



Investigation of proton affinities and gas phase vibrational spectra of protonated nucleosides, deoxynucleosides, and their analogs



Hou U. Ung^a, Kathy T. Huynh^b, John C. Poutsma^b, Jos Oomens^{c,d}, Giel Berden^c, Thomas Hellman Morton^{a,*}

^a Department of Chemistry, University of California, Riverside, CA 92521-0403, USA

^b Department of Chemistry, College of William & Mary, Williamsburg, VA 23185-2795, USA

^c Radboud University Nijmegen, Institute for Molecules and Materials, FELIX Laboratory, Toernooiveld 7, NL-6525ED Nijmegen, The Netherlands

^d van't Hoff Institute for Molecular Sciences, University of Amsterdam, NL-1098XH, The Netherlands

ARTICLE INFO

Article history:

Received 30 April 2014

Received in revised form

18 September 2014

Accepted 19 September 2014

Available online 12 October 2014

Dedicated to Veronica Bierbaum, pathfinder in ion chemistry.

Keywords:

i-motif

IR multiple photon dissociation (IRMPD) spectroscopy

Deuterium labeling

Ion cyclotron resonance (ICR)

Kinetic method (KM)

ABSTRACT

DNA nucleobases make use of hydrogen bonding, whether in associating to form the Watson–Crick double-helix or in producing alternative structures such as the G-quadruplex or the *i*-motif. Nucleoside proton-bound dimers provide an avenue for investigating characteristics that they possess within the *i*-motif and related non-Watson–Crick conformations. In addition, several nucleosides are approved antiviral and anticancer agents. The nucleosides under investigation (2'-deoxycytidine, gemcitabine, and decitabine) are capable of forming proton-bound dimers (PBDs) with their conjugate acids. Protonated monomers of 2'-deoxycytidine, gemcitabine, and decitabine and proton-bound dimers of gemcitabine and decitabine have been produced in the gas phase using electrospray ionization (ESI). This paper reports proton affinities of the neutral nucleosides as well as their infrared multiple photon dissociation (IRMPD) spectra from 2800 to 3800 cm⁻¹ collected using an Optical Parametric Oscillator (OPO) laser. In the case of the conjugate acid of 2'-deoxycytidine, a partial deuteration experiment suppresses overtones and combination bands, leading to the inference that a single tautomer predominates in the protonated monomer. IRMPD spectra of proton-bound dimers of gemcitabine and decitabine suggest furanose sugar ring puckering to be in the South orientation, as they are in the protonated monomers.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Pairings of DNA nucleobases make use of hydrogen bonding, regardless of whether neutral nucleobases associate to create the Watson–Crick double-helix or alternative structures form, such as the G-quadruplex or the *i*-motif. DNA conformations like the latter utilize non-Watson–Crick binding when guanine-rich [1–4] and the complementary cytosine-rich [5–7] strands, respectively, separate from one another to undergo single-strand self-association.

More attention has been devoted to stabilization of the G-quadruplex than of the *i*-motif, even though both secondary structures can be expected to form complementarily to one other when the two strands separate. The *i*-motif contains proton-bound dimers (PBDs) of cytosine (also called “hemiprotonated cytosine”) intercalated with one another, oriented 90° with respect to the

PBDs above and below. Double-strand association *via* cytosine PBDs requires a parallel orientation [8], but if a single strand bends into a hairpin three times, intercalation among the four regions within a single strand can form within the antiparallel orientation of duplex DNA [9,10]. Hydrogen bonding between a pair of cytosines (protonation of the nitrogen of one cytosine at the 3-position leading to three hydrogen bonds with the other, neutral cytosine) provides the driving force holding the PBDs together. Ionic hydrogen bonding between cytosine and its conjugate acid has an experimental dissociation enthalpy whose experimental value ranges from 160 kJ/mol [11] to 173 kJ/mol [5,6] in the gas phase. The latter value agrees with DFT calculations [5,6].

Promoter regions of several oncogenes are believed to adopt the aforementioned secondary structures while underwound. In other words, non-Watson–Crick binding may make a contribution in the time interval between dissociation from the nucleosome and restoration of the correct linkage number by topoisomerases. If this hypothesis is correct, it offers an approach for limiting the binding of promoters and suppressing transcription of potentially harmful mRNA.

* Corresponding author. Tel.: +1 951 827 4735; fax: +1 951 827 4713.
E-mail address: morton@citrus.ucr.edu (T.H. Morton).

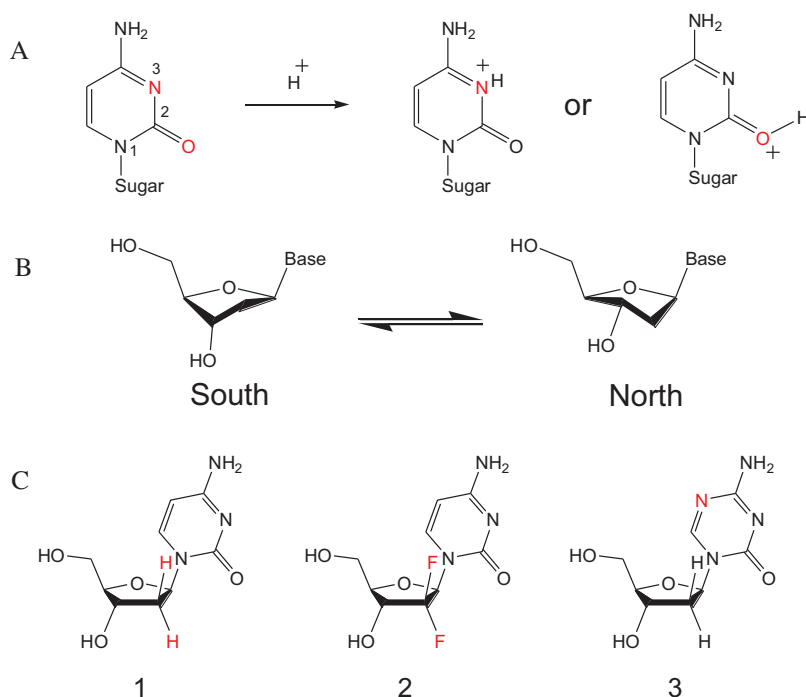


Fig. 1. (A) The two most basic protonation sites of cytosine shown in red. (B) Puckering of the deoxyribose sugar ring from the (South) to (North) orientation. The South orientation has the carbon at the 3'-position downward, while the North orientation has the carbon at the 3'-position up. (C) Nucleosides in this investigation: (1) 2'-deoxycytidine, (2) gemcitabine and (3) decitabine. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2. Background

NMR and gas phase vibrational studies of cytosine PBDs and their derivatives [12,13] have previously been reported, but nucleoside PBDs containing cytosine bases have not received equal attention. Yang and Rodgers have previously examined IRMPD spectra of protonated cytosine monomers and proton-bound dimers modified at the 5-position (with I, Br, F, and methyl) along with the base-pairing energies of cytosine heterodimers in the gas phase using threshold collision-induced dissociation [5,6]. They found that modification at the 5-position of cytosine affects the binding energies of the cytosine proton-bound heterodimers. The proton-bound heterodimer between cytosine and 5-methylcytosine, for example, has a greater binding energy than does the proton-bound homodimer of cytosine.

NMR studies of cytosine proton-bound dimers, modified at the 1-position, have been investigated in solution by Hooley et al. [12]. They conducted variable temperature NMR studies along with acid titration experiments on 1-octylcytosine, 5-fluoro-1-methylcytosine, and 5-fluoro-1-octylcytosine homodimers. The addition of acid favors the formation of proton-bound dimers, but excess acid causes them to revert to protonated monomers. They observed proton NMR absorptions that shift downfield with decreasing pH. However, with excess acid concentration, that peak reverses direction, shifting back slightly upfield (Fig. 1).

We have previously reported the gas phase IRMPD spectra of protonated 1-methylcytosine and its PBD in the gas phase from 600 to 1800 cm^{-1} , along with solid state NMR data for the iodide salt of $(1\text{-methylcytosine})_2\text{H}^+$ [13,14]. The gas phase spectrum of the PBD in the fingerprint domain from 300 to 1800 cm^{-1} shows a distinct band at 1570 cm^{-1} that disappears upon deuteration of the exchangeable hydrogens. The same band shows up in the IRMPD spectra of PBDs of the 5-fluoro- and 1,5-dimethylcytosine homodimer analogs, as well as in the three possible heterodimers. This band is therefore assigned to be the motion associated with the

bridging proton's transit from one N3-nitrogen to the other. A band at the same frequency also appears in powder and single crystal IR spectra, and it disappears with isotopic substitution. Solid state NMR of the crystalline iodide salt of $(1\text{-methylcytosine})_2\text{H}^+$ reveals that the bridging proton is not shared equally between the two 1-methylcytosines, but prefers to reside on one cytosine or the other, rendering all of the atoms inequivalent (consistent with gas phase vibrational spectra) [13].

Looking at the NH/OH stretching domain from 3200 to 3700 cm^{-1} in the experimental vibrational spectrum of the PBD of 1-methylcytosine, four bands are observed experimentally. Partial deuteration experiments, where four of the five hydrogens are exchanged for deuterium, indicate bands at 3240 and 3310 cm^{-1} to be overtones or combination bands [13]. By contrast, the experimental spectra in the fingerprint and NH/OH domain of the protonated monomer match a mixture of N3- and O-protonated 1-methylcytosines, suggesting both tautomers coexist in the gas phase, regardless of whether the ion arises from direct injection or from dissociation of the $(1\text{-methylcytosine})_2\text{H}^+$ dimer [14].

Speranza et al. previously reported IRMPD spectra of four protonated monomeric nucleosides: 2'-deoxycytidine, cytarabine, and gemcitabine in the fingerprint and CH/NH/OH stretching domains [15]. They examined tautomer distributions and concluded that N3- and O-protonated conjugate acid ions are both present in the gas phase to comparable extents. Their data show the N3-protonated tautomer contributing prominent bands in the fingerprint domain from 1400 to 1800 cm^{-1} , while the CH/NH/OH domain from 2800 to 3800 cm^{-1} , appears to match the predicted IR spectrum of the O-protonated tautomer.

The partial deuteration experiment described below suggests that the extra peaks in the previously reported IRMPD spectra in the CH/NH/OH domain [15] may come from overtones and combination bands, thus indicating the presence of only one tautomer of protonated monomeric 2'-deoxycytidine in the gas phase.

Isotopic substitution can have two consequences: shifting the band positions of fundamental vibrations and altering anharmonic interactions, which may change the intensities of overtones and combination bands. We recently described nearly complete suppression of overtones/combination bands from partial deuteration of the proton-bound dimer of 1-methylcytosine [13]. A recent report provides a mechanism – reduced intensity borrowing as a consequence of the larger difference in frequency – by which deuteration of the NH positions in G•C nucleoside pairs in solution diminishes the intensity of an overtone of the C=O stretch [16].

Comparable effects are seen in the gas phase. Dellepiane and Overend [17] have assigned the observed band at 3460 cm⁻¹ in the IR absorption spectrum of acetone and deuterated analogs to a C=O overtone and reproduce the spectra at different pressures. Based on the pressures they report, measuring the intensity of the overtone peak from undeuterated acetone gives an intensity 2.5% that of the C=O fundamental (and nearly half the intensity of the C–H stretch fundamentals). Assuming the reported pressures to be accurate, measurement of the corresponding overtone peak from acetone-*d*₃ (which appears to have the same frequency and bandwidth as does undeuterated acetone) gives an intensity for the C=O overtone of CD₃COCH₃ of only 1.5% of the C=O fundamental, though the effect of partial deuteration of acetone may not have a mechanism identical to that described for the G•C nucleoside pair in solution [16].

Here we report proton affinities (PA values) and vibrational spectra of selected nucleosides containing the cytosine moiety or an analog. The nucleosides under investigation (the anticancer drugs gemcitabine and decitabine) all have the capacity to hydrogen bond with their conjugate acid ions, analogous to the behavior of cytosine to form PBDs within the *i*-motif. Although there is no evidence to support PBDs of the aforementioned nucleosides occurring naturally, Infrared Multiple Photon Dissociation (IRMPD) spectroscopy provides a way to probe the PBDs as well as the protonated monomers.

3. Materials and methods

Commercially available 2'-deoxycytidine, cytidine, gemcitabine, and decitabine were used without further purification. All kinetic method (KM) experiments were performed in a Thermo TSQ Quantum Discovery triple quadrupole instrument. Full experimental details have been presented elsewhere [16]. Briefly, dilute solutions (*ca.* 1–5 × 10⁻⁴ M) of an analyte and one of a series of reference bases in slightly acidified (1% acetic acid) 50:50 methanol:water are directly infused (flow rates 5–15 μl/min) into the electrospray ionization source of the TSQ. Electrospray and ion focusing conditions were also varied to maximize the ion count for the proton-bound heterodimer [A–H⁺–B_{*i*}]⁺. These ions are isolated in Q1 at a resolution of 0.7 amu and are allowed to pass into the rf-only collision cell (q2). The isolated ions are allowed to undergo collision-induced dissociation with argon gas maintained at a pressure of 0.3 mTorr. Product ion spectra are recorded at collision energies between 0 and 30 V (lab). The ion intensities of each primary product and any secondary products are recorded and are analyzed using standard extended kinetic method (KM) techniques [17–19]. Secondary product ion intensities are added to the corresponding primary product intensities before undergoing KM analysis. Experiments are repeated on at least three days and are averaged to give the final ratios ln[B_{*i*}H⁺/AH⁺] for use in the KM workup. Data are given in the Supplementary Information Tables and in Figure S-15.

Proton affinities (PAs) are obtained from the extended kinetic method that has been described in detail elsewhere [18–21]. This method requires a plot of ln(I_{[B_{*i*}H⁺]/I_[AH⁺]) vs. PA_{B_{*i*}} – PA_{avg}, where I_[B_{*i*}H⁺] and I_[AH⁺] are the intensities of the protonated reference base}

and analyte products, PA_{B_{*i*}} is the proton affinity of the *i*th reference base, and PA_{avg} is the average proton affinity of the set of *i* reference acids. The Orthogonal Distance Regression (ODR) method, as implemented in the ODR-pack program of Ervin and co-workers, is used to extract PAs from the data [22]. This method gives a more realistic estimation of the errors in the derived proton affinities by using Monte Carlo simulations to determine isothermal points from randomly perturbed intensity ratios. For these studies, we used a window of ±8 kJ/mol in the reference proton affinity values and a window of ±0.05 for the ln(ratio) values. PA values are reported with error bars corresponding to ± one standard deviation, as determined from the Monte Carlo simulations, except when the simulations generated an uncertainty lower than the 8 kJ/mol uncertainty in the reference compounds. In these cases, we adopt ±8 kJ/mol as the uncertainty in our derived PAs. The ODR workup also generates effective temperature values for each activation voltage. All kinetic method values tabulated in this manuscript are generated using the ODR-derived effective temperatures rather than those obtained using the traditional best-fit method.

The use of IRMPD spectroscopy has also been described elsewhere [23,24]. Briefly, IRMPD spectra were obtained as “action spectra”, in which the laser-induced fragment ion intensities (as well as parent ion abundances) are plotted against the scanned laser frequency. IRMPD requires resonant absorption of the first photon but quasi-resonant absorption of subsequent photons, leading to ion dissociation. Ions were generated using an electrospray ionization (ESI) source, from a 50:50 MeOH/H₂O mixture containing 1% acetic acid. The gaseous cations were isolated into a Fourier Transform Ion Cyclotron (FT-ICR), where ions of a specific *m/z* value were isolated using SWIFT (stored waveform inverse Fourier Transform) ejection of unwanted ions. After *m/z* selection the ions were irradiated using a tunable bench-top optical parametric oscillator (OPO) laser capable of obtaining IRMPD spectra from 2600 to 3800 cm⁻¹, ideal for studying stretching vibrations associated with N–H and O–H bonds. The bandwidth of the laser amounts to about 3 cm⁻¹ and pulse energies reach up to 20 mJ toward the blue end of the scan range and gradually fall off to values around 7 mJ near 2600 cm⁻¹. The frequency step size was 5 cm⁻¹. IRMPD spectra are corrected for laser intensity assuming a linear dependence of the fragmentation rate on laser pulse energy. The *y*-axes of the experimental IRMPD spectra are fractional dissociation, *i.e.* Σ(fragment ion intensities)/[parent ion intensity + Σ(fragment ion intensities)]. The *y*-axes of predicted, scaled harmonic calculations are arbitrary, chosen to give visual clarity, although relative peak heights correspond to intensities calculated by Gaussian.

4. Theory/calculation

Geometry optimizations and harmonic vibrational frequencies were computed using the Gaussian 09 program suite at the B3LYP/6-311++G** level for both protonated monomers and proton-bound dimers (PBDs). Conformations of *N*3- and *O*-protonated analogs of gemcitabine and decitabine were also calculated at this level. Ring up puckering of the furanose ring at the 2'-(South) and 3'-position (North) of protonated gemcitabine and decitabine were calculated and their relative stabilities computed (see Supporting Information Figure S3 and Figure S9). Normal mode frequencies above 800 cm⁻¹ were scaled by 0.961.

5. Results

Proton affinities of modified nucleobases 1-methylcytosine and 5-fluoro-1-methylcytosine, along with deoxynucleosides 2'-deoxycytidine, gemcitabine, decitabine, and the nucleoside

Table 1

PA values of modified nucleobases and nucleosides from kinetic method experiments.

	Proton affinity (kJ/mol)	+/-
Nucleobase		
1-Methylcytosine	971	8
5-Fluoro-1-methylcytosine	961	8
Nucleoside		
2'-Deoxycytidine	1011	9
Cytidine	1001	16
Gemcitabine	989	21
Decitabine	969	8

cytidine (Table 1) were determined using the kinetic method as described in the Section 3.

2'-Deoxycytidine was found to have the highest proton affinity among the nucleobases/nucleosides/deoxynucleosides investigated. Illustrative ratios of protonated monomers upon collisional dissociation from various proton-bound nucleobase/nucleoside/deoxynucleoside heterodimers in the gas phase are tabulated in the Supporting Information (Table S1). The base having the higher proton affinity uniformly gave the higher proportion of protonated monomer upon dissociation of the corresponding proton-bound heterodimer.

To evaluate the various tautomers and conformations of protonated nucleosides by IRMPD spectroscopy, we compare the experimental IRMPD spectrum with the scaled normal mode frequencies predicted by DFT calculations using the Gaussian 09 suite. A Gaussian line shape with broadening of 10 cm^{-1} is applied to the theoretical spectra. The nucleosides of interest contain more than one protonation site, with the two principal basic atoms being the ring nitrogen (*N3*) and the carbonyl oxygen (*O*). Furanose ring puckering North and South are explored at B3LYP/6-311++G** level, and each of their predicted IR spectra are compared to the experimentally obtained spectra.

Before discussing IRMPD results, we must first mention the possibility of overtones and combination bands that could potentially be present in IRMPD spectra. Overtones represent multiple quantum transitions that are close to an integer multiple of a fundamental frequency. Combination bands result from simultaneous excitation of two fundamental bands. Overtones and combination bands are both potential issues when investigating the IR absorption profile in the CH/NH/OH stretch domain. To address this challenge, we employ the same partial deuteration methodology as previously discussed [13]. By dissolving the molecule of interest in a deuterated solvent (D_2O or MeOD), swapping the exchangeable hydrogens for deuterium need not replace all of the exchangeable Hs with Ds, for doing so would completely eliminate the NH/OH stretches. An advantage of using FT-ICR is that an ion of a specific m/z can be isolated and then be subsequently irradiated to obtain an experimental IRMPD spectrum. Partial deuteration experiments were done on the conjugate acid of 2'-deoxycytidine, where four of the five exchangeable protons are exchanged for deuterium. We believe that this procedure serves to suppress overtones and combination bands [13,16,17].

The pure d_4 ion should initially yield only m/z 114 and 115 upon dissociation, if the itinerant hydrogen comes from carbon and no randomization of hydrogens takes place, but the IRMPD experiment shows higher and lower mass ions, as well. The m/z 113 IRMPD fragment contributes a very low intensity and arises from the d_3 -ion containing one ^{13}C (which has the same nominal mass as the d_4 -ion). The m/z 114, 115, and 116 fragments contribute roughly equal intensity; m/z 116 comes from hydrogen exchange of the nascent fragment ion with the expelled neutral sugar in the course of dissociation.

Comparison of our experimental IRMPD spectrum of 2'-deoxycytidine in the CH/NH/OH domain from 2800 to 3800 cm^{-1} with that previously published by Speranza et al. [15] shows great similarity. Our IRMPD spectra show at least 4 distinct bands at 3430, 3540, 3570, and 3660 cm^{-1} and a small shoulder around 3400 cm^{-1} , as do theirs, but with different relative intensities. Partially deuteration the exchangeable position reduces the number of observed bands.

As previously mentioned, partial deuteration experiments help deconvolve the spectrum by suppressing combination bands and overtones. The partial deuteration experiment produces five d_4 -isotopomers, as Fig. 2A depicts. The IRMPD spectrum in Fig. 2D shows only three bands. We interpret the decrease in the number of observed bands as signifying the presence of only one tautomer in the gas phase. Should this prove correct, we suggest that only the *N3*-protonated tautomer of protonated 2'-deoxycytidine occurs to any appreciable extent. At first glance, the comparison of the experimental IRMPD spectrum of protonated 2'-deoxycytidine owes its fit to the *O*-protonated tautomer because four bands (plus a barely resolved shoulder) are observed, as depicted in Fig. 2C. However, the partial deuteration experiment of protonated 2'-deoxycytidine- d_4 in Fig. 2D shows that only three fundamental bands remain. This result agrees with the predicted IR spectrum for the *N3*-protonated tautomer, since it also predicts three bands in the NH/OH domain (dashed curve), while the *O*-protonated tautomer predicts four bands (red curve). Comparison of the experimental d_0 analog with the d_4 analog of 2'-deoxycytidine (Fig. 2B) shows two experimental bands disappearing. The IRMPD spectrum in the fingerprint domain previously obtained by Speranza et al. suggests that the strong fundamental band at 1767 cm^{-1} might give rise to the overtone observed at 3540 cm^{-1} . The shoulder at 3400 cm^{-1} may be a combination band of the fundamental bands at 1633 and 1767 cm^{-1} . The band observed at 3400 cm^{-1} in Fig. 2D is believed to correspond to the band seen in Fig. 2C at 3430 cm^{-1} , shifted due to deuteration (although contrary to the prediction of a harmonic DFT calculation). An anharmonic calculation could provide a rationale for this shift (or, alternatively, falsify our interpretation).

Based on the results for protonated 2'-deoxycytidine, we anticipate the same circumstance for gemcitabine. Gemcitabine, like 2'-deoxycytidine, has two principal protonation sites. Protonation at the *N3*-position of the aromatic ring with the sugar carbon puckering up at the 2'-position (South) was calculated to be the most stable (see Supporting Information Figure S3). The second most stable conformation is *N3*-protonation but having the sugar carbon puckering up at the 3'-position. Protonation at the carbonyl oxygen is possible, but is calculated to be less stable than either *N3*-protonated analog by about 7 kJ mol^{-1} . The least favorable protonation position is the carbonyl oxygen, but with the ring twisted to form an intramolecular ionic hydrogen bond with the oxygen of the alcohol at the 5'-position of the sugar.

Although partial deuteration experiments were not performed on gemcitabine, we conjecture the same overtones and combination bands may occur as in the case of 2'-deoxycytidine. Speranza et al. report the same fundamental bands at 1620 and 1765 cm^{-1} . Calculation predicts four fundamental bands in the NH/OH domain, two bands corresponding to the N–H symmetric stretch and N–H asymmetric stretch, and two bands corresponding to the O–H stretches, but seven bands are observed experimentally. Fig. 3A depicts the *N3*-protonated gemcitabine in its South orientation, which best fits the experimental IR spectrum.

A combination band involving the 1620 and 1765 cm^{-1} fundamental bands could result in the small shoulder observed near 3400 cm^{-1} . Two additional bands are seen at 3540 and 3560 cm^{-1} , while only one fundamental band is predicted in that domain. If our interpretation of the IRMPD spectrum of protonated 2'-deoxycytidine is correct, one would similarly expect corresponding

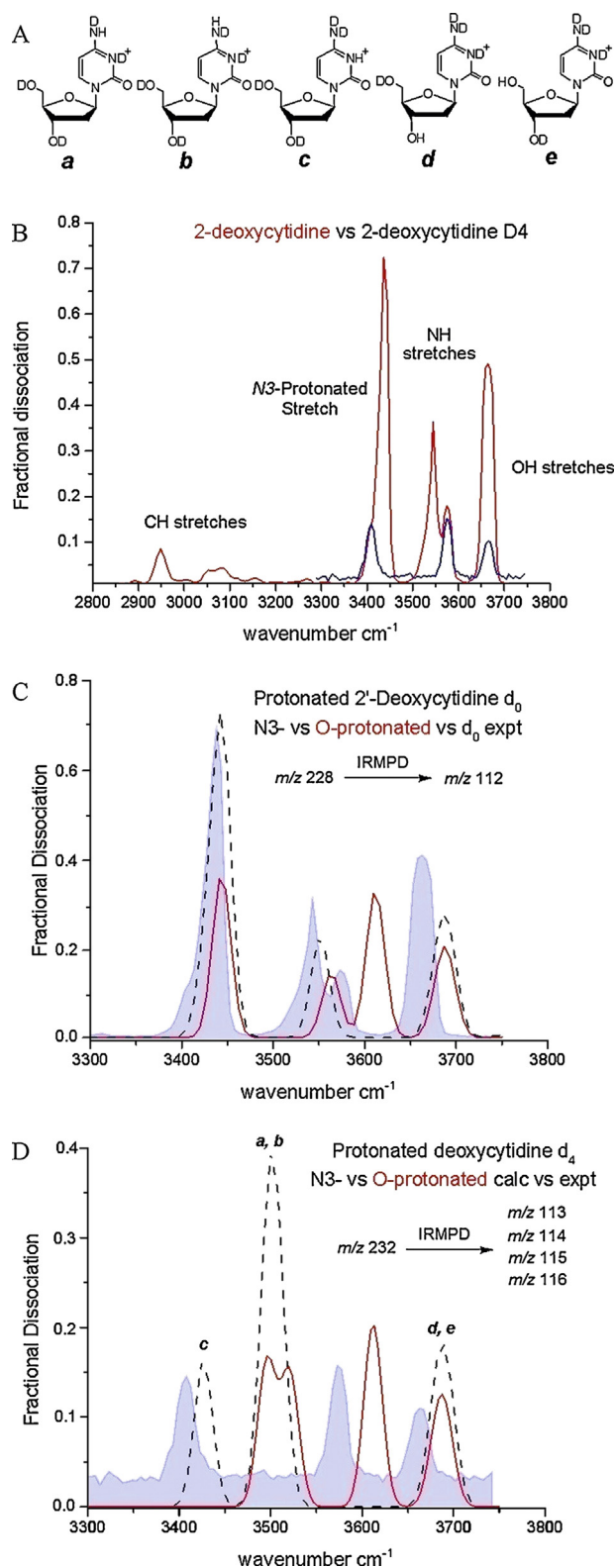


Fig. 2. (A) Five isomers of 2'-deoxycytidine, where four of the five exchangeable hydrogens are replaced with deuterium. (B) Experimental IRMPD spectrum of protonated 2'-deoxycytidine d_0 from 2800 to 3800 cm^{-1} (red) and the IRMPD spectrum of its d_4 analog (blue). (C) Comparison of the experimental d_0 spectrum of protonated 2'-deoxycytidine (blue) plotted against the calculated N3-protonated (dashed black) and O-protonated (red) spectra. (D) Experimental IRMPD of protonated 2'-deoxycytidine- d_4 (blue) plotted against predicted NH and OH stretches of the N3-protonated (dashed black) and O-protonated tautomer (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

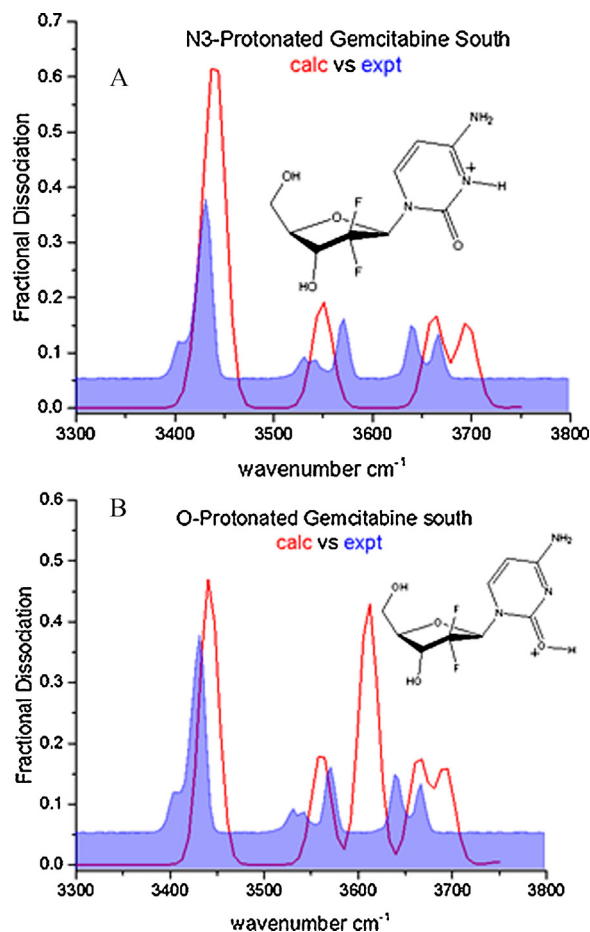


Fig. 3. (A) Experimental IRMPD spectrum of N3-protonated gemcitabine monomer (blue) plotted against theory calculated at B3LYP/6-311++G** (red). Sugar ring puckering at the 2'-position (South) was applied. (B) Experimental IRMPD spectrum of carbonyl-protonated gemcitabine monomer (blue) plotted against theory calculated at B3LYP/6-311++G** (red), also with ring puckering at the 2'-position (South). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

bands in monoprotonated gemcitabine, one an overtone of the fundamental at 1765 cm^{-1} and the other the aforementioned combination band. The N–H symmetric stretch is predicted at 3440 cm^{-1} and observed experimentally at 3430 cm^{-1} . The N–H asymmetric stretch is calculated to appear at 3550 cm^{-1} and seen experimentally at 3570 cm^{-1} . The O–H stretches are predicted to be separate at 3655 and 3690 cm^{-1} and are observed at 3640 and 3670 cm^{-1} .

While the experimental IRMPD spectrum best matches the calculated IR spectrum of the N3-protonated gemcitabine (South), as shown in Fig. 3A, more bands are present than can be accounted for by a single tautomer. Even allowing for an overtone and a combination band, the five remaining observed absorptions exceed the prediction for any mixture of N3-protonated tautomers. The N–H symmetric and asymmetric stretches predicted at 3440 and 3560 cm^{-1} , respectively, for the O-protonated tautomer, the protonated carbonyl oxygen stretch at 3610 cm^{-1} , and the O–H stretches at 3660 and 3690 cm^{-1} are consistent with its being a minor component. Other protonation sites of gemcitabine were also investigated (see Supporting Information Figure S4–Figure S8).

Like gemcitabine, two principal protonation sites are available for decitabine (see Supporting Information Figure S9). Protonated N3-decitabine (North) was calculated to be the most stable, followed closely by the South orientation. Protonation at the O-position can yield three different conformations, with the most

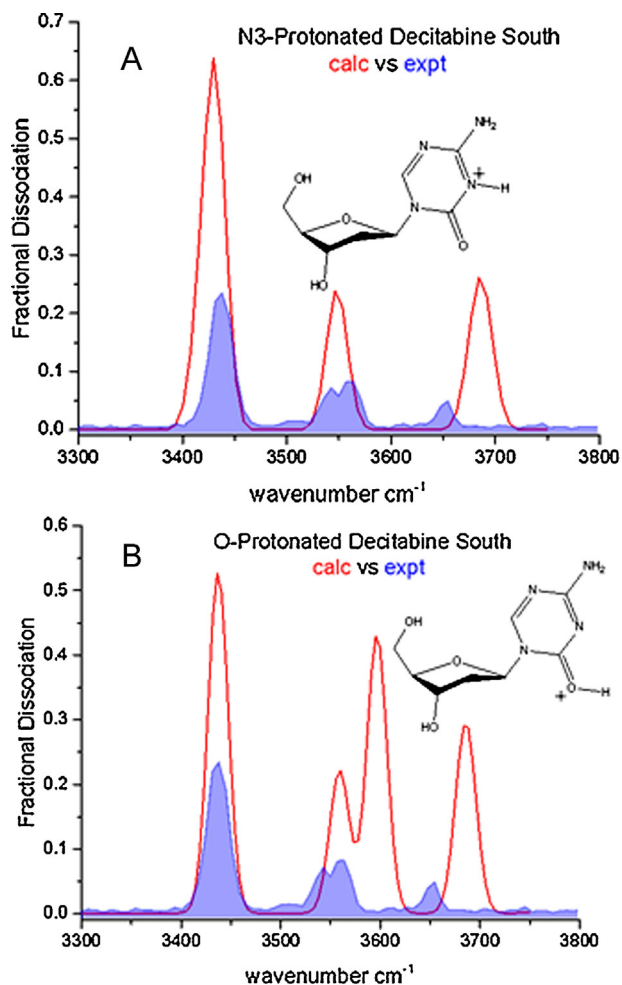


Fig. 4. (A) Experimental IRMPD spectrum of *N3*-protonated decitabine monomer (blue) plotted against theory calculated at B3LYP/6-311++G** (red). Sugar ring puckering at the 3'-position was applied. (B) Experimental IRMPD spectrum of carbonyl protonated decitabine monomer (blue) plotted against theory calculated at B3LYP/6-311++G**, also with ring puckering at the 2'-position (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

stable of the *O*-protonated tautomer oriented South, calculated to be about 12 kJ/mol less stable than the most stable *N*-protonated tautomer. The barrier from North to South conformation of the carbonyl protonated decitabine is calculated at <0.1 kJ/mol. The least stable conformation is protonation at the carbonyl position with the hydrogen facing "down" away from the sugar furanose ring and forming an intramolecular hydrogen bond with the alcohol oxygen at the 5'-position.

At first glance, the experimental IRMPD spectra of decitabine seems to fit the *O*-protonated tautomer in the South conformation best (Fig. 4B). The experimental spectrum displays four distinct bands, and three of the four bands can be fitted to harmonic calculations for *N3*-protonated South (Fig. 4A). While all four bands could be matched for the *O*-protonated South (Fig. 4B), it is conceivable that the smaller band around 3540 cm⁻¹ is an overtone, like that previously seen in 2'-deoxycytidine.

The predicted scaled harmonic IR spectra of the conjugate acids of *N3*-protonated 2'-deoxycytidine, and *N3*-protonated gemcitabine both show absorptions at 1758 and 1765 cm⁻¹, respectively, corresponding to the amide C=O stretches, and those bands are observed experimentally at 1767 and 1765 cm⁻¹, by Speranza et al. [15] The predicted IR spectrum of decitabine reveals a band at 1779 cm⁻¹ that also corresponds to the amide C=O stretch.

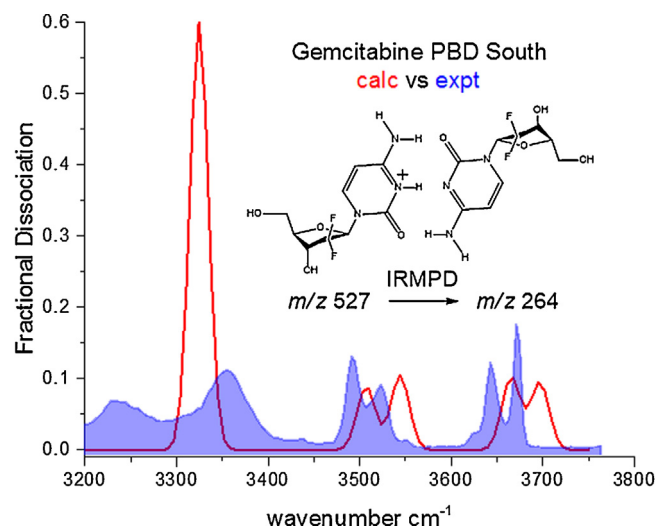


Fig. 5. Experimental IRMPD spectrum the proton-bound dimer of gemcitabine from 3300 to 3800 cm⁻¹ (blue) plotted against calculated spectrum at B3LYP/6-311++G** (red). The parent ion has *m/z* 527 and daughter ion *m/z* 264 is observed upon fragmentation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

An overtone of this band would appear around 3560 cm⁻¹. This would account for the extra band observed experimentally if *N3*-protonated decitabine is the only tautomer present.

The N–H symmetric stretch is predicted to appear at 3420 cm⁻¹ but is seen experimentally to be slightly shifted at 3440 cm⁻¹. The N–H asymmetric stretch is predicted at 3550 cm⁻¹, which places it on top of the two bands in that region. The two O–H stretches from the alcohol are predicted to cluster together at 3680 cm⁻¹, and shows up experimentally at 3650 cm⁻¹.

Fig. 4B shows the experimental IRMPD spectrum of decitabine plotted against the calculated carbonyl oxygen protonated decitabine. If the band at 3540 cm⁻¹ were a fundamental band and not an overtone (contrary to our interpretation), the carbonyl oxygen-protonated analog of decitabine would provide the best explanation for that band. The predicted N–H symmetric stretch of that tautomer lies directly on top of the experimental band at 3440 cm⁻¹. The predicted protonated carbonyl stretch lies near the domain of the two bands at 3540 and 3560 cm⁻¹. Like the *N3*-protonated analog, the predicted O–H stretches are clustered at 3690 cm⁻¹. Other protonation sites and ring puckering conformations of decitabine were also investigated (see Supporting Information Figure S10–Figure S14).

In summary, the partial deuteration experiment tends to indicate the presence of *N3*-protonated 2'-deoxycytidine, provided our assignments of an overtone and of a combination band are not in error. At the same time, the number of bands seen for the conjugate acid of gemcitabine indicates the presence of both tautomers, even assuming the occurrence of overtones and combination bands.

The foregoing analyses of protonated monomer spectra provide a prelude for presenting the IRMPD spectra in the NH/OH stretching domain of the proton-bound homodimers. The two unnatural nucleosides (gemcitabine and decitabine) were chosen for study because they possess fewer CH bonds than do natural nucleosides, simplifying assignments.

Fig. 5 depicts the comparison of the IRMPD of the PBD of gemcitabine with its predicted IR absorption spectrum having both of the furanose sugar rings in the South position. The two N–H symmetric stretches are predicted to be clustered at 3320 cm⁻¹, and is observed at 3350 cm⁻¹. The two predicted N–H asymmetric stretches are predicted at 3500 and 3540 cm⁻¹ and are seen at 3490 and 3525 cm⁻¹. Having both furanose rings in the South

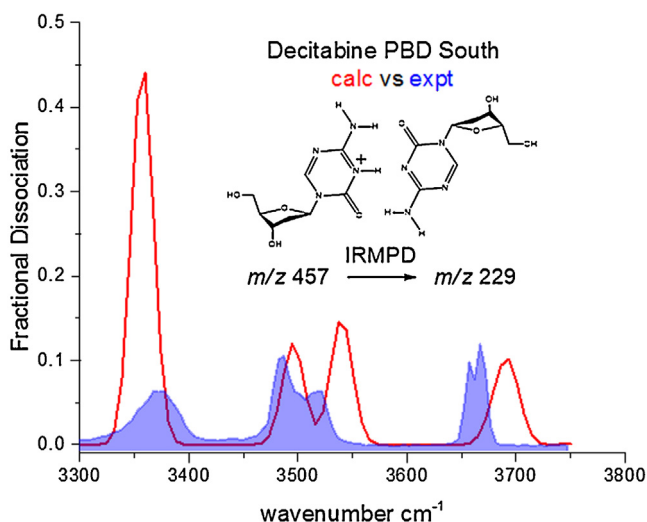


Fig. 6. Experimental IRMPD spectrum the proton-bound dimer of decitabine from 3300 to 3800 cm^{-1} (blue) plotted against calculated spectrum at B3LYP/6-311++G** (red). The parent ion has m/z 457 and daughter ion m/z 229 is observed upon fragmentation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

conformation gives the four O–H vibrations split into two bands, whereas for all other conformations they cluster under a single band. The O–H stretches predicted at 3660 and 3700 cm^{-1} are seen at 3640 and 3670 cm^{-1} .

No fundamental is predicted anywhere near the 3240 cm^{-1} band in Fig. 5. This band might be an overtone, since there is a fundamental band predicted at 1626 cm^{-1} in the gemcitabine PBD (South). Much smaller bands also occur near 3440, 3500, and 3620 cm^{-1} , which might result either from anharmonicity or from a minor isomer. If they are overtones/combination bands, then partial deuteration would be expected to remove these absorptions.

The experimental IRMPD spectrum of the PBD of decitabine is reproduced in Fig. 6. The N–H symmetric stretch is predicted at 3360 cm^{-1} , which overlaps the experimental N–H stretch at 3370 cm^{-1} . The two N–H asymmetric stretches are predicted to be separated at 3490 and 3540 cm^{-1} . Two resonances are observed experimentally near that domain, but the position of the bands is slightly shifted to lower frequencies at 3480 and 3520 cm^{-1} . The predicted band at 3700 cm^{-1} corresponds to four O–H stretches. In the calculated spectrum, the four O–H stretches are clustered so closely that they appear as one band, but in the experimental spectrum the absorptions separate into two bands.

6. Discussion

Recently Speranza et al. [15] published IRMPD spectra of the gaseous conjugate acid ions of cytidine, 2'-deoxycytidine, and gemcitabine (*inter alia*) both in the fingerprint region and the domain of CH/NH/OH stretching vibrations. The present work reproduces the IRMPD spectra of the latter two ions in the CH/NH/OH stretch domains and finds all the previously reported bands, albeit with different relative intensities. Because this form of action spectroscopy requires absorption of many photons in order to observe ion fragmentation, it is to be expected that a host of instrumental variables (laser fluence, for example) will affect the extent of absorption, such that relative intensities will vary from instrument to instrument.

A noticeable difference between our spectra and those previously published by Speranza et al. is the intensity of the sugar OH stretches of protonated gemcitabine. Also, three bands of very unequal intensity reported by Speranza et al. in the

3500–3600 cm^{-1} domain appear to have comparable intensities in Fig. 3. The band positions, however, match very well.

Having found the same bands as previously reported, the next experiments look at isotopic substitution and PBD formation, as well as extending IRMPD spectroscopy to conjugate acid ions of decitabine. The matrix-isolated 8K IR spectrum of neutral 1-methylcytosine reproduced in ref 25 shows a shoulder around 3410 cm^{-1} (nearly twice the fundamental at 1716 cm^{-1}) [25] that we assign as an overtone. We have proposed [13] that the method of partial deuteration provides a way to suppress overtones and combination bands in the 3 μm domain of gaseous ions that arise because of anharmonic couplings.

Applying this approach to the proton-bound dimer of 1-methylcytosine [13], two bands were identified as arising because of anharmonicity, since they disappear in the mixture of d_4 -isomers. As noted above, other workers have reported that partial deuteration can attenuate anharmonic effects, e.g. the difference between acetone and CD_3COCH_3 , even though the frequency and width of the C=O fundamental does not change appreciably. The IR absorption spectra reproduced in the paper by Dellepiane and Overend [17] show that the carbonyl overtone diminishes to 60% of its intensity in the partially deuterated analog, while the fully deuterated analog (CD_3COCD_3) has the same overtone intensity (relative to the fundamental) as undeuterated acetone.

A more recent study [16] attributes the result of deuteration the nitrogens in G•C nucleoside pairs (which also suppresses a C=O overtone) to intensity borrowing in the undeuterated heterodimer. Converting NH to ND stretches puts a substantial frequency difference between those fundamentals and the C=O overtone, which may be more important than physical proximity of the atoms involved. Those workers further report the intensity of the C=O overtone of acetone in CHCl_3 solution as being approximately 1% that of the fundamental, less than half the overtone intensity observed in gaseous acetone [17]. From these published results we conclude that anharmonic effects differ between solution and the gas phase, even for uncharged species, as well as differing between fully and partially deuterated samples. The determinants of overtone intensity, however, remain to be further explored.

Evidence for the presence of overtones/combination bands in the IRMPD spectra herein reported include the following. For the proton-bound dimers, little question arises as to the site of protonation, since hydrogen bonding between two molecules requires that the additional proton be situated on N3 of one of the partners. Hence, only one tautomer (albeit with different possible conformations) carries the signal. The band at 3240 cm^{-1} occurs nowhere near any of the bands predicted by harmonic DFT calculations, and all of the experimental bands are observed not far from their predicted positions. Since the 3240 cm^{-1} band has a frequency slightly less than twice that of an intense predicted fundamental, it seems plausible that the 3240 cm^{-1} corresponds to an overtone. This hypothesis could be tested by an anharmonic calculation.

The diminution of the number of IRMPD bands from conjugate acid ions of 2'-deoxycytidine from five to three as a consequence of partial deuteration admits of more than one interpretation. One possibility holds that absorptions fortuitously overlap. Another explanation holds that not all the isomers in Fig. 2A display observable bands. In our view, however, it seems most probable that at least one absorption of the undeuterated cation is an overtone, suppressed by partial deuteration, and that only a single tautomer of protonated 2'-deoxycytidine occurs to any great extent under electrospray conditions. The IR absorption frequencies calculated by harmonic DFT give a less satisfactory match to the experimental spectrum of a mixture of isomers **a–e** than does the fit for the undeuterated ion, for which we cannot yet account. But based on the number of IRMPD bands observed, we suggest that the signal carrier is the N3-protonated ion.

As noted in the foregoing paragraphs, other workers have presented a mechanism by which deuterating the exchangeable positions of a nucleobase pair in solution attenuates overtones/combination bands [16], and published gas phase spectra of acetone demonstrate that partial deuteration can have a more pronounced effect on overtone intensity than complete deuteration [17]. As mentioned above, we have previously reported reduced contribution from overtones and combination bands as a consequence of partial deuteration of the proton-bound dimer of 1-methylcytosine [13]. The overtone/combination bands that we hypothesize for protonated 2'-deoxycytidine arise from the nucleobase, and one should therefore expect the same behavior for protonated gemcitabine, which contains the same nucleobase. IRMPD spectroscopy confirms this expectation (although once again fortuitous overlap of stretching cannot be definitively ruled out). Even so, protonated gemcitabine exhibits more bands than does protonated 2'-deoxycytidine, which suggests the presence of more than one tautomer in the former (as Speranza et al. infer [15]).

Reasoning based on the number of stretching bands (assuming that all the observed bands are fundamentals) would lead to the inference that protonation of decitabine (which is the only nucleoside in Table 1 that is less basic than 1-methylcytosine) produces predominantly the *O*-protonated tautomer, even though DFT predicts the *N*-protonated tautomer to be preferred. Because decitabine contains a different nucleobase from gemcitabine or 2'-deoxycytidine, the overtone/combination bands of the protonated monomer of decitabine (Fig. 4) might not occur at the same frequencies hypothesized in Figs. 2 and 3. Alternatively, if the 3540 cm⁻¹ band from protonated decitabine were to prove to be an overtone (as we infer for the band from protonated 2'-deoxycytidine at the same frequency), then the assignment of the preferred tautomer would change. An anharmonic calculation would help in choosing the correct interpretation.

Fig. 2A depicts the five *d*₄-isomers of protonated 2'-deoxycytidine, and panel 2D in that figure compares the observed IRMPD spectrum (silhouette) with scaled harmonic calculations. In the undeuterated analog (Fig. 2B and C) four bands, as well as a shoulder, are seen, exactly as Speranza et al. have previously shown. Based on the number of bands observed in the region of NH/OH stretches, Speranza et al. [15] reasonably inferred the presence of two tautomers, *N3*-protonated and *O*-protonated. However, partial isotopic substitution (in which each of the ions in Fig. 2A ought to contribute no more than one band in the NH/OH stretch domain) diminishes the number of observed bands to only three.

The presence of overtones and combination bands in the CH/NH/OH stretching domain represents a commonly encountered problem in vibrational spectroscopy [16,17]. In the case of gemcitabine, as Speranza et al. [15] have noted, even if two bands are not fundamentals the number of absorptions in the NH/OH stretching domain indicates the presence of more than one tautomer.

One possible conclusion from the partial deuteration experiment is that the *O*-protonated tautomer must be largely absent, since (as the red curve in Fig. 2D represents) harmonic DFT calculations predict four bands for that tautomer. The positions calculated for the *N3*-tautomer (the dashed curve in Fig. 2D) do not match the experiment perfectly, but give the correct number of bands. It is worth noting that the absorption assigned as an overtone in the *d*₀ ion has twice the frequency of the very strong carbonyl stretch of the *N3*-tautomer, while the shoulder corresponds to a combination of the carbonyl stretch with the even stronger band due to HC=CH/HNH bending plus C=N/C=C stretching motions. As argued above, we believe that anharmonic interactions diminish as a consequence of partial deuteration. An alternative explanation holds that bands fortuitously overlap in the *d*₄ ion or that H/D exchange gives substantially unequal proportions of isomers

a–e (Fig. 2A). Such behavior would mask the presence of more than one tautomer, but (in our view) would require a set of fortuitous circumstances, not only in the present case, but also in our previously reported study [13] (where the presence of tautomers of the protonated 1-methylcytosine homodimer seems improbable).

The protonated monomer of decitabine displays the same 3540 cm⁻¹ band as do protonated 2'-deoxycytidine and gemcitabine. Because partial deuteration of the conjugate acid of 2'-deoxycytidine suggests that this band is an overtone of the intense carbonyl stretch of the *N3*-protonated tautomer, we surmise that the same holds for protonated decitabine. If this inference is correct, then the protonated decitabine monomer, like the protonated monomer of 2'-deoxycytidine, consists predominantly of the *N3*-protonated tautomer.

Theory and experiment demonstrate that PBDs of cytosine-derived nucleobases do not share the bridging proton equally [13]. Because of the constraints of ionic hydrogen bonding, these PBDs have to situate the proton between the two *N3* positions, as illustrated in Figs. 5 and 6. Thus, no question arises regarding the distribution of tautomers in the PBDs. Deoxyribonucleoside PBDs in addition have four sugar OH groups, but scaled harmonic calculations predict that they cannot all be resolved. In the case of the gemcitabine PBD, DFT harmonic calculations predict two broad bands separated by about 35 cm⁻¹, but the experimental spectrum in Fig. 5 shows three bands at 3625, 3645, and 3675 cm⁻¹ in the OH domain. Similarly, for the decitabine PBD DFT predicts all the OH stretches to be unresolved, but the experimental spectrum in Fig. 6 exhibits two bands at 3655 and 3670 cm⁻¹.

Apart from these failings, scaled DFT normal modes pretty much match observed vibrations for PBDs in the NH/OH domain, except for a band seen around 3240 cm⁻¹ in the gemcitabine PBD, which is not predicted as a fundamental nor observed in any of the protonated monomer spectra. Its frequency has a value almost exactly double that of an intense fundamental corresponding to the C=C/C–N stretch of the unprotonated nucleobase in the PBD combined with its H–N–H/HC=CH bends. Hence, we assign this band as an overtone. The net result of IRMPD spectroscopy in the CH/NH/OH stretch domain shows that in many cases overtones and combination bands contribute prominent experimental bands.

7. Conclusions

The proton affinities of modified nucleobases and nucleosides have been established using the Kinetic Method, as shown in Table 1. The examination of cytosine-derived nucleosides provides insight regarding the *i*-motif. IRMPD spectra of protonated monomers of 2'-deoxycytidine, gemcitabine, and decitabine and PBDs of the latter two show agreement with DFT at the B3LYP/6–311++G** level, except that overtones and combination bands appear in the CH/NH/OH stretch domain. If our interpretation of the data is correct, further partial deuteration studies have the ability to dissect fundamental vibrations away from multiple-quantum absorptions due to anharmonicity. Different conformations of the deoxyribose ring lead to small variations in the vibrational spectra, but the best fits in all cases (both protonated monomers and PBDs) are achieved with the South orientation, as drawn in Fig. 1.

The environment of cytosine PBDs within the *i*-motif shields them from solvent, suggesting that gas phase studies have relevance to their structure and behavior within DNA. In future, partial deuteration studies of proton-bound dimers of nucleosides can test the hypotheses presented here that overtones/combination bands

are present. Proton affinity and IRMPD spectroscopy therefore hold promise for further studies of non-Watson–Crick pairing.

Acknowledgment

This work was supported by the NSF, through both individual PI grants (CHE0848517 and CHE0911244) and a multi-investigator PIRE grant (0730072), and the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). HUU gratefully acknowledges a fellowship from a Department of Education Graduate Assistance in Areas of National Need (GAANN).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ijms.2014.09.017>.

References

- [1] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat. Chem.* 5 (2013) 182.
- [2] H.R. Nasiri, N.M. Bell, K.I.E. McLuckie, J. Husby, C. Abell, S. Neidle, S. Balasubramanian, *Chem. Commun.* 50 (2014) 1704.
- [3] H. Fernando, S. Sewitz, J. Darot, S. Tavare, D.L. Huppert, S. Balasubramanian, *Nucleic Acids Res.* 37 (2009) 6716.
- [4] J.L. Huppert, S. Balasubramanian, *Nucleic Acids Res.* 35 (2007) 406.
- [5] B. Yang, R.R. Wu, G. Berden, J. Oomens, M.T. Rodgers, *J. Phys. Chem. B* 117 (2013) 14191.
- [6] B. Yang, M.T. Rodgers, *J. Am. Chem. Soc.* 136 (2014) 282.
- [7] S. Kendrick, H. Kang, M.P. Alam, M.M. Madathil, P. Agrawal, V. Gokhale, D. Yang, D.S.M. Hecht, L.H. Hurley, *J. Am. Chem. Soc.* 136 (2014) 4161.
- [8] R. Langridge, A. Rich, *Nature* 198 (1963) 725.
- [9] J.L. Leroy, M. Gueron, C. Mergny, C. Helene, *Nucleic Acids Res.* 22 (1994) 1600.
- [10] J.L. Mergny, L. Lacroix, X. Han, J.L. Leroy, C. Helene, *J. Am. Chem. Soc.* 117 (1995) 8887.
- [11] M. Mautner, *Chem. Rev.* 112 (2012) PR22.
- [12] A.R. Moehlig, K.E. Djernes, V.M. Krishman, R.J. Hooley, *Org. Lett.* 14 (2012) 2560.
- [13] H.U. Ung, A.R. Moehlig, R.A. Kudla, L.J. Mueller, J. Oomens, G. Berden, T.H. Morton, *Phys. Chem. Chem. Phys.* 15 (2013) 19001.
- [14] J. Oomens, A.R. Moehlig, T.H. Morton, *J. Phys. Chem. Lett.* 1 (2010) 2891.
- [15] A. Filippi, C. Fraschetti, F. Rondino, S. Piccirillo, V. Steinmetz, L. Guidoni, M. Speranza, *Int. J. Mass Spectrom.* 354 (2013) 54.
- [16] H. Fidler, M. Yang, E.T.J. Nibbering, T. Elsaesser, K. Röttger, F. Temps, *J. Phys. Chem. A* 117 (2013) 845.
- [17] G. Dellepiane, J. Overend, *Spectrochim. Acta* 22 (1966) 593.
- [18] I. Webb, C. Muetterties, C.B. Platner, J.C. Poutsma, *Int. J. Mass Spectrom.* 126 (2012) 316.
- [19] Z. Wu, C. Fenselau, *Rapid Commun. Mass Spectrom.* 8 (1994) 777.
- [20] B.A. Cerda, C. Wesdemiotis, *J. Am. Chem. Soc.* 118 (1996) 11884.
- [21] P.B. Armentrout, *J. Am. Soc. Mass Spectrom.* 11 (2000) 371.
- [22] K.M. Ervin, P.B. Armentrout, *J. Mass Spectrom.* 39 (2004) 1004.
- [23] H.U. Ung, A.R. Moehlig, S. Khodaghlian, G. Berden, J. Oomens, T.H. Morton, *J. Phys. Chem. A* 117 (2013) 1360.
- [24] N.C. Polfer, J. Oomens, *Phys. Chem. Chem. Phys.* 9 (2007) 3804.
- [25] M. Szcześniak, M.J. Nowak, K. Szczepaniak, *J. Mol. Struct.* 115 (1984) 221.