Infrared Multiple Photon Dissociation Action Spectroscopy of Proton-Bound Dimers of Cytosine and Modified Cytosines: Effects of Modifications on Gas-Phase Conformations

Bo Yang,[†] R. R. Wu,[†] G. Berden,[‡] J. Oomens,^{‡,§} and M. T. Rodgers^{*,†}

[†]Department of Chemistry, Wayne State University, Detroit, Michigan 48202, United States

[‡]Radboud University Nijmegen, Institute for Molecules and Materials, FELIX Facility, Toernooiveld 7, 6525ED Nijmegen, The Netherlands

[§]van't Hoff Institute for Molecular Sciences, University of Amsterdam, Amsterdam, The Netherlands

Supporting Information

ABSTRACT: The gas-phase structures of proton-bound dimers of cytosine and modified cytosines and their d₆-analogues generated by electrospray ionization are probed via infrared multiple photon dissociation (IRMPD) action spectroscopy and theoretical electronic structure calculations. The modified cytosines examined include the 5-methyl-, 5-fluoro-, and 5-bromo-substituted species. IRMPD action spectra of seven proton-bound dimers exhibit both similar and distinctive spectral features over the range of ~2600–3700 cm⁻¹. The IRMPD spectra of all of these proton-bound dimers are relatively simple, but exhibit obvious shifts in the positions of several bands that correlate with the properties of the substituent. The measured IRMPD spectra are compared to linear IR spectra calculated for the stable low-energy tautomeric conformations, determined at the B3LYP/6-31G*



level of theory, to identify the conformations accessed in the experiments. Comparison of the measured IRMPD and calculated IR spectra indicates that only a single conformation, the ground-state structure, is accessed for all proton-bound homodimers, whereas the ground-state and a small population of the first-excited tautomeric conformations are accessed for all proton-bound heterodimers. In all cases, three hydrogen-bonding interactions in which the nucleobases are aligned in an antiparallel fashion analogous to that of the DNA *i*-motif are responsible for stabilizing the base pairing. Thus, base modifications such as 5-methyland 5-halo-substitution of cytosine should not alter the structure of the DNA *i*-motif.

INTRODUCTION

Double-stranded structures are the most commonly observed conformations for DNA molecules. However, DNA molecules can adopt a multiplicity of conformations that may be correlated with different functional roles they play in biological processes. Previous studies have shown that atypical DNA conformations are related to the expansions of repeated trinucleotide motifs, which lead to severe human diseases.^{1,2} For instance, $(CCG)_n \cdot (CGG)_n$ repeats are related to fragile-X syndrome, $(CAG)_n \cdot (CTG)_n$ repeats are related to Kennedy's and Huntington's diseases, spinocerebellar ataxia, and myotonic dystrophy, while $(GAA)_n$ $(TTC)_n$ repeats are associated with Friedrich's ataxia.³⁻⁶ The proposed structures of such triplet repeats always involve base-pair mismatches, and several of the proposed structures have been confirmed by either NMR spectroscopy or X-ray crystallography.⁷⁻¹¹ The two-dimensional NMR and gel electrophoresis studies of $(CCG)_n$. $(CGG)_n$ repeats show that both the G- and C-rich single strands form hairpins under physiological conditions. NMR data also show that the hairpins formed by the C-rich strands fold in such a way that at the CpG step of the stem, cytosine forms a (C⁺·C) mismatch.¹¹ The NMR studies of oligonucleotides $(GA)_4 \cdot (TC)_4$ indicate formation of a triple-stranded structure and Hoogsteen and Watson–Crick base-pairing for T- A-T and C⁺-G-C triplets.

Indeed, a variety of mismatches between neutral and protonated nucleobases can be formed. The proton-bound dimer of cytosine $(C^+ \cdot C)$ was first discovered in crystals of acetyl cytosine¹² and later in solutions of polycytidylic RNA^{13,14} and DNA.¹⁵ The proton bound-dimers of adenosine (A⁺·A) and cytosine $(C^+ \cdot C)$ form between both parallel and antiparallel strands in the $(C_4T_2)\,\cdot\,(A_2C_4)$ hairpin motif under acidic conditions.¹⁶ Studies by Rajabi et al. using infrared multiple photon dissociation (IRMPD) action spectroscopy and theoretical calculations to characterize the structures of the proton-bound dimer of adenine $(A^+ \cdot A)$ determined that two tautomeric conformations were accessed in the experiments. Computational results also show that for these small species, the solution-phase structure is preserved in the gas-phase,¹⁷ indicating that gas-phase studies can indeed provide insight into solution-phase structure and function. Recent studies have shown the formation of a G-quadruplex within double-stranded

```
Received:May 23, 2013Revised:September 28, 2013Published:October 23, 2013
```

ACS Publications © 2013 American Chemical Society

regions¹⁸⁻²³ with the corresponding tract of the complemen-tary C-rich strands forming an *i*-motif.^{22,23} Since the discovery of the *i*-motif, the biological role of *i*-motif structures as well as their potential in nanotechnological applications have drawn great attention. In fact, the *i*-motif conformation was first discovered in $(CCG)_n \cdot (CGG)_n$ trinucleotide repeats under acidic conditions. As discussed above, the $(CCG)_n (CGG)_n$ trinucleotide repeats are associated with fragile-X syndrome.² The DNA *i*-motif secondary structure is a four-stranded structure consisting of parallel-stranded DNA duplexes zipped together in an antiparallel orientation by intercalated protonbound dimers of cytosine (C⁺·C).²⁵ Recently studies have shown that the hydrogen bonds in the G-quadruplex²⁶ and the structure of the *i*-motif²⁷ are conserved in the gas-phase when electrospray ionization is used as the ionization technique, therefore enabling gas-phase studies of high-order DNA structure. The structure of the proton-bound dimer of cytosine is well established, and involves protonation at the N3 position of the canonical form of cytosine and formation of three hydrogen bonds. However, theoretical calculations indicate that both cytosine and protonated cytosine can adopt various tautomeric conformations of similar stability.²⁸⁻⁴⁵ Previous IRMPD studies on uracil and modified uracils including methyl-, thioketo-, and halo-substituted uracils have shown that protonation, and to a lesser extent sodium cationization, preferentially stabilizes rare tautomers of the nucleobases in the gas-phase.⁴⁶⁻⁴⁸ Therefore, a comprehensive study is needed to investigate the factors that influence the tautomeric equilibria of neutral and protonated cytosine, and whether alternative structures of comparable stability for the proton-bound dimers also exist. IRMPD studies of protonated adenosine and 9methyladenosine have shown that methylation at the C9 position alters the dominant protonated form observed in the gas-phase.⁴⁹ Modifications of cytosine, such as methylation and halogenation, especially at the C5 position, are commonly observed in DNA and RNA. Whether these modifications will cause formation of rare tautomers and alter the structure of the proton-bound dimer is presently unknown.

Recently, the structure of the proton-bound dimer of 1methylcytosine was studied by Oomens and co-workers using IRMPD action spectroscopy techniques.⁵⁰ Methylation at the N1 position does not alter the base-pairing interactions, and the confirmed structure of the proton-bound dimer is the same as that reported in condensed phase NMR studies.²⁵ To achieve a comprehensive understanding of how modifications influence the structural properties and stabilities of proton-bound cytosine dimers, we expand the complexes of interest to include 5-methyl- and 5-halo-substituted forms of cytosine in the present work. The conformations of the proton-bound dimers of cytosine and modified cytosines generated by electrospray ionization are characterized using IRMPD action spectroscopy and theory by a comparison of the measured IRMPD spectra to linear IR spectra derived from electronic structure calculations of the stable low-energy tautomeric conformations of the proton-bound dimers determined at the B3LYP/6-31G* level of theory.

EXPERIMENTAL AND COMPUTATION SECTION

Mass Spectrometry and Photodissociation. IRMPD action spectra of four proton-bound homodimers, (SXC)-H⁺(SXC), where X = H, F, Br, and Me, and three protonbound heterodimers, (C)H⁺(SXC), where X = F, Br, and Me, were measured using a 4.7 T Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS)⁵¹⁻⁵³ coupled to the output of a Nd:YAG-pumped (Innolas Spitlight 600) optical parametric oscillator (OPO, LaserVision, Bellevue, WA). The proton-bound homodimers of cytosine and the modified cytosines were generated using a Micromass "Z-spray" electrospray ionization (ESI) source from solutions containing 1 mM 5XC (X = H, F, Br, and Me) and 1% (v/v) acetic acid in an approximately 50%:50% MeOH:H₂O mixture. To generate the heterodimers, a mixture of 1 mM cytosine, 1 mM 5XC (X =F, Br, and Me) and 1% (v/v) acetic acid in an approximately 50%:50% MeOH:H₂O mixture was used. Cytosine (C) was purchased from Fluka (Zwijndrecht, The Netherlands); 5bromocytosine (5BrC) was purchased from Acros Organics (Geel, Belgium); 5-fluorocytosine (5FC) and 5-methylcytosine (5MeC) were purchased from TCI Europe (Zwijndrecht, Belgium). A solution flow rate of $10-30 \,\mu$ L/min was used, and the ESI needle was generally held at a voltage of \sim 3 kV. Ions emanating from the ESI source were accumulated in a hexapole trap for 6-10 s followed by pulsed extraction through a quadrupole bender and injection into the ICR cell by an rf octopole ion guide. To enable assignment of the IRMPD bands that are not (well) predicted in the theoretical calculations, the IRMPD spectra of the d₆-analogues of all proton-bound homoand heterodimers except those involving 5BrC were measured as well. The isotopes of Br contaminate the isotopic distribution such that the d₆-analogue also contains contributions from $d_4^{81}Br$ and $d_2^{81}Br_{2}$, and thus experiments of the Br containing dimers are less useful. The d₆-analogues were generated using similar solution conditions as employed for the proton-bound dimers, but instead the nucleobase and 1% acetic acid-d₄ are dissolved in a 50%:50% MeOD:D₂O mixture. The precursor ions were mass selected using stored waveform inverse Fourier transform (SWIFT) techniques and irradiated by the OPO laser at pulse energies of up to 17 mJ per pulse of 6 ns duration at 10 Hz for 4-8 s, corresponding to interaction with 40-80 pulses over the wavelength range extending from 3.85 μ m (2600 cm⁻¹) to 2.68 μ m (3735 cm⁻¹).

Computational Details. Theoretical calculations were first performed on the neutral and protonated forms of cytosine and modified cytosines, followed by calculations on the protonbound dimers of these species in order to extract structural and energetic information for the dimers. Briefly, geometry optimizations and frequency analyses of all plausible tautomeric conformations of the neutral and protonated forms of cytosine and the modified cytosines were performed using Gaussian 09⁵⁴ at the B3LYP/6-31G* level of theory. Single point energy calculations were carried out at the B3LYP/6-311+G(2d,2p)⁵⁵ level of theory. A neutral base 5XC, was then paired with a protonated base, H⁺(5YC), to generate a starting point for the proton-bound dimer, $(5XC)H^+(5YC)$, where X, Y = H, F, Br, and Me, for geometry optimization and vibrational frequency analyses at the B3LYP/3-21 level of theory. Plausible base-pair conformations that enable the formation of one, two, or three hydrogen bonds are carefully considered. Optimized structures obtained from this procedure were then used for higher level geometry optimizations and frequency analyses at the B3LYP/ 6-31G* level of theory. Single point energies were calculated at the B3LYP/6-311+ $G(2d,2p)^{55}$ level of theory. Zero-point energy (ZPE) corrections were determined using vibrational frequencies calculated at the B3LYP/6-31G* level and scaled by a factor of 0.9804.⁵⁶ For the analysis of the IRMPD spectra, linear IR spectra were generated from the computed vibrational frequencies and Raman intensities using the harmonic oscillator

approximation and analytical derivatives of the energyminimized Hessian calculated at the B3LYP/6-31G* level of theory. Frequencies were scaled by 0.958 to eliminate known systematic error.⁵⁶ For comparison to experiment, calculated vibrational frequencies were broadened using a 20 cm⁻¹ fwhm Gaussian line shape.

RESULTS AND DISCUSSION

IRMPD Action Spectroscopy. Photodissociation of the proton-bound homodimers, $(5XC)H^+(5XC)$, where X = H, F, Br, and Me, leads to loss of the intact neutral nucleobase 5XC, and detection of the protonated nucleobase, H⁺(5XC), for all four complexes. Photodissociation of the proton-bound heterodimers, $(C)H^+(5XC)$, where X = F, Br, and Me, exhibits two dissociation pathways. The bridging proton either leaves with cytosine, producing $H^+(C)$ and loss of neutral 5XC, or leaves with the modified cytosine, producing $H^+(5XC)$ and loss of neutral cytosine. The relative intensities of the $H^+(C)$ versus H⁺(5XC) products depend on the relative proton affinities (PAs) of cytosine and 5XC. The nucleobase with the larger PA is observed in greater intensity. On the basis of the relative intensities of the IRMPD products observed for the heterodimers, the PAs of the four nucleobases follow the order of 5MeC > C > 5FC, 5BrC. For all proton-bound dimers, the IRMPD spectra were plotted as the IRMPD yield of the product ions as a function of wavelength. An IRMPD yield was determined from the precursor ion intensity (I_p) and the fragment ion intensities $(I_{\rm f})$ after laser irradiation at each frequency as shown in eq 1.

IRMPD yield =
$$\sum_{i} I_{f_i} / (I_p + \sum_{i} I_{f_i})$$
 (1)

The IRMPD yield was normalized linearly with laser power to correct for changes in the laser power as a function of the photon energy $(h\nu = hc/\lambda)$ of the OPO laser.

The measured IRMPD spectra of the proton-bound homoand heterodimers are shown in Figures 1 and 2, respectively. The IRMPD action spectrum of the $(C)H^+(C)$ homodimer is also included in Figure 2 for comparison to facilitate determination of the effects of modifications on base-pairing interactions. As can be seen in Figure 1, the IR features observed in the spectrum of the $(C)H^+(C)$ homodimer are retained for the most part in the spectra of the other three proton-bound homodimers, except that the band at ~3525 cm⁻¹ disappears in the spectra of the (5FC)H⁺(5FC) and (5BrC)H⁺(5BrC) complexes. The most intense band appears at ~3450 cm⁻¹ for all four proton-bound homodimers. This band is increasingly broadened and red-shifted for the (5FC)H⁺(5FC) and (5BrC)H⁺(5BrC) complexes as compared to the $(C)H^+(C)$ complex, suggesting that more than one vibrational mode contributes to this IR feature. The two strong bands to the blue of the most intense band at 3490 and 3525 cm⁻¹ also red shift for the (5FC)H⁺(5FC) and (5BrC)- $H^+(5BrC)$ complexes. Because of the red shifting, the band at 3490 cm⁻¹ merges into the most intense band, leading to the broadened and asymmetric shape of this band. For the (5MeC)H⁺(5MeC) complex, the most intense band and the two strong bands to the blue are all blue-shifted by a few wavenumbers as compared to the $(C)H^+(C)$ complex. Very weak broad bands are observed for all proton-bound homodimers in the region between 2600 and 3000 cm^{-1} . Subtle differences include a rise in the intensity of the band



Figure 1. IRMPD action spectra of $(5XC)H^+(5XC)$ homodimers complexes, where X = H, F, Br, and Me.



Figure 2. IRMPD action spectra of $(C)H^+(SXC)$ heterodimer complexes, where X = H, F, Br, and Me.

near 3100 cm⁻¹ as well as a red shift of the band at 3230 cm⁻¹ for the (5FC)H⁺(5FC) and (5BrC)H⁺(5BrC) complexes, a rise in the intensity of the band at ~3320 cm⁻¹ for the (5FC)H⁺(5FC) complex, and a red shift of the band at 3370 cm⁻¹ (blue dashed line) for the (5FC)H⁺(5FC), (5BrC)-H⁺(5BrC), and (5MeC)H⁺(5MeC) complexes.

As can be seen in Figure 2, the measured IRMPD spectra of the proton-bound heterodimers exhibit great similarity to that of the $(C)H^+(C)$ complex. The most intense band at 3450

cm⁻¹ (black dashed line) and the strong band at 3490 cm⁻¹ (red dashed line) shift very little for the (C)H⁺(5FC) and $(C)H^{+}(5BrC)$ complexes, suggesting that vibrational modes of cytosine are the main contributors to these two bands. Red shifting is observed in the band at 3525 cm⁻¹ for the $(C)H^{+}(5FC)$ complex, and becomes greater for the (C)- $H^+(5BrC)$ complex, leading to the overlap of this band with one of the bands at 3450 and 3490 cm⁻¹. For the $(5MeC)H^+(C)$ complex, the three intense bands at 3450, 3490, and 3525 cm^{-1} are all blue-shifted by a few wavenumbers as compared to those of the $(C)H^+(C)$ complex. Subtle differences include a rise in the intensity of the band at 3320 cm^{-1} for the (C)H⁺(5FC) complex, and a rise in the intensity of the band near 3100 cm⁻¹ as well as red shifts of the bands at 3230 and 3370 cm^{-1} for the (C)H⁺(5FC) and (C)H⁺(5BrC) complexes.

Theoretical Results. As discussed above, the neutral and protonated nucleobases as well as the proton-bound dimers were calculated at the $B3LYP/6-311+G(2d,2p)//B3LYP/6-31G^*$ level of theory. The six most stable tautomeric conformations of 5MeC and H⁺(5MeC) are shown in Figures 3 and 4, respectively. All of the other neutral and protonated



5MeC

Figure 3. B3LYP/6-31G* optimized geometries of the six most stable tautomeric conformations of 5-methylcytosine, 5MeC.

nucleobases examined here exhibit structures similar to those of 5MeC and H⁺(5MeC). The geometry-optimized structures of all tautomeric conformations computed and their relative Gibbs free energies at 298 K of the neutral and protonated forms of each of the five nucleobases are included in Figures S1 and S2 of the Supporting Information, respectively. To differentiate the various stable low-energy tautomeric conformations of these species, Roman numerals are employed. Lowercase Roman numerals are used to describe the tautomeric conformations of the neutral base, while uppercase Roman numerals with a "+" sign are used to describe the tautomeric conformations of the protonated base, and both are ordered based on their relative free energies. The structures and Gibbs free energies at 298 K of the three most stable conformations of the (5MeC)- $H^+(5MeC)$ and $(5MeC)H^+(C)$ complexes are shown in Figure 5. The proton-bound homo- and heterodimers of the other nucleobases exhibit structures and relative stabilities similar to those of the (5MeC)H⁺(5MeC) and (5MeC)H⁺(C) complexes, respectively. As discussed above, base-pair conforma-





H⁺(5MeC)

Figure 4. B3LYP/6-31G* optimized geometries of the six most stable tautomeric conformations of protonated 5-methylcytosine, $H^+(5MeC)$.



Figure 5. B3LYP/6-31G* optimized geometries of the three most stable tautomeric conformations of the proton-bound (5MeC)- $\rm H^+(SMeC)$ homo- and (5MeC)H^+(C) heterodimers.

tions that enable formation of one, two, or three hydrogen bonds in different orientations were carefully examined. The structures and relative stabilities of all tautomeric conformations of each of the proton-bound dimers calculated to lie within 100 kJ/mol of the ground-state conformer are shown in Figures S3 and S4 of the Supporting Information. As can be seen in Figure 5, the most stable tautomeric conformations involve binding via three hydrogen bonds or a single hydrogen bond. The ground-state structure of the (5MeC)H⁺(5MeC) homodimer involves three hydrogen bonds and adopts an antiparallel configuration of the protonated and neutral bases, which is the most commonly observed conformation in multistranded DNAs. This conformer is designated as II+... i_3a to indicate that the first-excited II^+ tautomeric conformation of the protonated base, H⁺(5MeC), binds to the ground-state i tautomeric conformation of the neutral base, 5MeC. The underscore 3a designation indicates that the binding occurs via three hydrogen-bonding interactions and the

protonated and neutral bases are bound in an antiparallel configuration. Other stable conformations are designated in a similar fashion, where the lower case Roman numeral indicates the tautomeric conformation of the neutral base as summarized in Figure 3 and Figure S1 of the Supporting Information, the uppercase Roman numeral with the "+" indicates the tautomeric conformation of the protonated base as summarized in Figure 4 and Figure S2 of the Supporting Information, the number indicates the number of hydrogen bonding interactions, while the lowercase one or two letter designations indicate the relative orientations of the protonated and neutral bases, where \mathbf{a} = antiparallel, \mathbf{p} = parallel, \mathbf{at} = antiparallel and twisted, and **pt** = parallel and twisted. As mentioned above, the ground-state conformation of the proton-bound dimer involves an excited minor tautomer of the protonated base, the II+ tautomeric conformation of H⁺(5MeC), bound to the groundstate tautomer of the neutral base, i, as these species enable three nearly ideal (nearly linear) hydrogen bonds to stabilize the proton-bound dimer. It is unclear whether tautomerization of the protonated base to the ground-state species, I⁺, will occur during the dissociation process. Therefore, relaxed potential energy scans were performed to determine the height of the tautomerization barrier. The barriers for the (5XC)H⁺(5XC) tautomerization lie in the range of 194.4-204.8 kJ/mol, while simple cleavage of the three hydrogen bonds requires less energy, 160.3-173.4 kJ/mol, suggesting that tautomerization will not occur upon dissociation at threshold energies. A noteworthy observation is that the excess proton is not equally shared by the two nucleobases even in the proton-bound homodimers as seen in the ground-state structures. However, computational results indicate that transfer of the excess proton from one nucleobase to the other is a low barrier process in the homo- and heterodimers examined here, consistent with observations made by Han and co-workers for the $(C)H^{+}(C)$ complex.⁵⁷ Pairing of the ground-state conformations of the neutral and protonated nucleobase through a single hydrogen bonding interaction leads to a less stable tautomeric conformation, I+...i_1at, which lies 20.7 kJ/mol higher in Gibbs free energy than the ground-state tautomeric conformation. Attempts to optimize the corresponding I⁺···i 1 pt conformation always converged to the I+...i 1at tautomeric conformation due to steric effects. Pairing the ground-state conformation i of 5MeC with the third most stable conformation III⁺ of H⁺(5MeC) through a single hydrogen bond via the 2-hydroxyl group of H⁺(5MeC) and carbonyl group of neutral 5MeC leads to the third most stable tautomeric conformation, III⁺···i 1 pt, which lies 29.1 kJ/mol higher in Gibbs free energy than the ground-state conformer. In the I⁺···i 1at and III⁺···i 1 pt tautomeric conformations, a single hydrogen bond is formed between the 2-hydroxyl group of $H^+(5MeC)$ and the carbonyl oxygen atom of the i conformation of 5MeC. However, in order to obtain additional stabilization from the hydrogen bonding interactions between the carbonyl oxygen atom of the III+ conformation and the N4-H moiety of the i conformation, the protonated and neutral rings adopt a parallel twisted orientation in the III+... i_1 pt tautomeric conformation. Various types of single hydrogen bonds can occur between the protonated and neutral bases, but lead to tautomeric conformations that are higher in Gibbs free energy than the I⁺···i 1at and III⁺···i 1 pt tautomeric conformations, which involve an O2…H…O2 hydrogen bond. These other tautomeric conformations are not likely to be accessed in the experiments if the computed

energetics are reliable, and thus are not considered further. In summary, the most stable conformations for all proton-bound homodimers are the $II^+\cdots i_3a$ conformations. Assuming that the computed energetics are reliable, the first excited conformations lie high enough in free energy above the ground-state conformations that they are unlikely to be produced in measurable abundance at room temperature.

In the case of proton-bound heterodimers, the excess proton binds to the nucleobase with the higher proton affinity (PA) in the ground-state structure. In the most stable tautomeric conformation of the $(5MeC)H^+(C)$ complex, $II^+ \cdots i 3a$, the excess proton is bound to 5MeC (see Figure 5), suggesting that the N3 PA of 5MeC is greater than that of cytosine, as expected. On the basis of the ground-state structures of all of the proton-bound heterodimers, the PAs of these four nucleobases follow the order: 5MeC > C > 5BrC, 5FC. This order indicates that 5-methyl-substitution of cytosine increases the N3 PA, whereas 5-halo-substitution decreases the N3 PA, as expected based on the inductive effects of these substituents. Transfer of the excess proton to the other nucleobase, cytosine, produces the first-excited tautomeric conformation, i---II+ 3a. Calculations indicate that the i-II+ 3a conformation lies a mere 2.4-7.4 kJ/mol higher in free energy than the groundstate II⁺···i 3a structure, and that transfer of the excess proton is a low-barrier process that requires 19.6-21.3 kJ/mol for these three proton-bound heterodimers. This observation is consistent with previous theoretical and experimental studies of similar systems involving $N \cdots H^+ \cdots N$ hydrogen bonding interactions.⁵⁸⁻⁶⁰ Thus, under typical ESI conditions, both structures may be accessed in our experiments, but the groundstate conformation II+ ... i 3a should be the dominant species assuming that the calculated energetics are reliable. Upon dissociation of the (5MeC)H⁺(C) heterodimer, two competitive pathways leading to the generation of H⁺(5MeC) or $H^+(C)$, are observed. Pairing of the ground-state conformations of C and H⁺(5MeC) through a single O2…H…O2 hydrogen bond leads to a less stable tautomeric conformation, I⁺···i 1at, which lies 22.6 kJ/mol higher in Gibbs free energy than the ground-state structure. Therefore, the I⁺···i_1at conformation is not likely to be accessed under our experimental conditions assuming that the computed energetics are reliable. As can be seen in the Supporting Information, Figure S3, the three most stable tautomeric conformations of the other three protonbound heterodimers exactly parallel those found for the $(5MeC)H^+(C)$ complex. In summary, the ground-state tautomeric conformations of the proton-bound heterodimers are II+...i 3a in all cases. The first excited-state conformations are i---II+ 3a, which lie 2.4-7.4 kJ/mol higher in Gibbs free energies than the ground-state conformers. These results suggest that both the ground and the first-excited tautomeric conformations are likely to be accessed in the experiments with the ground-state II+...i_3a conformation being the dominant species, whereas all other stable conformations lie sufficiently high in free energy that they are unlikely to be accessed in the experiments.

Comparison of Measured IRMPD and Theoretical IR Spectra of Proton-Bound Dimers. As discussed above, weak and broad bands are observed for all proton-bound dimers in the region between 2600 and 3000 cm⁻¹. Theory indicates that the bands in this region correspond to in-plane stretches of the three bridging protons that are involved in the base-pair hydrogen bonds (in addition to weak aliphatic CH stretching modes for the species containing a methyl substituent).

Theoretical results suggest that the excess proton as well as the other two bridging protons can shift between the two nucleobases over a low energy barrier. Therefore, the stretches associated with these protons adopt double-well potential energy surfaces, and thus are anharmonic. However, the frequency analyses performed in the present work only provide reasonably accurate vibrational frequencies for near-harmonic vibrational modes such that those predicted by theory are generally less red-shifted than in reality. In addition, such anharmonic stretches are generally severely broadened such that they tend to appear as very weak and broad features in the IRMPD spectra and thus are often not very useful for diagnostic purposes. Advanced anharmonic computational approaches are required to accurately describe the band positions of such anharmonic proton stretches. 59-62 To examine the sensitivity of the computed band positions of the anharmonic bridging protons to the optimized structures, additional calculations were performed for the proton-bound dimer of cytosine where the hydrogen-bond length of the anharmonic bridging proton was fixed and the longer O2…H-N4 hydrogen bond was shortened, which resulted in a concomitant lengthening of the shorter O2…H-N4 hydrogen bond. As can be seen in Figure S5 of the Supporting Information, the band positions of the harmonic stretches in the region between 3400 and 3600 cm⁻¹ as well as that of the anharmonic bridging proton (N3-H⁺...N3) were shifted very little, whereas those of the two anharmonic O2…H-N4 bridging hydrogen bonds were displaced markedly and by as much as 300 cm⁻¹, clearly indicating that these stretches are very sensitive to the hydrogen bond length. In spite of the challenges associated with computing the band positions of the hydrogen stretches associated with the anharmonic bridging proton/hydrogen atoms, the current results are still sufficient to identify the structures of the proton-bound dimers that are accessed in the experiments due to differences in the computed band positions for the near harmonic bands observed. Further theoretical calculations using advanced methods are beyond the scope of the present work and are not pursued here, but will be the subject of future work as improvements in the interpretation of such systems will become increasingly important for larger more complex biological systems of interest. Thus, interpretation of the measured IRMPD and theoretical IR spectra focuses on the region between 3000 and 3700 cm⁻¹ for all proton-bound dimers. The calculated bands for the anharmonic shared-proton/hydrogen stretches that are poorly computed by the harmonic approximation are shown as dotted lines and are shaded in red in the figures that compare the measured IRMPD and calculated IR spectra, and are excluded from the discussion.

Homodimers. As discussed previously, the measured IRMPD action spectra for all four proton-bound homodimers, $(5XC)H^+(5XC)$ are remarkably similar, and thus are examined in parallel here. The measured IRMPD and the calculated IR spectra for the three most stable conformations found for the $(5MeC)H^+(5MeC)$ complex are shown in Figure 6 with the band predicted for the anharmonic shared-hydrogen stretch shaded in red. Analogous comparisons for the $(C)H^+(C)$, $(5FC)H^+(5FC)$, and $(5BrC)H^+(5BrC)$ complexes are shown in the Supporting Information as Figures S6 through S8. The calculated linear IR spectrum of the ground-state conformation, $II^+\cdots i_3a$, provides the best agreement with the measured IRMPD spectra for all homodimers, except that two weak bands observed at ~3230 and 3360 cm⁻¹ do not exhibit



Figure 6. Comparison of the measured IRMPD action spectrum of the $(5MeC)H^+(5MeC)$ complex and its d₆-analogue with the IR spectra of the three most stable tautomeric conformations of $(5MeC)H^+(5MeC)$ complex predicted at the B3LYP/6-31G* level of theory. The dashed line indicates that the IRMPD yield of the d₆-analogue has been multiplied by a factor of 1.5 over this region, while the dotted lines indicate that the IR intensities have been scaled down by a factor of 2.5 and 3.0 over this region.

comparable theoretical frequencies. To facilitate assignment of these two unexpected bands, the d₆-analogues of the protonbound homodimers were investigated. The IRMPD spectra of the d₆-analogues of each proton-bound homodimer are shown in the top panel of Figure 6 and Supporting Information, Figure S6 through S8, respectively. To facilitate comparison of the proton-bound dimers and their d₆-analogues, the IRMPD yields of the d_6 -analogues are scaled such that the yields of the most intense band for the deuterated complexes match the yields for the corresponding protic complexes. The H/D exchange rate of each hydrogen atom and the laser power both influence the yield of each vibrational mode, and thus overinterpretation of the scaling factor here is inappropriate. All proton-bound homodimers have seven exchangeable hydrogen atoms, and thus there are seven different combinations that may contribute to a d₆-analogue. Therefore, the mass selected d₆-analogue is a mixture of seven different species with one hydrogen and six deuteriums distributed over these seven positions. As a result, the vibrational modes that involve zero or one hydrogen atom will be unaffected, whereas the vibrational modes that involve two or more hydrogen atoms will shift to much lower frequencies. A decrease in the yield for the bands that involve only one hydrogen atom is expected (see eq 1), while the yield of bands that do not involve hydrogen atoms should be unaffected. The measured IRMPD spectra of the d₆-analogue of $(5MeC)H^+(5MeC)$ is included in the top panel of Figure 6, while the measured IRMPD spectra of the d₆-analogues of $(C)H^{+}(C)$ and $(5FC)H^{+}(5FC)$ are included in Figures S6 and S7 of the Supporting Information. Because of the isotopic distribution of bromine, experiments on the d₆-analogue of $(5BrC)H^+(5BrC)$ are not as useful because mixed contributions

to the isotopic envelope arise, and thus are not examined here. Again, the d₆-analogues of all proton-bound homodimers exhibit similar behavior, and thus are discussed in parallel here. As can be seen in Figure 6, all bands observed in the spectrum of the $(5MeC)H^+(5MeC)$ complex are retained for the d₆analogue except the band at 3230 cm⁻¹, indicating that the former vibrational modes involve zero or one hydrogen atom, whereas the 3230 cm⁻¹ vibrational mode involves motions of at least two exchangeable hydrogen atoms. The experimental bands observed for the d₆-analogue all decrease in yield except for the band at 3360 cm⁻¹, suggesting that this band may not involve hydrogen. Theoretical calculations do not predict bands at 3230 or 3360 cm⁻¹, suggesting that either theory does not predict accurate frequencies for the associated vibrational modes, or that these two bands are overtones of other vibrational modes. Theoretical calculations of the (5MeC)- $H^+(5MeC)$ complex predict a pair of intense bands at ~1600 and 1680 cm^{-1} (scaled by 0.958) that correspond to the coupling of the carbonyl stretch of the neutral base, NH₂ scissoring, and N3-H in-plane bending of the protonated base. The first overtones of these two bands should appear at 3200 and 3360 cm⁻¹, which match the unexpected bands in the measured IRMPD spectrum of (5MeC)H⁺(5MeC) reasonably well as shown in Figure S9 of the Supporting Information, which compares the measured and computed IR spectra of the homodimers over an expanded frequency range extending from 1500 to 3600 cm⁻¹. Upon deuteration, the carbonyl stretch is decoupled from NH₂ scissoring and N3-H in-plane bending, leading to a pure carbonyl stretch at 1659 cm⁻¹ (with an intensity similar to that of the band calculated at 1680 cm⁻¹ for the nondeuterated species) and a much lower frequency band at 1125 cm⁻¹ that is associated with ND₂ scissoring and N3–D in-plane bending. Therefore, the first overtones of these two bands after deuteration should appear at 3320 and 2250 cm⁻¹ such that the latter band lies outside of the frequency range of our experiments. The calculated IR behavior of the overtones of these two bands upon deuteration matches reasonably well with the observed behavior of the bands at \sim 3230 and 3360 cm⁻¹ in the measured IRMPD spectra. On the basis of these theoretical results, it is believed that the unexpected bands observed at \sim 3230 and 3360 cm⁻¹ in the measured IRMPD spectra of all of the proton-bound homodimers correspond to the first overtones of the coupled C=O stretch, NH₂ scissoring, and N3-H in-plane bending. Therefore, these two bands are excluded from the comparison of the measured IRMPD and calculated IR spectra. When the shaded band at 3320 cm⁻¹ and the unpredicted bands at 3230 and 3360 cm^{-1} are excluded, the IR spectrum calculated for the II+ ... i 3a conformation exhibits very good agreement with the measured IRMPD spectrum for the (5MeC)H⁺(5MeC) complex. All experimental bands have comparable theoretical frequencies, confirming that the groundstate structure, II+...i 3a, is accessed in the experiments. The most intense band, observed at \sim 3450 cm⁻¹, corresponds to the overlap of the free N1-H stretch of both rings. The strong bands to the blue of the most intense band at \sim 3490 and 3525 cm⁻¹ are assigned to the N4-H stretches of the neutral and protonated bases, respectively.

Comparison of the calculated IR spectrum of the first-excited tautomeric conformation, $I^+\cdots i_1at$, to the measured IRMPD action spectrum suggests that this conformation is not accessed in the experiments as there are several differences. First, the calculated IR spectrum for the $I^+\cdots i_1at$ conformation exhibits only two bands instead of three in the region between 3400 and

 3550 cm^{-1} . The most intense band now corresponds to the free N1–H and NH₂ symmetric stretches, and is red-shifted by \sim 10 cm^{-1} relative to the measure IRMPD band at ~3450 cm⁻¹. The free NH₂ symmetric stretch produces the band at 3565 cm⁻¹, which is blue-shifted by ~ 75 cm⁻¹ as compared to the measured IRMPD band observed at 3490 cm⁻¹. In addition, the overtone of the carbonyl stretch is red-shifted by 25 cm^{-1} , and is no longer coupled with NH₂ scissoring. The calculated IR spectrum for the III⁺····i 1 pt conformation exhibits very similar IR features to those observed for the I⁺···i 1at conformation in the region between 3400 and 3600 cm⁻¹ except that the most intense band is broadened and decreased in intensity. Again, the calculated IR spectrum for the III+ ... i 1 pt conformation exhibits only two bands instead of three as observed in the measured IRMPD spectrum in the region between 3400 and 3600 cm⁻¹. The most intense band is red-shifted by ~ 10 cm⁻¹ relative to the measure IRMPD band at \sim 3450 cm⁻¹, and the band to the blue of the most intense band is blue-shifted by \sim 75 cm⁻¹ as compared to the measured IRMPD band observed at 3490 cm⁻¹. In addition, computational results indicate that the overtones of the C=O stretch coupled with NH_2 scissoring and N3-H in-plane bending give rise to a pair of bands at 3167 and 3254 cm⁻¹, which do not agree with the experimentally observed bands in the measured IRMPD spectrum. On the basis of these comparisons, the III+...i 1 pt conformation is also not accessed in the experiments. Analogous comparisons for the other three proton-bound homodimers (see Supporting Information, Figure S6 through S8) suggest that neither the I⁺····i_1at or III⁺····i_1 pt conformations are accessed in the experiments.

Figure 7 compares of the measure IRMPD spectra with the calculated IR spectra of the ground-state II⁺···i_3a conformations, for all four proton-bound homodimers. The bands highlighted in red and blue correspond to the N4–H stretching of the protonated and neutral rings, respectively. Frequencies of the vibrational modes in the region between 3000 and 3600 cm⁻¹ for each proton-bound homodimer can be found in the Supporting Information, Table S1. As can be seen in Figure 7, the shifting of these two bands in the experimental spectra is nicely reproduced by the calculated IR spectra, confirming that the ground-state II⁺···i_3a conformations are accessed in the experiments. Thus, for all four proton-bound homodimers, only the ground-state II⁺···i_3a conformations are accessed in the experiments.

Heterodimers. As discussed previously, the measured IRMPD action spectra for all three proton-bound heterodimers, $(C)H^{+}(5XC)$ exhibit remarkably similar behavior, and thus are examined in parallel here. The measured IRMPD and the calculated IR spectra for the three most stable tautomeric conformations found for the $(5MeC)H^+(C)$ complex are shown in Figure 8 with the band predicted for the stretch of the anharmonic shared hydrogen/proton shaded in red. Analogous comparisons for the (C)H⁺(5FC) and (C)-H⁺(5BrC) complexes are shown in Figures S10 and S11 of the Supporting Information. As can be seen in Figure 8, the calculated linear IR spectrum of the ground-state II+---i 3a conformation provides the best agreement with the measured IRMPD spectrum for the proton-bound $(5MeC)H^+(C)$ heterodimer, except that two bands observed at ~3230 and 3360 cm^{-1} do not exhibit comparable theoretical frequencies. Similar results are found for the (C)H⁺(5FC) and (C)- $H^{*}(5BrC)$ complexes. To identify the unexpected bands observed at ${\sim}3230$ and $3360~{\rm cm}^{-1},$ the IRMPD spectra of



Figure 7. Comparison of the measured IRMPD action spectra of the $(5XC)H^+(5XC)$ homodimers, where X = H, F, Br, and Me, with the calculated IR spectra of the ground-state $II^+\cdots i_3a$ conformations predicted at the B3LYP/6-31G* level of theory. The solid line spectra on the left are the measured IRMPD action spectra, while the dashed line spectra on the right are the calculated IR spectra. The bands highlighted in red and blue are the N4–H stretching of the neutral and protonated rings, respectively.

the d_{6} -analogues of $(5MeC)H^{+}(C)$ and $(C)H^{+}(5FC)$ were measured and compared to the protic analogues in Figure 8 and Figure S10 of the Supporting Information, respectively. Again, because of the intrinsic isotopic distribution of bromine, IRMPD measurement on the $(C)H^{+}(5BrC)$ complex is not as useful, and thus is not included. As found for the proton-bound homodimers, the first overtones of the coupled C=O stretch, NH₂ scissoring, and N3-H in-plane bending should appear near 3200 and 3370 cm⁻¹, which match reasonably well with the unexpected bands at \sim 3230 and 3360 cm⁻¹ in the measured IRMPD spectra of the proton-bound heterodimers as shown in Figure S12 of the Supporting Information, which compares the measured IRMPD and computed IR spectra of the heterodimers over an expanded frequency range extending from 1500 to 3600 cm⁻¹. Thus, the unexpected bands at \sim 3230 and 3360 cm⁻¹ in the measured IRMPD spectra of all of the protonbound heterodimers are also assigned to the first overtones of the coupled C=O stretch, NH₂ scissoring, and N3-H in-plane bending. As a result, the theoretical IR spectrum calculated for the ground-state conformation, II+...i 3a, exhibits very good agreement with the measured IRMPD spectrum when the shaded bands and two overtones are excluded, suggesting that the ground-state conformations, II+...i_3a, are indeed accessed in the experiments. Transfer of the excess proton from 5MeC to C (or from C to 5FC or 5BrC in the other heterodimers) leads to the first excited conformers, i…II⁺_3a, where the excess proton is bound to the nucleobase with the lower N3 PA, which lie 2.4–7.4 kJ/mol higher in Gibbs free energies than the ground-state conformers. The calculated IR spectrum of the first-excited i--II⁺ 3a conformer of $(5MeC)H^+(C)$ shares



Figure 8. Comparison of the measured IRMPD action spectrum of the $(5MeC)H^+(C)$ complex and its d₆-analogue with the IR spectra of the three most stable tautomeric conformations of $(5MeC)H^+(C)$ complex predicted at the B3LYP/6-31G* level of theory. The dashed line indicates that the IRMPD yield of the d₆-analogue has been multiplied by a factor of 1.7 over the region, while the dotted lines indicate that the IR intensities have been scaled down by a factor of 2.5 over this region.

similarity with that of the ground-state II+...i 3a conformer. The calculated band for the unresolved free N1-H stretches of both rings agrees well with the measured IRMPD band observed at \sim 3450 cm⁻¹. However, the calculated band for the free N4–H stretch of protonated C is red-shifted by 10 cm^{-1} , whereas the calculated band for the free N4-H stretch of neutral 5MeC is blue-shifted by 10 cm⁻¹ as compared to the measured IRMPD bands at \sim 3490 and 3525 cm⁻¹, respectively. On the basis of this comparison, the first-excited i--II+ 3a conformer may also be accessed in the experiments, but is likely only present in low abundance. For the third most stable conformers of the $(5MeC)H^+(C)$ complex, I⁺···i 1at, the calculated IR spectrum exhibits only two bands instead of three in the region between 3400 to 3550 cm⁻¹. The most intense band again corresponds to the free N1-H and NH₂ symmetric stretches, which is red-shifted by a few wavenumbers relative to the most intense IRMPD band at \sim 3450 cm⁻¹. The free NH₂ symmetric stretch produces the band at 3565 cm⁻¹, which is blue-shifted by ~ 75 cm⁻¹ as compared to the measured IRMPD band observed at 3490 cm⁻¹. In addition, the two overtones at 3230 and 3360 cm⁻¹ should be red-shifted to ${\sim}3211$ and 3330 $\rm cm^{-1}\!$, and the carbonyl stretch is no longer coupled with NH₂ scissoring. These differences indicate that the I⁺···i 1at conformer is not accessed in the experiments. Analogous comparisons for the other two proton-bound heterodimers (see Supporting Information, Figures S10 and S11) provide highly parallel results, suggesting that the I^+ ... i 1at conformations of the proton-bound heterodimers are not accessed in the experiments.

Figure 9 compares the measure IRMPD spectra with the calculated IR spectra of the ground-state II⁺···i 3a and first



Figure 9. Comparison of the measured IRMPD action spectra of the $(C)H^+(5XC)$ heterodimers, where X = F, Br, and Me, with the calculated IR spectra of the ground-state $II^+\cdots i_3a$ and first excited $i\cdots II^+_3a$ conformations predicted at the B3LYP/6-31G* level of theory. The measure IRMPD and calculated IR spectra of the $(C)H^+(C)$ complex is included for comparison. The solid line spectra on the left are the measured IRMPD action spectra, while the dashed and dotted line spectra on the right are the calculated IR spectra. The bands highlighted in red and blue are the N4–H stretching of the neutral and protonated rings, respectively.

excited i…II+_3a conformations for the three proton-bound heterodimers and the $(C)H^+(C)$ complex. The bands highlighted in red and blue correspond to the N4-H stretches of the neutral and protonated rings, respectively. Frequencies and assignments of the vibrational modes in the region of 3400 and 3600 cm⁻¹ for each proton-bound heterodimer are provided in Table S2 of the Supporting Information. As found for the proton-bound homodimers, the calculated linear IR spectra of the ground-state II+...i 3a conformations accurately reproduce the shifts associated with these two bands observed in the experiments. Thus, for all three heterodimers, the ground-state II+...i 3a conformers are accessed in the experiments. Comparisons of the measured IRMPD spectra with the calculated IR spectra of first excited i-II+ 3a conformations show that the predicted band positions of the N4-H stretches do not match well with the bands observed in the experiments. The differences include a red-shift of the N4-H stretch of protonated ring, a blue-shift of the N4-H stretch of neutral ring, and therefore a larger frequency difference between these two bands. Because of the broadening of the measure IRMPD bands, it is difficult to definitely establish the presence or absence of the first-excited i-II+ 3a conformers. However, as discussed previously, the shifts of the N4-H stretches of the protonated and neutral bases suggest that the contributions of the first-excited i…II+_3a conformers are likely very minor.

CONCLUSIONS

IRMPD action spectra of four proton-bound homodimers, $(SXC)H^+(SXC)$, where X = H, F, Br, and Me, and three

Article

proton-bound heterodimers, $(C)H^+(5XC)$, where X = F, Br, and Me, were measured in the 3 μ m spectral range. The measured IRMPD of the four homodimers share similarities, but also exhibit shifts in the bands positions due to the influence of the modifications. Comparison of the measured IRMPD spectra to the IR spectra calculated at B3LYP/6-31G* level of theory for the three most stable $(5XC)H^{+}(5XC)$ tautomeric conformations, II+...i 3a, I+...i 1at, and III+...i 1 pt, are made to identify the species accessed under our experimental conditions. In all cases, it is clear that the only tautomeric conformation accessed in the experiments is the II+...i 3a conformation, in agreement with the predicted ground-state structures for these complexes and the large difference in relative free energies for the excited conformers. In the case of the proton-bound heterodimers, the measured IRMPD spectra exhibit similar IR behavior compared to that of the homodimers, and also exhibit shifts in the bands positions as a function of the modifications. Comparison of the measured IRMPD spectra to the IR spectra calculated at the B3LYP/6- $31G^*$ level of theory for the three most stable (C)H⁺(5XC) tautomeric conformations, II+...i 3a, i...II+ 3a, and I+...i 1at, are made to identify the conformations accessed under our experimental conditions. In all cases, the ground-state structures, II⁺···i 3a, which involve an excited minor tautomer II⁺ of the protonated base binding to the ground-state tautomer i of the neutral base, are accessed in the experiments. The firstexcited conformers of the proton-bound heterodimers, i--- II^+_3a , where the excess proton is now bound to the base with the lower N3 PA, and which lie 2.4-7.4 kJ/mol higher in free energy, may also be accessed in the experiments, but are likely only present in low abundance. On the basis of the combination of experimental and theoretical results presented here, it is clear that the modifications alter the relative stabilities of the various conformations of the proton-bound dimers. However, their effects are small enough that the preferred tautomeric conformations and binding modes are not significantly altered. Quantitative determination of the strength of the base pairing energies in the ground-state proton-bound dimers remains elusive. Calculations performed here suggest that this binding is quite strong, 160.3 to 173.4 kJ/mol, which is much stronger than typical Watson-Crick G·C base pairing. Therefore, it would be useful to re-examine the proton-bound dimers using instrumentation with the capability to determine thermochemical properties, such as a guided ion beam tandem mass spectrometer equipped with an ESI source, so that the strength of the hydrogen-bonding interactions in the groundstate conformers can be determined.

ASSOCIATED CONTENT

S Supporting Information

Figures showing all plausible conformations of neutral and protonated forms of each nucleobase, tautomeric conformations that lie within 100 kJ/mol in Gibbs free energy for each of the proton-bound homo- and heterodimers; figures comparing the calculated IR spectra predicted at the B3LYP/6-31G* level of theory at different positions along the potential energy scan associated with twisting of the protonated and neutral nucleobases such that the O2…H-N4 hydrogen bonds shorten/lengthen; figures comparing the measured IRMPD action spectra of the proton bound homo- and heterodimers and their d_6 -analogues with IR spectra for the three most stable tautomeric conformations of theory; tables listing frequencies of the

vibrational modes in the region between 3400 and 3600 cm⁻¹ for each of the proton-bound homo- and heterodimers calculated at B3LYP/6-31G* level of theory and scaled using a scaling factor of 0.958. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: mrodgers@chem.wayne.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Financial support for this work is provided by the National Science Foundation, Grants OISE-0730072 and CHE-0911191. We would like to thank WSU C&IT for computer time and support. This work is part of the research program of FOM, which is financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). The skillful assistance of the FELIX staff is gratefully acknowledged.

REFERENCES

(1) Wells, R. D. Molecular Basis of Genetic Instability of Triplet Repeats. J. Biol. Chem. **1996**, 271, 2875–2878.

(2) Mitas, M. Trinucleotide Repeats Associated with Human Disease. *Nucleic Acids Res.* **1997**, *25*, 2245–2253.

(3) Richards, R. I.; Sutherland, G. R. Dynamic Mutations: A New Class of Mutations Causing Human Disease. *Cell* **1992**, *70*, 709–712.

(4) Sinden, R. R.; Wells, R. DNA Structure, Mutations, and Human Genetic Disease. *Curr. Opin. Biotechnol.* **1992**, *3*, 612–622.

(5) Pearson, C. E.; Sinden, R. R. Trinucleotide Repeat DNA Structures: Dynamic Mutations from Dynamic DNA. *Curr. Opin. Struct. Biol.* **1998**, *8*, 321–330.

(6) Ashley, C. T., Jr; Warren, S. T. Trinucleotide Repeat Expansion and Human Disease. *Annu. Rev. Genet.* **1995**, *29*, 703–728.

(7) Kiliszek, A.; Kierzek, R.; Krzyzosiak, W. J.; Rypniewski, W. Structural Insights into CUG Repeats Containing the 'Stretched U–U Wobble': Implications for Myotonic Dystrophy. *Nucleic Acids Res.* **2009**, *37*, 4149–4156.

(8) Mooers, B. H. M.; Logue, J. S.; Berglund, J. A. The Structural Basis of Myotonic Dystrophy from the Crystal Structure of CUG Repeats. *Proc. Natl Acad. Sci. U.S.A.* **2005**, *102*, 16626–16631.

(9) Kiliszek, A.; Kierzek, R.; Krzyzosiak, W. J.; Rypniewski, W. Atomic Resolution Structure of CAG RNA Repeats: Structural Insights and Implications for the Trinucleotide Repeat Expansion Diseases. *Nucleic Acids Res.* **2010**, *38*, 8370–8376.

(10) Kiliszek, A.; Kierzek, R.; Krzyzosiak, W. J.; Rypniewski, W. Crystal Structures of CGG RNA Repeats with Implications for Fragile X-Associated Tremor Ataxia Syndrome. *Nucleic Acids Res.* **2011**, *16*, 7308–7315.

(11) Chen, X.; Mariappan, S. V.; Catasti, P.; Moyzis, R. K.; Laayoun, A.; Smith, S. S.; Bradbury, E. M.; Gupta, G. Hairpins are Formed by the Single DNA Strands of the Fragile X Triplet Repeats: Structure and Biological Implications. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5199–5203.

(12) Marsh, R. E.; Bierstedt, R.; Eichhorn, E. L. The Crystal Structure of Cytosine-5-Acetic Acid. *Acta Crystallogr.* **1962**, *15*, 310–316.

(13) Akrinimiski, E. O.; Sander, C.; Ts'o, P. O. P. Properties of Helical Polycytidylic Acid. *Biochemistry* 1963, 2, 340-344.

(14) Hartman, K. A., Jr; Rich, A. The Tautomeric Form of Helical Polycytidylic Acid. J. Am. Chem. Soc. 1965, 87, 2033–2039.

(15) Inman, R. B. Transitions of DNA Homopolymers. J. Mol. Biol. 1964, 9, 624–637.

(16) Weiner, A. M. Eukaryotic Nuclear Telomeres: Molecular Fossils of the RNP World? *Cell* **1988**, *52*, 155–157.

(17) Rajabi, K.; Theel, K.; Gillis, E. A. L.; Beran, G.; Fridgen, T. D. The Structure of the Protonated Adenine Dimer by Infrared Multiple Photon Dissociation Spectroscopy and Electronic Structure Calculations. *J. Phys. Chem. A.* **2009**, *113*, 8099–8107.

(18) Simonsson, T.; Kubista, M. DNA Tetraplex Formation in the Control Region of c-MYC. Nucleic Acids Res. **1998**, 26, 1167–1172.

(19) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. Direct Evidence for a G-Quadruplex in a Promoter Region and Its Targeting with a Small Molecule to Repress c-MYC Transcription. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11593–11598.

(20) Sun, D.; Hurley, L. H. Biochemical Techniques for the Characterization of G-Quadruplex Structures: EMSA, DMS Footprinting, and DNA Polymerase Stop Assay. *Methods Mol. Biol.* **2010**, 608, 65–79.

(21) Guo, K.; Pourpak, A.; Beetz-Rogers, K.; Gokhale, V.; Sun, D.; Hurley, L. H. Formation of Pseudosymmetrical G-Quadruplex and i-Motif Structures in the Proximal Promoter Region of the *RET* Oncogene. J. Am. Chem. Soc. **2007**, 129, 10220–10228.

(22) Sun, D.; Hurley, L. H. The Importance of Negative Superhelicity in Inducing the Formation of G-quadruplex and I-Motif Structures in the *c-Myc* Promoter: Implications for Drug Targeting and Control of Gene Expression. *J. Med. Chem.* **2009**, *52*, 2863–2874.

(23) Dai, J.; Ambrus, A.; Hurley, L. H.; Yang, D. A Direct and Nondestructive Approach To Determine the Folding Structure of the I-Motif DNA Secondary Structure by NMR. *J. Am. Chem. Soc.* **2009**, *131*, 6102–6104.

(24) Fotjík, P.; Vorlícková, M. The Fragile X Chromosome (GCC) Repeat Folds into a DNA Tetraplex at Neutral pH. *Nucleic Acids Res.* **2001**, *29*, 4684–4690.

(25) Gehring, K.; Leroy, J.-L.; Guéron, M. A Tetrameric DNA Structure with Protonated Cytosine-Cytosine Base Pairs. *Nature* **1993**, 363, 561–565.

(26) Gabelica, V.; Rosu, F.; Pauw, E. D.; Lemaire, J.; Gillet, J.-C.; Poully, J.-C.; Lecomte, F.; Grégoire, G.; Schermann, J.-P.; Desfrançois, C. Infrared Signature of DNA G-Quadruplexes in the Gas Phase. J. Am. Chem. Soc. **2008**, 130, 1810–1811.

(27) Rosu, F.; Gabelica, V.; Joly, L.; Grégoire, G.; Pauw, E. D. Zwitterionic *I*-Motif Structures are Preserved in DNA Negatively Charged Ions Produced by Electrospray Mass Spectrometry. *Phys. Chem. Chem. Phys.* **2010**, *12*, 13448–13454.

(28) Šponer, J.; Hobza, P. Bifurcated Hydrogen Bonds in DNA Crystal Structures: An Ab Initio Quantum Chemical Study. J. Am. Chem. Soc. **1994**, 116, 709-714.

(29) Šponer, J.; Hobza, P. Nonplanar Geometries of DNA Bases: Ab Initio Second-Order Moeller-Plesset Study. *J. Phys. Chem.* **1994**, *98*, 3161–3164.

(30) Hall, R. J.; Burton, N. A.; Hillier, I. H.; Young, P. E. Tautomeric Equilibria in 2-Hydroxypyridine and in Cytosine: An Assessment of Density Functional Methods, Including Gradient Corrections. *Chem. Phys. Lett.* **1994**, 220, 129–132.

(31) Sobolewski, A. L.; Adamowicz, L. Theoretical Investigations of Proton Transfer Reactions in a Hydrogen Bonded Complex of Cytosine with Water. J. Chem. Phys. **1995**, *102*, 5708–5718.

(32) Florián, J.; Leszczyński, J.; Johnson, B. G. On the Intermolecular Vibrational Modes of the Guanine…Cytosine, Adenine…Thymine, and Formamide…Formamide H-Bonded Dimers. *J. Mol. Struct.* **1995**, *349*, 421–426.

(33) Fülscher, M. P.; Roos, B. O. Theoretical Study of the Electronic Spectrum of Cytosine. J. Am. Chem. Soc. **1995**, 117, 2089–2095.

(34) Hobza, P.; Šponer, J.; Polásšek, M. H-Bonded and Stacked DNA Base Pairs: Cytosine Dimer. An ab Initio Second-Order Moeller-Plesset Study. J. Am. Chem. Soc. **1995**, 117, 792–798.

(35) Colominas, C.; Luque, F. J.; Orozco, M. Tautomerism and Protonation of Guanine and Cytosine: Implication in the Formation of Hydrogen-Bonded Complexes. *J. Am. Chem. Soc.* **1996**, *118*, 6811–6821.

(36) Florián, J.; Baumruk, V.; Leszczyński, J. IR and Raman Spectra, Tautomerism, and Scaled Quantum Mechanical Force Fields of Protonated Cytosine. *J. Phys. Chem.* **1996**, *100*, 5578–5589.

(37) Kwiatkowski, J. S.; Leszczyński, J. Molecular Structure and Vibrational IR Spectra of Cytosine and Its Thio and Seleno Analogues by Density Functional Theory and Conventional ab Initio Calculations. J. Phys. Chem. **1996**, 100, 941–953.

(38) Fogarasi, G. High-Level Electron Correlation Calculations on Some Tautomers of Cytosine. J. Mol. Struct. 1997, 413, 271–278.

(39) Kobayashi, R. A CCSD(T) Study of the Relative Stabilities of Cytosine Tautomers. J. Phys. Chem. A **1998**, 102, 10813–10817.

(40) Russo, N.; Toscano, M.; Grand, A. Lithium Affinity for DNA and RNA Nucleobases: The Role of Theoretical Information in the Elucidation of the Mass Spectrometry Data. *J. Phys. Chem. B* **2001**, *105*, 4735–4741.

(41) Russo, N.; Toscano, M.; Grand, A. Bonded Energies and Attachment Sites of Sodium and Potassium Cations to DNA and RNA Nucleic Acid Bases in the Gas Phase. *J. Am. Chem. Soc.* **2001**, *123*, 10272–10279.

(42) Fogarasi, G. Relative Stabilities of Three Low-Energy Tautomers of Cytosine. J. Phys. Chem. A. 2002, 106, 1381–1390.

(43) Trygubenko, S. A.; Bogdan, T. V.; Rueda, M.; Orozco, M.; Luque, F. J.; Šponer, J.; Slavíček, P.; Hobza, P. Correlated ab Initio Study of Nucleic Acid Bases and Their Tautomers in the Gas Phase, in a Microhydrated Enviroment and in Aqueous Solution. Part 1. Cytosine. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4192–4203.

(44) Yang, Z.; Rodgers, M. T. Theoretical Studies of the Unimolecular and Biomolecular Tautomerization of Cytosine. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2749–2757.

(45) Mazzuca, D.; Marino, T.; Russo, N.; Toscano, M. A Theoretical Study on Tautomerization Process of Dehydrated and Monohydrated Cytosine. *THEOCHEM*. **2007**, *811*, 161–167.

(46) Nei, Y.-w.; Akinyemi, T. E.; Steill, J. D.; Oomens, J.; Rodgers, M. T. Inrared Multiple Photon Dissociation Action Spectroscopy of Protonated Uracil and Thiouracils: Effects of Thioketo-Substitution on Gas-Phase Conformation. *Int. J. Mass. Spectrom* **2010**, *297*, 139–151.

(47) Crampton, K. T.; Rathur, A. I.; Nei, Y.-w.; Berden, G.; Oomens, J.; Rodgers, M. T. Protonation Preferentially Stabilizes Minor Tautomers of the Halouracils: IRMPD Action Spectroscopy and Theoretical Studies. J. Am. Soc. Mass. Spectrom. 2012, 23, 1469–1478.

(48) Nei, Y.-w.; Akinyemi, T. E.; Kaczan, C. M.; Steill, J. D.; Berden, G.; Oomens, J.; Rodgers, M. T. Infrared Multiple Photon Dissociation Action Spectroscopy of Sodiated Uracil and Thiouracils: Effects of Thioketo-Substitution on Gas-Phase Conformation. *Int. J. Mass. Spectrom.* **2011**, 308, 191–202.

(49) van Zundert, G. C. P.; Jaeqx, S.; Berden, G.; Bakker, J. M.; Kleinermanns, K.; Oomens, J.; Rijs, A. M. IR Spectroscopy of Isolated Neutral and Protonated Adenine and 9-Methyladenine. *Chem. Phys. Chem.* **2011**, *12*, 1921–1927.

(50) Oomens, J.; Moehlig, A. R.; Morton, T. H. Infrared Multiple Photon Dissociation (IRMPD) Spectroscopy of the Proton-Bound Dimer of 1-Methylcytosine in the Gas Phase. *J. Phys. Chem. Lett.* **2010**, *1*, 2891–2897.

(51) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. Free Electron Laser-Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Facility for Obtaining Infrared Multiphoton Dissociation Spectra of Gaseous Ions. *Rev. Sci. Instrum.* **2005**, *76*, 023103–0230109.

(52) Polfer, N. C.; Oomens, J.; Moore, D. T; von Helden, G.; Meijer, G.; Dunbar, R. C. Infrared Spectroscopy of Phenylalanine Ag(I) and Zn(II) Complexes in the Gas Phase. *J. Am. Chem. Soc.* **2006**, *128*, 517–525.

(53) Polfer, N. C.; J. Oomens, J. Reaction Products in Mass Spectrometry Elucidated with Infrared Spectroscopy. *Phys. Chem. Chem. Phys.* 2007, 9, 3804–3817.

(54) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.

(55) McLean, A. D.; Chandler, G. S. Contracted Gaussian Basis Sets for Molecular Calculations. I. Second row atoms, Z = 11-18. *J. Chem. Phys.* **1980**, *72*, 5639–5648.

(56) Foresman, J. B.; Frisch, A. E.. *Exploring Chemistry with Electronic Structures Methods*, 2nd ed.; Gaussian, Inc.: Pittsburg, PA, 1996; p 64. (57) Han, S. Y.; Oh, H. B. Theoretical Study of the Ionic Hydrogen Bond in the Isolated Proton-Bound Dimer of Cytosine. *Chem. Phys. Lett.* **2006**, 432, 269–274.

(58) Majerz, I.; Olovsson, I. The Shape of the Potential Energy Curves for NHN⁺ Hydrogen Bonds and the Influence of Non-Linearity. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3043–3051.

(59) Beran, G. J. O.; Chronister, E. L.; Daemen, L. L.; Moehlig, A. R.; Mueller, L. J.; Oomens, J.; Rice, A.; Santiago-Dieppa, D. R.; Tham, F. S.; Theel, K.; Taghmaei, S.; Morton, T. H. Vibrations of a Chelated Proton in a Protonated Tertiary Diamine. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20380–20392.

(60) Horbatenko, Y.; Vyboishchikov, S. V. Hydrogen Motion in Proton Sponge Cations: A Theoretical Study. *Chem. Phys. Chem.* **2011**, *12*, 1118–1129.

(61) Asmis, K. R.; Yang, Y.; Santambrogio, G.; Dipl.-Ing, M. B.; Roscioli, J. R.; McCunn, L. R.; Johnson, M. A.; Kühn, O. Gas-Phase Infrared Spectroscopy and Multidimensional Quantum Calculations of the Protonated Ammonia Dimer $N_2H_7^+$. Angew. Chem., Int. Ed. 2007, 46, 8691–8694.

(62) Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. D. Spectral Signatures of Hydrated Proton Vibrations in Water Clusters. *Science* **2005**, *308*, 1765–1769.