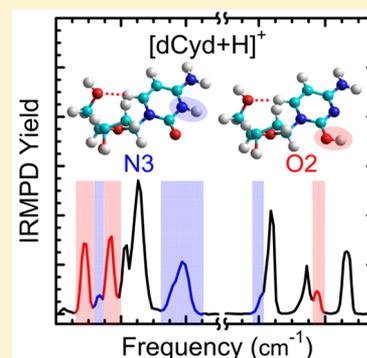


## N3 and O2 Protonated Tautomeric Conformations of 2'-Deoxycytidine and Cytidine Coexist in the Gas Phase

R. R. Wu,<sup>†</sup> Bo Yang,<sup>†</sup> C. E. Frier,<sup>†</sup> G. Berden,<sup>‡</sup> J. Oomens,<sup>‡,§</sup> and M. T. Rodgers<sup>\*,†</sup><sup>†</sup>Department of Chemistry, Wayne State University, Detroit, Michigan 48202, United States<sup>‡</sup>Institute for Molecules and Materials, FELIX Facility, Radboud University Nijmegen, Toernooiveld 7, 6525 ED, Nijmegen, The Netherlands<sup>§</sup>van't Hoff Institute for Molecular Sciences, University of Amsterdam, Amsterdam, The Netherlands

## Supporting Information

**ABSTRACT:** Infrared multiple photon dissociation action spectra of the protonated forms of the cytidyl nucleosides, 2'-deoxycytidine, [dCyd+H]<sup>+</sup>, and cytidine, [Cyd+H]<sup>+</sup>, are acquired over the IR fingerprint and hydrogen-stretching regions. Electronic structure calculations are performed at the B3LYP/6-311+G(d,p) level to determine the stable low-energy tautomeric conformations of these species generated upon electrospray ionization (ESI) and to generate the linear IR absorption spectra of these protonated nucleosides. Comparison between the experimental and theoretical spectra allows the tautomeric conformations of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> populated by ESI to be determined. B3LYP predicts N3 as the preferred protonation site for both [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup>, whereas MP2 suggests that protonation at O2 is more favorable. The 2'-hydroxyl substituent does not significantly alter the structures of the B3LYP N3 and MP2 O2 protonated ground tautomeric conformations of [dCyd+H]<sup>+</sup> vs [Cyd+H]<sup>+</sup>. [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> exhibit very similar spectral signatures in both regions. Nonetheless, the 2'-hydroxyl does affect the relative intensities of the IRMPD bands of [dCyd+H]<sup>+</sup> vs [Cyd+H]<sup>+</sup>. The spectral features observed in the hydrogen-stretching region complement those of the fingerprint region and allow the N3 and O2 protonated tautomeric conformations to be readily distinguished. Comparison between the measured and computed spectra indicates that both N3 and O2 protonated tautomeric conformations coexist in the experiments, and the populations are dominated by the most stable N3 and O2 protonated tautomeric conformations. Least-squares fitting of the IRMPD spectra to the IR spectra for these most stable conformers suggests relative populations of ~55% N3 vs 45% O2 protonated conformers of [dCyd+H]<sup>+</sup>, whereas ~47% N3 vs 53% O2 protonated conformers of [Cyd+H]<sup>+</sup>. This change in the preferred site of protonation indicates that the 2'-hydroxyl substituent plays an important role in controlling the reactivity of the cytidyl nucleosides.



## INTRODUCTION

The natural DNA and RNA nucleosides, composed of the various nucleobases and 2'-deoxyribose and ribose sugar moieties, are the fundamental building blocks of nucleic acids. Despite the limited numbers of naturally occurring DNA and RNA nucleoside building blocks, the biological information storage in nucleic acids, such as the genetic code and expression of proteins, is tremendous.<sup>1</sup> A variety of weak noncovalent molecular interactions<sup>2,3</sup> contribute to the stabilization of nucleic acids, including van der Waals, hydrophobic, hydrogen bonding,  $\pi$ -stacking, and ion-dipole interactions, and are responsible for the complexity of the structural and biochemical functions of DNA and RNA. Both hydrogen-bond donor and acceptor moieties on the nucleobases allow the formation of a complex network in nucleic acids through hydrogen-bonding interactions<sup>4,5</sup> that modulate their structure and function. Because of the importance of noncovalent interactions to the higher-order structure of nucleic acids, the 2'-hydroxyl substituents of RNA provide increased possibilities for hydrogen-bonding interactions, including sugar edge, Watson–Crick edge, and Hoogsteen edge interactions,<sup>6,7</sup> that

increase the structural complexity of RNA. The three-dimensional structural features of nucleic acids are primarily determined by the sugar configuration, which is typically C2'-endo or C3'-endo, as well as the nucleobase orientation about the glycosidic bond, anti or syn,<sup>1,8</sup> of the constituent nucleosides. For instance, double-stranded RNA helices adopt the A-form with C3'-endo sugar pucker rather than the B-form geometry that generally occurs in double-stranded DNA that involves C2'-endo sugar pucker.<sup>9</sup> Therefore, a more detailed understanding of the structures of the nucleosides as the fundamental building blocks of nucleic acids, and the role that the 2'-hydroxyl substituent plays in differentiating the DNA vs RNA nucleosides, would provide improved insight into the structures and functions of more complex nucleic acids.

Among the nucleobases, cytosine is the most basic in aqueous environments.<sup>1</sup> This property is of great interest, and crystal diffraction data<sup>10</sup> suggest that protonation often occurs

Received: December 31, 2014

Revised: April 14, 2015

Published: April 15, 2015

at the N3 position of cytosine in aqueous media or in the absence of solvent. The excess proton has been shown to participate in a variety of noncanonical base-pairing interactions involving cytosine.<sup>11–16</sup> The C<sup>+</sup>·C base pair has been found in DNA tetrameric *i*-motif structures.<sup>11,12,16</sup> The protonated C<sup>+</sup>·G·C base triplet contributes to DNA triple helices.<sup>13–15,17–19</sup> N3 protonation of cytosine has been identified in noncanonical protonated cytosine–cytosine<sup>11,12,16,20–23</sup> and cytosine–2-aminopurine base pairs.<sup>24,25</sup> In contrast, the O2 position of cytosine participates in the hydrogen-bonding interactions that stabilize the noncanonical protonated A·C<sup>+</sup> base pair observed in crystalline oligonucleotide duplexes.<sup>26</sup> However, NMR results<sup>27</sup> suggest that adenine not cytosine is protonated in the A·C<sup>+</sup> base pair. Because of the participation of protonated cytosine in these noncanonical base pairs that ultimately alter the structure and function of nucleic acids, elucidating the effects of protonation on the structures of the canonical cytidyl nucleosides is of great interest. <sup>15</sup>N, <sup>14</sup>N NMR, and NQR<sup>28</sup> have successfully examined the effects of protonation on cytosine (Cyt), cytidine (Cyd), and cytidine monophosphate (Cyd5'p) in aqueous and organic solvents.<sup>29–33</sup> These studies found that N3 protonation is preferred in solution. However, the nitrogen NMR spectroscopy techniques are only able to elucidate information regarding the nitrogen resonances, N1, N3, and NH<sub>2</sub> of cytosine, such that the effect of O2 protonation remains unclear. The intrinsic properties of protonated nucleosides are difficult to elucidate from condensed phase studies because they are masked by effects induced by the solvent and the local environment. In order to distinguish the intrinsic behavior of these systems from external effects, investigations of the isolated nucleosides in the gas phase are necessary. Numerous experimental and theoretical studies<sup>34–46</sup> have probed the gas-phase structural and thermochemical properties of Cyt, dCyd, and Cyd. These studies suggested that N3 and O2 have very similar proton affinities.<sup>30,31,34,35</sup> In particular, N3 and O2 protonated cytosine are very close in energy in the gas phase.<sup>38</sup> However, detailed information regarding the actual structures of the protonated forms of dCyd and Cyd remains ambiguous. Speranza and co-workers studied [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> by infrared multiple photon dissociation (IRMPD) action spectroscopy and complementary theoretical calculations.<sup>46</sup> *Ab initio* molecular dynamics simulations combined with calculations at the B3LYP/6-311++G(d,p) level suggest that N3 protonated conformers are the most stable for both [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup>. On the basis of comparison between the measured and computed spectra of the most stable N3 and O2 protonated conformers of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> from 1400–1800 and 2800–3800 cm<sup>-1</sup>, they found that N3 and O2 protonated tautomeric conformations with the nucleobases in anti configurations coexist in the experimental populations. Their measured IRMPD spectra are of limited utility for differentiating C2'- vs C3'-endo sugar configurations, and thus they concluded that the furanose rings may adopt either of these two conformations without evidence for a change in the puckering of the sugar moieties in these gas-phase measurements.

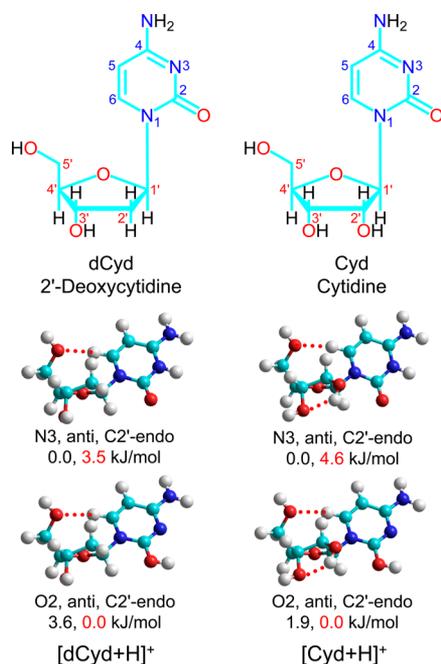
In the current work, IRMPD action spectroscopy techniques are employed to probe the structures and stabilities of the protonated forms of 2'-deoxycytidine [dCyd+H]<sup>+</sup> and cytidine [Cyd+H]<sup>+</sup>. Complementary quantum-chemical calculations are performed to characterize the stable low-energy tautomeric conformations, their relative stabilities, and their linear IR

absorption spectra. The resonant vibrational modes that lead to photofragmentation, the preferred sites of protonation, and the low-energy tautomeric conformations populated by electrospray ionization are elucidated by comparison between the measured and predicted IR spectra. The effects of the 2'-hydroxyl on the measured IRMPD spectra and structures and relative stabilities of the low-energy conformers are elucidated by comparing [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup>. To complement and enhance the previous work of Speranza and co-workers on these systems,<sup>46</sup> we have carried out computations at two levels of theory in an effort to more carefully elucidate the relative stabilities of the most stable conformers. In addition, our measured IRMPD spectra cover a larger portion of the IR fingerprint region and are slightly better resolved. These enhancements in the experimental and theoretical approaches enable the sugar puckering to be elucidated and the specific tautomeric conformations populated to be determined. Least-squares fitting of the IRMPD data also enables approximate populations of the conformers accessed to be determined. Calculations were also performed to characterize the stable low-energy tautomeric conformations of neutral dCyd and Cyd, which are compared to the protonated species to decipher the effect of protonation on the structures of these cytidyl nucleosides. The structures of neutral dCyd and Cyd are also compared to results from condensed phase studies to elucidate the effects of solvent on their structures.

## ■ EXPERIMENTAL AND COMPUTATIONAL APPROACHES

**Experimental Protocols.** IRMPD action spectra of the protonated forms of 2'-deoxycytidine, [dCyd+H]<sup>+</sup>, and its RNA analogue, cytidine, [Cyd+H]<sup>+</sup>, were acquired using a 4.7 T Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS)<sup>47–49</sup> in combination with a widely tunable free electron laser (FEL)<sup>50</sup> source or an optical parametric oscillator/amplifier (OPO/OPA) laser setup to examine both the fingerprint and the hydrogen-stretching regions. Both nucleosides were procured from Sigma-Aldrich. The protonated nucleosides were generated from a solution of 0.5 mM dCyd or Cyd with 3 mM acetic acid dissolved in a 1:1 methanol–water mixture using an electrospray ionization (ESI) source of the *z*-geometry infused at a rate in the range 2.5–8.5 μL/min. The protonated nucleosides were accumulated for a few seconds in an rf hexapole ion trap and then pulse extracted through a quadrupole bender and injected into the trapping cell of the FT-ICR MS via an rf octopole ion guide. Collisional heating of the ions is avoided as previously described.<sup>48</sup> The ions were cooled to room temperature in the ICR trapping cell by radiative emission. Stored waveform inverse Fourier transform (SWIFT) techniques were used to isolate the protonated nucleoside, [dCyd+H]<sup>+</sup> or [Cyd+H]<sup>+</sup>. The isolated protonated nucleoside was irradiated for 2.5–3 s by the FEL or for 4–8 s by the OPO/OPA laser system to induce photodissociation. Spectral features in the fingerprint region were examined using the FEL from ~550 to 1920 cm<sup>-1</sup> (corresponding to wavelengths extending from 18.0 to 5.2 μm). The features in the hydrogen-stretching region were measured using the OPO/OPA laser set up from ~2800 to 3800 cm<sup>-1</sup> (or ~3.6 to 2.6 μm).

**Computational Protocols.** Chemical structures of neutral dCyd and Cyd are displayed in Figure 1. Cytosine is illustrated in an anti orientation in both structures. The most favorable sites of protonation, N3 and O2, of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup>



**Figure 1.** Chemical structures of neutral 2'-deoxycytidine (dCyd) and cytidine (Cyd). The numbering of the nucleobase and sugar is also shown. Ground-state structures of  $[\text{dCyd}+\text{H}]^+$  and  $[\text{Cyd}+\text{H}]^+$  predicted at the B3LYP/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) (shown in black) and MP2(full)/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) (shown in red) levels of theory. The site of protonation, nucleobase orientation, and sugar pucker are also indicated for each protonated nucleoside.

were probed in detail. HyperChem software<sup>51</sup> using the Amber 2 force field was first employed to generate candidate structures for  $[\text{dCyd}+\text{H}]^+$  and  $[\text{Cyd}+\text{H}]^+$  via simulated annealing procedures exactly analogous to those described previously.<sup>52,53</sup> Based on the relative stabilities predicted in the simulated annealing procedures and taking into account all combinations of the N3 and O2 protonation sites examined, C2'-endo and C3'-endo puckering of the sugar moieties, and syn and anti nucleobase orientations, 30 contending structures for each protonation site were subjected to more advanced theoretical methods. The Gaussian 09 suite of programs was used to perform geometry optimizations, frequency analyses, and energy calculations for all contending structures.<sup>54</sup> To speed up convergence of the geometry optimization, each structure was optimized at the B3LYP/6-31G(d) level, followed by reoptimization at the B3LYP/6-311+G(d,p) level to improve the geometric description of these protonated nucleosides. Frequency analyses of each stable tautomeric conformation again at the B3LYP/6-311+G(d,p) level were performed to elucidate the vibrational frequencies and IR intensities associated with each stable conformer. Energies were computed at both the B3LYP/6-311+G(2d,2p) and MP2/6-311+G(2d,2p) levels to determine the relative stabilities of all of the stable low-energy tautomeric conformations found. Standard formulas without scaling of the frequencies were employed to make zero-point energy (ZPE) and thermal corrections to 298 K. The IR spectrum for each tautomeric conformation was generated using the IR intensities and B3LYP/6-311+G(d,p) frequencies, which were multiplied by 0.98 and convoluted with a 20  $\text{cm}^{-1}$  fwhm Gaussian line shape for the FELIX region, and

a 0.957 scaling factor with a 15  $\text{cm}^{-1}$  broadening for the OPO region.

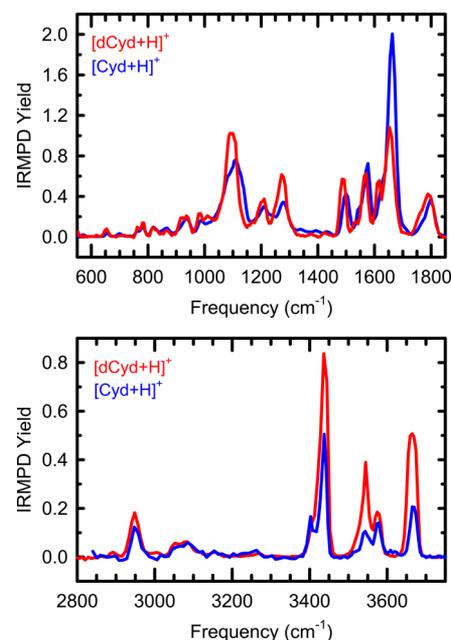
## RESULTS

**IRMPD Action Spectroscopy.** In both the IR fingerprint and hydrogen-stretching regions, cleavage of the N-glycosidic bond producing protonated cytosine,  $[\text{Cyt}+\text{H}]^+$ , was the only fragmentation pathway observed for both  $[\text{dCyd}+\text{H}]^+$  and  $[\text{Cyd}+\text{H}]^+$ . The IRMPD yield of the protonated nucleoside,  $[\text{Nuo}+\text{H}]^+ = [\text{dCyd}+\text{H}]^+$  and  $[\text{Cyd}+\text{H}]^+$ , was calculated from the measured ion intensities after irradiation at each wavelength as demonstrated in eq 1

$$\text{IRMPD yield} = I_{[\text{Cyt}+\text{H}]^+} / (I_{[\text{Cyt}+\text{H}]^+} + I_{[\text{Nuo}+\text{H}]^+}) \quad (1)$$

The computed IRMPD yield was normalized to correct for variations in the laser output as a function of the photon energy, i.e., the frequency of irradiation. IRMPD spectra were measured for  $[\text{dCyd}+\text{H}]^+$  and  $[\text{Cyd}+\text{H}]^+$  over the ranges extending from  $\sim 550$  to 1900 and  $\sim 2800$  to 3800  $\text{cm}^{-1}$  and are compared in Figure 2. Very similar spectral features in both regions of the IRMPD spectra of  $[\text{dCyd}+\text{H}]^+$  and  $[\text{Cyd}+\text{H}]^+$  are observed. The relative intensities of many of the observed bands are strongly impacted by the 2'-hydroxyl substituent. In the IR fingerprint region above  $\sim 1500$   $\text{cm}^{-1}$ ,  $[\text{Cyd}+\text{H}]^+$  exhibits a greater IRMPD yield than  $[\text{dCyd}+\text{H}]^+$ , whereas  $[\text{dCyd}+\text{H}]^+$  produces a higher yield than  $[\text{Cyd}+\text{H}]^+$  in the hydrogen-stretching region above  $\sim 3400$   $\text{cm}^{-1}$ . This behavior is consistent with that observed for the protonated guanidyl and adenosyl nucleosides,  $[\text{dGuo}+\text{H}]^+$  vs  $[\text{Guo}+\text{H}]^+$ <sup>52</sup> and  $[\text{dAdo}+\text{H}]^+$  vs  $[\text{Ado}+\text{H}]^+$ , respectively.<sup>53</sup>

**Theoretical Results.** The most stable tautomeric conformations of  $[\text{dCyd}+\text{H}]^+$  and  $[\text{Cyd}+\text{H}]^+$  found at the B3LYP/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) and MP2(full)/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) levels are compared in Figure 1. B3LYP predicts N3, whereas MP2 predicts O2 as the most stable site of protonation for both  $[\text{dCyd}+\text{H}]^+$  and  $[\text{Cyd}+\text{H}]^+$ . In the B3LYP and MP2 ground structures of both



**Figure 2.** Infrared multiple photon dissociation (IRMPD) action spectra of  $[\text{dCyd}+\text{H}]^+$  and  $[\text{Cyd}+\text{H}]^+$  in the FELIX and OPO regions.

Table 1. Relative Enthalpies and Free Energies at 0 and 298 K in kJ/mol of Stable Low-Energy Conformers of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup><sup>a</sup>

species	conformer	B3LYP			MP2(full)		
		$\Delta H_0$	$\Delta H_{298}$	$\Delta G_{298}$	$\Delta H_0$	$\Delta H_{298}$	$\Delta G_{298}$
[dCyd+H] <sup>+</sup>	N3A	0.9	0.8	<b>0.0</b>	4.4	4.9	3.5
	N3B	0.0	0.0	0.2	5.4	5.9	5.6
	O2A	3.6	3.1	3.6	0.0	0.0	<b>0.0</b>
	O2B	4.4	4.0	5.2	3.0	3.1	3.8
	O2C	7.5	5.8	10.3	2.4	1.1	5.1
	N3C	17.8	18.0	15.0	24.4	25.2	21.7
	O2D	20.7	20.5	18.7	20.6	21.0	18.6
	N3D	25.1	24.5	25.6	28.5	28.4	29.2
	N3E	30.2	31.1	26.5	37.1	38.6	33.4
	O2E	33.2	33.5	31.3	33.7	34.6	31.9
	O2F	34.2	33.6	33.6	31.8	31.8	31.5
	O2G	36.4	36.3	35.6	33.0	33.5	32.2
	O2H	39.6	40.0	39.0	38.4	39.3	37.8
O2I	55.5	55.8	52.6	55.8	56.7	52.9	
[Cyd+H] <sup>+</sup>	N3A	0.4	0.5	<b>0.0</b>	5.2	5.8	4.6
	N3i	0.0	0.0	0.6	5.0	5.4	5.4
	O2A	1.5	1.2	1.9	0.0	0.0	<b>0.0</b>
	N3B	2.3	2.4	2.5	7.3	7.8	7.2
	N3ii	3.4	2.9	4.0	6.6	6.4	6.9
	O2B	3.9	3.5	4.4	2.7	2.7	3.0
	O2i	4.6	3.2	5.8	3.4	2.4	4.3
	N3iii	6.1	6.1	6.3	9.8	10.1	9.6
	O2ii	6.2	5.8	6.9	3.3	3.2	3.7
	O2C	5.4	4.2	7.4	3.1	2.2	4.7
	O2iii	7.9	7.7	8.9	6.0	6.1	6.6
	O2iv	15.1	14.2	14.1	20.7	20.1	19.4
	N3C	18.4	18.7	16.3	24.3	24.9	21.9
	N3D	16.4	15.9	16.7	22.3	22.2	22.4
	O2D	18.4	18.2	16.9	18.6	18.7	16.7
	N3E	22.1	22.6	19.2	30.9	31.8	27.7
	O2E	26.5	26.8	24.2	29.2	29.9	26.5
O2F	27.0	26.8	25.9	27.6	27.7	26.1	

<sup>a</sup>Single point energy calculations of B3LYP/6-311+G(d,p) optimized structures are performed at the B3LYP/6-311+G(2d,2p) and MP2(full)/6-311+G(2d,2p) levels of theory including ZPE and thermal corrections without scaling of the B3LYP/6-311+G(d,p) frequencies.

[dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup>, the orientation of the cytosine residues are both anti, and the 2'-deoxyribose and ribose sugars both exhibit C2'-endo pucker. The nucleobase orientation and sugar pucker of the B3LYP N3 and MP2 O2 protonated ground conformers of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> are virtually identical such that the only significant distinction between the B3LYP and MP2 ground structures is the location of the excess proton. The 2'-hydroxyl substituent does not significantly affect the tautomeric conformations of the B3LYP and MP2 ground conformers of [Cyd+H]<sup>+</sup> versus those of [dCyd+H]<sup>+</sup>, except that the O2'H...O3' intramolecular hydrogen-bonding interaction of [Cyd+H]<sup>+</sup> leads to change in the rotation of the 3'-hydroxyl versus that of [dCyd+H]<sup>+</sup> where this hydrogen bond cannot occur. For both N3 and O2 protonation, the presence of the 2'-hydroxyl substituent leads to a small contraction, ~0.02 Å of the glycosidic bond C1'–N1 in [Cyd+H]<sup>+</sup>, suggesting that the 2'-hydroxyl substituent slightly increases the intrinsic glycosidic bond stability. The hydrogen bond between the 2'- and 3'-hydroxyls of [Cyd+H]<sup>+</sup> results in a small shift in the rotation of cytosine relative to the sugar that lengthens the very weak H6...O5' hydrogen bond by ~0.04 Å in [Cyd+H]<sup>+</sup>. The hydrogen bond between the 2'- and 3'-hydroxyl substituents in [Cyd+H]<sup>+</sup> ever so slightly enhances

C2'-endo sugar pucker as seen by the 2.0° increase in the  $\angle C1'C2'C3'C4'$  dihedral angle. The hydrogen bond between the 2'- and 3'-hydroxyl substituents also slightly alters the orientation of the 5'-hydroxylmethyl substituent in [Cyd+H]<sup>+</sup> and leads to an ~1.5° increase in the  $\angle O5'C5'C4'O4'$  dihedral angle.

The relative enthalpies and free energies at 298 K of the N3 and O2 protonated low-energy tautomeric conformations of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> are given in Table 1, whereas the corresponding optimized structures of the low-energy conformers of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> along with their relative free energies at 298 K are shown in Figures S1 and S2 of the Supporting Information. The nomenclature used to describe the low-energy conformers is based on the site of protonation (N3 or O2) along with a capital letter for conformers that are found for both [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> or a lowercase Roman numeral for conformers that are only found for [Cyd+H]<sup>+</sup>. In both cases, the ordering of the low-energy conformers is based on the B3LYP stabilities at 298 K. The low-energy conformers chosen for display are based primarily on their 298 K relative free energies, but also include all combinations of the protonation sites (N3 and O2), nucleobase orientations (anti and syn), and puckering of the sugar moieties

(C2'-endo and C3'-endo). Because the 2'-hydroxyl substituent enhances the possibilities for hydrogen-bonding interactions with the cytosine residue, and multiple favorable rotational orientations of the 2'- and 3'-hydroxyls exist, more low-energy tautomeric conformations are available to [Cyd+H]<sup>+</sup> and displayed in Figure S2. The low-energy tautomeric conformations found and shown in these figures are described in detail in the text of the Supporting Information.

**Influence of the 2'-Hydroxyl on [Cyd+H]<sup>+</sup> versus [dCyd+H]<sup>+</sup>.** Comparison of the low-energy conformers of [dCyd+H]<sup>+</sup> to those of [Cyd+H]<sup>+</sup> enables the structural impact of the 2'-hydroxyl substituent to be elucidated. The 2'-hydroxyl enables a hydrogen-bonding interaction to occur between the 2'- and 3'-hydroxyls, with several favorable orientations possible, such that more stable low-energy tautomeric conformations are found for [Cyd+H]<sup>+</sup>, such as the N3i, N3iii, O2ii, and O2iii conformers. The 2'-hydroxyl also provides possibilities for hydrogen bonding to the nucleobase, such as in the N3ii, O2i, and O2iv conformers, which involve either the O3'H...O2'H...O2 or O2H...O2'H...O3' bridging hydrogen bonds. However, in all cases the nucleobase in these latter structures changes conformation significantly leading to loss of the O5'...H6 noncanonical hydrogen bonding interaction and destabilization of the system. Conformations analogous to the O2G, O2H, and O2I conformers of [dCyd+H]<sup>+</sup> are not found for [Cyd+H]<sup>+</sup> because the presence of the 2'-hydroxyl substituent allows it to interact with O2H<sup>+</sup>, which leads to the O2i or O2iv conformers instead.

**Orientations of the 2'- and 3'-Hydroxyls in [Cyd+H]<sup>+</sup>.** In general, C2'-endo puckering of the ribose sugar is favored over C3'-endo. For N3 protonation, C2'-endo puckering favors both hydroxyl groups to be rotated down and away from the cytosine residue, parallel to that found for [Guo+H]<sup>+</sup> and [Ado+H]<sup>+</sup>,<sup>52,53</sup> whereas C3'-endo puckering favors both hydroxyl groups to be rotated down and toward the cytosine residue, opposite to that found for [Guo+H]<sup>+</sup> and [Ado+H]<sup>+</sup>.<sup>52,53</sup> In contrast, for O2 protonation, C2'-endo puckering prefers both hydroxyls to be rotated down and away from the cytosine residue, whereas C3'-endo puckering prefers both hydroxyl groups to be rotated up and away from the nucleobase, consistent with that found for [Guo+H]<sup>+</sup> and [Ado+H]<sup>+</sup>.<sup>52,53</sup>

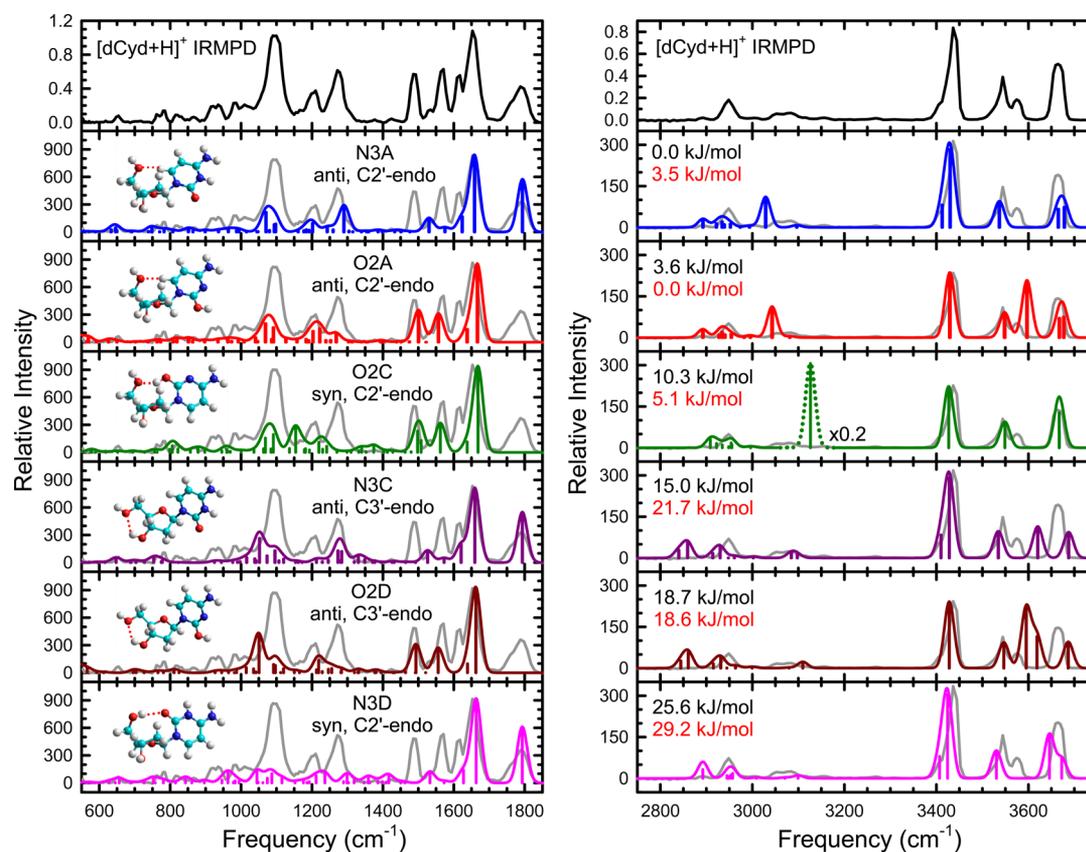
## DISCUSSION

**IRMPD vs IR Spectra.** Comparison between experimental and calculated spectra has proven to be a potent and robust method for characterizing the gas-phase conformations of ionic species.<sup>55</sup> Because multiple photons are involved in IRMPD fragmentation processes and vibrational modes exhibit varying degrees of anharmonicity, the measured IRMPD and linear IR spectra are not identical. However, discrepancies in the vibrational frequencies and relative intensities of the linear vs nonlinear IR spectra are generally sufficiently limited that the IRMPD spectrum provides a good reflection of the IR absorption spectrum.<sup>56</sup> For the protonated nucleosides investigated in this work, ~5–15 photons must be absorbed to cleave the glycosidic bond to produce the protonated nucleobase, which requires ~1.0–1.2 eV.<sup>57</sup> Because of the modest number of photons involved in the dissociation processes examined here, the differences between the measured and predicted spectra ascribed to the multiple-photon effect should be limited. The measured IR bandwidth is related to the linewidth of the laser employed. Frequency analyses using the

B3LYP hybrid density functional have proven to be adequate for a variety of ions.<sup>58</sup> In related work, the 6-311+G(d,p) basis set employed here has been found to reproduce the measured IRMPD spectra of the protonated adenosyl and guanidyl DNA and RNA nucleosides.<sup>52,53</sup> Frequency scaling factors and the fwhm for broadening of the computed line spectra are carefully chosen to best reproduce the band shapes and positions of the measured spectra. Therefore, agreement between the measured and computed spectra is used as the criterion for determining the absence or presence of the low-energy tautomeric conformations that contribute to the experimental measurements.

**Experimental IRMPD versus Theoretical IR Spectra of [dCyd+H]<sup>+</sup>.** The IRMPD and predicted IR spectra of the most illustrative low-energy tautomeric conformations of [dCyd+H]<sup>+</sup>, protonated at N3 and O2, N3A, O2A, O2C, N3C, O2D, and N3D, are overlaid in Figure 3. In the FELIX region, the IR spectra predicted for N3A and O2A exhibit good agreement with the experimental spectrum. In particular, the IR spectra predicted for the B3LYP ground conformer, N3A, and the MP2 ground conformer, O2A, complement each other, indicating that each make unique contributions to the measured spectrum. The calculated bands at ~1790, ~1530, and ~1290 cm<sup>-1</sup> for N3A, which are not predicted for O2A, contribute to the measured bands at ~1790, ~1530, and ~1280 cm<sup>-1</sup>. The calculated bands at ~1555 and ~1500 cm<sup>-1</sup> for O2A, which are not predicted for N3A, contribute to the measured bands at ~1560 and ~1490 cm<sup>-1</sup>. The band predicted at ~1155 cm<sup>-1</sup> for O2C is strongly shifted to lower frequencies relative to the band observed at ~1210 cm<sup>-1</sup>. For both N3C and O2D, the 5'-hydroxymethyl moiety is staggered away from the sugar and base and hydrogen-bonded to the 3'-hydroxyl moiety. This unique interaction leads to a predicted band at ~1050 cm<sup>-1</sup> for N3C and O2D, which would broaden the relatively sharp band measured at ~1095 cm<sup>-1</sup> if either of these two tautomeric conformations are populated by electrospray ionization. The band predicted at ~960 cm<sup>-1</sup> for N3D is not seen in the experimental spectrum. Very weak C–H stretches occur below ~3300 cm<sup>-1</sup> in the OPO region but are not usefully diagnostic. Above 3350 cm<sup>-1</sup>, the IR features predicted for N3A and O2A agree very well with the observed IRMPD features except for the small band observed at ~3575 cm<sup>-1</sup>. Both of these conformers contribute to the measured bands at ~3660, ~3545, and ~3435 cm<sup>-1</sup>, whereas only the O2A conformer contributes to the minor band that appears at ~3575 cm<sup>-1</sup>. However, theory overestimates the frequency of this mode by ~20 cm<sup>-1</sup>. The IR spectrum predicted for O2C agrees well with the IRMPD spectrum in the OPO region, but its presence in the experiments is ruled out based on differences in the FELIX region. Large discrepancies are found between the measured and calculated spectra in the OPO region above 3475 cm<sup>-1</sup> for N3C and O2D. The band at ~3645 cm<sup>-1</sup> for N3D is shifted and broadened as compared to the measured band at ~3660 cm<sup>-1</sup>. The calculated bands at ~3530 and ~3420 cm<sup>-1</sup> for N3D are also shifted to lower frequencies versus the bands observed at ~3545 and ~3435 cm<sup>-1</sup>, respectively. Therefore, based on comparison between experiment and theory, both the N3A and O2A conformers are populated by electrospray ionization.

The measured and calculated spectra predicted for the N3B, O2B, N3E, and O2E conformers of [dCyd+H]<sup>+</sup> are overlaid in Figure S3. In the FELIX region, the IR spectra predicted for N3B, O2B, N3E, and O2E exhibit nice agreement with the



**Figure 3.** Comparison of the measured IRMPD action spectrum of  $[\text{dCyd}+\text{H}]^+$  with the theoretical linear IR spectra for the ground and select stable low-energy conformers of  $[\text{dCyd}+\text{H}]^+$  and the corresponding optimized structures calculated at the B3LYP/6-311+G(d,p) level of theory. Features shown as dotted lines are multiplied by a factor of 0.2 for clarity. Also shown are the B3LYP/6-311+G(2d,2p) (shown in black) and MP2(full)/6-311+G(2d,2p) (shown in red) relative Gibbs free energies at 298 K. The site of protonation, nucleobase orientation, and sugar pucker are also indicated for each conformer. To facilitate comparison of the measured and computed spectra, the IRMPD spectrum is overlaid (in gray) with each computed spectrum and scaled to match the intensity of the most intense feature in each region.

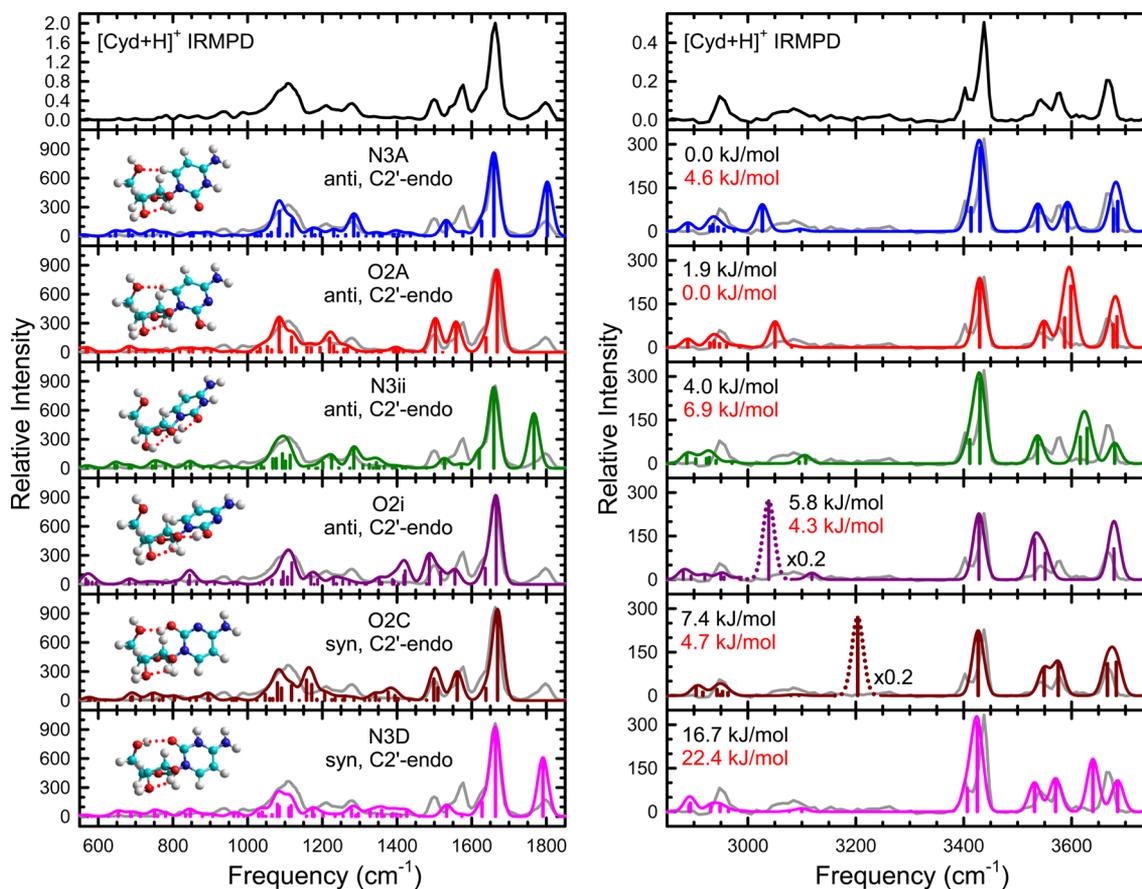
IRMPD spectrum. The calculated IR features in the FELIX region of N3B and O2B are very similar to those of N3A and O2A (Figure 3), respectively, suggesting that the change in the sugar pucker does not significantly influence the IR features in the FELIX region. In the OPO region, the bands predicted at  $\sim 3680\text{ cm}^{-1}$  for N3B and O2B are pushed to a higher frequency versus the band observed at  $\sim 3660\text{ cm}^{-1}$ , a likely result of the C3'-endo pucker in these conformers. The computed bands at  $\sim 3695\text{ cm}^{-1}$  for N3E and  $\sim 3690\text{ cm}^{-1}$  for O2E are also at higher frequencies than the band measured at  $\sim 3660\text{ cm}^{-1}$ . The band at  $\sim 3530\text{ cm}^{-1}$  predicted for N3E is at a lower frequency than the band measured at  $\sim 3545\text{ cm}^{-1}$ . These significant differences in the measured and predicted spectra suggest that these four low-energy tautomeric conformations do not contribute to the experiments.

The experimental and IR spectra predicted for the O2F, O2G, O2H, and O2I conformers of  $[\text{dCyd}+\text{H}]^+$  are overlaid in Figure S4. In the FELIX region, the computed bands at  $\sim 1230$  and  $\sim 965\text{ cm}^{-1}$  for O2F do not appear in the experimental spectrum. The calculated band at  $\sim 1230\text{ cm}^{-1}$  would broaden the bands observed at  $\sim 1280$  and  $1210\text{ cm}^{-1}$  if O2F is accessed. The orientation of the excess proton of O2G, O2H, and O2I results in a band being predicted at  $\sim 1550\text{ cm}^{-1}$  for all three conformers, which is shifted to lower a frequency versus the band measured at  $\sim 1570\text{ cm}^{-1}$ . Similar to N3C and O2D (Figure S3), the O2I conformer also has the 5'-hydroxymethyl group staggered away from the sugar and base and hydrogen-

bonded to the 3'-hydroxyl, which results in a band predicted at  $\sim 1060\text{ cm}^{-1}$ , which would broaden the relatively sharp band observed at  $\sim 1095\text{ cm}^{-1}$  if O2I is accessed. In the OPO region, the IR spectrum predicted for O2F agrees nicely with the experimental spectrum, but its presence is ruled out based on differences in the FELIX region. Obvious discrepancies above  $\sim 3600\text{ cm}^{-1}$  between experiment and theory are seen for O2G and O2I. The IR feature calculated at  $\sim 3680\text{ cm}^{-1}$  for O2H would broaden the band observed at  $\sim 3660\text{ cm}^{-1}$  if O2H were populated. These comparisons indicate these four tautomeric conformations are not populated by electrospray ionization.

In summary, comparison between the measured IRMPD spectrum with the theoretical IR spectra indicates that only the B3LYP ground conformer, N3A, and the MP2 ground conformer, O2A, contribute to the experimental measurements. Thus, regardless of the protonation site, only C2'-endo puckered and anti oriented tautomeric conformations are populated by electrospray ionization.

**Vibrational Assignments of  $[\text{dCyd}+\text{H}]^+$ .** Comparison of the IRMPD spectrum with the IR spectra of the N3A and O2A conformers allows vibrational assignments to be made. In the FELIX region, the measured band at  $\sim 1790\text{ cm}^{-1}$  representing carbonyl stretching clearly establishes the presence of N3A in the experiments. The strong IR band measured at  $\sim 1650\text{ cm}^{-1}$  with a shoulder on the low-frequency side primarily reflects  $\text{NH}_2$  scissoring coupled with stretching of the nucleobase ring and nucleobase-H in-plane bending. The two sharp IR bands



**Figure 4.** Comparison of the measured IRMPD action spectrum of  $[\text{Cyd}+\text{H}]^+$  with the theoretical linear IR spectra for the ground and select stable low-energy conformers of  $[\text{Cyd}+\text{H}]^+$  and the corresponding optimized structures calculated at the B3LYP/6-311+G(d,p) level of theory. Features shown as dotted lines are multiplied by a factor of 0.2 for clarity. Also shown are the B3LYP/6-311+G(2d,2p) (shown in black) and MP2(full)/6-311+G(2d,2p) (shown in red) relative Gibbs free energies at 298 K. The site of protonation, nucleobase orientation, and sugar pucker are also indicated for each conformer. To facilitate comparison of the measured and computed spectra, the IRMPD spectrum is overlaid (in gray) with each computed spectrum and scaled to match the intensity of the most intense feature in each region.

observed at  $\sim 1560$  and  $\sim 1490$   $\text{cm}^{-1}$  are contributed by the **O2A** conformer and represent stretches of nucleobase ring coupled with nucleobase-H in-plane bending. The weak IR feature measured at  $\sim 1530$   $\text{cm}^{-1}$ , which appears as a shoulder to the red of the sharp band at  $\sim 1560$   $\text{cm}^{-1}$ , is contributed by the **N3A** conformer and arises from N3-H in-plane bending coupled with stretches of the nucleobase ring. The band observed at  $\sim 1280$   $\text{cm}^{-1}$ , also arising from the **N3A** conformer, arises from coupled N3-H, C6-H, and C1'-H in-plane bending. The band observed at  $\sim 1210$   $\text{cm}^{-1}$  primarily reflects coupled bends of hydrogen atoms of the 2'-deoxyribose moiety. The strong IR feature measured at  $\sim 1095$   $\text{cm}^{-1}$  results from stretching of the ring of the 2'-deoxyribose moiety. In the OPO region, the band observed at  $\sim 3660$   $\text{cm}^{-1}$  reflects coupled O3'-H and O5'-H stretching. The weak feature observed at  $\sim 3575$   $\text{cm}^{-1}$ , contributed by **O2A**, results from O2-H stretching. The band measured at  $\sim 3545$   $\text{cm}^{-1}$  reflects asymmetric  $\text{NH}_2$  stretching, and the strong IR feature observed at  $\sim 3435$   $\text{cm}^{-1}$  results from symmetric  $\text{NH}_2$  stretching. The very weak shoulder (observed at  $\sim 3405$   $\text{cm}^{-1}$ ) on the low-frequency side of the most intense band is contributed by N3-H stretching, which is predicted at  $\sim 3410$   $\text{cm}^{-1}$  for **N3A**.

**Experimental IRMPD versus Theoretical IR Spectra of  $[\text{Cyd}+\text{H}]^+$ .** The IRMPD and IR spectra predicted of the most illustrative low-energy tautomeric conformations of  $[\text{Cyd}+\text{H}]^+$ , protonated at N3 and O2, **N3A**, **O2A**, **N3ii**, **O2i**, **O2C**, and

**N3D**, are overlaid in Figure 4. In the IR fingerprint region above  $1400$   $\text{cm}^{-1}$ , the IR spectra determined for the B3LYP ground conformer, **N3A**, and the MP2 ground conformer, **O2A**, are highly complementary except for the band predicted at  $\sim 1660$   $\text{cm}^{-1}$ , which is predicted for both conformers. The IR band predicted at  $\sim 1800$   $\text{cm}^{-1}$  of **N3A** contributes to the IRMPD feature observed at  $\sim 1800$   $\text{cm}^{-1}$ , which is not predicted for **O2A**. The weak IR feature predicted at  $\sim 1530$   $\text{cm}^{-1}$  for **N3A** is absent for **O2A**. This calculated IR feature of **N3A** contributes to the shoulder to the red of the band measured at  $\sim 1577$   $\text{cm}^{-1}$ . Two moderately strong IR features predicted at  $\sim 1560$  and  $\sim 1500$   $\text{cm}^{-1}$  for **O2A**, but absent for **N3A**, contribute to the two IRMPD bands measured at  $\sim 1575$  and  $\sim 1500$   $\text{cm}^{-1}$ . The calculated IR features below  $\sim 1400$   $\text{cm}^{-1}$  of **N3A** and **O2A** also agree nicely with the IRMPD spectrum over this region. Therefore, the lowest energy N3 and O2 protonated conformers **N3A** and **O2A** coexist in the population accessed by ESI. The IR band predicted at  $\sim 1765$   $\text{cm}^{-1}$  for **N3ii** is shifted to a lower frequency than the band observed at  $\sim 1800$   $\text{cm}^{-1}$ , which may be due to stabilization of the O2 carbonyl group provided by the O3'H...O2'H...O2 bridging hydrogen-bonding interaction. The IR features predicted at  $\sim 1515$  and  $\sim 1485$   $\text{cm}^{-1}$  for **O2i** would tend to broaden the band at  $\sim 1500$   $\text{cm}^{-1}$  if **O2i** was populated in the experiments. The band predicted at  $\sim 1410$   $\text{cm}^{-1}$  for **O2i** does not contribute to the measured spectrum. The calculated band at  $\sim 1160$   $\text{cm}^{-1}$

for O2C would broaden the measured band at  $\sim 1115\text{ cm}^{-1}$  if O2C was accessed in the experiments. Therefore, these differences indicate that N3ii, O2i, and O2C are not populated in the experiments. The IR spectrum of N3D exhibits nice agreement with the IRMPD spectrum in the FELIX region. Both N3A and O2A are clearly produced by electrospray ionization and contribute to the IRMPD spectrum in the FELIX region. Therefore, the experimental spectrum in the OPO region must also have contributions from both N3A and O2A. Theory tends to very slightly overestimate the IR frequencies above  $\sim 3550\text{ cm}^{-1}$ , but the calculated IR features of N3A and O2A generally match the observed IR features quite nicely. The calculated band at  $\sim 3620\text{ cm}^{-1}$  for N3ii deviates markedly from the band observed at  $\sim 3575\text{ cm}^{-1}$ . The broad band computed at  $\sim 3530\text{ cm}^{-1}$  for O2i would significantly broaden the band observed at  $\sim 3540\text{ cm}^{-1}$  if this conformer was populated by electrospray ionization. The IR spectrum predicted for O2C shows good agreement with the measured spectrum in the OPO region, but its presence is eliminated based on the absence of the band calculated at  $\sim 1160\text{ cm}^{-1}$  in the IRMPD spectrum in the FELIX region. The band predicted at  $\sim 3685\text{ cm}^{-1}$  for N3D is shifted to higher frequency compared to the measured band at  $\sim 3665\text{ cm}^{-1}$ , and the band predicted at  $\sim 3640\text{ cm}^{-1}$  is not observed in the experiments. Therefore, among the N3A, O2A, N3ii, O2i, O2C, and N3D conformers, only the N3A and O2A tautomeric conformations are accessed upon electrospray ionization.

The measured and the IR spectra predicted for N3i, N3B, O2B, N3iii, O2ii, and O2iii tautomeric conformations of [Cyd+H]<sup>+</sup> are overlaid in Figure S5. In the FELIX region, the C3'-endo sugar configuration of N3i, N3B, and O2B does not significantly influence the calculated IR features as compared to those of N3A and O2A (Figure 4), which exhibit C2'-endo puckering of the sugar. The change in the orientations of the hydroxyl groups of the sugar moieties of N3iii and O2ii do not influence the calculated IR features as compared to those of N3A and O2A, respectively (Figure 4). Overall, the calculated IR spectra of N3i, N3B, O2B, N3iii, and O2ii exhibit nice matches with the experimental spectrum in the FELIX region. The band predicted at  $\sim 1070\text{ cm}^{-1}$  for O2iii would tend to broaden the band at  $\sim 1115\text{ cm}^{-1}$  if O2iii was contributing to the experiments. In the OPO region, the calculated IR bands at  $\sim 3605\text{ cm}^{-1}$  for N3i and at  $\sim 3615\text{ cm}^{-1}$  for N3iii are significantly shifted relative to the band observed at  $\sim 3575\text{ cm}^{-1}$ . The IR features predicted for N3B and O2B above  $\sim 3300\text{ cm}^{-1}$  are very similar to those of N3A and O2A (Figure 4) and thus agree well with the measured bands, although theory tends to slightly overestimate the IR frequencies above  $\sim 3550\text{ cm}^{-1}$ . The IR features predicted for O2ii and O2iii at  $\sim 3620$  and  $\sim 3605\text{ cm}^{-1}$ , respectively, broaden the calculated IR band at  $\sim 3600\text{ cm}^{-1}$  and also lead to a shift to higher frequencies than the measured band at  $\sim 3575\text{ cm}^{-1}$ . Therefore, among the N3i, N3B, O2B, N3iii, O2ii, and O2iii tautomeric conformations, only N3B and O2B, which have C3'-endo sugar puckering with both hydroxyls rotated up and away from the nucleobase, may be populated in the experiments. B3LYP predicts that N3B and O2B lie higher in energy than N3i, but the comparison between the measured and predicted spectra eliminates the presence of N3i. This suggests that either B3LYP overestimates the stability of N3i or that because the IR spectra predicted for N3B and O2B are virtually identical to those of N3A and O2A (Figure 4), N3B and O2B may not be present.

The measured and IR spectra predicted for the O2iv, N3C, O2D, N3E, O2E, and O2F conformers of [Cyd+H]<sup>+</sup> are compared in Figure S6. In the FELIX region, the calculated band at  $\sim 1425\text{ cm}^{-1}$  for O2iv is not observed in the experimental spectrum. The orientation of the 5'-hydroxymethyl substituent and its interaction with the 3'-hydroxyl moiety in N3C and O2iv produce the calculated bands at  $\sim 1050\text{ cm}^{-1}$  for both conformers, which would broaden the band measured at  $\sim 1115\text{ cm}^{-1}$  if these conformers were populated upon electrospray ionization. The IR spectra predicted for N3E, O2E, and O2F agree nicely with the experimental spectrum in the FELIX region. However, obvious differences between the experimental and predicted spectra are observed for these six conformers in the OPO region. Therefore, all six of these conformers are not accessed upon electrospray ionization.

In summary, comparison between the IRMPD spectrum and the calculated IR spectra indicates that both N3 and O2 protonated tautomeric conformations coexist in the experimental population. Both N3 and O2 protonation exhibit a preference for the nucleobase to be in the anti orientation with C2'-endo puckering of the ribose moiety and with both hydroxyls oriented down and away from the cytosine residue. The B3LYP and the MP2 ground conformers, N3A and O2A, which adopt C2'-endo puckering of the sugar with both hydroxyl groups rotated down and away from the cytosine residue, clearly dominate the experimental population. In addition, C3'-endo puckered N3B and O2B with both hydroxyls rotated up and away from the cytosine residue may be accessed upon electrospray ionization as well.

**Vibrational Assignments of [Cyd+H]<sup>+</sup>.** Comparison of the experimental IRMPD spectrum with the IR spectra of the N3A, O2A, N3B, and O2B conformers enables vibrational assignments to be made. Given the very similar spectral features observed in the IRMPD spectra of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup>, it should not come as much of a surprise that the vibrational assignments are also highly parallel. In the FELIX region, the carbonyl stretch is observed at  $\sim 1800\text{ cm}^{-1}$ , corresponding to an  $\sim 10\text{ cm}^{-1}$  shift to higher frequencies versus that observed for [dCyd+H]<sup>+</sup>, and again establishes the presence of N3 protonated tautomeric conformations in the experiments. The strong IR absorption observed at  $\sim 1660\text{ cm}^{-1}$  and the small shoulder to the red are also shifted by  $\sim 10\text{ cm}^{-1}$  to the blue versus that observed for [dCyd+H]<sup>+</sup> and again represent NH<sub>2</sub> scissoring coupled with stretches of nucleobase ring and nucleobase-H in-plane bending. The two moderate bands observed at  $\sim 1580$  and  $\sim 1500\text{ cm}^{-1}$  are contributed by the O2 protonated conformers and represent stretches of nucleobase ring coupled with nucleobase-H in-plane bending. In this case, these bands are blue-shifted by 20 and 10  $\text{cm}^{-1}$ , respectively, relative to those observed for [dCyd+H]<sup>+</sup>. The small shoulder to the red of the band at  $\sim 1580\text{ cm}^{-1}$  is again contributed by the N3 protonated conformers and again also arises from N3-H in-plane bending coupled with stretches of nucleobase ring. The small IRMPD feature observed at  $\sim 1285\text{ cm}^{-1}$  represents N3-H, C6-H, and C1'-H in-plane bending, whereas the broad weak IR band measured at  $\sim 1220\text{ cm}^{-1}$  results from coupled bending of the sugar H atoms. These bands are again blue-shifted relative the analogous bands observed for [dCyd+H]<sup>+</sup> by 5 and 10  $\text{cm}^{-1}$ , respectively. The moderate IR feature measured at  $\sim 1115\text{ cm}^{-1}$  represents stretches of sugar ring and is blue-shifted versus that observed for [dCyd+H]<sup>+</sup> by 20  $\text{cm}^{-1}$ . In the OPO region, the IR band

measured at  $\sim 3665\text{ cm}^{-1}$  comprises coupled 3'- and 5'-hydroxyl stretches, which is again blue-shifted versus that observed for  $[\text{dCyd+H}]^+$  by  $\sim 5\text{ cm}^{-1}$ . The band observed at  $\sim 3575\text{ cm}^{-1}$  represents the O2–H stretch (from O2 protonated conformers) coupled with the O2'–H stretch (from both N3 and O2 protonated tautomeric conformations), whereas the analogous band observed at this same frequency for  $[\text{dCyd+H}]^+$  only shows contributions from O2–H stretching. The slightly smaller IR band at  $\sim 3540\text{ cm}^{-1}$  arises from asymmetric  $\text{NH}_2$  stretches and is only one of two bands that are shifted to lower frequencies versus that observed for  $[\text{dCyd+H}]^+$ , by  $\sim 5\text{ cm}^{-1}$  in this case. The sharp band at  $\sim 3440\text{ cm}^{-1}$  represents symmetric  $\text{NH}_2$  stretching and is again blue-shifted (by  $\sim 5\text{ cm}^{-1}$ ) versus that observed for  $[\text{dCyd+H}]^+$ . The shoulder on the low-frequency side of this band at  $\sim 3402\text{ cm}^{-1}$  represents the N3–H stretch and exhibits a very small red-shift (of  $\sim 3\text{ cm}^{-1}$ ) versus that observed for  $[\text{dCyd+H}]^+$ .

**IRMPD Yields of  $[\text{dCyd+H}]^+$  versus  $[\text{Cyd+H}]^+$ .** As discussed above, the IRMPD spectra of  $[\text{dCyd+H}]^+$  and  $[\text{Cyd+H}]^+$  exhibit a high degree of similarity and only differ by minor shifts ( $\sim 3\text{--}20\text{ cm}^{-1}$ ) in the frequency of the measured bands and relative IRMPD yields. In particular, the bands observed at  $\sim 1650\text{--}1660\text{ cm}^{-1}$  arising from stretches of the nucleobase exhibit greater IRMPD yield for  $[\text{Cyd+H}]^+$  than  $[\text{dCyd+H}]^+$ . Hydrogen-bond stabilization arising from interaction of the 2'- and 3'-hydroxyl substituents of the ribose moiety of  $[\text{Cyd+H}]^+$  leads to the protonated nucleobase of  $[\text{Cyd+H}]^+$  having greater conformational freedom such that IVR is more efficient for  $[\text{Cyd+H}]^+$ . In contrast, the 3'-hydroxyl of  $[\text{dCyd+H}]^+$  is not restricted by such hydrogen bonding. Therefore, the IRMPD yield of  $[\text{dCyd+H}]^+$  above  $\sim 3300\text{ cm}^{-1}$  is greater than that of  $[\text{Cyd+H}]^+$ . The much higher yield of the band at  $\sim 3535\text{--}3540\text{ cm}^{-1}$  (asymmetric  $\text{NH}_2$  stretching) relative to that of the band at  $\sim 3575\text{ cm}^{-1}$  (O2–H and O2'–H stretching) may indicate that N3 protonation is slightly more favorable than O2 protonation for  $[\text{dCyd+H}]^+$ , whereas the opposite behavior observed for  $[\text{Cyd+H}]^+$  may indicate that O2 protonation is slightly more favorable than N3 protonation for  $[\text{Cyd+H}]^+$ , consistent with the computed relative stabilities.

**Experimental IRMPD versus Maxwell–Boltzmann Weighted IR Spectra.** Based on comparison between the IRMPD and calculated IR spectra for the low-energy tautomeric conformations of  $[\text{dCyd+H}]^+$  and  $[\text{Cyd+H}]^+$ , we deduce that the N3A and O2A conformers of  $[\text{dCyd+H}]^+$ , and the N3A, O2A, N3B, and O2B conformers of  $[\text{Cyd+H}]^+$ , are populated upon electrospray ionization. However, the free energies of these low-energy tautomeric conformations differ somewhat depending upon the level of theory at which they are computed, B3LYP or MP2. In particular, B3LYP favors N3, whereas MP2 favors O2 protonation. We compared the IRMPD spectra of  $[\text{dCyd+H}]^+$  and  $[\text{Cyd+H}]^+$  to the linear IR spectra predicted based on the B3LYP and MP2 Maxwell–Boltzmann average of the conformers populated by electrospray ionization at 298 K to determine which theory more accurately predicts the relative stabilities of the low-energy tautomeric conformations of  $[\text{dCyd+H}]^+$  and  $[\text{Cyd+H}]^+$ . Results for  $[\text{dCyd+H}]^+$  and  $[\text{Cyd+H}]^+$  are shown in Figures S7 and S8, respectively. As can be surmised from the figures, the B3LYP and MP2 Maxwell–Boltzmann weighted spectra both exhibit reasonable agreement with the experimental IRMPD spectra. The measured band positions are reasonably well reproduced by theory excluding the band observed at  $\sim 3575\text{ cm}^{-1}$ , where theory predicts this band to occur  $\sim 20$

$\text{cm}^{-1}$  higher in frequency for both protonated nucleosides. Because this band is primarily associated with O2–H stretching of the O2 protonated conformers, this discrepancy is likely the result of the anharmonicity of this mode due to the interaction between the O2 proton and adjacent N3 atom. In particular, for both  $[\text{dCyd+H}]^+$  and  $[\text{Cyd+H}]^+$ , the IR features arising from the protonated carbonyl observed in the region of  $\sim 1450$  to  $\sim 1600\text{ cm}^{-1}$  are better reproduced by the MP2 level of theory than B3LYP, suggesting that the O2 protonated conformers are more stable than predicted by B3LYP. Conversely, the band observed at  $\sim 3575\text{ cm}^{-1}$  and representing the O2–H stretch is overemphasized by MP2 and suggests that the O2 protonated conformers are less stable than predicted by MP2. To provide an independent assessment of the relative stabilities of the conformers accessed upon electrospray ionization, least-squares fitting of the IRMPD spectra based on the IR spectra of the most stable N3 and O2 protonated tautomeric conformations, N3A and O2A, was performed and finds that population ratios of  $55 \pm 2\%:45 \pm 2\%$  and  $47 \pm 2\%:53 \pm 2\%$  for  $[\text{dCyd+H}]^+$  and  $[\text{Cyd+H}]^+$ , respectively, best reproduce the measured spectra. These least-squares fits are compared to the experimental data in Figures S9 and S10. When the N3B and O2B conformers are included in the least-squares fitting, the solutions are not as stable, and many local minima are found that exhibit varying populations of these conformers. However, all local minima find that the total population of the N3 conformers (N3A + N3B) is  $45 \pm 4\%$  and that of the O2 protonated conformers (O2A + O2B) is  $55 \pm 4\%$ . The residuals associated with the various local minima suggest that N3A is present in larger abundance than N3B and that O2A is present in greater abundance than O2B. These results suggest that N3 protonation is slightly preferred for  $[\text{dCyd+H}]^+$ , whereas O2 protonation is slightly more favorable for  $[\text{Cyd+H}]^+$ , indicating that the 2'-hydroxyl substituent preferentially stabilizes the O2 protonated conformers. These populations suggest that N3A is  $\sim 0.5\text{ kJ/mol}$  more stable than O2A for  $[\text{dCyd+H}]^+$ , whereas O2A is  $\sim 0.3\text{ kJ/mol}$  more favorable than N3A for  $[\text{Cyd+H}]^+$ . Overall, these results are consistent with the stability trends for both the B3LYP and MP2 theories, whereas the absolute values are most consistent with the average of the MP2 and B3LYP relative stabilities. Interestingly, the N3A and O2A conformers of  $[\text{dCyd+H}]^+$  and of  $[\text{Cyd+H}]^+$ , as well as the N3B and O2B conformers of  $[\text{Cyd+H}]^+$ , differ only in the position of the excess proton. We calculated the barrier for tautomerization (i.e., proton transfer between N3 and O2) of these conformers and find that  $\sim 150\text{ kJ/mol}$  is needed, indicating that interconversion between these N3 and O2 protonated conformers is not possible under our experimental conditions.

**Comparison to Condensed-Phase Structures.** NMR spectroscopy<sup>59,60</sup> and crystallography<sup>61,62</sup> are useful techniques for characterization of the structures of nucleic acids in the condensed phase. However, investigations of the canonical DNA and RNA nucleosides in the condensed phase are limited. Studies of the structures of neutral dCyd and Cyd in the condensed phase have been reported<sup>63–65</sup> and have demonstrated that both dCyd and Cyd prefer the nucleobase in the anti orientation and C3'-endo sugar puckering. Roberts and co-workers<sup>30</sup> determined that N3 is the preferred site of protonation for Cyd using  $^{15}\text{N}$  NMR spectroscopy. However, other structural details such as the nucleobase orientation and sugar puckering were not reported in their work. We also performed quantum chemical calculations for neutral dCyd and Cyd. The optimized low-energy tautomeric conformations of

dCyd and Cyd and their relative 298 K free energies are shown in Figure S11 of the Supporting Information. The relative enthalpies and free energies at 0 and 298 K of these low-energy tautomeric conformations are listed in Table S1. The **neuD** conformer of dCyd and the **neuE** conformer of Cyd calculated here are similar to the ground conformers observed for dCyd<sup>64</sup> and Cyd<sup>63</sup> in the condensed phase, suggesting that the anti orientation of dCyd and Cyd may be stabilized by interactions of N4–H, O2, and N3 with the local environment in the condensed phase. In contrast, in the gas-phase ground conformer of dCyd, cytosine rotates into a syn orientation to form a hydrogen bond between the 5'-hydroxyl hydrogen and O2 atoms, and the sugar is C2'-endo puckered. The 2'-hydroxyl substituent helps to stabilize Cyd such that cytosine remains in an anti orientation and forms an O3'H...O2'H...O2 dual hydrogen-bonding interaction with C2'-endo puckering. Except for the excess proton on N3, these ground conformers of neutral dCyd and Cyd resemble the excited **N3D** and **N3ii** conformers of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup>, respectively, indicating that protonation changes the preferred orientation of the nucleobase, puckering of the sugar moiety, sugar hydroxyl orientations, and the relative stabilities of the low-energy tautomeric conformations of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> vs those of dCyd and Cyd.

**Comparison to Speranza's Work.** Speranza and co-workers<sup>46</sup> compared the measured IRMPD and B3LYP/6-311++G(d,p) IR spectra predicted for the most stable N3 and O2 protonated conformers of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> in the ranges of 1400–1800 and 2800–3800 cm<sup>-1</sup>. The relative stabilities of the low-energy conformers computed in their work via *ab initio* molecular dynamics simulations predict that the most stable conformations for both [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> are N3 protonated, anti oriented, and C3'-endo puckered. The most stable O2 protonated conformations they reported for [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> are also anti oriented, but prefer C2'-endo puckering and lie 6.3 and 1.2 kJ/mol, respectively, above their N3 protonated analogues. These results differ from that found here for both B3LYP or MP2, possibly the result of the different method used for energetics in their work. They concluded that anti oriented N3 and O2 protonated low-energy conformers coexist, consistent with the present results. However, their calculations and analysis were not sufficiently detailed to specifically determine the sugar puckering and the specific conformer(s) populated in the experiments. Careful examination of the IR features measured in the region of 900–1400 cm<sup>-1</sup> in the present work, however, exhibit greater variation and enable higher energy conformers that share parallel IR features with the ground conformers above 1400 cm<sup>-1</sup> to be eliminated. In addition, the IR bands measured above 3500 cm<sup>-1</sup> enable C2'- and C3'-endo puckering to be distinguished as well as the ribose hydroxyl orientations of [Cyd+H]<sup>+</sup> to be determined. Our more detailed analyses find that only C2'-endo puckered **N3A** and **O2A** of [dCyd+H]<sup>+</sup>, and C2'-endo puckered with sugar hydroxyls oriented down and away from the nucleobase, **N3A** and **O2A**, C3'-endo puckered with sugar hydroxyls rotated up and away from the nucleobase, **N3B** and **O2B** of [Cyd+H]<sup>+</sup>, are populated upon electrospray ionization. Comparison between the experimental IRMPD and Maxwell–Boltzmann weighted vs least-squares fitted spectra of these conformers enables the accuracy of B3LYP and MP2 theories for describing the relative energetics to be evaluated. Both B3LYP and MP2 theories exhibit very modest deficiencies in predicting the relative

stabilities as B3LYP underestimates the stabilities of the O2 protonated conformers and MP2 underestimates the stabilities of the N3 protonated tautomeric conformations. Overall, the experimental results suggest that O2 and N3 protonation are nearly isoenergetic, and the least-squares analysis suggests that [dCyd+H]<sup>+</sup> prefers N3 protonation, whereas [Cyd+H]<sup>+</sup> prefers O2 protonation. Moreover, the effects of the 2'-hydroxyl substituent are elucidated from the relative IRMPD yields and conformational differences between [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup>.

## CONCLUSIONS

The IRMPD action spectra of the protonated forms of 2'-deoxycytidine, [dCyd+H]<sup>+</sup>, and cytidine, [Cyd+H]<sup>+</sup>, have been measured and compared with linear IR spectra for the stable low-energy conformers predicted at the B3LYP/6-311+G(d,p) level to determine the structures populated by electrospray ionization. Comparisons between the experimental IRMPD and predicted IR spectra enable thorough interpretation of the tautomeric conformations accessed upon electrospray ionization. For both [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup>, B3LYP predicts N3 as the preferred site of protonation, whereas MP2 favors O2. Present results suggests the coexistence of both N3 and O2 protonated low-energy tautomeric conformations of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> in the experimental population. In particular, the B3LYP N3 protonated and the MP2 O2 protonated ground-state conformers, **N3A** and **O2A**, of [dCyd+H]<sup>+</sup> are present. Both **N3A** and **O2A** exhibit C2'-endo puckering with the cytosine residue in an anti orientation. For [Cyd+H]<sup>+</sup>, the B3LYP N3 protonated and the MP2 O2 protonated ground conformers, **N3A** and **O2A** are present, whereas the **N3B** and **O2B** excited conformers cannot be excluded, but if present should exist in lower population. These four conformers all have the nucleobase in the anti orientation. When the sugar is C2'-endo puckered, as in **N3A** and **O2A**, both hydroxyl groups on the sugar are rotated down and away from the nucleobase, whereas when the sugar is C3'-endo puckered, as in **N3B** and **O2B**, both hydroxyl groups are rotated up and away from the nucleobase. Comparison between the experimental IRMPD and Maxwell–Boltzmann statistically weighted and least-squares fitted IR spectra of the conformers accessed suggests that the N3 and O2 protonated conformers exist in nearly equal populations in the experiments, with N3 slightly favored for [dCyd+H]<sup>+</sup> and O2 for [Cyd+H]<sup>+</sup>. Thus, the hydrogen-bond stabilization between the two hydroxyl groups of the ribose sugar of [Cyd+H]<sup>+</sup> leads to a slight preferential stabilization of the O2 protonated conformers.

## ASSOCIATED CONTENT

### Supporting Information

Full Gaussian 09 citation, ref 54; figures showing the B3LYP/6-311+G(d,p) low-energy tautomeric conformations of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup> and their 298 K relative Gibbs free energies determined at B3LYP/6-311+G(2d,2p) and MP2(full)/6-311+G(2d,2p) levels along with a detailed description of these species; comparisons of the experimental IRMPD and computed linear IR spectra of select low-energy tautomeric conformations of [dCyd+H]<sup>+</sup> and [Cyd+H]<sup>+</sup>; structures of the B3LYP/6-311+G(d,p) low-energy tautomeric conformations of neutral dCyd and Cyd and their relative 298 K free energies computed at the B3LYP/6-311+G(2d,2p) and MP2(full)/6-311+G(2d,2p) levels of theory; comparisons of the experimental IRMPD and Maxwell–Boltzmann weighted

IR and least-squares fitted spectra calculated for the low-energy conformers of  $[d\text{Cyd}+\text{H}]^+$  and  $[\text{Cyd}+\text{H}]^+$  populated by electrospray ionization and based on the B3LYP and MP2 computed energetics. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jp5130316.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail mrodgers@chem.wayne.edu (M.T.R.).

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work is financially supported by the National Science Foundation, under Grants OISE-0730072 and CHE-1409420. R.R.W. also received financial support via a Thomas C. Rumble Graduate Fellowship from Wayne State University. We also thank WSU C&IT for the computational infrastructure and support. As part of the research program of FOM, financial support for this work was also derived from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). The FELIX staff is gratefully acknowledged for their excellent support of the facility where these measurements were made.

## REFERENCES

- (1) Saenger, W. *Principles of Nucleic Acid Structure*; Springer: New York, 1984; pp 1–556.
- (2) Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737–738.
- (3) Giglio, E. Calculation of van der Waals Interactions and Hydrogen Bonding in Crystals. *Nature* **1969**, *222*, 339–341.
- (4) Crisp, G. T.; Jiang, Y. L. Intramolecular Hydrogen Bonding of Nucleobases. *Tetrahedron Lett.* **2002**, *43*, 3157–3160.
- (5) Baglioni, P.; Berti, D. Self-Assembly in Micelles Combining Stacking and H-Bonding. *Curr. Opin. Colloid Interface Sci.* **2003**, *8*, 55–61.
- (6) Batey, R. T.; Rambo, R. P.; Doudna, J. A. Tertiary Motifs in RNA Structure and Folding. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 2326–2343.
- (7) Westhof, E.; Fritsch, V. RNA Folding: Beyond Watson–Crick Pairs. *Struct. Fold Des.* **2000**, *8*, R55–R65.
- (8) Altona, C.; Sundaralingam, M. Conformational Analysis of the Sugar Ring in Nucleosides and Nucleotides. A New Description Using the Concept of Pseudorotation. *J. Am. Chem. Soc.* **1972**, *94*, 8205–8212.
- (9) Salazar, M.; Fedoroff, O. Y.; Miller, J. M.; Ribeiro, N. S.; Reid, B. R. The DNA Strand in DNA•RNA Hybrid Duplexes is Neither B-form nor A-form in Solution. *Biochemistry* **1993**, *32*, 4207–4215.
- (10) Taylor, R.; Kennard, O. The Molecular Structures of Nucleosides and Nucleotides. 1. The Influence of Protonation on the Geometries of Nucleic-Acid Constituents. *J. Mol. Struct.* **1982**, *78*, 1–28.
- (11) Gehring, K.; Leroy, J.-L.; Gueron, M. A Tetrameric DNA Structure with Protonated Cytosine•Cytosine Base Pairs. *Nature* **1993**, *363*, 561–565.
- (12) Berger, I.; Kang, C.; Fredian, A.; Ratliff, R.; Moyzis, R.; Rich, A. Extension of the Four-Stranded Intercalated Cytosine Motif by Adenine•Adenine Base Pairing in the Crystal Structure of d-(CCCAAT). *Nat. Struct. Biol.* **1995**, *2*, 416–425.
- (13) Roberts, R. W.; Crothers, D. M. Prediction of the Stability of DNA Triplexes. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 4320–4325.
- (14) Scaria, P. V.; Shafer, R. H. Calorimetric Analysis of Triple Helices Targeted to the  $d(\text{G}_3\text{A}_4\text{G}_3)\bullet d(\text{C}_3\text{T}_4\text{C}_3)$  Duplex. *Biochemistry* **1996**, *35*, 10985–10994.
- (15) Asensio, J. L.; Lane, A. N.; Dhesi, J.; Bergqvist, S.; Brown, T. The Contribution of Cytosine Protonation to the Stability of Parallel DNA Triple Helices. *J. Mol. Biol.* **1998**, *275*, 811–822.
- (16) Armentano, D.; De Munno, G.; Di Donna, L.; Sindona, G.; Giorgi, G.; Salvini, L. Self-Assembling of Cytosine Nucleoside into Triply-Bound Dimers in Acid Media. A Comprehensive Evaluation of Proton-Bound Pyrimidine Nucleosides by Electrospray Tandem Mass Spectrometry, X-Ray Diffractometry, and Theoretical Calculations. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 268–279.
- (17) Williams, L. D.; Shaw, B. R. Protonated Base-Pairs Explain the Ambiguous Pairing Properties of O-6-Methylguanine. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 1779–1783.
- (18) Leitner, D.; Schroder, W.; Weisz, K. Direct Monitoring of Cytosine Protonation in an Intramolecular DNA Triple Helix. *J. Am. Chem. Soc.* **1998**, *120*, 7123–7124.
- (19) Powell, S. W.; Jiang, L. H.; Russu, I. M. Proton Exchange and Base Pair Opening in a DNA Triple Helix. *Biochemistry* **2001**, *40*, 11065–11072.
- (20) Yang, B.; Wu, R. R.; Berden, G.; Oomens, J.; Rodgers, M. T. Infrared Multiple Photon Dissociation Action Spectroscopy of Proton-Bound Dimers of Cytosine and Modified Cytosines: Effects of Modifications on Gas-Phase Conformations. *J. Phys. Chem. B* **2013**, *117*, 14191–14201.
- (21) Yang, B.; Wu, R. R.; Rodgers, M. T. Base-Pairing Energies of Proton-Bound Homodimers Determined by Guided Ion Beam Tandem Mass Spectrometry: Application to Cytosine and 5-Substituted Cytosines: Implications for the Stability of DNA *i*-Motif Conformations. *Anal. Chem.* **2013**, *85*, 282–290.
- (22) Yang, B.; Rodgers, M. T. Base-Pairing Energies of Proton-Bound Heterodimers of Cytosine and Modified Cytosines: Implications for the Stability of DNA *i*-Motif Conformations. *J. Am. Chem. Soc.* **2014**, *136*, 282–290.
- (23) 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.
- (24) Sowers, L. C.; Fazakerley, G. V.; Eritja, R.; Kaplan, B. E.; Goodman, M. F. Base-Pairing and Mutagenesis—Observation of a Protonated Base Pair Between 2-Aminopurine and Cytosine in an Oligonucleotide by Proton NMR. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 5434–5438.
- (25) Sowers, L. C.; Shaw, B. R.; Veigl, M. L.; Sedwick, W. D. DNA-Base Modification—Ionized Base-Pairs and Mutagenesis. *Mutat. Res.* **1987**, *177*, 201–218.
- (26) Hunter, W. N.; Brown, T.; Anand, N. N.; Kennard, O. Structure of an Adenine•Cytosine Base Pair in DNA and its Implications for Mismatch Repair. *Nature* **1986**, *320*, 552–555.
- (27) Wang, C.; Gao, H.; Gaffney, B. L.; Jones, R. A. Nitrogen-15-Labeled Oligodeoxy-nucleotides. 3. Protonation of the Adenine N1 in the A•C and A•G Mispairs of the Duplexes  $\{d[\text{CG}(^{15}\text{N}^1)\text{AGAAT-TCCCG}]\}_2$  and  $\{d[\text{CGGGAATTC}(^{15}\text{N}^1)\text{ACG}]\}_2$ . *J. Am. Chem. Soc.* **1991**, *113*, 5486–5488.
- (28) Suits, B. H. *Nuclear Quadrupole Resonance Spectroscopy, Handbook of Applied Solid State Spectroscopy*; Springer: New York, 2006; pp 65–96.
- (29) Roberts, B. W.; Lambert, J. B.; Roberts, J. D. Nitrogen-15 Magnetic Resonance Spectroscopy. VI. Pyrimidine Derivatives. *J. Am. Chem. Soc.* **1965**, *87*, 5439–5441.
- (30) Markowski, V.; Sullivan, G. R.; Roberts, J. D. Nitrogen-15 Nuclear Magnetic Resonance Spectroscopy of Some Nucleosides and Nucleotides. *J. Am. Chem. Soc.* **1977**, *99*, 714–718.
- (31) Buchner, P.; Maurer, W.; Ruterjans, H. N-15 Nuclear Magnetic Resonance Spectroscopy of N-15-Labeled Nucleotides. *J. Magn. Reson.* **1978**, *29*, 45–63.
- (32) Burgar, M. I.; Dhawan, D.; Fiat, D. O-17 and N-14 Spectroscopy of O-17-Labeled Nucleic-Acid Bases. *Org. Magn. Reson.* **1982**, *20*, 184–190.

- (33) Garcia, M. L. S.; Smith, J. A. S. N-14 and H-2 Quadrupole Double-Resonance in Salts of Cytosine and Adenine. *J. Chem. Soc., Perkin Trans. 2* **1983**, *9*, 1401–1408.
- (34) Wilson, M. S.; McCloskey, J. A. Chemical Ionization Mass-Spectrometry of Nucleosides-Mechanisms of Ion Formation and Estimations of Proton Affinity. *J. Am. Chem. Soc.* **1975**, *97*, 3436–3444.
- (35) Greco, F.; Liguori, A.; Sindona, G.; Uccella, N. Gas-Phase Proton Affinity of Deoxyribonucleosides and Related Nucleobases By Fast-Atom-Bombardment Tandem Mass-Spectrometry. *J. Am. Chem. Soc.* **1990**, *112*, 9092–9096.
- (36) McCloskey, J. A. Structural Characterization of Natural Nucleosides By Mass Spectrometry. *Acc. Chem. Res.* **1991**, *24*, 81–88.
- (37) Florian, J.; Baumruk, V.; Leszczynski, J. IR and Raman Spectra, Tautomeric Stabilities, and Scaled Quantum Mechanical Force Field of Protonated Cytosine. *J. Phys. Chem.* **1996**, *100*, 5578–5589.
- (38) Russo, N.; Toscano, M.; Grand, A.; Jolibois, F. Protonation of Thymine, Cytosine, Adenine, and Guanine DNA Nucleic Acid Bases: Theoretical Investigation Into the Framework of Density Functional Theory. *J. Comput. Chem.* **1998**, *19*, 989–1000.
- (39) Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, T. Theoretical Study of the Protonation and Deprotonation of Cytosine. Implications For the Interaction of Cytosine with Water. *J. Mol. Struct.* **2000**, *519*, 1–11.
- (40) Yang, Z.; Rodgers, M. T. Theoretical Studies of the Unimolecular and Bimolecular Tautomerization of Cytosine. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2749–2757.
- (41) Lee, J. K. Insights into Nucleic Acid Reactivity Through Gas-Phase Experimental and Computational Studies. *Int. J. Mass Spectrom.* **2005**, *240*, 261–272.
- (42) Raczynska, E. D.; Gal, J.-F.; Maria, P.-C.; Zientara, K.; Szelag, M. Application of FT-ICR-MS For the Study of Proton-Transfer Reactions Involving Biomolecules. *Anal. Bioanal. Chem.* **2007**, *389*, 1365–1380.
- (43) Salpin, J.-Y.; Guillaumont, S.; Tortajada, J.; MacAleese, L.; Lemaire, J.; Maitre, P. Infrared Spectra of Protonated Uracil, Thymine and Cytosine. *ChemPhysChem* **2007**, *8*, 2235–2244.
- (44) Liu, M.; Li, T. T.; Amegayibor, F. S.; Cardoso, D. S.; Fu, Y. L.; Lee, J. K. Gas-Phase Thermochemical Properties of Pyrimidine Nucleobases. *J. Org. Chem.* **2008**, *73*, 9283–9291.
- (45) Wincel, H. Microhydration of Protonated Nucleic Acid Bases and Protonated Nucleosides in the Gas Phase. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1900–1905.
- (46) Filippi, A.; Frascchetti, C.; Rondino, F.; Piccirillo, S.; Steinmetz, V.; Guidoni, L.; Speranza, M. Protonated Pyrimidine Nucleosides Probed by IRMPD Spectroscopy. *Int. J. Mass Spectrom.* **2013**, *354*, 54–61.
- (47) 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*, 023103-1–023103-7.
- (48) 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.
- (49) Polfer, N. C.; Oomens, J. Reaction Products in Mass Spectrometry Elucidated with Infrared Spectroscopy. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3804–3817.
- (50) Oepts, D.; van der Meer, A. F. G.; van Amersfoort, P. W. The Free-Electron-Laser User Facility FELIX. *Infrared Phys. Technol.* **1995**, *36*, 297–308.
- (51) HyperChem Computational Chemistry Software Package, Version 5.0; Hypercube, Inc., Gainesville, FL, 1997.
- (52) Wu, R. R.; Yang, B.; Berden, G.; Oomens, J.; Rodgers, M. T. Gas-Phase Conformations and Energetics of Protonated 2'-Deoxyguanosine and Guanosine: IRMPD Action Spectroscopy and Theoretical Studies. *J. Phys. Chem. B* **2014**, *118*, 14774–14784.
- (53) Wu, R. R.; Yang, B.; Berden, G.; Oomens, J.; Rodgers, M. T. Gas-Phase Conformations and Energetics of Protonated 2'-Deoxyadenosine and Adenosine: IRMPD Action Spectroscopy and Theoretical Studies. *J. Phys. Chem. B* **2015**, *119*, 2795–2805.
- (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.; et al. *Gaussian 09, Revision C.01*; Gaussian, Inc.: Wallingford, CT, 2009. See Supporting Information for full reference.
- (55) Oomens, J.; Sartakov, B. G.; Meijer, G.; von Helden, G. Gas-Phase Infrared Multiple Photon Dissociation Spectroscopy of Mass-Selected Molecular Ions. *Int. J. Mass Spectrom.* **2006**, *254*, 1–19.
- (56) Oomens, J.; Tielens, A. G. G. M.; Sartakov, B. G.; von Helden, G.; Meijer, G. Laboratory Infrared Spectroscopy of Cationic Polycyclic Aromatic Hydrocarbon Molecules. *Astrophys. J.* **2003**, *591*, 968–985.
- (57) Wu, R. R.; Rodgers, M. T. Work in progress.
- (58) MacAleese, L.; Maitre, P. Infrared Spectroscopy of Organometallic Ions in the Gas Phase: From Model to Real World Complexes. *Mass Spectrom. Rev.* **2007**, *26*, 583–605.
- (59) Wijmenga, S. S.; van Buuren, B. N. M. The Use of NMR Methods for Conformational Studies of Nucleic Acids. *Prog. NMR Spectrosc.* **1998**, *32*, 287–387.
- (60) Flinders, J.; Dieckmann, T. NMR Spectroscopy of Ribonucleic Acids. *Prog. NMR Spectrosc.* **2006**, *48*, 137–159.
- (61) Wimberly, B. T.; Brodersen, D. E.; Clemons, W. M.; Morgan-Warren, R. J.; Carter, A. P.; Vornrhein, C.; Hartsch, T.; Ramakrishnan, V. Structure of the 30S Ribosomal Subunit. *Nature* **2000**, *407*, 327–339.
- (62) Schuwirth, B. S.; Borovinskaya, M. A.; Hau, C. W.; Zhang, W.; Vila-Sanjurjo, A.; Holton, J. M.; Cate, J. H. D. Structures of the Bacterial Ribosome at 3.5 Angstrom Resolution. *Science* **2005**, *310*, 827–834.
- (63) Furberg, S.; Petersen, C. S.; Romming, C. A Refinement of Crystal Structure of Cytidine. *Acta Crystallogr.* **1965**, *18*, 313–320.
- (64) van Dam, L.; Ouwerkerk, N.; Brinkmann, A.; Raap, J.; Levitt, M. H. Solid-State NMR Determination of Sugar Ring Pucker in C-13-Labeled 2'-Deoxynucleosides. *Biophys. J.* **2002**, *83*, 2835–2844.
- (65) Yildirim, I.; Stern, H. A.; Kennedy, S. D.; Tubbs, J. D.; Turner, D. H. Reparameterization of RNA  $\chi$  Torsion Parameters For the AMBER Force Field and Comparison to NMR Spectra For Cytidine and Uridine. *J. Chem. Theory Comput.* **2010**, *6*, 1520–1531.