of biological and synthetic polymers.[38,42−45,48−52] A scatter plot of cross-section vs relative energy of the 100 theoretical structures is then used to identify conformational families and determine the most likely conformations of the ions responsible for the measured ATDs. Comparison of calculated and experimental cross-sections can be determined two ways. In most cases, the average cross-section of the lowest 2 kcal/mol structures in a given conformational family was used for comparison to experiment. In the other method, molecular dynamics simulations are used to compare theory with experiment. In this case, the lowest energy structure in a given conformational family is used as the starting structure for a 300 K dynamics simulation run for 1000 ps. Every 1 ps, the structure is saved and its cross-section calculated. The average of the 1000 cross-section calculations is then used in comparison with experiment. For the mononucleotides discussed here, the conformers stayed in their respective "families" throughout the 300 K dynamics and the resulting average cross-section was the same as the average obtained from the lowest 2 kcal/mol structures.

The projection model was also used to calculate the cross-sections of theoretical structures generated by DFT calculations. For the deprotonated mononucleotides, structures were obtained from E. Vorpagel at PNNL who used the hybrid B3LYP exchange-correlation functional and the TZVP basis set with additional diffuse functions.[53] For the protonated mononucleotides, the calculations were performed at UCSB using the Gaussian98 package[54] with the standard B3LYP/6-31G* basis set. In both cases, the initial starting structures were obtained from the AMBER calculations. The angle-averaged cross-section of each DFT structure was calculated 10 separate times and the

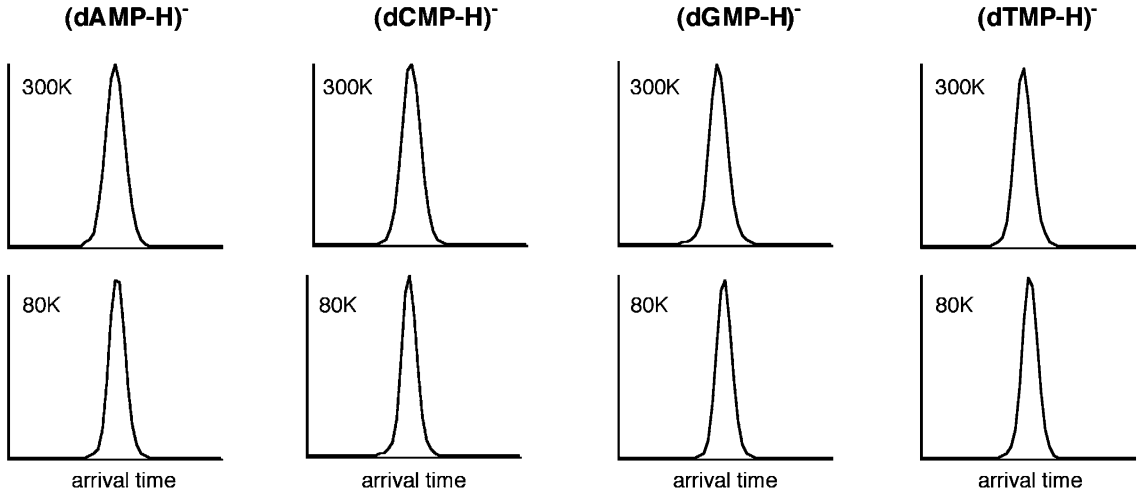

Figure 1. Arrival time distributions (ATDs) for the deprotonated mononucleotides $(M - H)^-$ measured at 300 and 80 K. The time scale for each ATD is 100 μs.

TABLE 1: Experimental and Calculated Collision Cross-Sections (Å^2) for the deprotonated mononucleotides $(M - H)^-$

		Amber		DFT	
	experiment	$C_{3'}$[a]	$C_{2'}$[a]	syn[b]	anti
dAMP	106 ± 1	106 ± 2	109 ± 2	109 ± 2	107 ± 2
dGMP	103 ± 1	104 ± 2	113 ± 2	105 ± 2	115 ± 2
dCMP	98 ± 1	98 ± 2	102 ± 2	104 ± 2	100 ± 2
dTMP	100 ± 1	102 ± 2	108 ± 2	110 ± 2	103 ± 2

[a] Sugar is in a $C_{2'}$- or $C_{3'}$-endo conformation (see Figures 2 and 3). [b] When the base is in the "syn" position for dAMP, dCMP, and dTMP, the phosphate is below the plane of the sugar, similar to the $C_{2'}$-endo conformer shown in Figure 2.

overall average used for comparison to the experimental and AMBER results.

Results and Discussion

A. Deprotonated Mononucleotides. Figure 1 shows typical arrival time distributions (ATDs) for the deprotonated mononucleotides measured at drift cell temperatures of 300 and 80 K. Single, symmetric peaks appear in all of the ATDs at both temperatures. This indicates that either one family of conformers is present or, if multiple conformers exist, they either have very similar cross-sections or rapidly interconvert in the drift cell. Previous ion mobility studies have shown that multiple conformers differing in cross-section by ≥5% can be separated in the drift cell and appear as multiple peaks in the ATDs.[35−45] Additionally, temperature-dependent ion mobility experiments on deprotonated dinucleotides and trinucleotides showed that conformers rapidly interconverting at 300 K could be separated at 80 K (where the isomerization significantly slowed or stopped) as long as the barrier to isomerization was ≥0.8 kcal/mol.[44,45] Because single peaks are present in the 300 and 80 K ATDs shown in Figure 1, the mononucleotide ions most likely have just one family of conformers.

Collision cross-sections of the ions can be determined from the ATDs using eqs 1 and 2 and are listed in Table 1 for the 300 K data. As expected, the cross-sections of the pyrimidine systems (dCMP and dTMP) are smaller than the purine systems (dAMP and dGMP). However, the cross-section for dGMP is smaller than that of dAMP, even though guanine is a slightly larger base than adenine. Thus, dGMP may fold somewhat differently than the other mononucleotides.

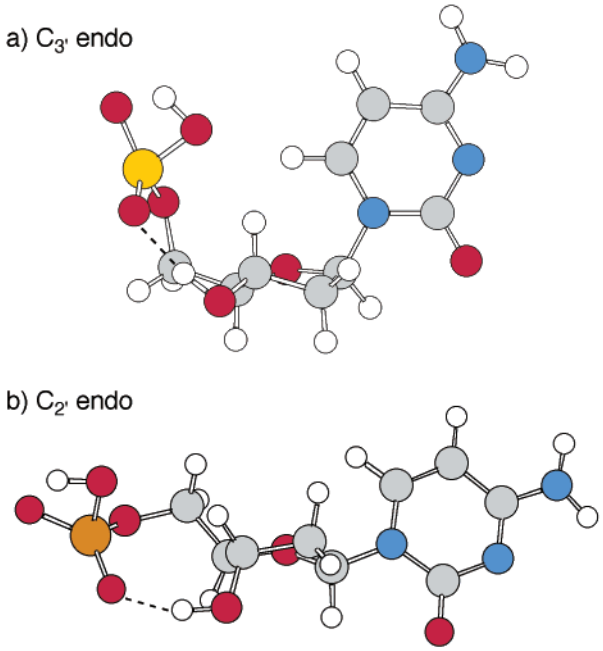

Figure 2. Examples of the two families of conformers predicted by AMBER for the deprotonated mononucleotides. Carbons are gray, oxygens are red, nitrogens are blue, hydrogens are white, and the phosphorus atom is orange. In (a) the sugar is twisted into a $C_{3'}$-endo conformation with the base and phosphate above the plane of the sugar. In (b) the sugar is twisted into a $C_{2'}$-endo conformation and the phosphate is now below the plane of the sugar.

1. AMBER Structures. Also listed in Table 1 are calculated cross-sections of the theoretical models of the mononucleotide ions predicted by AMBER. For each mononucleotide, 100 candidate structures were generated and their cross-sections calculated. AMBER predicts two low-energy conformational families for each mononucleotide, depending on the twist configuration of the sugar. Examples of each family are shown in Figure 2. These conformers differ in energy by 1−4 kcal/mol and in cross-section by 3−9 Å^2, depending on the base.

In the smaller cross-section family (Figure 2a), the sugar is twisted into a $C_{3'}$-endo conformation in which the 3′ carbon is bent up from the plane of the sugar. The 5′ phosphate group and the base are also above the plane of the sugar. Because the 3′ carbon in the sugar bends up toward the 5′ end of the nucleotide, the 3′ OH group is pulled in that direction as well

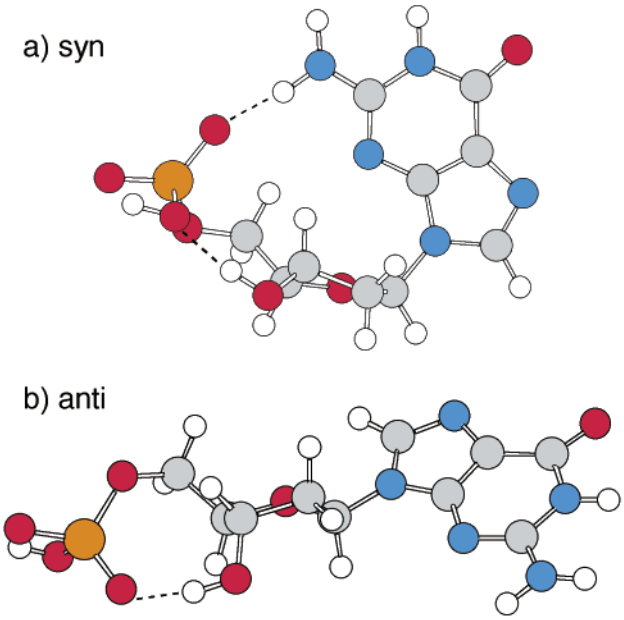

Figure 3. Examples of the (a) "syn" and (b) "anti" conformers predicted for $(dGMP - H)^-$ by AMBER. In the "syn" conformer, the sugar is in a $C_{3'}$-endo conformation and the base is in a syn position about the glycosidic bond with the amino group on guanine hydrogen bonded to the deprotonated phosphate. In the "anti" conformer, the sugar is in a $C_{2'}$-endo conformation and guanine is in an anti position about the glycosidic bond.

and hydrogen bonds to the deprotonated phosphate. For dAMP, dCMP, and dTMP, the base is in an "anti" position with respect to the glycosidic bond but is tilted toward the phosphate group so that the H6 (dCMP, dTMP) or H8 (dAMP) atom on the base is directly over the sugar. For dGMP, however, the guanine base rotates into a "syn" position so that the N3 atom is directly over the sugar and the amino group hydrogen bonds to the deprotonated phosphate. This "syn" conformation is shown in Figure 3a and explains why the cross-section for dGMP is smaller than that of dAMP.

In the larger cross-section family (Figures 2b and 3b), the sugar is twisted into a $C_{2'}$-endo conformation with the 2′ carbon bent out of the plane of the sugar. As a result, the phosphate and base are no longer above the plane of the sugar. The 5′ phosphate still hydrogen bonds to the terminal 3′ OH group, but, in this case, the phosphate folds toward the 3′ end. All of the bases are in an "anti" position (even G) but extend away from the phosphate so that no part of the base is directly over the sugar.

Although AMBER predicts similar energies for the $C_{3'}$- and $C_{2'}$-endo conformers, the ATDs for the deprotonated mononucleotides (Figure 1) indicate that only one family of conformers is present (or the conformers are rapidly isomerizing in the drift cell). The resulting experimental collision cross-sections agree very well with those calculated for the $C_{3'}$ conformers with differences falling between 1 and 2%. However, the calculated cross-sections for the $C_{2'}$ conformers are consistently larger (3−9%) than experiment. Thus, either the mononucleotide ions are only in a $C_{3'}$ conformation or, if both conformers are present and isomerize in the drift cell, the $C_{3'}$ conformers are the dominant form.

For example, the cross-section difference between the "syn" ($C_{3'}$) and "anti" ($C_{2'}$) conformers of $(dGMP - H)^-$ is 10%, so they should easily be separated in the drift cell and appear as two distinct peaks in the ATD if they do not interconvert. If the conformers do interconvert as they drift, the ATD should show a single peak with an average drift time that falls between the drift times of pure "syn" and "anti" ions (weighted by the relative abundance of each conformer). The resulting collision cross-section obtained from the ATD measurements should then fall somewhere between the calculated values of the "syn" and "anti" conformers.[42,44] From Table 1, the experimental cross-section for $(dGMP - H)^-$ agrees very well with the theoretical value of the "syn" conformer but not the "anti" conformers. Therefore, if both conformers were present in the drift cell, the "syn" conformers must be the dominant form.

Additionally, if the $C_{3'}$ and $C_{2'}$ conformers (or "syn" and "anti") are interconverting in the drift cell, the isomerization barrier between the two forms must be very small because one peak appears in the 80 K ATDs. Ion mobility studies on deprotonated dinucleotides indicated that conformers with isomerization barriers of at least 1 kcal/mol can be separated in the drift cell at 80 K (or at least yield an unsymmetrical ATD). Hence, if both $C_{3'}$ and $C_{2'}$ conformers were present in substantial amounts and a significant barrier existed between the two forms (>1 kcal/mol), two separate peaks, or an unsymmetrical single peak, should have been observed in the 80 K ATDs.

For $(dTMP - H)^-$, the cross-sections of the $C_{2'}$ conformers are 6% larger than those of the $C_{3'}$ conformers, which should at least yield an ATD peak with a significant shoulder if not two distinct peaks. However, like the dGMP system, the ATDs for $(dTMP - H)^-$ show single, symmetric peaks and yield cross-sections that agree very well with the theoretical value of the $C_{3'}$ conformer and not that of the $C_{2'}$ conformer. Hence, if both conformers are present in the drift cell and rapidly interconvert, the $C_{3'}$ conformers are the dominant forms.

For dAMP and dCMP, the cross-section difference between the $C_{3'}$ and $C_{2'}$ conformers is 3−4%. Although this difference in size will not yield two separate peaks in the ATD, the ATD peak should not be symmetric if the two conformers do not interconvert. But, like dGMP and dTMP, if both conformers of $(dAMP - H)^-$ and $(dCMP - H)^-$ are present and isomerizing in the drift cell, the $C_{3'}$ conformers must be the dominant form because their cross-sections agree much better with experimental values.

2. Density Functional Theory Structures. Table 1 also lists the calculated cross-sections of theoretical models of the deprotonated mononucleotides predicted by DFT calculations. For each calculated structure, its cross-section was calculated 10 times and averaged to give the value listed in Table 1. The lowest energy DFT structures are very similar to the $C_{3'}$-endo conformers predicted by AMBER. The 5′ phosphate group and the base are above the plane of the sugar (which is in a $C_{3'}$-endo conformation) with the 3′ OH group hydrogen-bonded to the deprotonated phosphate. For dAMP, dCMP, and dTMP, the bases are in an "anti" conformation about the glycosidic bond whereas in dGMP guanine is rotated into a "syn" position with the amino group hydrogen-bonded to the deprotonated phosphate (see Figures 2 and 3). The cross-sections of these structures agree well with experimental values but not as well as those calculated for the AMBER structures. The source of the difference most likely lies in the fact that the cross-section of a single, frozen DFT structure is being compared to experiment rather than the average cross-section of a family of structures as done with the AMBER structures. The ions certainly undergo thermal motion as they drift through the cell and the resulting experimental cross-section is really an average for a conformational family, not a single conformation. With the AMBER results, this motion can be accounted for by averaging the cross-sections of a conformational family (consisting of 10−50

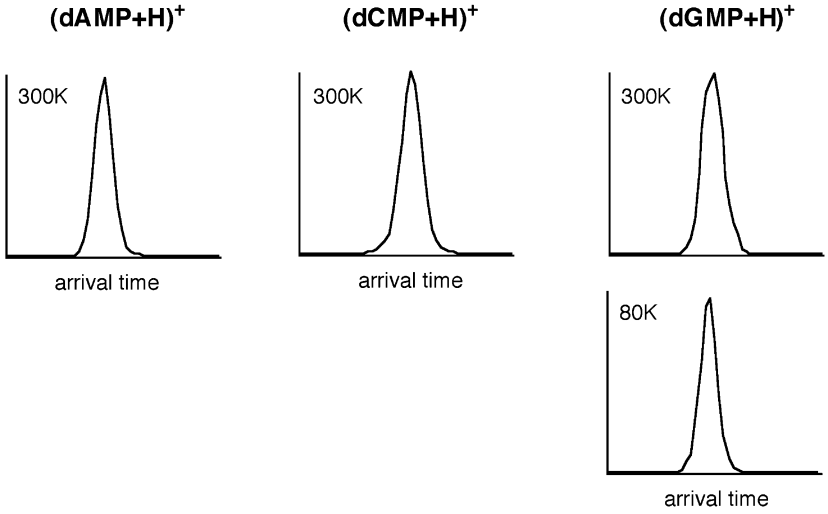

Figure 4. Arrival time distributions (ATDs) for the protonated mononucleotides $(M + H)^+$ measured at 300 and 80 K (for dGMP). The time scale for each ATD is 100 μs.

TABLE 2: Relative Energy Differences (kcal/mol) between the "Syn" and "Anti" Forms of the Deprotonated Mononucleotides $(M - H)^-$ Predicted by DFT Calculations

	dAMP	dGMP	dCMP	dTMP
syn	5.3	0.0	12.0	5.8
anti	0.0	11.4	0.0	0.0

structures) that have slightly different bond and dihedral angles or by running molecular dynamics simulations that allow a given structure to undergo thermal motion. These options are not available for the DFT structures.

Both "syn" and "anti" conformers were optimized in the DFT calculations. Except for dGMP, the "syn" conformers resemble the $C_{2'}$-endo conformers predicted by AMBER in that the phosphate group has moved below the plane of the sugar to hydrogen bond to the 3′ OH group. However, the sugar remains in a $C_{3'}$-endo conformation. The relative differences in energy between the "syn" and "anti" conformers are listed in Table 2. For dAMP, dCMP, and dTMP, the "syn" conformers are higher in energy than their "anti" counterparts whereas in dGMP the trend is reversed, a result consistent with the experimentally observed cross-sections.

B. Protonated Mononucleotides. Figure 4 shows the ATDs obtained for the protonated mononucleotides at 300 and 80 K. Like the deprotonated systems, single symmetric peaks appear in the spectra, indicating that only one conformational family is likely present for each mononucleotide. Collision cross-sections obtained from the 300 K ATDs listed in Table 3. Except for dAMP, the cross-sections of the protonated mononucleotides are consistently larger than their deprotonated counterparts. The largest change occurs for dGMP, where the cross-section for $(dGMP + H)^+$ is 8% larger than that of $(dGMP - H)^-$. For dTMP, only sodiated species were observed in the mass spectra and thus no experimental data are available for $(dTMP + H)^+$.

In the protonated mononucleotides, several sites on the bases are available for protonation. Ab initio calculations on free nucleobases indicate that the most favorable site of protonation is N1 on adenine, N7 on guanine, O2 or N3 on cytosine, and O4 on thymine.[55,56] H/D exchange studies and semiempirical calculations (AM1) on protonated mononucleotides indicate that the most probable site of protonation is N3 on dAMP, N3 or N7 on dGMP, N3 or O2 on dCMP, and the phosphate group on dTMP.[25,26] In this study, the AMBER and DFT calculations were performed for dAMP protonated at N1, N3, and N7, dGMP protonated at N3, O6, and N7, dCMP protonated at O2 and N3, and dTMP protonated at P═O, O2, and O4.

TABLE 3: Experimental and Calculated Collision Cross-Sections (Å^2) for the Protonated Mononucleotides $(M + H)^+$

		AMBER		DFT	
	experiment	syn	anti	syn	anti
dAMP	106 ± 1				
N1		104 ± 2	111 ± 2	106 ± 2	112 ± 2
N3		105 ± 2	*a*	107 ± 2[b]	110 ± 2
N7		*a*	110 ± 2	106 ± 2	110 ± 2
dGMP	111 ± 1				
N3		107 ± 2	*a*	108 ± 2	113 ± 2
O6		106 ± 2	112 ± 2	107 ± 2	112 ± 2
N7		*a*	110 ± 2	108 ± 2	113 ± 2[b]
dCMP	102 ± 1				
O2		101 ± 2	100 ± 2	103 ± 2	104 ± 2
N3		101 ± 2	101 ± 2	103 ± 2	103 ± 2[b]
dTMP	*c*				
P═O		106 ± 2	*a*	105 ± 2[b]	104 ± 2
O2		106 ± 2	*a*	107 ± 2	106 ± 2
O4			105 ± 2	105 ± 2	106 ± 2

[a] No "syn" or "anti" conformers observed in calculations. [b] Lowest energy structure (see Table 4). [c] Only sodiated dTMP ions were observed.

In the molecular mechanics/dynamics studies, the theoretical models of each mononucleotide protonated at a given site were generated via the annealing process described earlier. AMBER predicts one or two families of conformers depending on the base and the site of protonation. Unlike the deprotonated mononucleotides, where the difference between the two conformational families depended on the twist configuration of the sugar, the two families in the protonated series depend on the orientation of the base. In each mononucleotide, regardless of protonation site, the sugar is always twisted into a $C_{2'}$-endo configuration and the 5′ phosphate is always above the plane of the sugar and not hydrogen bonded to the terminal 3′ OH group. The differences lie in whether the base is in a "syn" or "anti" orientation about the glycosidic bond.

1. dAMP. For $(dAMP + H)^+$ protonated at N1, AMBER predicts two families of conformers: "syn" and "anti". In the lowest energy and smaller cross-section family, the adenine base is in a "syn" position, so that N3 is directly over the sugar, but it is not hydrogen bonded to the phosphate or sugar. The "anti" conformers, with H6 on adenine directly over the sugar, are 8

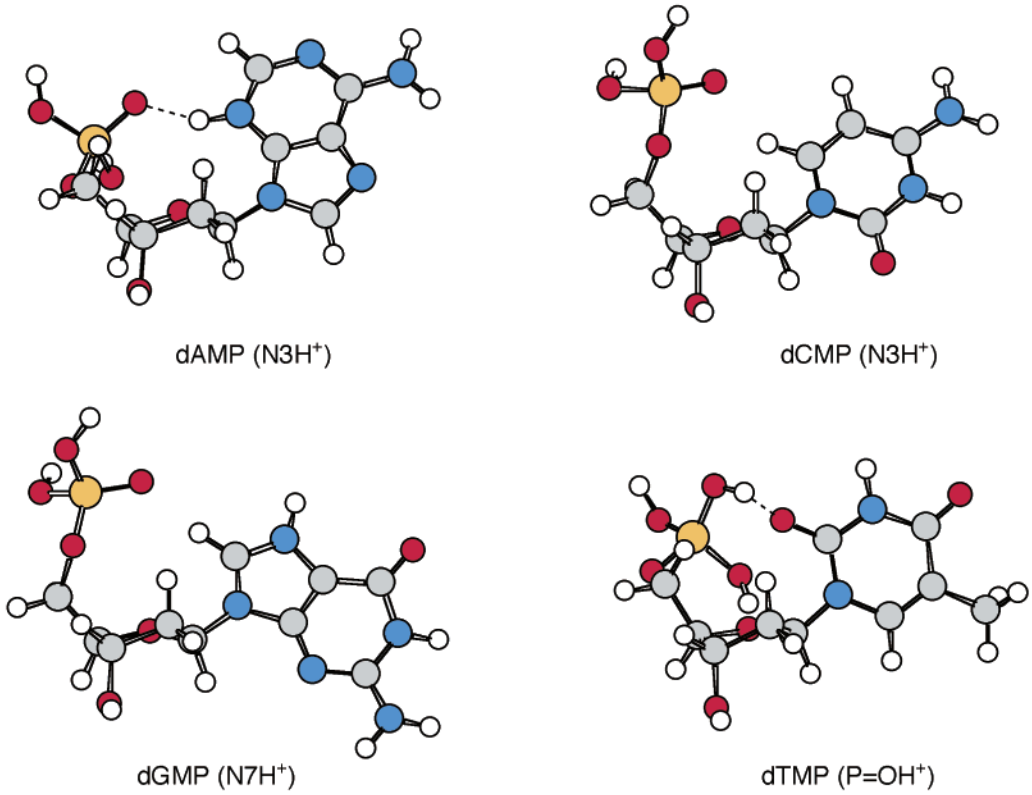

Figure 5. Lowest energy structures for the protonated mononucleotides $(M + H)^+$ as determined by DFT calculations. The most favorable protonation sites are N3 in dAMP, N7 in dGMP, N3 in dCMP, and the phosphate group (P═O) in dTMP. The bases are in "syn" positions in dAMP and dTMP and in "anti" positions in dGMP and dCMP. Except for dCMP, this is opposite to the preferred base positions in the corresponding deprotonated series.

TABLE 4: Relative Energy Differences (kcal/mol) between the Mononucleotides with the Base in a "Syn" or "Anti" Position, Protonated at Various Sites

	dAMP			dGMP			dCMP		dTMP		
	N1	N3	N7	N3	O6	N7	O2	N3	P═O	O2	O4
syn	10.7	0.0	20.5	5.2	8.3	1.8	9.2	5.4	0.0	6.6	11.4
anti	14.2	12.2	9.3	26.5	13.5	0.0	5.3	0.0	15.5	17.8	5.7

kcal/mol higher in energy and have cross-sections 7 Å^2 larger than the "syn" conformers. For $(dAMP + H)^+$ protonated at N3, AMBER predicts only "syn" conformers with the H^+ on N3 hydrogen bonded to the phosphate. An example is shown in Figure 5. For $(dAMP + H)^+$ protonated at N7, only "anti" conformers are predicted by AMBER.

The calculated cross-sections of each family are compared to experimental values in Table 3. The cross-sections of the "syn" conformers agree very well with experiment (<2% error) whereas those of the "anti" conformers are 4−5% larger than experiment. Thus, the $(dAMP + H)^+$ mononucleotides are most likely in a "syn" conformation with a $C_{2'}$-endo sugar twist. This is the exact opposite of the deprotonated $(dAMP - H)^-$ mononucleotides, in which adenine is in an "anti" conformation and the sugar is twisted into a $C_{3'}$-endo configuration.

Unfortunately, the collision cross-sections of the "syn" conformers of $(dAMP + H)^+$ protonated at N1 and N3 are the same so that ion mobility experiments alone cannot definitively determine the site of protonation (although it does rule out protonation at N7, which yield only "anti" conformers). Therefore, DFT calculations were used to determine the most energetically favorable site of protonation. The results are listed in Table 4. Conformers with the base in a "syn" and "anti" orientation were optimized.

Our DFT calculations on $(dAMP + H)^+$ agree with the semiempirical results that N3 is the most favorable site of protonation. This structure is shown in Figure 5. The adenine base is in a "syn" position where the H^+ on N3 is hydrogen bonded to the 5′ phosphate. Protonation at N3 is the only site that allows this type of interaction to occur. If the adenine base is forced into an "anti" position, the resulting structure is 12 kcal/mol higher in energy, which may be why AMBER never predicts this structure.

For protonation at N1 (where AMBER predicts two families), the "anti" conformer is only 3.5 kcal/mol higher in energy than the "syn" conformer. However, both conformers are significantly less stable (11-14 kcal/mol) than the "syn" conformer protonated at N3. For protonation at N7, DFT predicts that the "anti" conformer is 11 kcal/mol more stable than the "syn" conformer, in agreement with AMBER, which predicted only "anti" conformers. Thus, from the results presented in Tables 3 and 4, $(dAMP + H)^+$ is most likely protonated at N3 with the adenine base in a "syn" position and hydrogen bonded to the phosphate.

2. dGMP. Three base protonation sites are also available in dGMP, at N3, O6, and N7. For $(dGMP + H)^+$ protonated at N3, AMBER predicts only "syn" conformers with the amino group and the H^+ on N3 hydrogen bonded to the phosphate. For $(dGMP + H)^+$ protonated at O6, AMBER predicts two families of structures with the base in a "syn" or "anti" position. In the lowest energy structures, the guanine base is in a "syn" position with the amino group hydrogen bonded to the phosphate and the H^+ on O6 bent toward N7. The "anti" conformers are 7−8 kcal/mol higher in energy and have 6 Å^2 larger cross-sections. For $(dGMP + H)^+$ protonated at N7, AMBER predicts only "anti" conformers. An example is shown in Figure 5.

The calculated cross-sections of each family are listed in Table 3 along with experimental values. In this case, the cross-sections of the "anti" conformers agree very well with experiment (1% error) whereas those of the "syn" conformers are 4−5% smaller than experiment. Hence, like the dAMP series, the conformations of protonated dGMP mononucleotides appear to be opposite that of the deprotonated mononucleotides (where G is in a "syn" position). Also, like the dAMP mononucleotides, the cross-sections of the "anti" conformers of dGMP protonated at O6 and N7 are the same so that the site of protonation cannot be determined from a conformational standpoint. Therefore, DFT calculations were again used to determine the most energetically favored site of protonation.

The results of the DFT calculations on $(dGMP + H)^+$ protonated at N3, O6, and N7 are listed in Table 4. The calculations predict that N7 is the most favorable site. This structure is shown in Figure 5. In this case, the guanine base prefers to be in an "anti" position. However, if the guanine base is forced into a "syn" conformation, the resulting structure is only 2 kcal/mol higher in energy. In this case, the amino group on guanine hydrogen bonds to the phosphate. No "syn" conformers were observed in the 100 structures of $(dGMP + H)^+$ protonated at N7 predicted by AMBER.

Protonation at O6, with guanine in an "anti" position, yields a structure that is 13 kcal/mol higher in energy than the one shown in Figure 5. The "syn" conformer is predicted to be 6 kcal/mol more stable than the "anti" form, in line with the AMBER results. For $(dGMP + H)^+$ protonated at N3, the "syn" conformer is predicted to be the most stable with the "anti" conformer 21 kcal/mol higher in energy.

Even though DFT calculations predict that the "syn" and "anti" conformers of $(dGMP + H)^+$ protonated at N7 are energetically similar, the AMBER and ion mobility results appear to indicate that only the "anti" conformers are present. ATDs for $(dGMP + H)^+$ measured at 80 K (Figure 4) show just one peak in the spectra, indicating that either one conformer is present or if "syn" and "anti" conformers exist, they rapidly interconvert in the drift cell. The cross-section difference between the "syn" and "anti" forms is large enough (10%) that both conformers should be easily separated. As in the deprotonated system, if "syn" and "anti" conformers were present, the isomerization barrier must be extremely small. In this case, however, the "anti" conformers would be the dominant form because the experimental cross-section agrees very well with the calculated value for the "anti" form and not the "syn" form. Molecular dynamics simulations on $(dGMP + H)^+$ protonated at N7 run for 1000 ps at 800 K indicate that the guanine base never converts from an "anti" position into a "syn" position. If the structure is started with guanine in a "syn" position, it immediately converts into an "anti" position. Hence, all results are consistent with an N7 protonated base in the "anti" position.

3. dCMP. Two sites are available for protonation on dCMP, at N3 and O2. AMBER predicts two families of structures, "syn" and "anti", for each case. In the lowest energy family, the cytosine base is in an "anti" position where the H6 atom on the base is directly over the sugar. However, the "syn" conformers, with O2 directly over the sugar, are predicted to be only 2−3 kcal/mol less stable and they have the same cross-section as the "anti" conformers, as shown in Table 3. The calculated cross-sections of each family in each protonated system agree very well with experimental values. Unfortunately, because all of the calculated cross-sections are the same, the ion mobility results cannot be used to determine the site of protonation or distinguish between "syn" and "anti" conformers.

The relative energies of the "syn" and "anti" conformers of $(dCMP + H)^+$ protonated at N3 and O2 as determined by DFT calculations are shown in Table 4. The results indicate that N3, with cytosine in an "anti" position, is the preferred protonation site. This lowest energy structure is shown in Figure 5. Like dAMP and dGMP, the sugar is twisted into a $C_{2'}$-endo configuration instead of a $C_{3'}$-endo configuration, as observed in deprotonated $(dCMP - H)^-$. However, the cytosine base in $(dCMP + H)^+$ prefers to be in an "anti" position, just as it is in the deprotonated system. dCMP is the only mononucleotide in which the base is predicted to be in an "anti" position in the lowest energy structures of both the deprotonated and protonated systems. If the base is rotated into a "syn" position, the resulting structure is 5 kcal/mol less stable. For $(dCMP + H)^+$ protonated at O2, DFT calculations also predict that the "anti" conformer is the lowest energy structure, although this structure is 5 kcal/mol higher in energy than the "anti" conformer protonated at N3. When cytosine is rotated into a "syn" position, the resulting structure is 4 kcal/mol less stable than the "anti" conformer.

4. dTMP. Three protonation sites need to be considered for $(dTMP + H)^+$, O2, O4, or the phosphate group.[23−26] The results for $(dTMP + H)^+$ protonated at P═O, O2, and O4 as determined by AMBER and DFT calculations are given in Tables 3 and 4.

For $(dTMP + H)^+$ protonated on the phosphate, AMBER predicts only one family of conformers. An example is shown in Figure 5. Like the other three mononucleotides, the sugar is in a $C_{2'}$ endo twist configuration. The thymine base is in a "syn" position with O2 on thymine hydrogen bonded to the protonated phosphate. Another OH group on the phosphate is hydrogen bonded to O4′ atom on the sugar. For $(dTMP + H)^+$ protonated at O2, AMBER also predicts only "syn" conformers. In this case, the H^+ on O2 is hydrogen bonded to the phosphate. For $(dTMP + H)^+$ protonated at O4, only "anti" conformers are predicted by AMBER.

Like dCMP, the calculated cross-sections of $(dTMP + H)^+$ are the same regardless of protonation site or whether the base is in a "syn" or "anti" position. Thus, the site of protonation and orientation of the base cannot be determined on the basis of conformation alone. The relative energies of $(dTMP + H)^+$ protonated at the various sites with thymine in a "syn" or "anti" position determined by DFT calculations are listed in Table 4.

The DFT calculations agree with the semiempirical calculations[25,26] in that P═O is the most favored protonation site. Protonation at P═O is favored over O2 by 6.6 kcal/mol and over O4 by 5.7 kcal/mol. The lowest energy structure is shown in Figure 5. In this case, the thymine is in a "syn" position with O2 hydrogen bonded to the phosphate, similar to the structure predicted by AMBER. Like dAMP and dGMP, the conformations of the sugar and base in the lowest energy structure of $(dTMP + H)^+$ are opposite to that found in the lowest energy structure of $(dTMP - H)^-$. The sugar has switched from a $C_{3'}$ endo conformation in $(dTMP - H)^-$ to a $C_{2'}$-endo conformation in $(dTMP + H)^+$, and the base has switched from an "anti" position in $(dTMP - H)^-$ to a "syn" position in $(dTMP + H)^+$. If thymine is rotated into an "anti" position in $(dTMP + H)^+$ protonated at P═O, the resulting structure is 15.5 kcal/mol higher in energy.

For $(dTMP + H)^+$ protonated at O4, DFT predicts that the "anti" conformer is more stable than the "syn" conformer by 5.7 kcal/mol and is less stable than the lowest energy structure by 5.7 kcal/mol. For $(dTMP + H)^+$ protonated at O2, AMBER predicts only "syn" conformers with the O2 atom over the sugar and hydrogen bonded to the phosphate. However, in the DFT

calculations, when thymine is in a "syn" position, the H^+ jumps from O2 to the phosphate during the optimization process.

Summary

The preferred gas-phase conformations of a series of deprotonated and protonated mononucleotides were investigated by ion mobility experiments and theoretical modeling calculations. In the deprotonated systems, AMBER predicts two low-energy families of conformers on the basis of the twist configuration of the sugar. In the lower energy, smaller cross-section family, the sugar is in a $C_{3'}$-endo conformation with the phosphate and base above the plane of the sugar. For dAMP, dCMP, and dTMP, the base is in an "anti" orientation about the glycosidic bond whereas in dGMP, the base is in a "syn" orientation. In the higher energy, larger cross-section family, the sugar is in a $C_{2'}$-endo conformation with the phosphate below the plane of the sugar and the base, which is in an "anti" orientation for each mononucleotide, tilted away from the phosphate. The lowest energy DFT structures are similar to the $C_{3'}$-endo conformers predicted by AMBER. The ion mobility results indicate that only the $C_{3'}$-endo conformers are present (or if both conformers exist and rapidly isomerize at 80 K, the $C_{3'}$ conformers dominate).

In the protonated systems, the ion mobility results also indicate that only one family of conformers is present. AMBER predicts two families of conformers on the basis of the orientation of the base. In the lower energy family, the base is in a "syn" position for dAMP, dCMP, and dTMP and in an "anti" position for dGMP. This is opposite to that observed for the deprotonated systems. Regardless of base or protonation site, the sugars are always in a $C_{2'}$-endo conformation (again, opposite to that observed in the deprotonated systems). The cross-sections of these structures agree very well with experimental values (1−2% difference). DFT calculations predict that the most energetically favored sites for protonation are N3 on dAMP, N7 on dGMP, N3 on dCMP, and the phosphate group on dTMP. Unfortunately, the cross-sections for each mononucleotide protonated at various sites are similar and so absolute assignment of the protonation site could not be made based solely on the experimental results.

Acknowledgment. The support of the National Science Foundation under grant CHE-0140215 is gratefully acknowledged. We also thank Dr. Erich Vorpagel at PNNL for providing theoretical results prior to publication.

References and Notes

(1) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.

(2) Fuller, W.; Wilkins, M. H. F.; Wilson, H. R.; Hamilton, L. D.; Arnott, S. D. *J. Mol. Biol.* **1965**, *12*, 60.

(3) Sutherland, J. C.; Griffen, K. P. *Biopolymers* **1983**, *22*, 1445.

(4) Pohl, F. M.; Jovin, T. M. *J. Mol. Biol.* **1972**, *67*, 375.

(5) Rich, A.; Nordheim, A.; Wang, A. H. J. *Annu. Rev. Biochem.* **1984**, *53*, 791.

(6) Marini, J. C.; Levene, S. D.; Crothers, D. M.; Englund, P. T. *Proc. Natl. Acad. Sci U.S.A.* **1982**, *79*, 7664.

(7) Koo, H. S.; Wu, H. M.; Crothers, D. *Nature* **1986**, *320*, 501.

(8) Wu, H. M.; Crothers, D. *Nature* **1984**, *308*, 509.

(9) Schultz, S. C.; Shields, G. C.; Steitz, T. A. *Science* **1991**, *253*, 1001.

(10) Schleif, R. *Annu. Rev. Biochem.* **1992**, *61*, 119.

(11) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023.

(12) Arnott, S.; Selsing, E. *J. Mol. Biol.* **1974**, *88*, 509.

(13) Morgan, A. R. *Nature* **1970**, *227*, 1310.

(14) McGavin, S. *J. Mol. Biol.* **1971**, *55*, 293.

(15) Sen, D.; Gilbert, W. *Nature* **1988**, *334*, 364.

(16) Sniden, R. R. *DNA Structure and Function*; Academic Press: San Diego, 1994.

(17) Mirkin, S. M.; Lyamichev, V. I.; Drushlyak, K. N.; Dobrynin, V. N.; Filippov, S. A.; Frank-Kamenetskii, M. D. *Nature* **1987**, *330*, 495.

(18) Kohwi, Y.; Kohwi-Shigematsu, T. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3781.

(19) Guo, B. C. *Anal. Chem.* **1999**, *71*, 333R.

(20) Koomen, J. M.; Russel, W. K.; Tichy, S. E.; Russel, D. H. *J. Mass Spectrom.* **2002**, *37*, 357.

(21) Kirpekar, F.; Nordhoff, E.; Larsen, L. K.; Kristiansen, K.; Roepstorff, P.; Hillenkamp, F. *Nucl. Acids Res.* **1998**, *26*, 2554.

(22) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377.

(23) Kirpekar, F.; Berkenkamp, S.; Hillenkamp, F. *Anal. Chem.* **1999**, *71*, 2334.

(24) Veenstra, T. D. *Biochem. Biophys. Res. Commun.* **1999**, *257*, 1.

(25) Green-Church, K. B.; Limbach, P. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 24.

(26) Green-Church, K. B.; Limbach, P. A.; Freitas, M. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 268.

(27) Freitas, M. A.; Shi, S. D.-H.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Chem. Soc.* **1998**, *120*, 10187.

(28) Robinson, J. M.; Greig, M. J.; Griffey, R. H.; Mohan, V.; Laude, D. A. *Anal. Chem.* **1998**, *70*, 3566.

(29) Phillips, D. R.; McCloskey, J. A. *Int. J. Mass Spectrom. Ion Processes* **1993**, *137*, 121.

(30) Rodgers, M. T.; Cambell, S.; Marzluff, E. M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Processes* **1995**, *137*, 121.

(31) Wan, K. X.; Gross, J.; Hillenkamp, F.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 193.

(32) Habibi-Goudarzi, S.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 102.

(33) Giese, B. *Curr. Opin. Chem. Biol.* **2002**, *6*, 612.

(34) Wagenknecht, H. A. *Angew. Chem., Intl. Ed.* **2003**, *42*, 2454.

(35) Bowers, M. T.; Kemper, P. R.; vonHelden, G.; vanKoppen, P. A. M. *Science* **1993**, *260*, 1446.

(36) Clemmer, D. E.; Jarrold, M. F. *J. Mass Spectrom.* **1997**, *32*, 377.

(37) Kemper, P. R.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 197.

(38) vonHelden, G.; Wyttenbach, T.; Bowers, M. T. *Int. J. Mass Spectrom. Ion Processes* **1995**, *146/147*, 349.

(39) vonHelden, G.; Hsu, M.-T.; Kemper, P. R.; Bowers, M. T. *J. Chem. Phys.* **1991**, *95*, 3835.

(40) vonHelden, G.; Hsu, M.-T.; Gotts, N.; Bowers, M. T. *J. Phys. Chem.* **1993**, *97*, 8182.

(41) Gotts, N. G.; vonHelden, G.; Bowers, M. T. *Int. J. Mass Spectrom. Ion Processes* **1995**, *149/150*, 217.

(42) Gidden, J.; Wyttenbach, T.; Batka, J. J.; Weis, P.; Jackson, A. T.; Scrivens, J. H.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 883.

(43) Gidden, J.; Bushnell, J. E.; Bowers, M. T. *J. Am. Chem. Soc.* **2001**, *123*, 5610.

(44) Gidden, J.; Bowers, M. T. *Eur. Phys. J. D* **2002**, *20*, 409.

(45) Gidden, J.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 161.

(46) Mason, E. A.; McDaniel, E. W. *Transport Properties of Ions in Gases*; Wiley: New York 1988.

(47) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Wang, J.; Ross, W. S.; Simmerling, C. L.; Dearden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pittera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman P. A. AMBER 7, University of California, San Francisco, 2002.

(48) Wyttenbach, T.; vonHelden, G.; Batka Jr, J. J.; Carlat, D.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 275.

(49) Wyttenbach, T.; Witt, M.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 3458.

(50) Gidden, J.; Jackson, A. T.; Scrivens, J. H.; Bowers, M. T. *Int. J. Mass Spectrom.* **1999**, *188*, 121.

(51) Gidden, J.; Jackson, A. T.; Scrivens, J. H.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 4692.

(52) Gidden, J.; Jackson, A. T.; Scrivens, J. H.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 499.

(53) Wang, L. S.; Wang, X.; Yang, X.; Vorpagel, E. Private communication.

(54) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.;

Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A.*Gaussian 98*, revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.

(55) Smets, J.; Houben, L.; Schoone, K.; Maes, G.; Adamowicz, L. *Chem. Phys. Lett.* **1996**, *262*, 789.

(56) Russo, N.; Toscano, M.; Grand, A.; Jolibis, F. *J. Comput. Chem.* **1998**, *9*, 989.