
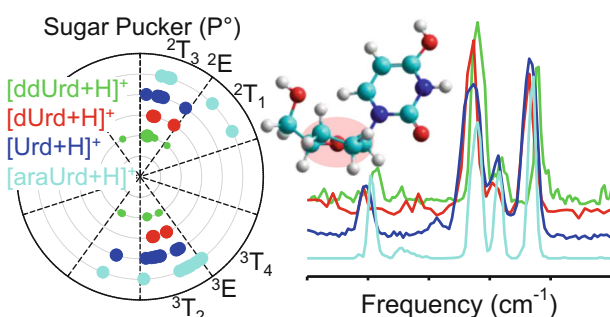


Impact of the 2'- and 3'-Sugar Hydroxyl Moieties on Gas-Phase Nucleoside Structure

L. A. Hamlow,¹ Zachary J. Devereaux,¹ H. A. Roy,¹ N. A. Cunningham,¹ G. Berden,² J. Oomens,² M. T. Rodgers¹ 

¹Department of Chemistry, Wayne State University, 5101 Cass Ave, Detroit, MI 48202, USA

²Institute for Molecules and Materials, FELIX Laboratory, Radboud University, Toernooiveld 7c, 6525ED, Nijmegen, The Netherlands



Abstract. Modified nucleosides have been an important target for pharmaceutical development for the treatment of cancer, herpes simplex virus, and the human immunodeficiency virus (HIV). Amongst these nucleoside analogues, those based on 2',3'-dideoxyribose sugars are quite common, particularly in anti-HIV applications. The gas-phase structures of several protonated 2',3'-dideoxyribose nucleosides are examined in this work and compared

with those of the analogous protonated DNA, RNA, and arabinose nucleosides to elucidate the influence of the 2'- and combined 2',3'-hydroxyl groups on intrinsic structure. Infrared multiple photon dissociation (IRMPD) action spectra are collected for the protonated 2',3'-dideoxy forms of adenosine, guanosine, cytidine, thymidine and uridine, [ddAdo+H]⁺, [ddGuo+H]⁺, [ddCyd+H]⁺, [ddThd+H]⁺, and [ddUrd+H]⁺, in the IR fingerprint and hydrogen-stretching regions. Molecular mechanics conformational searching followed by electronic structure calculations generates low-energy conformers of the protonated 2',3'-dideoxynucleosides and corresponding predicted linear IR spectra to facilitate interpretation of the measured IRMPD action spectra. These experimental IRMPD spectra and theoretical calculations indicate that the absence of the 2'- and 3'-hydroxyls largely preserves the protonation preferences of the canonical forms. The spectra and calculated structures indicate a slight preference for C3'-*endo* sugar puckering. The presence of the 3'- and further 2'-hydroxyl increases the available intramolecular hydrogen-bonding opportunities and shifts the sugar puckering modes for all nucleosides but the guanosine analogues to a slight preference for C2'-*endo* over C3'-*endo*.

Keywords: 2',3'-Dideoxyadenosine, 2',3'-Dideoxyguanosine, 2',3'-Dideoxycytidine, 2',3'-Dideoxythymidine, 2',3'-Dideoxyuridine, Infrared multiple photon dissociation action spectroscopy, IRMPD, Computational chemistry

Received: 14 January 2019/Revised: 12 February 2019/Accepted: 12 February 2019/Published Online: 8 March 2019

Introduction

Two of the most important biopolymers, DNA and RNA, are composed of mononucleotides consisting of a 2'-deoxyribose or ribose pentose sugar moiety, respectively, a nucleobase attached to the 1'-carbon, and a phosphate moiety attached to the 5'-hydroxyl moiety. The structures of the individual nucleotides can have a significant effect on the secondary structure of the nucleic acid polymer and are important to

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13361-019-02155-0>) contains supplementary material, which is available to authorized users.

Correspondence to: M. Rodgers; e-mail: mrodders@chem.wayne.edu

the phosphorylation enzymes involved in replication. For instance, the primary form of DNA, B-DNA, forms a right-handed double-helix structure stabilized by hydrogen-bonding interactions between opposing nucleobases with the sugar and phosphate moieties responsible for the backbone of the structure. B-DNA displays C2'-*endo* puckering of the sugars [1], resulting in a longer distance between phosphate groups at the 5'- and 3'-positions [2]. In contrast, A-DNA generally displays C3'-*endo* sugar puckering, which reduces the distance between neighboring phosphate groups and leads to a more compact, wider double-helix [1]. RNA, on the other hand, can adopt a wide variety of secondary structures. RNA, both single stranded and double stranded, preferentially adopts C3'-*endo* sugar puckering, though other puckering modes are accessible to facilitate other secondary structures [3, 4].

Synthetic modification of nucleosides has been a common target for pharmaceutical application [5–7]. Such modified nucleosides might take advantage of the cellular uptake and phosphorylation processes already present in the cell in order to alter cell function [8–12]. Most nucleoside analogues approved for use by the United States Food and Drug Administration (FDA) involve modification of the sugar moiety, particularly at the 2'- and/or 3'-position of the sugar, but may also involve modification of the base [13]. Two such nucleoside analogue drugs take advantage of stereochemical inversion at the 2'-position to form the arabinose-based nucleosides (arabino-sides), adenine arabinoside (araAdo, Vidarabine) [14], and cytosine arabinoside (araCyd, Cytarabine) [15], whereas several nucleoside analogues used as reverse transcriptase inhibitors targeting HIV have 2',3'-deoxyribose sugars [13]. Structures of the RNA (Nuo), DNA (dNuo), arabinoside (araNuo), and 2',3'-dideoxynucleosides (ddNuo) are shown in Figure 1. Zidovudine (ZDV or azidothymidine, AZT or 3'-azido-2',3'-dideoxythymidine) is the most well known of the 2',3'-dideoxyribose analogues. Zalcitabine (ddCyd) [16] and didanosine (ddIno) [9] are the 2',3'-deoxyribose-based nucleoside analogues of cytosine and inosine, respectively. Extensive study on the conformations of several 2',3'-dideoxynucleosides by crystallography and NMR has sought to understand the effect of conformation on pharmaceutical activity [17–21]. Despite changes to the local environment, sugar puckering, namely the observation of N-type (C2'-*exo*/C3'-*endo*) sugar pucker by NMR, appears to be conserved in the gas phase.

The intrinsic conformations adopted by the protonated gas-phase DNA and RNA nucleosides have been studied previously [22–29], allowing for isolation of conformational preference of these nucleosides in the absence of intermolecular interactions such as crystal packing forces or solvation environment. A recent study of the protonated gas-phase arabinosides of adenine, guanine, cytosine, and uracil has also examined the impact of stereochemical inversion at the 2'-position of the sugar moiety on the intrinsic conformational preferences [30]. In this work, the study of the impact of the 2'- and 3'-hydroxyls on intrinsic conformational preference is extended to include the 2',3'-dideoxynucleosides of adenine, guanine, cytosine, thymine, and uracil.

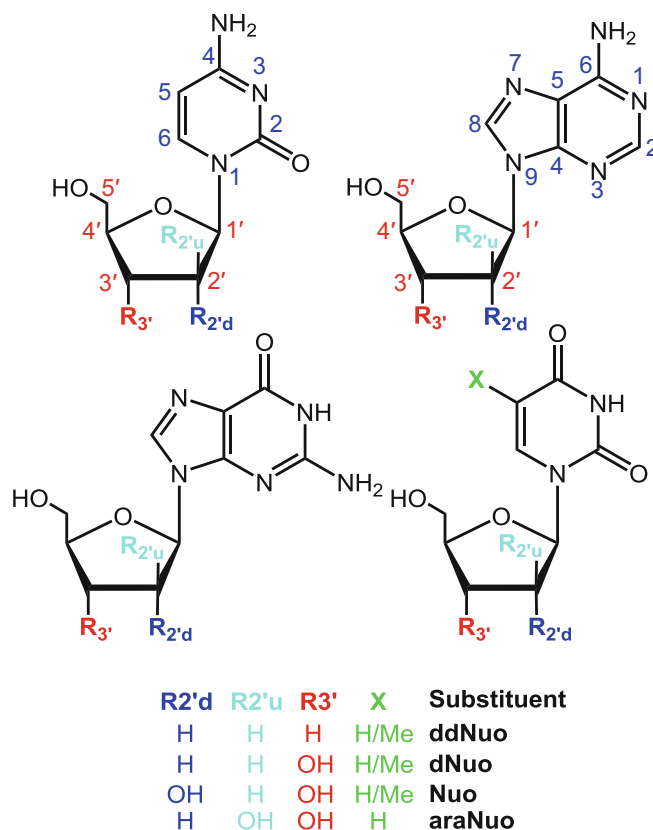


Figure 1. Structures of the five base nucleosides studied here with labels indicating atom numbers as well as identification of the RNA, arabinoside, DNA, and 2',3'-dideoxynucleoside modifications

Methods

Materials

The 2',3'-dideoxynucleosides of adenine, guanine, cytosine, and uracil were purchased from Chem Impex (Wood Dale, IL, USA), whereas 2',3'-dideoxythymidine was purchased from Carbosynth (San Diego, CA, USA). The HPLC grade methanol and water, and acetic acid used for the experiments in the IR fingerprint region were purchased from Sigma Aldrich (Zwijndrecht, The Netherlands). The HPLC grade methanol, and glacial acetic acid used for the experiments in the hydrogen-stretching region were purchased from Fischer Scientific (Waltham, MA, USA), whereas the HPLC grade water for these experiments was purchased from Sigma Aldrich (St. Louis, MO, USA).

IRMPD Action Spectroscopy

IRMPD action spectra of the protonated 2',3'-dideoxynucleosides of adenine, guanine, cytosine, thymine, and uracil in the IR fingerprint region were measured using a custom-built 4.7-T Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) [31, 32] coupled to a FELIX free electron laser (10 Hz repetition rate, bandwidth 0.3% of the central frequency, energy up to 70 mJ/pulse) [33]. The

protonated $[\text{ddNuo}+\text{H}]^+$ ions were generated by diluting a stock solution of the corresponding ddNuo to approximately 1 mM in 50%:50% methanol:water acidified with $\sim 1\%$ acetic acid, and infusing that solution into a Micromass “Z-spray” electrospray ionization source at 6 $\mu\text{L}/\text{min}$. Ions were accumulated in a hexapole ion guide before extraction through a quadrupole bender and octopole ion guide into the FT-ICR cell. The $[\text{ddNuo}+\text{H}]^+$ ions of interest were isolated using stored waveform inverse Fourier transform (SWIFT) techniques and excited via absorption of IR photons in the range of 700 to 1900 cm^{-1} for 1.5 s. Each protonated 2',3'-dideoxynucleoside dissociated via cleavage of the glycosidic bond resulting in the protonated nucleobase, $[\text{ddNuo}+\text{H}]^+ + n \text{ h}\nu \rightarrow [\text{Nua}+\text{H}]^+ + (\text{ddNuo}-\text{Nua})$, where Nua is the relevant nucleobase. The IRMPD yield was calculated using Eq. (1) below and linearly corrected for the frequency dependent FEL pulse energy.

$$\text{IRMPD yield} = \sum_i I_{\text{product}_i} / \left(I_{[\text{ddNuo}+\text{H}]^+} + \sum_i I_{\text{product}_i} \right) \quad (1)$$

IRMPD action spectra in the hydrogen-stretching region were measured in a modified Bruker amaZon ETD quadrupole ion trap (Bruker, Billerica, MA, USA) coupled to a Nd:YAG (Continuum Lasers, San Jose, CA, USA) pumped OPO laser system (LaserVision, Bellevue, WA, USA; repetition rate 10 Hz, bandwidth 3 cm^{-1} , energy up to 15 mJ/pulse) described in detail elsewhere [34]. Solutions of approximately 25 μM ddNuo were prepared in 50%:50% methanol:water, again acidified with $\sim 1\%$ acetic acid. These solutions were introduced to the Apollo II electrospray ionization source at 3 $\mu\text{L}/\text{min}$. The $[\text{ddNuo}+\text{H}]^+$ ions were isolated in the trap and irradiated for 0.1 to 0.5 s; the irradiation time was chosen to produce roughly 50% dissociation at the most intense feature. The IR spectra in the hydrogen-stretching region were collected between 3300 and 3800 cm^{-1} . IRMPD yields were calculated as described above, and the pulse energy was sufficiently independent of frequency so as to not require power correction.

Computational Approach

Previous studies of protonated canonical and modified nucleosides determined the most favorable protonation sites for each nucleoside. Structures representing each of these protonation sites were submitted to a simulated annealing procedure in AMBER [35], using parameters generated by the Antechamber [36] utility using the *ff14SB*, and *GAFF* force fields [37, 38]. Each manually generated structure was submitted to an equilibration at 300 K for a random number of steps between 1000 and 3000, to generate a random starting structure. Each annealing cycle started at 300 K, linearly ramped up to 1000 K over 30,000 steps, with a step size of 0.01 fs, was held at 1000 K for 150,000 steps, and cooled to 0 K over 30,000 steps. After each annealing cycle, a molecular dynamics (MD) minimization was performed and the resulting structure was saved for later analysis. After 3000 cycles, the

dihedral angles that determine the primary structural parameters (sugar puckering: $\angle\text{C4}'\text{O4}'\text{C1}'\text{C2}'$, $\angle\text{O4}'\text{C1}'\text{C2}'\text{C3}'$, $\angle\text{C1}'\text{C2}'\text{C3}'\text{C4}'$, $\angle\text{C2}'\text{C3}'\text{C4}'\text{O4}'$, and $\angle\text{C3}'\text{C4}'\text{O4}'\text{C1}'$; 5'-hydroxyl: $\angle\text{O4}'\text{C4}'\text{C5}'\text{O5}'$; and nucleobase orientation: $\angle\text{O4}'\text{C1}'\text{N9C6}$ for purines and $\angle\text{O4}'\text{C1}'\text{N1C2}$ for the pyrimidines) were extracted from the minimized structures and used to calculate root-mean squared deviations (RMSD) between the structures using Eq. (2), where X and Y are the corresponding dihedral angles of the two structures being compared.

$$\text{RMSD} = \sum_{i=0}^N |X_i - Y_i|^2 / N \quad (2)$$

A RMSD cutoff value of 0.45 was used to determine if the conformer structure was *unique* amongst those conformers already identified and, if so, was set aside for further comparisons. The calculated energies of the conformers were used to select the lowest energy example of each *unique* conformer found. This process resulted in approximately 100–200 *unique* structures from the MD conformational search for each protonation site of each ddNuo examined. These structures were submitted to DFT optimization using Gaussian 09 [39] at the B3LYP/6-31+G(d) level of theory to further refine these structures. The same RMSD filtering process was then used for these refined structures. Those structures highlighted as *unique* after the initial DFT optimization were subjected to further optimization and frequency analysis with a larger basis set, B3LYP/6-311+G(d,p). Finally, the single point energies of these structures were determined with a still larger basis set, B3LYP/6-311+G(2d,2p), to provide better relative energetics. These levels of theory were chosen as they were found in previous studies of the protonated canonical DNA and RNA nucleosides to perform well both spectroscopically and energetically [22–26]. Manual structure manipulation was used to generate structures that the previous studies of the canonical DNA and RNA nucleosides indicated may be important, but was missed in the MD-driven conformational search. Such manual structure building was largely unnecessary in the present work due to the comprehensive nature of the conformational search employed.

The vibrational frequencies calculated via the frequency analysis are scaled by an arbitrary scaling factor to correct for the anharmonicity of the resonant vibrational modes, and thus present in the experiment, but not included in the calculation. The degree to which the anharmonicity affects peak position differs by the type of vibration, though similar vibrations will typically exhibit the same scaling factor. Different scale factors are used for the fingerprint and hydrogen-stretching regions in each system. Scale factors are chosen by first considering all of the conformers calculated for a given ion and selecting a scale factor that provides the best overall agreement amongst them and the experimental IRMPD spectrum. This scale factor is further refined by comparison with only those predicted spectra that exhibit reasonable agreement with the measured spectrum and are relatively stable amongst those calculated. Following this correction, the calculated frequencies and corresponding intensities are convoluted with a Gaussian peak shape with a

full width at half maximum of 20 cm^{-1} in the IR fingerprint region and 15 cm^{-1} in the hydrogen-stretching region, to better reproduce the experimental peak shapes and facilitate comparisons.

The same series of dihedral angles used during the conformational search are gathered from each conformer following the electronic structure calculations and used to determine three important structural parameters as outlined by Sundaralingam and Altona [4]. A pseudorotation angle (P) is calculated from the series of five dihedral angles within the sugar ring starting at $\angle\text{C4}'\text{O4}'\text{C1}'\text{C2}'$ and continuing through $\angle\text{O4}'\text{C1}'\text{C2}'\text{C3}'$, $\angle\text{C1}'\text{C2}'\text{C3}'\text{C4}'$, $\angle\text{C2}'\text{C3}'\text{C4}'\text{O4}'$, and $\angle\text{C3}'\text{C4}'\text{O4}'\text{C1}'$, and is used to assign a specific sugar puckering mode to the conformer. These specific envelope (E) or twisted (T) puckering modes are inclusive of the classical C2'-endo/C3'-endo designations. In the C2'-endo/C3'-endo designations, the sugar atom specified, C2' and C3' in these examples, lies either above (*endo*) or below (*exo*) the plane of the ring. The glycosidic bond angle (χ) is used to determine the nucleobase orientation and is measured as $\angle\text{O4}'\text{C1}'\text{N1C2}$ for the pyrimidine-based nucleosides, and $\angle\text{O4}'\text{C1}'\text{N9C6}$ for the purine-based nucleosides. The nucleobase orientation is generally described as either *anti* where the nucleobase is oriented to facilitate Watson-Crick base pairing or *syn* facilitating Hoogsteen base pairing. The third important structural parameter is the position of the 5'-hydroxyl, which often acts as an acceptor in the hydrogen-bonding interactions that stabilize the nucleobase orientation in gas-phase protonated nucleosides. The position of the 5'-hydroxyl is characterized by $\angle\text{O4}'\text{C4}'\text{C5}'\text{O5}'$ (τ) and falls into three categories. The *gauche*⁺ conformation is characterized by the O5' moiety lying between the O4' and C3', i.e., directly above the sugar ring upon projection down the C4'–C5' bond. The *trans* conformation is characterized by the 5'-hydroxyl group pointing away from C3', whereas the *gauche*[−] conformation positions the 5'-hydroxyl pointing away from O4'. Specific hydrogen-bonding interactions between the nucleobase and sugar are described as $\text{XH}\cdots\text{Y}$ with groups on the nucleobase followed by a subscript *a* or *s* denoting the *anti* or *syn* orientation of the nucleobase. The 5'-hydroxyl, if involved, is followed by a subscript *g*⁺, *g*[−], or *t* denoting its *gauche*⁺, *gauche*[−], or *trans* orientation. Alongside structural labels, each conformer is identified by a designation based upon its site of protonation followed by a letter incremented alphabetically according to the relative Gibbs energies at 298 K for all conformers within that designation.

Results and Discussion

[ddAdo+H]⁺

Theoretical Results

Low-energy conformers of [ddAdo+H]⁺ are shown in Figure S1 and structural parameters alongside relative energetics are listed in Table S1. As found for other protonated adenine-based nucleosides [23, 30, 40], the most stable gas-

phase conformers calculated for [ddAdo+H]⁺ are all protonated at the N3 position. The five conformers calculated to be the most stable, **N3A_{ddAdo}** through **N3E_{ddAdo}**, all exhibit *syn* nucleobase orientations stabilized by $\text{N3H}^+\cdots\text{O5}'$ hydrogen-bonding interactions ($\sim 1.8\text{ \AA}$, $\angle\text{N3H}^+\text{O5}' 163^\circ\text{--}169^\circ$), whereas the *syn* nucleobase of the sixth most stable conformer, **N3F_{ddAdo}**, is stabilized by a $\text{N3H}^+\cdots\text{O4}'$ hydrogen-bonding interaction ($\sim 2.153\text{ \AA}$, $\angle\text{N3H}^+\text{O4}' 118^\circ$). **N3A_{ddAdo}** and **N3B_{ddAdo}** have *gauche*⁺ 5'-hydroxyls with C3'-endo and C2'-endo sugar puckering, respectively. In contrast, *trans* 5'-hydroxyls are found in **N3C_{ddAdo}**, **N3D_{ddAdo}**, and **N3E_{ddAdo}** with C2'-endo (**N3C_{ddAdo}** and **N3D_{ddAdo}**) and C2'-exo (**N3E_{ddAdo}**) sugar puckering. Conformers protonated at the N1 position, **N1A_{ddAdo}** and **N1B_{ddAdo}**, are found 24.0 and 25.8 kJ/mol higher in Gibbs energy, respectively, and exhibit a $\text{C8H}_a\cdots\text{O5}'_{g+}$ hydrogen-bonding interaction (2.1 \AA , $\angle\text{C8HO5}' 117^\circ$). Previous studies of [Ado+H]⁺, [dAdo+H]⁺ [23], protonated 2'-O-methyladenosine [Adom+H]⁺ [40], and [araAdo+H]⁺ [30] indicate that although the N1 protonated conformers are calculated to be $\geq 20\text{ kJ/mol}$ less stable than N3 protonated conformers in the gas phase, in a polarizable continuum N1 and N3 protonation are much closer in energy, and in some cases, the preference is inverted. In a polarizable continuum, the most stable N1 protonated conformer of [ddAdo+H]⁺, **N1A_{ddAdo}**, is calculated to be only 6.8 kJ/mol less stable than **N3A_{ddAdo}**. In gas-phase calculations, the most stable N7 protonated conformer, **N7A_{ddAdo}**, is found 27.2 kJ/mol higher in Gibbs energy and is stabilized by the same noncanonical $\text{C8H}_a\cdots\text{O5}'_{g+}$ hydrogen-bonding interaction as **N1A_{ddAdo}**. Conformers exhibiting *gauche*[−] 5'-hydroxyls lie higher still in Gibbs energy, with the most stable amongst them, **N3G_{ddAdo}**, which is stabilized by an $\text{N3H}^+\cdots\text{O4}'$ hydrogen-bonding interaction, lying 28.4 kJ/mol above the calculated ground conformer. The most stable conformers for each site of protonation, **N3A_{ddAdo}**, **N1A_{ddAdo}**, and **N7A_{ddAdo}**, prefer C3'-endo over C2'-endo sugar puckering with the former preferred by 0.4 kJ/mol for N3 conformers, 1.8 kJ/mol for N1 conformers, and 8.1 kJ/mol for N7 protonated conformers.

IRMPD Action Spectroscopy

The measured IRMPD spectrum of [ddAdo+H]⁺ is displayed with those measured in previous work of [dAdo+H]⁺ and [Ado+H]⁺ [23] and [araAdo+H]⁺ [30] in Figure 2. Predicted linear IR spectra of the conformers primarily responsible for the experimental populations of [Ado+H]⁺, [araAdo+H]⁺, [dAdo+H]⁺, and [ddAdo+H]⁺ are also shown in Figure 2 for comparison. The predicted linear IR spectra of [ddAdo+H]⁺ were scaled by 0.9810 in the IR fingerprint region and 0.9603 in the hydrogen-stretching region to best match the measured spectrum. Comprehensive comparisons of the predicted spectra of the low-energy conformers calculated for [ddAdo+H]⁺ with the measured spectrum of [ddAdo+H]⁺ are found in Figure S2 with highlighted regions indicating areas of disagreement that limit or prohibit the conformation contribution to the experimental population. In brief, the predicted spectra for **N3A_{ddAdo}**

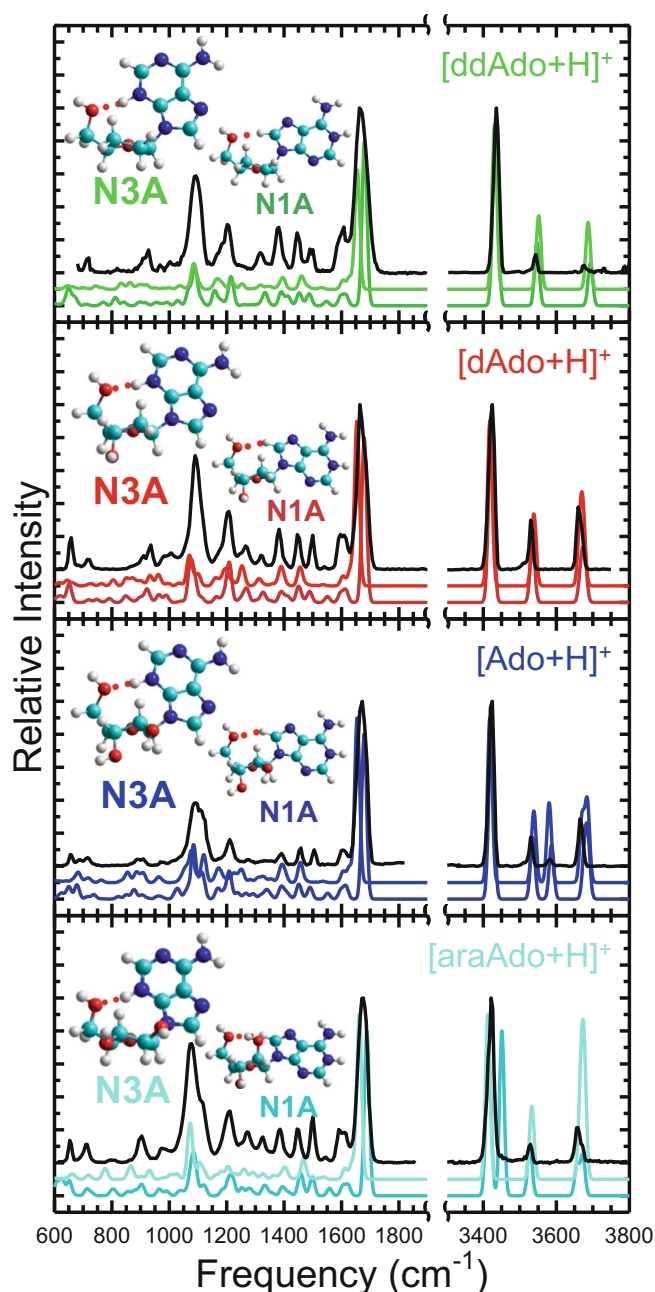


Figure 2. Comparison of the experimental IRMPD spectra of 2',3'-dideoxyadenosine, 2'-deoxyadenosine, adenosine, and adenosine arabinoside. Also compared are the predicted linear IR spectra and images of the conformers identified as the primary contributors to the experimental spectrum as calculated at the B3LYP/6-311+G(d,p) level of theory. The intensities of the experimental and predicted spectra are normalized to facilitate comparison. The dominant, typically more stable, conformer is represented by the larger image to the left and middle spectrum in each panel, whereas the smaller image to the right and bottom spectrum represents the minor conformer that is calculated to be higher in energy, but still contributes to the experimental spectrum

through **N3E_{ddAdo}** all demonstrate reasonable agreement with the measured IRMPD spectrum of $[\text{ddAdo}+\text{H}]^+$, with

N3A_{ddAdo} providing the best agreement amongst the N3 protonated conformers. The width of the feature at $\sim 1670\text{ cm}^{-1}$ in each of the measured IRMPD spectra indicates the presence of both N3 and N1 protonated conformers in the experiments. For $[\text{ddAdo}+\text{H}]^+$, **N1A_{ddAdo}** provides good agreement with the measured spectrum and is very complementary to **N3A_{ddAdo}** in the feature at $\sim 1670\text{ cm}^{-1}$. Conformers protonated at N7 provide a poor spectral match and thus do not contribute measurably to the experimental population. The consistent appearance of the spectral features in the measured IRMPD spectra indicates that the protonation sites and overall structures are preserved in the experimentally accessed conformers of $[\text{Ado}+\text{H}]^+$, $[\text{dAdo}+\text{H}]^+$, and $[\text{ddAdo}+\text{H}]^+$. The feature measured for $[\text{Ado}+\text{H}]^+$ at $\sim 3582\text{ cm}^{-1}$ was attributed to stretching of the 2'-hydroxyl, which explains its absence in the spectra of $[\text{dAdo}+\text{H}]^+$ and $[\text{ddAdo}+\text{H}]^+$. Likewise, the small shoulder at 3450 cm^{-1} measured for $[\text{araAdo}+\text{H}]^+$ indicates the presence of an alternate nucleobase–sugar hydrogen-bonding interaction, but the primary structure responsible for the measured spectrum remains consistent with those measured for $[\text{Ado}+\text{H}]^+$, $[\text{dAdo}+\text{H}]^+$, and $[\text{ddAdo}+\text{H}]^+$. The **N3A_{ddAdo}** and **N1A_{ddAdo}** conformers primarily responsible for the experimental $[\text{ddAdo}+\text{H}]^+$ population measured here are consistent with the agreement amongst the experimental spectra. These conformers are highly parallel to the **N3A** and **N1A** conformers of $[\text{Ado}+\text{H}]^+$ and $[\text{dAdo}+\text{H}]^+$ that provide excellent agreement with the spectrum measured in that work as seen in Figure 2. The primary difference between the conformers majorly populated for $[\text{ddAdo}+\text{H}]^+$ versus those for $[\text{Ado}+\text{H}]^+$ and $[\text{dAdo}+\text{H}]^+$ is the preference for C3'-endo sugar puckering, which can be observed spectroscopically in the better alignment and reproduction of the complex sugar feature between 1050 and 1150 cm^{-1} by **N3A_{ddAdo}** over **N3B_{ddAdo}**.

$[\text{ddGuo}+\text{H}]^+$

Theoretical Results

Structural parameters and relative energetics of $[\text{ddGuo}+\text{H}]^+$ conformers calculated are provided in Table S2 with images of the corresponding conformers displayed in Figure S3. As found for $[\text{Guo}+\text{H}]^+$, $[\text{araGuo}+\text{H}]^+$, and $[\text{dGuo}+\text{H}]^+$ [22], protonation at the N7 position of $[\text{ddGuo}+\text{H}]^+$ is heavily preferred. The calculated ground conformer, **N7A_{ddGuo}**, displays C3'-endo sugar puckering stabilized by a noncanonical $\text{C8H}_a\cdots\text{O5}'_{g+}$ hydrogen-bonding interaction. **N7B_{ddGuo}** displays a very similar structure to **N7A_{ddGuo}**, the primary difference being C2'-endo sugar puckering, which is calculated to be 7.6 kJ/mol less stable. **N7C_{ddGuo}** and **N7D_{ddGuo}** lie just 16.8 and 17.4 kJ/mol higher in Gibbs energy and display C3'-exo sugar puckering alongside *trans* 5'-hydroxyls, both stabilized by noncanonical $\text{C8H}_a\cdots\text{O5}'_t$ hydrogen-bonding interactions. The most stable O6 protonated conformer, **O6A_{ddGuo}**, displays C3'-endo sugar puckering stabilized by a noncanonical $\text{C8H}_a\cdots\text{O5}'_{g+}$ hydrogen-bonding interaction, but lies 44.2 kJ/mol higher in Gibbs energy than **N7A_{ddGuo}**. The most stable conformer populated at

the N3 position, **N3A_{ddGuo}**, lies 54.1 kJ/mol higher in Gibbs energy with C2'-*endo* sugar pucker stabilized by an N3H⁺...O5'_{g+} hydrogen-bonding interaction. Conformers with *gauche*⁻ 5'-hydroxyls, such as **N7E_{ddGuo}**, exhibit no strong nucleobase-sugar hydrogen-bonding interactions, and therefore lie higher in Gibbs energy, ≥ 20.9 kJ/mol above **N7A_{ddGuo}**.

IRMPD Action Spectroscopy

Measured IRMPD spectra of [ddGuo+H]⁺, [dGuo+H]⁺, [Guo+H]⁺, and [araGuo+H]⁺ [22, 30] are compared in Figure 3. The predicted IR spectra of the major conformers determined to be populated in the experiments are shown for comparison. The predicted linear IR spectra of [ddGuo+H]⁺ were scaled by 0.9770 in the IR fingerprint region and 0.9610 in the hydrogen-stretching region. Comparisons of the predicted IR spectra for the conformers calculated for [ddGuo+H]⁺ are shown in Figure S4 with highlighting to indicate regions of disagreement precluding the conformer from contributing significantly to the measured spectrum. Altogether, **N7A_{ddGuo}** and **N7C_{ddGuo}** provide the best reproduction of the measured IRMPD spectrum of [ddGuo+H]⁺. Their predicted spectra are highly parallel to one another and have no notable disagreements with the measured spectrum that would suggest conformers protonated at another site are necessary to describe the experimental population. That said, the predicted spectrum of **N7B_{ddGuo}** is also highly parallel with only marginally worse representation of the feature of moderate intensity at ~ 1230 cm⁻¹. It is interesting to note that these three conformers represent C3'-*endo*, C2'-*endo*, and C3'-*exo* sugar pucker and both *gauche*⁺ and *trans* 5'-hydroxyls yet exhibit only relatively minor spectroscopic variation, which makes differentiation difficult. The **N7A** conformers of [Guo+H]⁺, [dGuo+H]⁺, and [ddGuo+H]⁺, as seen in Figure 3, are extremely similar to one another, stabilized by C8H_a...O5'_{g+} hydrogen-bonding interactions with C3'-*endo* sugar pucker. The **N7A_{araGuo}** conformer of [araGuo+H]⁺ exhibits the unique O2'H...O5'_{g+} hydrogen-bonding dictating C2'-*endo* sugar pucker, whereas **N7C_{araGuo}** is more similar in hydrogen-bonding and sugar pucker to the **N7A_{ddGuo}** conformer of [ddGuo+H]⁺.

[ddCyd+H]⁺

Theoretical Results

Relative energetics and structural parameters of the calculated conformers of [ddCyd+H]⁺ can be found in Table S3, and images of these conformers are displayed in Figure S5. Protonation at the N3 position in **N3A_{ddCyd}** is calculated to be preferred by 4.8 kJ/mol over protonation at the O2 position in **O2A_{ddCyd}** for [ddCyd+H]⁺. N3 protonation is also preferred over O2 protonation for [dCyd+H]⁺, [Cyd+H]⁺ [25] and [araCyd+H]⁺ [30]. **N3A_{ddCyd}** and **O2A_{ddCyd}** exhibit *anti* nucleobases stabilized by noncanonical C6H_a...O5'_{g+} hydrogen-bonding interactions. C3'-*endo* sugar pucker, as found in **N3A_{ddAdo}** and **O2A_{ddAdo}**, is preferred over C2'-*endo* pucker in **N3B_{ddCyd}** and **O2B_{ddCyd}** by 6.8 and 6.1 kJ/mol

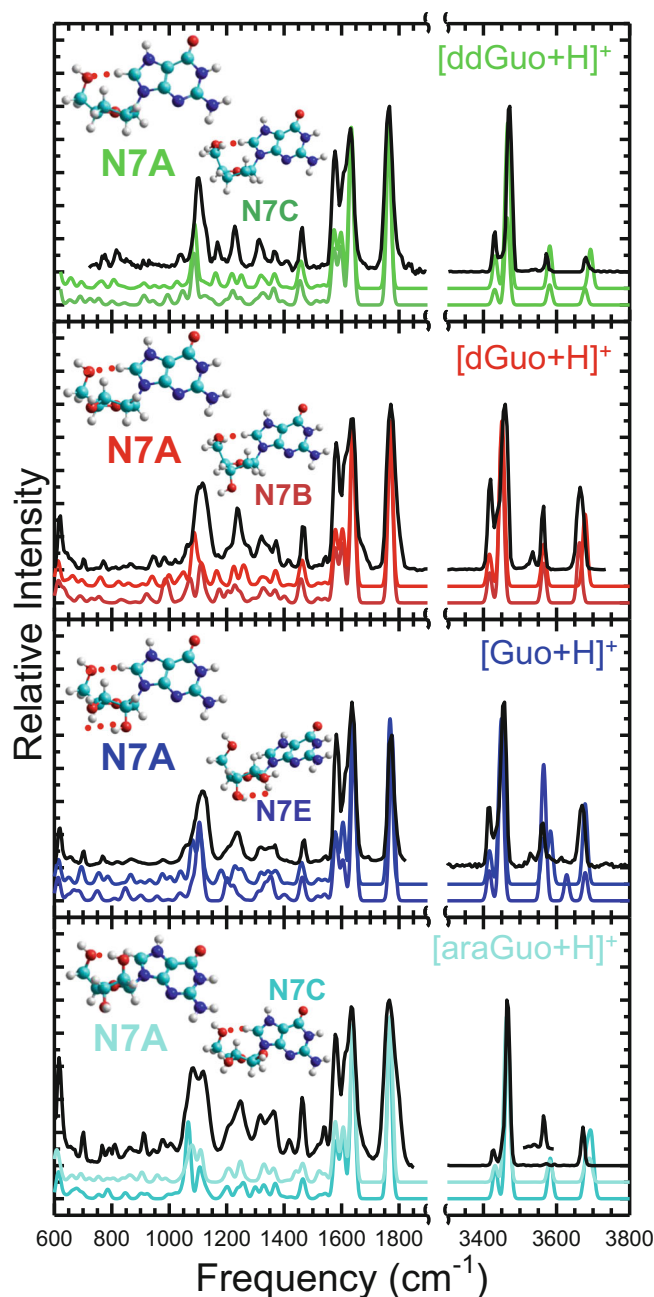


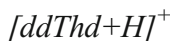
Figure 3. Comparison of the experimental IRMPD spectra of 2',3'-dideoxyguanosine, 2'-deoxyguanosine, guanosine, and guanosine arabinoside. Also compared are the predicted linear IR spectra and images of the conformers identified as the primary contributors to the experimental spectrum as calculated at the B3LYP/6-311+G(d,p) level of theory. The intensities of the experimental and predicted spectra are normalized to facilitate comparison. The dominant, typically more stable, conformer is represented by the larger image to the left and middle spectrum in each panel, whereas the smaller image to the right and bottom spectrum represents the minor conformer that is calculated to be higher in energy, but still contributes to the experimental spectrum

respectively. **N3C_{ddCyd}**, lying just 14.2 kJ/mol above **N3A_{ddCyd}** in Gibbs energy, has C3'-*exo* sugar pucker with a noncanonical C6H_a...O5'_t hydrogen-bonding interaction.

O2C_{ddCyd}, lying just 12.4 kJ/mol higher in Gibbs energy than **O2A_{ddCyd}**, is the most stable conformer displaying a *syn* nucleobase orientation, which is stabilized by an $\text{O2H}^+_{\text{s}} \cdots \text{O5}^+_{\text{g}}$ hydrogen-bonding interaction. Other conformers displaying *syn* nucleobase orientations lie much higher in Gibbs energy, >28.4 kJ/mol above the calculated **N3A_{ddCyd}** ground conformer. Conformers with *gauche*[−] 5'-hydroxyls, such as **N3E_{ddCyd}**, lie ≥ 18.2 kJ/mol above the calculated ground conformer and, as expected, provide no hydrogen-bonding stabilization with the 5'-hydroxyl.

IRMPD Action Spectroscopy

Measured IRMPD spectra of $[\text{dCyd}+\text{H}]^+$, $[\text{Cyd}+\text{H}]^+$ [25], and $[\text{araCyd}+\text{H}]^+$ [30] from previous work are compared with that measured for $[\text{ddCyd}+\text{H}]^+$ in Figure 4. Linear IR spectra predicted for the conformers primarily responsible for the measured IRMPD spectra of these protonated cytosine nucleoside analogues are also displayed for comparison, accompanied by their structures. Predicted linear IR spectra of $[\text{ddCyd}+\text{H}]^+$ are scaled by 0.9750 in the IR fingerprint and 0.9630 in the hydrogen-stretching regions. Comparisons of the predicted linear IR spectra for a representative set of conformers calculated of $[\text{ddCyd}+\text{H}]^+$ are displayed in Figure S6. In summary, the conformers providing the best agreement with the measured IRMPD spectrum are **N3A_{ddCyd}** and **O2A_{ddCyd}**, the most stable conformers of each favorable protonation site. The two predicted spectra are highly complementary, with **N3A_{ddCyd}** contributing the distinctive major features at ~ 1790 and 1280 cm^{-1} , whereas **O2A_{ddCyd}** contributes the features at ~ 1215 , 1490 , and 3585 cm^{-1} . The C2'-*endo* puckered versions of each of these conformers, **N3B_{ddCyd}** and **O2B_{ddCyd}**, provide similar predicted spectra but worse agreement for the sugar-stretching modes at $\sim 1100\text{ cm}^{-1}$. The relative intensities of the features unique to the N3 versus O2 protonated conformers and their similar energetics suggest that they are present in roughly equal proportion in the experiments. This result for $[\text{ddCyd}+\text{H}]^+$ is consistent with that found previously for $[\text{dCyd}+\text{H}]^+$, $[\text{Cyd}+\text{H}]^+$, and $[\text{araCyd}+\text{H}]^+$, where roughly equal proportions of N3 and O2 protonated conformers were populated. Similar to that found for $[\text{ddAdo}+\text{H}]^+$, $[\text{ddCyd}+\text{H}]^+$ demonstrates a spectroscopic preference for C3'-*endo* sugar puckering, in contrast to the C2'-*endo* puckering that is marginally preferred by $[\text{dCyd}+\text{H}]^+$, $[\text{Cyd}+\text{H}]^+$, and $[\text{araCyd}+\text{H}]^+$. The sugar pucker preference of $[\text{ddCyd}+\text{H}]^+$ is the only notable structural change observed spectroscopically from the RNA and DNA analogues.



Theoretical Results

Relative energetics and structural parameters of the conformers calculated for $[\text{ddThd}+\text{H}]^+$ are listed in Table S4 with images of these structures displayed in Figure S7. Parallel to that found for the DNA and RNA analogues, protonation of ddThd preferentially results in tautomerization to form the minor 2,4-

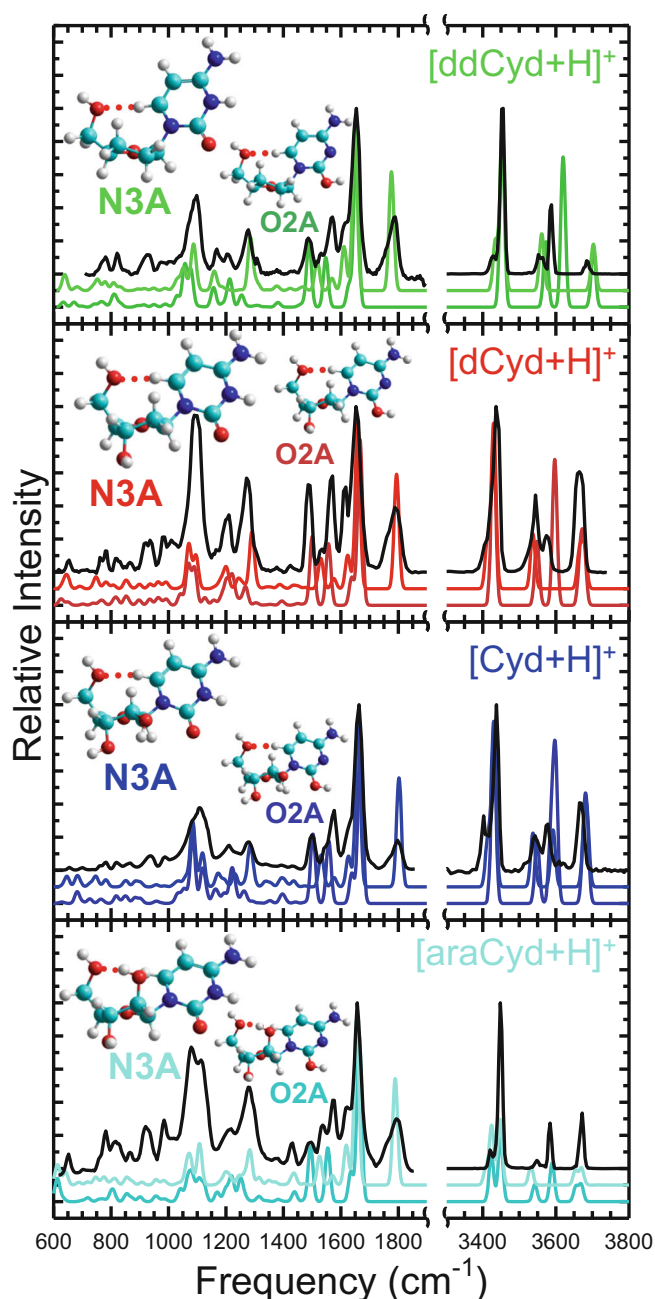


Figure 4. Comparison of the experimental IRMPD spectra of 2',3'-dideoxycytidine, 2'-deoxycytidine, cytidine and cytosine arabinoside. Also compared are the predicted linear IR spectra and images of the conformers identified as the primary contributors to the experimental spectrum as calculated at the B3LYP/6-311+G(d,p) level of theory. The intensities of the experimental and predicted spectra are normalized to facilitate comparison. The dominant, typically more stable, conformer is represented by the larger image to the left and middle spectrum in each panel, whereas the smaller image to the right and bottom spectrum represents the minor conformer that is calculated to be higher in energy, but still contributes to the experimental spectrum

dihydroxy tautomer (T) [26]. The most stable conformer, the **TA_{ddThd}** minor tautomer, exhibits C3'-*endo* sugar puckering

and is stabilized by a noncanonical $C6H_a \cdots O5'_g$ hydrogen-bonding interaction. **O4A_{ddThd}** lies just 2.5 kJ/mol higher in Gibbs energy and retains a highly parallel C3'-*endo* puckered structure with the same $C6H_a \cdots O5'_g$ hydrogen-bonding interaction. **TB_{ddThd}** and **O4B_{ddThd}** are highly parallel to **TA_{ddThd}** and **O4A_{ddThd}**, but with C2'-*endo* sugar puckering and lie just 6.0 and 7.0 kJ/mol higher in Gibbs energy than **TA_{ddThd}** and **O4A_{ddThd}**, respectively. The most stable conformers exhibiting *syn* nucleobase orientations, **TC_{ddThd}** and **O2A_{ddThd}**, lie just 9.5 and 9.7 kJ/mol above **TA_{ddThd}** and have C2'-*endo* sugar puckering stabilized by an $O2H^+ \cdots O5'_g$ hydrogen-bonding interaction. **O2B_{ddThd}**, which lies another 2.4 kJ/mol Gibbs energy above **O2A_{ddThd}** displays the same $O2H^+ \cdots O5'_g$ interaction, but with C4'-*exo* sugar puckering. The most stable conformers displaying *trans* 5'-hydroxyls, **TE_{ddThd}** and **O4C_{ddThd}**, also display C3'-*exo* sugar puckers and $C6H_a \cdots O5'_t$ hydrogen-bonding interactions and lie 17.0 and 17.7 kJ/mol above the ground **TA_{ddThd}** conformer. **TG_{ddThd}**, the most stable conformer displaying a *gauche*⁻ 5'-hydroxyl, displays no notable nucleobase–sugar hydrogen-bonding interactions and lies 18.5 kJ/mol above **TA_{ddThd}**.

IRMPD Action Spectroscopy

Measured IRMPD spectra of $[ddThd+H]^+$, $[dThd+H]^+$, and $[Thd+H]^+$ [26] are compared in Figure 5. The predicted IR spectra of the major conformers determined to be populated in the experiments are shown for comparison. The predicted linear IR spectra of $[ddThd+H]^+$ are scaled by 0.9780 in the IR fingerprint and 0.9550 in the hydrogen-stretching regions. Comparisons of the predicted IR spectra for the conformers calculated for $[ddThd+H]^+$ are shown in Figure S8 with highlighted regions to indicate disagreements that preclude the conformers from contributing to the experiments. The most stable minor tautomer, **TA_{ddThd}**, and O2 protonated conformer, **O2A_{ddThd}**, predict complementary spectra that combined provide good agreement with the measured IRMPD spectrum of $[ddThd+H]^+$. **TA_{ddThd}** provides excellent representation of the features measured at ~ 1200 , 1380 , and 3580 cm^{-1} and the two major features near 1500 cm^{-1} , whereas **O2A_{ddThd}** provides excellent representation of the features measured at ~ 1150 , 1600 , 1790 , and 3405 cm^{-1} . The relative intensities of the distinctive features in both conformers indicate that the minor tautomer is preferred in the experiments, but conformers exhibiting O2 protonation are also present. Conformers protonated at the O4 position provide similar reproduction of the measured feature at $\sim 1790\text{ cm}^{-1}$, but predict a major feature at $\sim 1570\text{ cm}^{-1}$ that is not present in the measured spectrum. This consistent feature precludes all O4 protonated conformers from being significant contributors to the experimentally accessed ions. The *syn* **TC_{ddThd}** conformer provides a similarly located feature at $\sim 1560\text{ cm}^{-1}$ that precludes it from a measurable presence in the experiments. **TB_{ddThd}** and **O2B_{ddThd}** lie only 2.5 and 2.4 kJ/mol above **TA_{ddThd}** and **O2A_{ddThd}** respectively, and both also provide decent agreement with the measured spectrum. However,

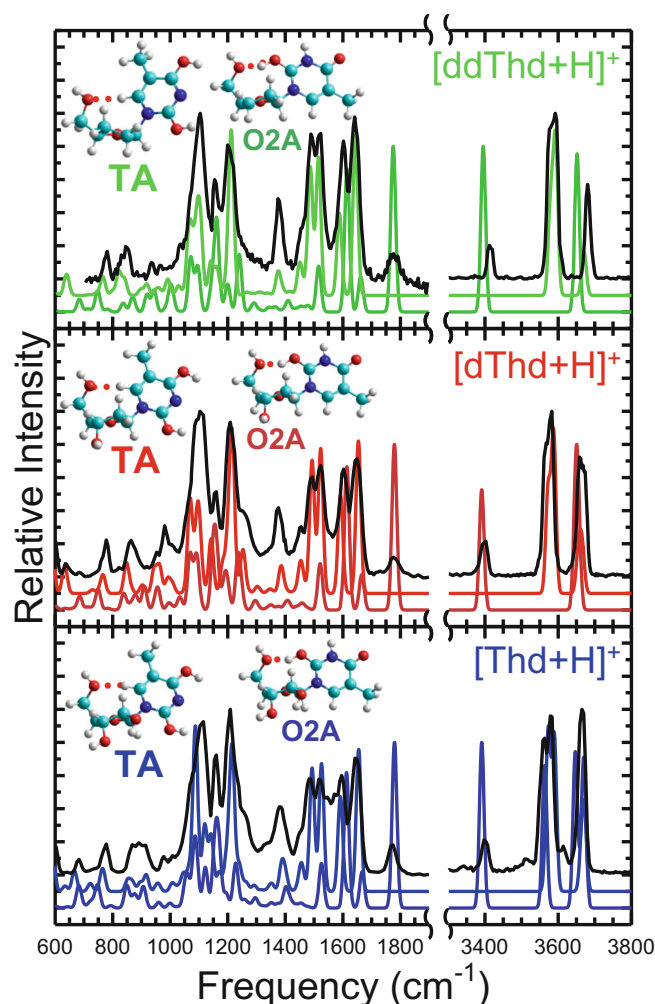


Figure 5. Comparison of the experimental IRMPD spectra of 2',3'-dideoxythymidine, 2'-deoxythymidine, and thymidine. Also compared are the predicted linear IR spectra and images of the conformers identified as the primary contributors to the experimental spectrum as calculated at the B3LYP/6-311+G(d,p) level of theory. The intensities of the experimental and predicted spectra are normalized to facilitate comparison. The dominant, typically more stable, conformer is represented by the larger image to the left and middle spectrum in each panel, whereas the smaller image to the right and bottom spectrum represents the minor conformer that is calculated to be higher in energy, but still contributes to the experimental spectrum

the spectra of these conformers both have intense features around $\sim 1075\text{ cm}^{-1}$ that exhibit poorer alignment with the measured feature at $\sim 1105\text{ cm}^{-1}$ indicating that they are not major contributors to the experimentally populated ions. Both $[dThd+H]^+$ and $[Thd+H]^+$ displayed a preference for formation of the minor tautomer upon protonation, with a very similar structure to that found in **TA_{ddThd}**, but displaying C2'-*endo* sugar puckering instead of the C3'-*endo* puckering observed in **TA_{ddThd}**. The equivalent **O2A** conformer is also populated for each of the three protonated thymidine nucleoside analogues, these conformers

are highly parallel in structure, and all are C2'-*endo* puckered structures with $\text{O2H}_s^+ \cdots \text{O5}'_{g+}$ hydrogen-bonding interactions.

$[\text{ddUrd}+\text{H}]^+$

Theoretical Results

Structural parameters and relative energetics of conformers calculated for $[\text{ddUrd}+\text{H}]^+$ are listed in Table S5 with images of these structures shown in Figure S9. Protonation of ddUrd is preferred at the O4 position as in **O4A_{ddUrd}** over formation of the minor 2,4-dihydroxy tautomer **TA_{ddUrd}** by 1.4 kJ/mol. This is in agreement with the preference found for the DNA and arabinoside analogues, but the RNA analogue indicates a preference for the minor tautomer by 2.9 kJ/mol over O4 protonation [24]. Both **O4A_{ddUrd}** and **TA_{ddUrd}** exhibit C3'-*endo* sugar puckering with *anti* nucleobase orientations and noncanonical $\text{C6H}_a \cdots \text{O5}'_{g+}$ hydrogen-bonding interactions. **O4B_{ddUrd}** and **TB_{ddUrd}** exhibit highly parallel structures to their more stable counterparts, except with C2'-*endo* sugar puckering, and lie 7.1 and 6.8 kJ/mol higher in Gibbs energy than **O4A_{ddUrd}** and **TA_{ddUrd}**, respectively. **TC_{ddUrd}** is the most stable structure to exhibit a *syn* nucleobase orientation and C2'-*endo* sugar puckering stabilized by an $\text{O2H}_s \cdots \text{O5}'_{g+}$ hydrogen-bonding interaction. The most stable conformer protonated at the O2 position, **O2A_{ddUrd}**, has C2'-*endo* sugar puckering stabilized by an $\text{O2H}_s^+ \cdots \text{O5}'_{g+}$ hydrogen-bonding interaction and lies only 14.8 kJ/mol higher in Gibbs energy than **O4A_{ddUrd}**. **O4C_{ddUrd}** is the most stable conformer with a *trans* 5'-hydroxyl, with C3'-*exo* sugar puckering stabilized by a noncanonical $\text{C6H}_a \cdots \text{O5}'_t$ hydrogen-bonding interaction. 5'-Hydroxyls exhibiting a *gauche*⁻ orientation lie higher in energy, with the most stable conformer, **O4E_{ddUrd}**, lying 20.2 kJ/mol above **O4A_{ddUrd}**, and displaying no hydrogen-bonding interaction with the 5'-hydroxyl.

IRMPD Action Spectroscopy

The measured IRMPD spectra of $[\text{dUrd}+\text{H}]^+$, $[\text{Urd}+\text{H}]^+$ [24], and $[\text{araUrd}+\text{H}]^+$ [30] are compared with that of $[\text{ddUrd}+\text{H}]^+$ in Figure 6; also included are the predicted spectra of the two primary conformers for each system needed to reproduce the measured spectrum. The predicted linear IR spectra of $[\text{ddUrd}+\text{H}]^+$ are scaled by 0.9770 in the IR fingerprint region and 0.9550 in the hydrogen-stretching region. Comparisons and images of the conformers calculated for $[\text{ddUrd}+\text{H}]^+$ are shown in Figure S10 with regions of disagreement highlighted. The two conformers displayed in Figure 6, **O4A_{ddUrd}** and **TA_{ddUrd}**, provide several characteristic features crucial for representing the measured spectrum. **O4A_{ddUrd}** and other O4 protonated conformers predict features at ~ 1430 , 1805 , 3390 , and 3615 cm^{-1} that match well with the measured spectrum and are not predicted for the minor tautomer conformers. Conversely, **TA_{ddUrd}** and other T conformers predict characteristic features at 1210 , 1380 , 1640 , and 3580 cm^{-1} that also match well with the measured spectrum. The predicted spectra of O2 protonated conformers present several stark

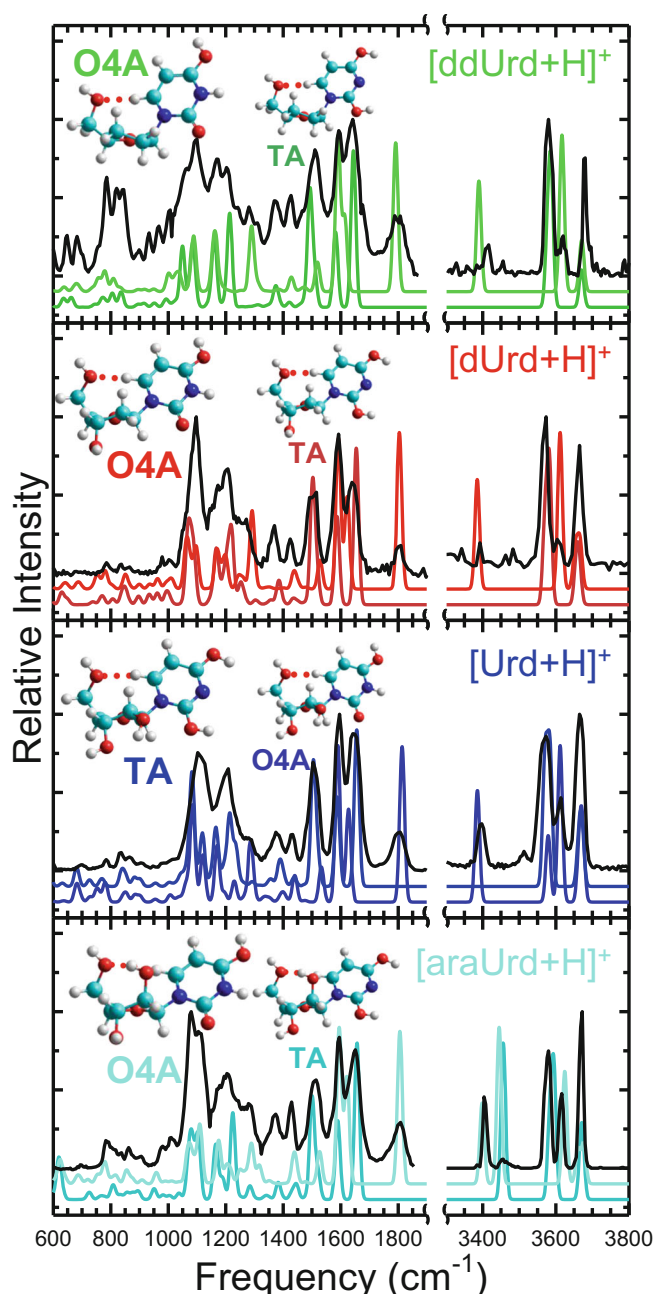


Figure 6. Comparison of the experimental IRMPD spectra of 2',3'-dideoxyuridine, 2'-deoxyuridine, uridine, and uracil arabinoside. Also compared are the predicted linear IR spectra and images of the conformers identified as the primary contributors to the experimental spectrum as calculated at the B3LYP/6-311+G(d,p) level of theory. The intensities of the experimental and predicted spectra are normalized to facilitate comparison. The dominant, typically more stable, conformer is represented by the larger image to the left and middle spectrum in each panel, whereas the smaller image to the right and bottom spectrum represents the minor conformer that is calculated to be higher in energy, but still contributes to the experimental spectrum

disagreements with the measured spectrum, indicating that they are not present in the experiments. The measured features of $[\text{ddUrd}+\text{H}]^+$ in the IR fingerprint region are very

broad due to poor ionization efficiency (and thus ion signal), a result of the low basicity of ddUrd. This makes confident differentiation between the predicted and measured spectra difficult. As such, conformers **O4A_{ddUrd}**, **O4B_{ddUrd}**, **O4D_{ddUrd}**, **O4E_{ddUrd}**, **TA_{ddUrd}**, **TB_{ddUrd}**, and **TE_{ddUrd}** all present reasonable agreement with the measured spectrum and cannot be excluded from the experimental population. However, as previously stated, combined **O4A_{ddUrd}** and **TA_{ddUrd}** provide excellent agreement with the measured spectrum, and the higher-energy conformers do not provide any additional major features that definitively establish their presence in the experiments.

Impact of 2'- and 2',3'-Hydroxyls on Structure of Purine Nucleosides

The parameters most indicative of structure (glycosidic bond angle, 5'-hydroxyl orientation, and pseudorotation angle) for all conformers calculated within 25 kJ/mol Gibbs energy of the respective ground conformer for the 2',3'-dideoxyribose, 2'-deoxyribose, ribose, and arabinose nucleosides of the purines, adenine and guanine are shown on polar plots in Figure 7. As discussed previously but not displayed in Figure 7, the preferred protonation sites for the adenine and guanine nucleoside analogues are not altered by the presence of the 2'- or 2',3'-hydroxyls or inversion of the stereochemistry at the 2'-position. The presence of both 2'- and 3'-hydroxyls in the RNA and arabinoside analogues results in more stable conformers with very similar structural parameters, most visible in the glycosidic bond and pseudorotation angles due to the relative flexibility of these modes. This clustering is partly due to rotamers of the 2'- and 3'-hydroxyls, which can be stabilized by either O2'H...O3' or O3'H...O2' sugar-sugar hydrogen-bonding interactions that tend to be of similar Gibbs energy. The presence of the 2'-hydroxyl in the RNA analogues also allows for a number of unique nucleobase-sugar hydrogen-bonding interactions, such as the dual N3H⁺_a...O2'H...O3' hydrogen-bonding interaction found in some stable conformations of [Ado+H]⁺ or the O3'H...O2'H...N3_s dual hydrogen-bonding interaction found in some conformers of [Guo+H]⁺. In the arabinosides, the 2'-hydroxyl has a propensity for an O2'H...O5' sugar-sugar hydrogen-bonding interaction, which competes heavily with formation of nucleobase-sugar hydrogen-bonding interactions, further increasing the number of relatively stable conformers calculated. The absence of the 2'-hydroxyl in the DNA analogues eliminates the O2'H...O3' and O2'...HO3' sugar-sugar hydrogen-bonding interactions found in the RNA analogues, which effectively reduces the number of “duplicate” conformers that only differ by the orientation of that interaction. However, a sugar-sugar hydrogen-bonding interaction between the 3'- and 5'-hydroxyls is still available to the DNA analogues. O3'H...O5' hydrogen-bonding interactions are found for [dGuo+H]⁺, [Guo+H]⁺, and [araGuo+H]⁺ at about the same relative stability. In [Ado+H]⁺, an O3'H...O5' interaction is found in a relatively stable conformer as a triple N3H⁺_a...O2'H...O3'H...O5'_t hydrogen-bonding interaction,

whereas in [dAdo+H]⁺, the hydrogen-bonding interaction O3'H...O5' is found only in conformers much higher in Gibbs energy, and in [araAdo+H]⁺, a dual O2'H...O3'H...O5' hydrogen-bonding interaction is found just within the 25 kJ/mol range shown in the figure. The absence of dual or triple hydrogen-bonding interactions due to removal of the 2'-hydroxyl leads to a lack of conformers exhibiting glycosidic bond angles between ~165 and 180° that are stabilized by such interactions.

The absence of both 2'- and 3'-hydroxyls results in an apparent increase in structurally unique conformers, but this is likely due to more 5'-rotamers and alternate sugar puckering modes than calculated for the DNA species. Beyond the rotamers and different puckering modes, [ddAdo+H]⁺ is highly parallel to [dAdo+H]⁺, but with an N1 protonated *anti* conformer very slightly more stable and just making it inside the 25 kJ/mol Gibbs energy range. The prominent outliers of [ddAdo+H]⁺, with glycosidic bond angles around ~338° and *exo* sugar puckering modes, each gains some stabilization by N3H⁺_s...O4' interactions. This stabilization mode is also observed in [dAdo+H]⁺, but provides less relative stability. Several conformers are found for [ddGuo+H]⁺ that contain no significant nucleobase-sugar hydrogen-bonding interactions. Those 2',3'-dideoxy conformers with *gauche*⁻ 5'-hydroxyls are very similar to those stabilized by O3'H...O5' hydrogen-bonding interactions in the DNA analogues.

Study of the solution conformations of ddAdo and ddGuo [19] by NMR also indicated a preference for N-type sugar puckering modes (C2'-*exo*/C3'-*endo*). This is in good agreement with the primarily populated conformers of [ddAdo+H]⁺ and [ddGuo+H]⁺ preferring C3'-*endo* sugar puckering. The *anti* nucleobase orientation of ddGuo found by NMR in solution agrees with that found here for [ddGuo+H]⁺ in the gas phase; however, for [ddAdo+H]⁺, the strong N3H⁺_s...O5' hydrogen-bonding interaction dictates a *syn* nucleobase orientation in contrast to the *anti* nucleobase observed in solution. Crystallographic study of ddAdo indicated a preference for S-type (C2'-*endo*/C3'-*exo*) sugar puckering [21], in contrast to that observed here or by NMR.

Altogether, the purine nucleosides are quite resistant to notable structural changes induced by the 2'- or 2',3'-hydroxyls. The absence of the 2'-hydroxyl results in the largest shift in accessible conformers by eliminating a sugar-sugar hydrogen-bonding interaction, either between the 2'- and 3'-hydroxyls (RNA) or between the 2'- and 5'-hydroxyls (arabinosides). Absence of the 2'-hydroxyl also eliminates a reasonably competitive alternative nucleobase-sugar hydrogen-bonding interaction between N3 and O2'. The absence of both the 2'- and 3'-hydroxyls removes the next most accessible sugar-sugar hydrogen-bonding interaction between 3'- and 5'-hydroxyls, which competes directly with the most common nucleobase-5'-hydroxyl interaction. The primary intramolecular hydrogen-bonding interactions remain consistent between the 2',3'-dideoxy and the DNA, and RNA nucleosides, N3H⁺_s...O5'_{g+} for adenine analogues and N7H_a...O5'_{g+} for guanine

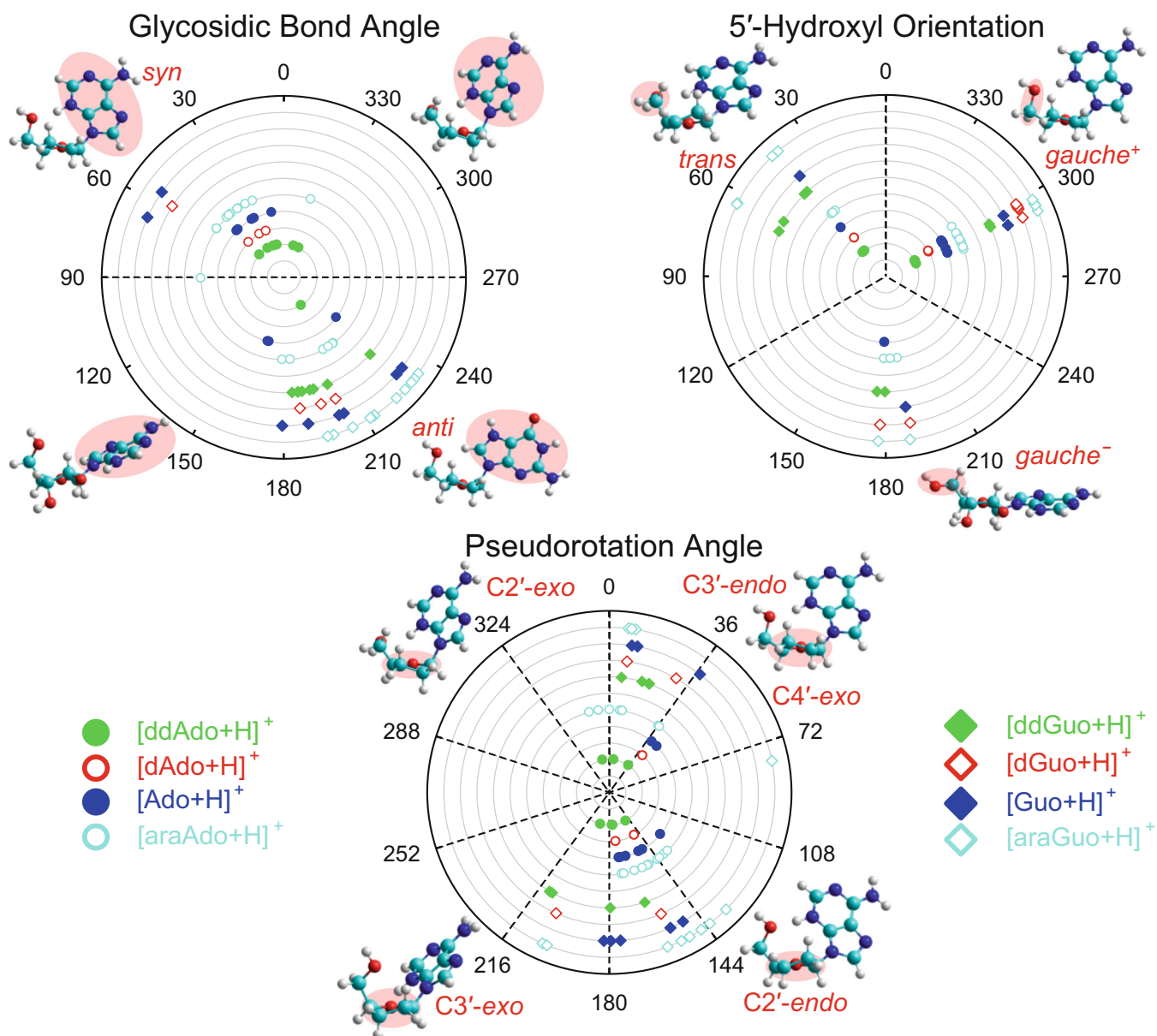


Figure 7. The three primary parameters that govern the structure of the conformers of the protonated purine nucleoside analogues calculated, glycosidic bond angle, 5'-hydroxyl orientation, and pseudorotation angle (sugar pucker), are displayed for all conformers with relative stabilities predicted within 25 kJ/mol Gibbs energy of the corresponding ground conformers

analogues, resulting in largely conserved low-energy conformations. Inversion of the stereochemistry at the 2'-position results in competition between a unique sugar–sugar hydrogen-bonding interaction and the nucleobase–sugar hydrogen-bonding interactions observed in the other analogues. In the absence of both the 2'- and 3'-hydroxyls, a shift in the calculated sugar preference, supported by some spectroscopic evidence, is observed for $[\text{ddAdo}+\text{H}]^+$, from C2'-endo in the DNA and RNA analogues to C3'-endo. The preference for C3'-endo sugar pucker in the guanosine analogues is maintained across the 2',3'-dideoxyribose, DNA, and RNA analogues, whereas the arabinose analogue displays a preference for C2'-endo sugar pucker due to its unique sugar–sugar hydrogen-bonding interaction.

Impact of 2'- and 2',3'-Hydroxyls on Structure of Pyrimidine Nucleosides

Structural parameters for the calculated conformers of the protonated 2',3'-dideoxyribose, 2'-deoxyribose, ribose, and arabinose nucleosides of the pyrimidine nucleobases cytosine, thymine, and uracil with relative stability within 25 kJ/mol Gibbs energy of their respective ground conformers are shown in Figure 8. The most notable change as a function of the 2' or 2',3'-hydroxyl is the shift in protonation preference of the uracil nucleosides discussed earlier. However, this change has relatively little impact on the structural parameters discussed here as the resulting rearrangement of relative stabilities is well within the stability range displayed in Figure 8. Unlike the protonated

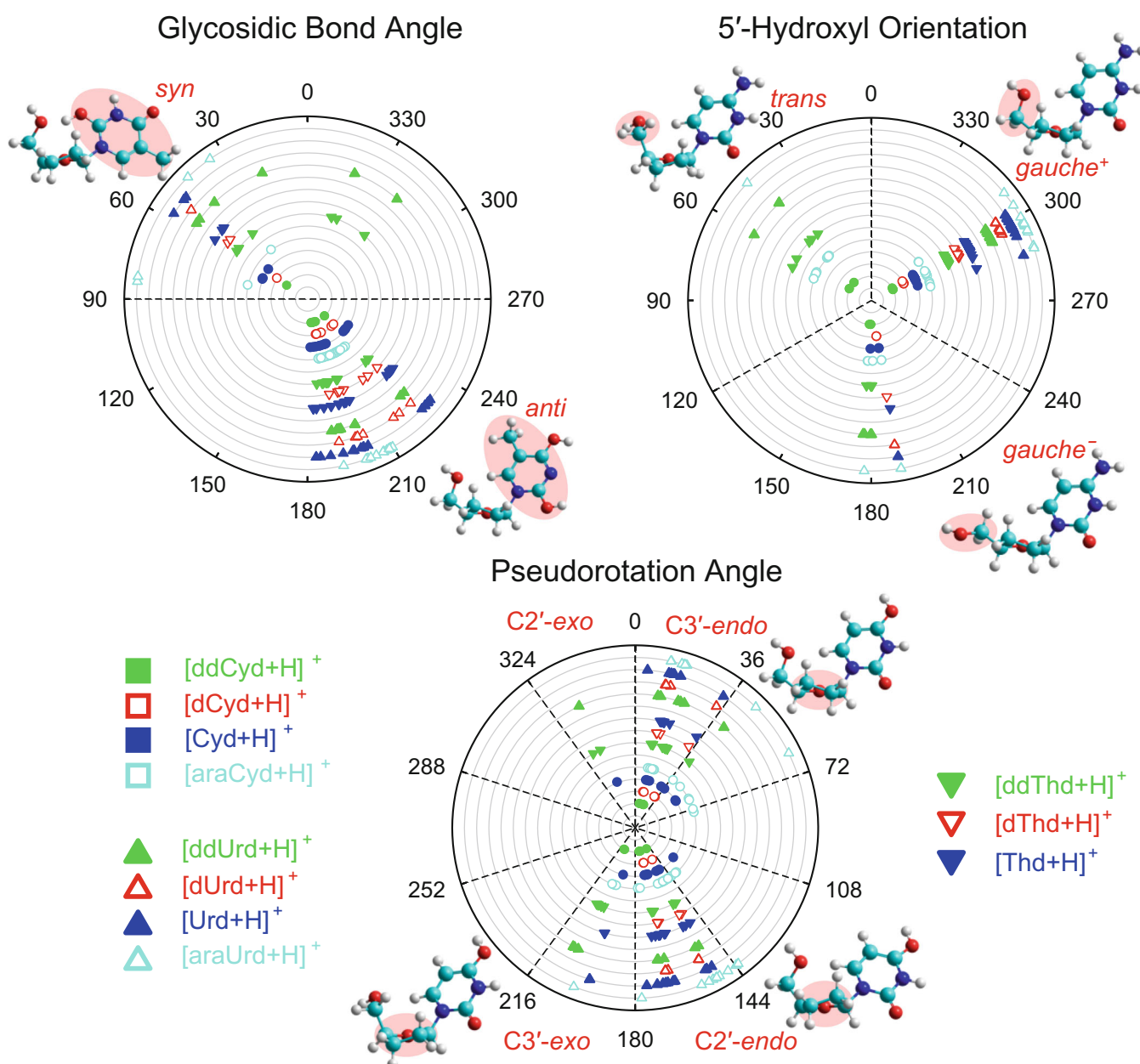


Figure 8. The three primary parameters that govern the structure of the conformers of the protonated pyrimidine nucleoside analogues calculated, glycosidic bond angle, 5'-hydroxyl orientation, and pseudorotation angle (sugar puckering), are displayed for all conformers with relative stabilities predicted within 25 kJ/mol Gibbs energy of the corresponding ground conformers

purine nucleosides, the protonated pyrimidine nucleosides generally adopt highly parallel conformations due to the prevalence of the $C6H_a \cdots O5'_{g+}$ and $O2H^+_{s+} \cdots O5'_{g+}$ nucleobase–sugar hydrogen-bonding interactions and protonation modes that facilitate the latter. Two trends in the RNA analogues can be readily discerned in Figure 8, with $[Nuo+H]^+$ conformers generally having broader distributions around the highly conserved structural parameter values. Also observed for the $[Nuo+H]^+$ species is the subtle shift of one of the *anti* glycosidic bond angle distributions from $\sim 224^\circ$ in $[ddNuo+H]^+$ and $[dNuo+H]^+$ to $\sim 230^\circ$, which locates the O2 atom in these $[Nuo+H]^+$ conformers slightly farther from the 2'-hydroxyl than found in similar $[ddNuo+H]^+$ or $[dNuo+H]^+$ conformers.

Within the most stable *anti* conformers of each species lies a trend between glycosidic bond angle and sugar puckering, conformers with C2'-endo sugar puckering having glycosidic bond angles in the range of ~ 224 – 230° and those with C3'-endo sugar puckering having glycosidic bond angles of $\sim 198^\circ$. This trend is generally broken by conformers stabilized by $O2H \cdots O2'H$ or $O2'H \cdots O2$ nucleobase–sugar hydrogen-bonding interactions as they tend to prefer C2'-endo sugar puckering and glycosidic bond angles of $\sim 190^\circ$. Outliers of $[ddUrd+H]^+$ and $[ddThd+H]^+$ with glycosidic bond angles of $\sim 330^\circ$ display $O2H^+_{s+} \cdots O4'$ hydrogen-bonding interactions, which were not found in the conformers calculated for their protonated DNA and RNA analogues. Similar conformers are

also calculated for $[\text{ddCyd}+\text{H}]^+$ but lie higher in energy and just outside of the Gibbs energy range shown in Figure 8.

Study of the conformations of ddCyd in solution by NMR [19] indicated a preference for N-type ($\text{C2}'\text{-exo}/\text{C3}'\text{-endo}$) sugar puckering, in agreement with these results for $[\text{ddCyd}+\text{H}]^+$. As with the purine nucleosides, crystallographic study of ddCyd indicated a preference for S-type ($\text{C3}'\text{-exo}/\text{C2}'\text{-endo}/\text{C4}'\text{-endo}$) sugar puckering [21]. The crystal structure of ddUrd was found with $\text{C3}'\text{-endo}/\text{C2}'\text{-exo}$ sugar puckering [21]. Agreement between the previous studies of 2',3'-dideoxy nucleoside conformations by NMR and crystallography is sparse, but the gas-phase results for sugar puckering in the protonated 2',3'-dideoxypyrimidines presented here are generally in agreement with those observed for their counterparts in solution by NMR.

The general preference for *gauche*⁺ 5'-hydroxyls offering strong nucleobase–sugar hydrogen-bonding interactions is seen quite clearly, with each system presenting several unique conformers in a relatively confined range of 5'-hydroxyl orientation. 5'-Hydroxyls in the *gauche*[−] orientation are observed for each system as well, generally amongst the less stable conformers in the Gibbs energy range displayed. As also observed for $[\text{ddGuo}+\text{H}]^+$, these *gauche*[−] conformers lack significant nucleobase–sugar or sugar–sugar hydrogen-bonding interactions. Within the pyrimidine analogues, conformers with *trans* 5'-hydroxyl orientations were only computed for the 2',3'-dideoxynucleosides. These conformers are also primarily responsible for the $\text{C3}'\text{-exo}$ puckered conformers observed. $\text{C3}'\text{-exo}$ puckered conformers were also found for the protonated RNA species with *gauche*⁺ 5'-hydroxyls and the $\text{O2}'\text{H}\cdots\text{O2}_a$ or $\text{O2}_a\text{H}\cdots\text{O2}'$ hydrogen-bonding interaction unique to the RNA species.

Conclusions

IRMPD action spectroscopy of the protonated 2',3'-dideoxyribose analogues of adenosine, guanosine, cytidine, thymidine, and uridine reveal a generally consistent trend in protonation preference and primary nucleobase–sugar intramolecular hydrogen-bonding modes with their DNA and RNA counterparts. Comparison of both the experimentally accessed conformations of these protonated gas-phase nucleosides and their low-energy computed structures begins to reveal the impact of the 2'- and 2',3'-hydroxyls on intrinsic nucleoside structure, namely, the absence of sugar–sugar hydrogen-bonding interactions that systematically reduces the variety of low-energy structures accessible across the RNA to DNA to ddNuo analogue series. In the case of the uridine analogues, the 2'-hydroxyl also impacts protonation preference, with the RNA analogue preferring the minor 2,4-dihydroxyl tautomer over protonation at the O4 position, whereas the DNA and ddNuo analogues prefer O4 protonation over the minor tautomer. This shift is primarily observed in computed energetics as the experimental IRMPD spectra reveal little difference in the site of protonation of the experimentally populated conformers, with both the minor tautomer and O4 protonated conformers

present. Interestingly, preferences in sugar puckering appear to be altered between the protonated RNA/DNA analogues and 2',3'-dideoxy analogues, with the former generally preferring $\text{C2}'\text{-endo}$ sugar puckering, except for the guanosine analogues that prefer $\text{C3}'\text{-endo}$ puckering, and the latter preferring $\text{C3}'\text{-endo}$ sugar puckering across the series. Also observed, specifically for the pyrimidine 2',3'-dideoxynucleoside analogues, is the presence of low-energy structures with *trans* 5'-hydroxyls, in addition to the *gauche*⁺ orientation that is stabilized by the nucleobase–sugar hydrogen-bonding interaction preferred by the RNA, DNA, and ddNuo analogues. These conformers are not thought to be populated in these experiments, but their presence amongst the low-energy conformers computed indicates a greater relative stability for the *trans* 5'-hydroxyl in the absence of other sugar–sugar hydrogen-bonding interactions. Inversion of the stereochemistry at the 2'-position provides a greater impact, introducing a unique sugar–sugar hydrogen-bonding interaction that directly competes with the nucleobase–sugar interactions present in the other analogues, and displays a generally stronger preference for $\text{C2}'\text{-endo}$ sugar puckering. 2'-Stereochemical inversion produces a more notably different set of structural parameters than the absence of the 2'- or 2',3'-hydroxyls, but also helps demonstrate the largely conserved preferences in intrinsic structure, primarily the site of protonation, nucleobase orientation, and intramolecular hydrogen-bonding interactions between nucleosides that differ by the 2'- and 3'-sugar hydroxyls.

Acknowledgements

This work was financially supported by the National Science Foundation, grant numbers OISE-1357887 (for the FEL IRMPD measurements and international travel), DBI-0922819 (for the Bruker amazon ETD quadrupole ion trap mass spectrometer), and CHE-01709789 (for other research costs). The authors gratefully acknowledge the financial support of the FELIX facility by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). L.A.H and Z.J.D. gratefully acknowledge support from Wayne State University Thomas C. Rumble Graduate Fellowships and Z.J.D. also acknowledges support from the Joseph Jasper Scholarship for Graduate Students in Chemistry. Computational resources for this work were provided by Wayne State University C&IT. The assistance of the skilled FELIX staff is immensely appreciated.

References

1. Dickerson, R.E., Drew, H.R., Conner, B.N., Wing, R.M., Fratini, A.V., Kopka, M.L.: The anatomy of A-DNA, B-DNA, and Z-DNA. *Science*. **216**, 475–485 (1982)
2. Murray, L.J.W., Arendall, W.B., Richardson, D.C., Richardson, J.S.: RNA backbone is rotameric. *P Natl Acad Sci USA*. **100**, 13904–13909 (2003)

3. Levitt, M., Warshel, A.: Extreme conformational flexibility of Furanose ring in DNA and RNA. *J. Am. Chem. Soc.* **100**, 2607–2613 (1978)
4. Altona, C., Sundaralingam, M.: Conformational-analysis of sugar ring in nucleosides and nucleotides: new description using concept of pseudorotation. *J. Am. Chem. Soc.* **94**, 8205–8212 (1972)
5. De Clercq, E.: Dancing with chemical formulae of antivirals: a personal account. *Biochem. Pharmacol.* **86**, 711–725 (2013)
6. Sommadossi, J.P.: Nucleoside analogs - similarities and differences. *Clin. Infect. Dis.* **16**, S7–S15 (1993)
7. Mathews, C.K.: DNA synthesis as a therapeutic target: the first 65 years. *FASEB J.* **26**, 2231–2237 (2012)
8. Mackey, J.R., Baldwin, S.A., Young, J.D., Cass, C.E.: Nucleoside transport and its significance for anticancer drug resistance. *Drug Resist. Updat.* **1**, 310–324 (1998)
9. Perry, C.M., Noble, S.: Didanosine - an updated review of its use in HIV infection. *Drugs*. **58**, 1099–1135 (1999)
10. Gray, J.H., Owen, R.P., Giacomini, K.M.: The concentrative nucleoside transporter family, SLC28. *Pflugers Arch. - Eur. J. Physiol.* **447**, 728–734 (2004)
11. Marquez, V.E., Ezzitouni, A., Russ, P., Siddiqui, M.A., Ford, H., Feldman, R.J., Mitsuya, H., George, C., Barchi, J.J.: HIV-1 reverse transcriptase can discriminate between two conformationally locked carbocyclic AZT triphosphate analogues. *J. Am. Chem. Soc.* **120**, 2780–2789 (1998)
12. Pastor-Anglada, M., Felipe, A., Casado, F.J.: Transport and mode of action of nucleoside derivatives used in chemical and antiviral therapies. *Trends Pharmacol. Sci.* **19**, 424–430 (1998)
13. De Clercq, E., Li, G.D.: Approved antiviral drugs over the past 50 years. *Clin. Microbiol. Rev.* **29**, 695–747 (2016)
14. Cullis, P.A., Cushing, R.: Vidarabine encephalopathy. *J. Neurol Neurosurg Ps.* **47**, 1351–1354 (1984)
15. Absalon, M.J., Smith, F.O.: Treatment strategies for pediatric acute myeloid leukemia. *Expert. Opin. Pharmacother.* **10**, 57–79 (2009)
16. Veal, G.J., Barry, M.G., Back, D.J.: Zalcitabine (ddC) phosphorylation and drug-interactions. *Antivir. Chem. Chemother.* **6**, 379–384 (1995)
17. Everaert, D.H., Peeters, O.M., Deranter, C.J., Bleton, N.M., Vanaerschot, A., Herdewijn, P.: Conformational-analysis of substituent effects on the sugar puckering mode and the anti-HIV activity of 2',3'-dideoxypyrimidine nucleosides. *Antivir. Chem. Chemother.* **4**, 289–299 (1993)
18. Nasr, M., Litterst, C., McGowan, J.: Computer-assisted structure-activity correlations of dideoxynucleoside analogs as potential anti-HIV drugs. *Antivir. Res.* **14**, 125–148 (1990)
19. Plavec, J., Koole, L.H., Chattopadhyaya, J.: Structural-analysis of 2',3'-dideoxyinosine, 2',3'-dideoxyadenosine, 2',3'-dideoxyguanosine and 2',3'-dideoxycytidine by 500-MHz H-1-NMR spectroscopy and ab-initio molecular-orbital calculations. *J. Biochem. Biophys. Methods.* **25**, 253–272 (1992)
20. Taylor, E.W., Van Roey, P., Schinazi, R.F., Chu, C.K.: A stereochemical rationale for the activity of anti-HIV nucleosides. *Antivir. Chem. Chemother.* **1**, 163–173 (1990)
21. Vanroey, P., Salemo, J.M., Chu, C.K., Schinazi, R.F.: Correlation between preferred sugar ring conformation and activity of nucleoside analogs against human immunodeficiency virus. *P Natl Acad Sci USA.* **86**, 3929–3933 (1989)
22. 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.* **118**, 14774–14784 (2014)
23. 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.* **119**, 2795–2805 (2015)
24. Wu, R.R., Yang, B., Frieler, C.E., Berden, G., Oomens, J., Rodgers, M.T.: Diverse mixtures of 2,4-dihydroxy tautomers and O4 protonated conformers of uridine and 2'-deoxyuridine coexist in the gas phase. *Phys. Chem. Chem. Phys.* **17**, 25978–25988 (2015)
25. Wu, R.R., Yang, B., Frieler, C.E., Berden, G., Oomens, J., Rodgers, M.T.: N3 and O2 protonated tautomeric conformations of 2'-deoxycytidine and cytidine coexist in the gas phase. *J. Phys. Chem. B.* **119**, 5773–5784 (2015)
26. Wu, R.R., Yang, B., Frieler, C.E., Berden, G., Oomens, J., Rodgers, M.T.: 2,4-Dihydroxy and O2 protonated tautomers of dThd and Thd coexist in the gas phase: methylation alters protonation preferences versus dUrd and Urd. *J. Am. Soc. Mass Spectrom.* **27**, 410–421 (2016)
27. Filippi, Antonello, Frascchetti, Caterina, Rondino, Flaminia, Piccirillo, Susanna, Steinmetz, Vincent, Guidoni, Leonardo, Speranza, Maurizio: *Protonated Pyrimidine Nucleosides Probed by IRMPD Spectroscopy*. *Int. J. Mass Spectrom.* 354–355, 54–61 (2013)
28. Salpin, J.-Y., Scuderi, D.: Structure of protonated thymidine characterized by infrared multiple photon dissociation and quantum calculations. *Rapid Commun. Mass Spectrom.* **29**, 1898–1904 (2015)
29. Ung, H.U., Huynh, K.T., Poutsma, J.C., Oomens, J., Berden, G., Morton, T.H.: Investigation of proton affinities and gas phase vibrational spectra of protonated nucleosides, deoxynucleosides, and their analogs. *Int. J. Mass Spectrom.* **378**, 294–302 (2015)
30. Hamlow, L. A., He, C. C., Devereaux, Zachary J., Roy, H. A., Cunningham, N. A., Soley, Erik O., Berden, G., Oomens, J., Rodgers, M. T.: Gas-phase structures of protonated arabinoside nucleosides. *Int. J. Mass Spectrom.* **438**, 124–134 (2019)
31. 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.* **128**, 517–525 (2006)
32. Polfer, N.C., Oomens, J.: Reaction products in mass spectrometry elucidated with infrared spectroscopy. *Phys. Chem. Chem. Phys.* **9**, 3804–3817 (2007)
33. Oepts, D., van der Meer, A.F.G., van Amersfoort, P.W.: The Free-Electron-Laser user facility Felix. *Infrared Phys. Technol.* **36**, 297–308 (1995)
34. Hamlow, L.A., Zhu, Y., Devereaux, Z.J., Cunningham, N.A., Berden, G., Oomens, J., Rodgers, M.T.: Modified quadrupole ion trap mass spectrometer for infrared ion spectroscopy: application to protonated thiated uridines. *J. Am. Soc. Mass Spectrom.* **29**, 2125–2137 (2018)
35. Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.R., Cheatham, T.E., DeBolt, S., Ferguson, D., Seibel, G., Kollman, P.: AMBER, a computer program for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to elucidate the structures and energies of molecules. *Comput. Phys. Commun.* **91**, 1–41 (1995)
36. Wang, J.M., Wang, W., Kollman, P.A., Case, D.A.: Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph Model.* **25**, 247–260 (2006)
37. Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., Simmerling, C.: ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015)
38. Wang, J.M., Wolf, R.M., Caldwell, J.W., Kollman, P.A., Case, D.A.: Development and testing of a general Amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004)
39. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, Jr., J. A., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, O., Foresman, J. B., Ortiz, J. V., Cioslowski, J., Fox, D. J., *Gaussian 09, Revision A.02*. (2009), Gaussian Inc.: Pittsburgh, PA
40. He, C.C., Hamlow, L.A., Devereaux, Z.J., Zhu, Y., Nei, Y.W., Fan, L., McNary, C.P., Maitre, P., Steinmetz, V., Schindler, B., Compagnon, I., Armentrout, P.B., Rodgers, M.T.: Structural and energetic effects of O2'-ribose methylation of protonated purine nucleosides. *J. Phys. Chem. B.* **122**, 9147–9160 (2018)