**ASMS**

CrossMark

**RESEARCH ARTICLE**

# 2,4-Dihydroxy and O2 Protonated Tautomers of dThd and Thd Coexist in the Gas Phase: Methylation Alters Protonation Preferences versus dUrd and Urd

R. R. Wu,[1] Bo Yang,[1] C. E. Frieler,[1] G. Berden,[2] J. Oomens,[2,3] M. T. Rodgers[1]

[1]Department of Chemistry, Wayne State University, Detroit, MI 48202, USA
[2]FELIX Laboratory, Institute for Molecules and Materials, Radboud University, Toernooiveld 7, 6525 ED, Nijmegen, The Netherlands
[3]van't Hoff Institute for Molecular Sciences, University of Amsterdam, 1090 GD, Amsterdam, The Netherlands
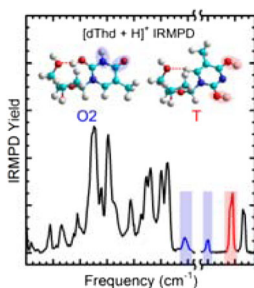
**Abstract.** The gas-phase structures of protonated thymidine, $[dThd + H]^+$, and its modified form, protonated 5-methyluridine, $[Thd + H]^+$, are examined by infrared multiple photon dissociation (IRMPD) action spectroscopy combined with electronic structure calculations. IRMPD action spectra are measured over the ranges extending from ~600 to 1900 $cm^{-1}$ and ~2800 to 3800 $cm^{-1}$ using the FELIX free electron laser and an optical parametric oscillator/amplifier (OPO/OPA) laser system, respectively. Comparisons between the B3LYP/6-311+G(d,p) linear IR spectra calculated for the stable low-energy conformers and the measured IRMPD spectra are used to determine the most favorable tautomeric conformations of $[dThd + H]^+$ and $[Thd + H]^+$ and to identify those populated in the experiments. Both B3LYP and MP2 levels of theory predict a minor 2,4-dihydroxy tautomer as the ground-state conformer of $[dThd + H]^+$ and $[Thd + H]^+$ indicating that the 2'-hydroxyl substituent of Thd does not exert a significant impact on the structural features. $[dThd + H]^+$ and $[Thd + H]^+$ share parallel IRMPD spectral profiles and yields in both the FELIX and OPO regions. Comparisons between the measured IRMPD and calculated IR spectra suggest that minor 2,4-dihydroxy tautomers and O2 protonated conformers of $[dThd + H]^+$ and $[Thd + H]^+$ are populated in the experiments. Comparison of this work to our previous IRMPD spectroscopy study of protonated 2'-deoxyuridine and uridine suggests that the 5-methyl substituent alters the preferences of O2 versus O4 protonation.

**Keywords:** 5-Methyluridine, Electronic structure theory, IRMPD action spectroscopy, Protonation, Thymidine

## Introduction

Thymine (5-methyluracil, T) is one of the naturally occurring DNA pyrimidine nucleobases. Thymine has drawn a great deal of attention due to its enhanced photoreactivity, which results from its slow dissipation of electronic energy upon optical excitation [1, 2]. Thymine absorbs ultraviolet light to yield a cyclobutane photodimer that may lead to carcinogenic and mutagenic effects [3]. Therefore, studies of the photochemistry of thymine may provide insight into the nature

of UV-induced carcinogenesis and mutagenesis [4]. Another intriguing characteristic of thymine is its ability to undergo keto-enol tautomerization, which may alter its base pairing properties, form mispairs, and eventually lead to point mutations and molecular based diseases [5]. The tautomeric behavior of thymine has been examined both experimentally and theoretically [6–9]. These studies suggest that the canonical diketo-tautomer of neutral thymine predominates, whereas upon protonation, keto-enol tautomerization takes place rapidly, and the minor 2,4-dihydroxy tautomer is predominant in the population with the O4 protonated canonical tautomer also present in low abundance. Thymine has been found to be involved in conventional *i*-motif structures containing protonated cytosine (C) base pairs, C·C$^+$. The *i*-motif structures may incorporate T·T pairs to form various *i*-motif tetramers, such as $[d(TCCCCC)]_4$, $[d(TCC)]_4$, $[d(5mCCT)]_4$, $[d(T5mCC)]_4$,

*Correspondence to:* M. T. Rodgers; *e-mail:* mrodgers@chem.wayne.edu

[d(CCTCC)]$_4$, or [d(5mCCTCC)]$_4$ [10] (5m = 5-methyl). Thymine also participates in T·A·T triplets (A = adenine) to form a DNA asymmetric triple helix via Hoogsteen hydrogen-bonding interactions [11].

The photochemical properties and the tautomeric behavior of thymine have motivated a great deal of experimental and theoretical studies during the past two decades [6–9, 12–14]. However, much less attention has been paid to the thymine nucleosides, thymidine (dThd) and its ribonucleoside counterpart, 5-methyluridine (Thd). Because Thd is not one of the naturally occurring canonical nucleosides, studies of isolated Thd are even more limited [15, 16]. Fragmentation of gas-phase dThd has been investigated via electron impact [17], and valence/core photoionization [18]. Structural aspects of dThd have been investigated via theoretical calculations in the presence of the first hydration shell [19]. Therefore, the gas-phase conformation(s) of the isolated nucleosides remain unknown. The intrinsic gas-phase conformations of the protonated forms of dThd and Thd provide insight into the effects of protonation on the local structure (i.e., the nucleobase orientation relative to the glycosidic bond and sugar puckering), which are the two dominant factors that determine the overall structures of DNA and RNA [20]. Therefore, determination of the structural consequences of protonation of the nucleosides may help elucidate the effects that pH changes induce in complex macromolecular structures of DNA and RNA. Because keto-enol tautomerization prevails in isolated protonated thymine, the influence of the 2'-deoxyribose and ribose moieties on the tautomeric conformation(s) of protonated thymine can be determined by comparing the structures of protonated thymine to its nucleoside analogues.

In this work, the gas-phase conformations and relative stabilities of protonated thymidine, [dThd + H]$^+$, and protonated 5-methyluridine, [Thd + H]$^+$, are examined by infrared multiple photon dissociation (IRMPD) action spectroscopy techniques and electronic structure calculations. A Fourier transform ion cyclotron resonance mass spectrometer is coupled to the FELIX free electron laser or an OPO laser to measure the IRMPD action spectra in the IR fingerprint and hydrogen-stretching regions. The vibrational modes, preferred tautomeric conformations of the thymine moiety, and the stable gas-phase conformations of [dThd + H]$^+$ and [Thd + H]$^+$ are determined based on comparisons between the measured IRMPD spectra and theoretically computed linear IR spectra. Comparisons between the conformations and the IRMPD spectra of [dThd + H]$^+$ and [Thd + H]$^+$ elucidate the effect of the 2'-hydroxyl substituent. Comparison of present results to the previous IRMPD study of protonated thymine [9] elucidates the influence of the sugar moieties on the tautomeric conformations and the IR signatures of protonated thymine. Moreover, methylation is found to play a significant role in various biological processes [21], and is epigenetically natural to living organisms for gene expression. Methylation may also induce negative effects when it arises from external chemicals such that the sequence may be misread upon DNA transcription, leading to undesirable genetic codes. dThd and Thd are the 5-methyl derivatives of 2'-deoxyuridine (dUrd) and uridine (Urd), respectively. Therefore, comparisons between this work and the previous IRMPD study of [dUrd + H]$^+$ and [Urd + H]$^+$ [22] provide insight into the effects of methylation on the preferred tautomeric conformations of these nucleosides. Most recently, and virtually simultaneous with this work, Salpin et al. [23] have characterized the gas-phase conformations of [dThd + H]$^+$ by IRMPD spectroscopy in the IR fingerprint region using CLIO free electron laser. Comparison of the present results to those just reported by Salpin et al. show that the general conclusions remain valid. However, the additional information provided in the current work for [Thd + H]$^+$ as well as the spectral data in the hydrogen-stretching region allow the number of conformations that may be populated in the experiments to be reduced as the IR signatures in the hydrogen-stretching region are somewhat more discriminating. Thus, the present work complements and enhances that reported by Salpin et al.

## Experimental and Computational

### Mass Spectrometry and Photodissociation

IRMPD action spectra of [dThd + H]$^+$ and its modified form [Thd + H]$^+$ were measured using a 4.7 T Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) [24–26]. Photodissociation was induced by the widely-tunable free electron laser for infrared experiments (FELIX) [27] and an OPO/OPA laser system to examine the IR fingerprint and hydrogen-stretching regions, respectively. Approximately 1 mM of the nucleosides, purchased from Sigma-Aldrich (Zwijndrecht, the Netherlands), and 10 mM hydrochloric acid were dissolved in 50%:50% MeOH/H$_2$O solutions. The solutions were delivered to a Micromass "Z-spray" electrospray ionization (ESI) source at a flow rate in the range of 3.8−8.0 μL/min. [dThd + H]$^+$ or [Thd + H]$^+$ were accumulated in an rf hexapole ion trap for several seconds to affect efficient thermalization of the ions prior to pulsed extraction through a quadrupole deflector. [dThd + H]$^+$ or [Thd + H]$^+$ were injected into the FT-ICR MS via a 1-m long rf octopole ion guide. Electrostatic switching of the dc bias of the octopole allows ion capturing without collisional heating of the ions [25]. Although significant heating of the ions during transmission from the hexapole to the ICR cell was not expected, the ions were stored in the ICR cell at ~10$^{-8}$ Torr for at least 300 ms to enable any excess internal energy to be dissipated by radiative emission. [dThd + H]$^+$ and [Thd + H]$^+$ were isolated using stored waveform inverse Fourier transform (SWIFT) techniques. The selected ions were irradiated for 2.5−3 s by the free electron laser to induce IR photodissociation over the wavelength range between ~16.7 μm (~600 cm$^{-1}$) and ~5.3 μm (~1900 cm$^{-1}$), and in separate experiments for 5–10 s using a benchtop optical parametric oscillator/amplifier (OPO/OPA) system over the hydrogen-stretching region, from ~3.6 μm (~2800 cm$^{-1}$) to ~2.6 μm (~3800 cm$^{-1}$). The IRMPD features measured between ~2800 and 3300 cm$^{-1}$ exhibit very low yields and extensive broadening such that this

region is of poor diagnostic quality. Therefore, this region is not shown in most figures or discussed further.

## Computational Details

The chemical structures of neutral dThd and Thd are shown in Figure 1. In these structures, the thymine nucleobases are depicted in the anti-orientation relative to the glycosidic bond. The favorable states of protonation for dThd and Thd were investigated including O2, O4, and the minor 2,4-dihydroxy tautomers. Candidate structures for each protonated form of dThd and Thd were generated by simulated annealing using HyperChem software [28] with the Amber 2 force field. Detailed descriptions of the simulated annealing procedures can be found in our previous IRMPD studies of the protonated forms of the DNA and RNA nucleosides of guanine, adenine, and cytosine [29–31]. Similar to these studies [29–31] as well as our previous IRMPD study of the protonated uridine and 2'-deoxyuridine [22] approximately 20 to 30 candidate structures from the 300 generated by simulated annealing for each protonated form were chosen for higher level optimization. The structures chosen were primarily based on their relative stabilities determined by simulated annealing, but also include all possible combinations of protonation site (O2, O4, and the 2,4-dihydroxy tautomers), nucleobase orientation (anti and syn), and sugar puckering (C2'-endo, C3'-endo, C2'-exo, and C3'-



**Figure 1.** Chemical structures of thymidine (dThd) and 5-methyluridine (Thd). Ground-state structures of [dThd + H]$^+$ and [Thd + H]$^+$ predicted at the B3LYP/6-311+G(2d,2p)// B3LYP/6-311+G(d,p) and MP2(full)/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) levels of theory

exo) to ensure a comprehensive examination of conformational space.

Geometry optimizations, frequency analyses, and single point energy calculations of all candidate structures were carried out using the Gaussian 09 suite of programs [32]. All candidate structures were first optimized at the B3LYP/6-31G(d) level of theory to facilitate convergence of the geometry optimization. These structures were re-optimized using the 6-311+G(d,p) basis set to improve the structural and spectral description of these systems and, in particular, the hydrogen-bonding interactions. Frequency analyses of these structures were also performed using the 6-311+G(d,p) basis set. Single point energies were calculated at the B3LYP and MP2 levels of theory using the 6-311+G(2d,2p) basis set to determine the relative stabilities of the low-energy conformers, including zero-point energy (ZPE) and thermal corrections. Vibrational frequencies calculated at the B3LYP/6-311+G(d,p) level of theory are scaled by a factor of 0.98 over the FELIX region, and by a factor of 0.954 over the OPO region [33, 34]. Before comparison to the experimental IRMPD spectra, the calculated vibrational frequencies are broadened using 20 cm$^{-1}$ and 15 cm$^{-1}$ FWHM Gaussian line shapes over the FELIX and OPO regions, respectively.

# Results

## IRMPD Action Spectroscopy

When photodissociation of [dThd + H]$^+$ and [Thd + H]$^+$ is induced by the free electron laser (FEL), the primary dissociation pathway involves $N$-glycosidic bond cleavage, producing protonated thymine, [Thy + H]$^+$ at $m/z$ =127, as the ionic product detected. In addition to the primary photodissociation pathway, two minor products are also observed for [dThd + H]$^+$ at $m/z$ = 207 and 81, corresponding to loss of two H$_2$O molecules and loss of neutral thymine and two H$_2$O molecules, respectively. In contrast, [Thy + H]$^+$ is the only ionic product observed for [dThd + H]$^+$ and [Thd + H]$^+$ when the OPO laser is used due to the lower photon output associated with this setup. The IRMPD yield was determined for each protonated nucleoside, [Nuo + H]$^+$ = [dTh + H]$^+$ or [Thd + H]$^+$, from its intensity and the sum of the intensities of the product ions after laser irradiation at each frequency as shown in Equation 1:

$$\text{IRMPD Yield} = \sum_i \text{I}_{\text{product}_i} \bigg/ \left( \sum_i \text{I}_{\text{product}_i} + \text{I}_{[\text{Nuo}+\text{H}]^+} \right) \quad (1)$$

The IRMPD yield was normalized linearly with laser power to correct for changes in the laser power as a function of photon energy for both the fingerprint and hydrogen-stretching regions. IRMPD action spectra were obtained for [dThd + H]$^+$ and [Thd + H]$^+$ over the IR fingerprint region from ~600 to 1900 cm$^{-1}$, and the hydrogen-stretching region from ~2800 to 3800 cm$^{-1}$, and are shown in Figure 2. Although minor differences are seen, the 2'-hydroxyl substituent does not exert a
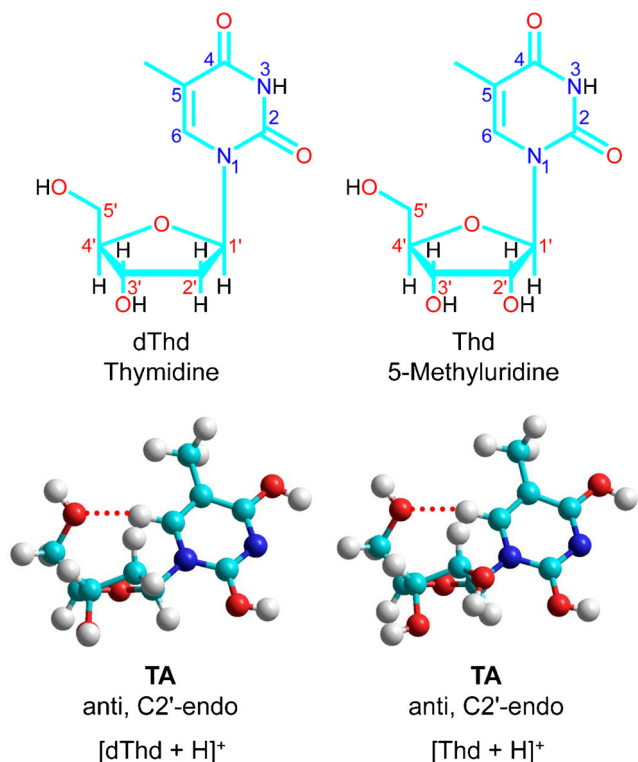
**Figure 2.** Infrared multiple photon dissociation (IRMPD) action spectra of [dThd + H]$^+$ and [Thd + H]$^+$ in the FELIX and OPO regions

significant influence on the measured IRMPD spectral features of [dThd + H]$^+$ versus [Thd + H]$^+$ as seen in the highly parallel spectral profiles and IRMPD yields in both regions.

## Theoretical Results

The ground-state conformers of [dThd + H]$^+$ and [Thd + H]$^+$ calculated at the B3LYP/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) and M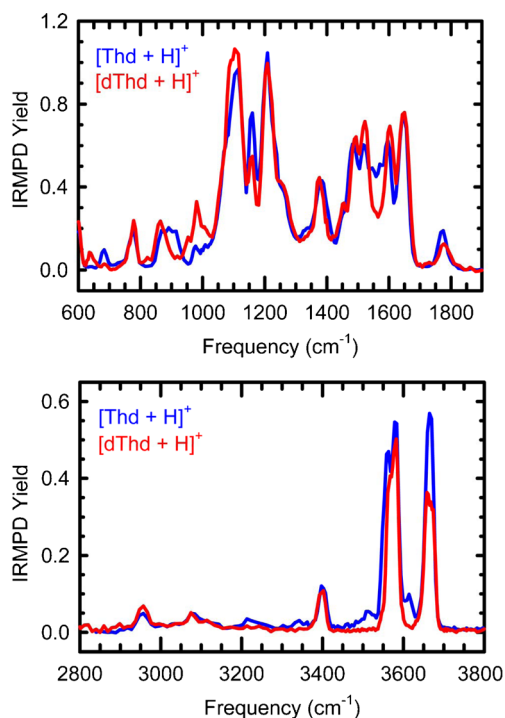P2(full)/6-311+G(2d,2p)//B3LYP/6-311+G(d,p) levels of theory are shown in Figure 1. Both B3LYP and MP2 predict the same ground-state structures for [dThd + H]$^+$ and [Thd + H]$^+$. The 2'-hydroxyl substituent does not affect the conformational features of the ground-state conformer of [Thd + H]$^+$ compared with those of [dThd + H]$^+$ except that the 3'-hydroxyl of [Thd + H]$^+$ rotates to enable formation of a hydrogen-bonding interaction between the 2'- and 3'-hydroxyl substituents. Protonation of dThd and Thd results in preferential stabilization of a minor 2,4-dihydroxy tautomer where thymine is in an anti orientation relative to the glycosidic bond, both the 2- and 4-hydroxyl hydrogen atoms are oriented toward the adjacent N3 atom, and the sugar exhibits C2'-endo puckering.

Table 1 lists the relative 0 and 298 K enthalpies and Gibbs free energies of the low-energy conformers of [dThd + H]$^+$ and [Thd + H]$^+$ found. The optimized structures of these low-energy conformations along with their relative free energies at 298 K are shown in Figures S1 and S2 of the Supporting Information. The nomenclature used to describe the low-energy conformers is based on the tautomeric conformation

or site of protonation. **T** is used for the 2,4-dihydroxy tautomers, **O2** or **O4** for the canonical tautomers protonated at these sites, and followed by a letter, and are ordered based on their B3LYP 298 K Gibbs free energies. Uppercase letters are used for conformers found for both [dThd + H]$^+$ and [Thd + H]$^+$, lowercase letters for conformers that are only found for [dThd + H]$^+$, and lowercase Roman numerals for conformers that are only found for [Thd + H]$^+$. Parallel conformations of [dThd + H]$^+$ and [Thd + H]$^+$ are named in the same fashion even though the absolute relative stabilities may vary. A greater number of low-energy conformers are found for [Thd + H]$^+$ than [dThd + H]$^+$ because the 2'-hydroxyl substituent increases the opportunities for hydrogen-bonding interactions with the nucleobase and allows multiple favorable orientations of the 2'- and 3'-hydroxyl substituents to exist. The lack of a 2'-hydroxyl substituent enables the **O2a** and **O2b** conformers to exist as stable conformers of [dThd + H]$^+$, whereas the analogous conformers of [Thd + H]$^+$ are not found because when the O2 proton is pointed away from the adjacent N3 atom, it can interact with the 2'-hydroxyl substituent leading to the **O2i** conformer. A detailed description of the low-energy conformers shown in Figures S1 and S2 is given in the Supporting Information.

### 2,4-Dihydroxy Tautomers

The most stable 2,4-dihydroxy tautomers of both [dThd + H]$^+$ and [Thd + H]$^+$, **TA**, prefer the anti-orientation of the nucleobase with both the 2- and 4-hydroxyl hydrogen atoms pointed toward N3, and C2'-endo sugar puckering. The **Ti** conformer of [Thd + H]$^+$ also exhibits the anti-orientation of the nucleobase, but is preferentially stabilized by rotating the 2'-hydroxyl hydrogen atom to enable an O2H···O2'H···O3'H dual hydrogen-bonding interaction and C2'-endo sugar puckering. **Ti** lies only 0.1 kJ/mol (both B3LYP and MP2) above **TA**. For both [dThd + H]$^+$ and [Thd + H]$^+$, the **TC** and **TD** conformers adopt the syn orientation of the nucleobase and are stabilized by either an O2H···O5' or O5'H···O2 hydrogen-bonding interaction. They lie higher in free energy than the anti-oriented **TA**. The relative stabilities of **TC** and **TD** suggest that the O2H···O5' hydrogen-bonding interaction is > 20 kJ/mol more favorable than the O5'H···O2 hydrogen-bonding interaction, which is easily understood based on the additional stabilization provided by ion-dipole and ion-induced dipole interactions in the former.

### O2 Protonation

The most stable O2 protonated conformers of both [dThd + H]$^+$ and [Thd + H]$^+$, **O2A**, prefer a syn orientation of the nucleobase with the 2-hydroxyl proton hydrogen bonded to O5' and C2'-endo sugar puckering. In contrast, **O2B**, **O2a**, **O2C**, and **O2b** of [dThd + H]$^+$, which all exhibit the anti-orientation of the nucleobase are >35 kJ/mol less favorable than **O2A**. The **O2i** conformer of [Thd + H]$^+$ adopts the anti-orientation of the nucleobase with the excess proton rotated away from N3 and hydrogen bonded to the 2'-hydroxyl oxygen atom, enabling

**Table 1.** Relative Enthalpies and Free Energies at 0 and 298 K in kJ/mol of Stable Low-Energy Conformers of [dThd + H]$^+$ and [Thd + H]$^{+a}$

| Conformer | Nucleobase orientation | Sugar puckering | B3LYP | | | MP2(full) | | |
|---|---|---|---|---|---|---|---|---|
| | | | $\Delta H_0$ | $\Delta H_{298}$ | $\Delta G_{298}$ | $\Delta H_0$ | $\Delta H_{298}$ | $\Delta G_{298}$ |
| [dThd + H]$^+$ | | | | | | | | |
| **TA** | anti | C2'-endo | **0.1** | **1.5** | **0.0** | **1.0** | **2.5** | **0.0** |
| **TB** | anti | C3'-endo | **0.5** | **2.2** | **0.7** | **2.7** | **4.4** | **2.0** |
| **TC** | syn | C2'-endo | **0.0** | **0.0** | **3.0** | **0.0** | **0.0** | **2.1** |
| **O4A** | anti | C3'-endo | 3.9 | 6.0 | 3.3 | 13.7 | 15.8 | 12.2 |
| **O4B** | anti | C2'-endo | 5.6 | 7.6 | 3.8 | 15.0 | 17.0 | 12.2 |
| **O2A** | syn | C2'-endo | **1.4** | **1.5** | **4.1** | **3.7** | **3.8** | **5.5** |
| **O4C** | anti | C3'-endo | 16.2 | 18.5 | 15.3 | 26.3 | 28.6 | 24.4 |
| **O4D** | anti | C2'-endo | 17.4 | 19.6 | 15.5 | 26.7 | 28.9 | 23.9 |
| **O4E** | anti | C3'-endo | 23.2 | 25.5 | 20.3 | 35.4 | 37.7 | 31.6 |
| **TD** | syn | C2'-endo | **31.0** | **32.8** | **29.9** | **34.0** | **35.7** | **32.0** |
| **O4F** | syn | C2'-endo | 31.2 | 32.8 | 31.3 | 40.6 | 42.2 | 39.8 |
| **O2B** | anti | C2'-endo | 42.0 | 44.3 | 41.5 | 43.8 | 46.1 | 42.3 |
| **O2a** | anti | C2'-endo | 43.4 | 45.6 | 42.6 | 45.2 | 47.3 | 43.4 |
| **O2C** | anti | C3'-endo | 45.9 | 48.8 | 44.2 | 51.4 | 54.3 | 48.7 |
| **O2b** | anti | C3'-endo | 51.3 | 54.2 | 49.0 | 57.3 | 60.2 | 54.1 |
| [Thd + H]$^+$ | | | | | | | | |
| **TA** | anti | C2'-endo | **0.6** | **1.8** | **0.0** | **0.7** | **1.8** | **0.0** |
| **Ti** | anti | C2'-endo | **0.0** | **0.0** | **0.1** | **0.0** | **0.0** | **0.1** |
| **TC** | syn | C2'-endo | **1.0** | **1.0** | **2.2** | **0.8** | **0.9** | **2.1** |
| **TB** | anti | C3'-endo | **3.3** | **4.5** | **2.6** | **3.2** | **4.6** | **2.4** |
| **Tii** | anti | C2'-endo | **5.4** | **6.6** | **5.2** | **4.0** | **5.2** | **3.7** |
| **O2A** | syn | C2'-endo | **4.1** | **4.3** | **5.2** | **6.1** | **6.3** | **7.1** |
| **O2i** | anti | C2'-endo | **5.5** | **5.6** | **5.8** | **7.8** | **7.9** | **8.0** |
| **O4i** | anti | C3'-endo | 6.9 | 8.6 | 6.2 | 13.7 | 15.4 | 12.9 |
| **Tiii** | anti | C3'-endo | **7.7** | **9.2** | **7.3** | **6.6** | **8.1** | **6.1** |
| **O4B** | anti | C2'-endo | 8.9 | 10.7 | 7.3 | 16.5 | 18.3 | 14.8 |
| **O4A** | anti | C3'-endo | 9.9 | 11.7 | 8.5 | 16.6 | 18.4 | 15.1 |
| **O4ii** | anti | C2'-endo | 10.5 | 11.6 | 10.1 | 15.6 | 16.6 | 15.1 |
| **O4iii** | anti | C2'-endo | 15.0 | 16.6 | 13.9 | 21.3 | 23.0 | 20.2 |
| **O4D** | anti | C2'-endo | 20.0 | 21.9 | 18.3 | 27.5 | 29.5 | 25.8 |
| **O4iv** | anti | C3'-endo | 19.6 | 21.6 | 18.6 | 26.5 | 28.5 | 25.4 |
| **O4C** | anti | C3'-endo | 21.9 | 23.9 | 20.3 | 29.2 | 31.2 | 27.5 |
| **O4v** | anti | C2'-endo | 22.9 | 24.2 | 22.3 | 28.0 | 29.3 | 27.3 |
| **O4F** | syn | C2'-endo | 24.8 | 26.1 | 23.4 | 33.8 | 35.2 | 32.4 |
| **O4E** | anti | C3'-endo | 27.4 | 29.3 | 24.1 | 35.8 | 37.8 | 32.5 |
| **O4vi** | anti | C2'-endo | 26.0 | 27.8 | 24.9 | 32.2 | 34.0 | 31.0 |
| **TD** | syn | C2'-endo | **27.3** | **28.7** | **25.1** | **30.2** | **31.6** | **27.9** |
| **O2B** | anti | C2'-endo | 40.4 | 42.6 | 38.7 | 41.3 | 43.5 | 39.6 |
| **O2ii** | anti | C2'-endo | 45.0 | 47.2 | 43.8 | 44.3 | 46.4 | 43.0 |
| **O2C** | anti | C3'-endo | 47.1 | 49.4 | 44.6 | 50.6 | 52.9 | 48.1 |
| **O2iii** | anti | C3'-endo | 55.0 | 57.9 | 52.1 | 56.9 | 59.8 | 54.0 |

[a]Single point energy calculations of B3LYP/6-311+G(d,p) optimized structures are performed at the B3LYP/6-311+G(2d,2p) and MP2(full)/6-311+G(2d,2p) levels of theory including ZPE and thermal corrections.
Note: The stabilities of all conformers that may be populated in the experiments are indicated in boldface.

O2H···O2'H···O3' dual hydrogen-bonding interactions and C2'-endo sugar puckering. **O2i** lies only 0.6 kJ/mol (B3LYP) and 0.9 kJ/mol (MP2) higher in energy than **O2A**. The other O2 protonated conformers of [Thd + H]$^+$ that exhibit an anti-oriented nucleobase, but without the dual hydrogen-bonding interaction, lie >30 kJ/mol higher in free energy than **O2i**.

## O4 Protonation

The most stable O4 protonated conformers of [dThd + H]$^+$ and [Thd + H]$^+$, **O4A** and **O4i**, respectively, exhibit the anti orientation of the nucleobase with the 4-hydroxyl hydrogen atom pointed away from N3 and C3'-endo sugar puckering. For both [dThd + H]$^+$ and [Thd + H]$^+$, when

the 4'-hydroxyl hydrogen atom of the anti-oriented nucleobase points toward the N3 atom, the stability of the conformers is reduced by >10 kJ/mol. The **O4F** conformers of both [dThd + H]$^+$ and [Thd + H]$^+$ that adopt a syn orientation of the nucleobase with the 5'-hydroxyl hydrogen atom oriented toward O2 to form an O5'H···O2 hydrogen-bonding interaction are >15 kJ/mol less favorable than the most stable anti-oriented O4 protonated conformers.

## Effects of the 2'-Hydroxyl Substituent

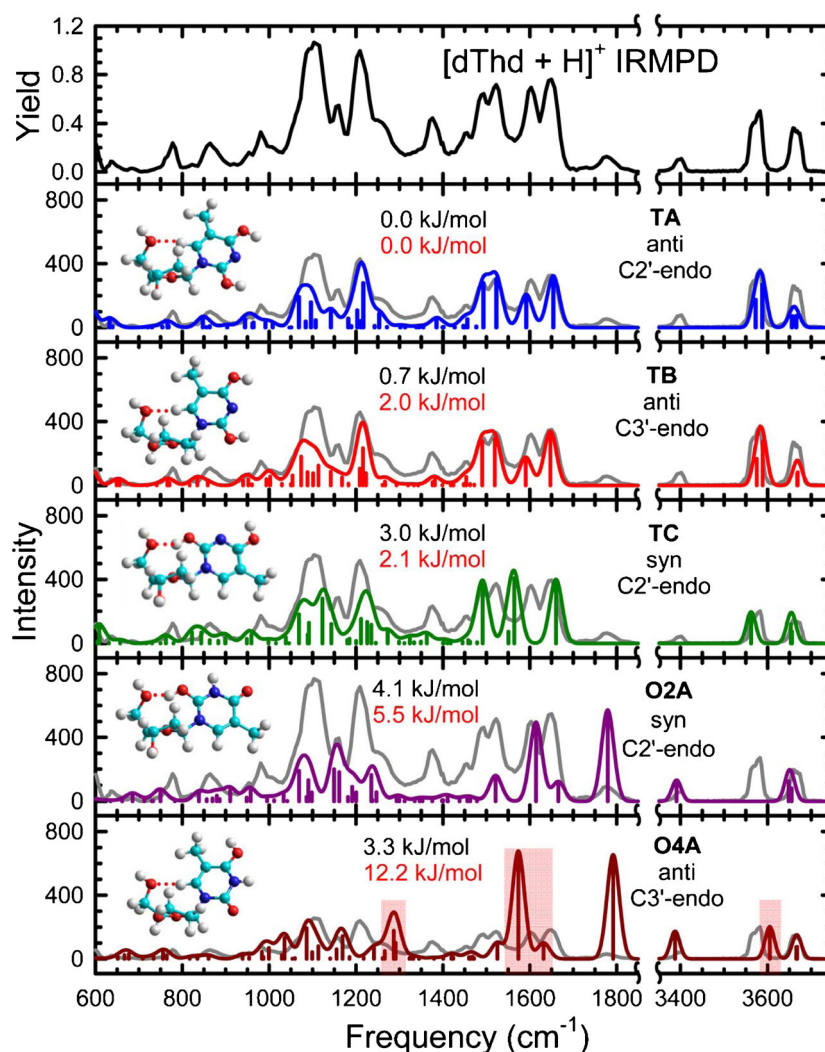The presence of the 2'-hydroxyl substituent allows a hydrogen-bonding interaction with the 3'-hydroxyl

**Figure 3.** Comparison of the measured IRMPD action spectrum of [dThd + H]$^+$ with the theoretical linear IR spectra for the select low-energy conformers of [dThd + H]$^+$ and the corresponding optimized structures calculated at the B3LYP/6-311+G(d,p) level of theory. Also shown are the B3LYP/6-311+G(2d,2p) (in black) and MP2(full)/6-311+G(2d,2p) (in red) relative Gibbs free energies at 298 K. The state of protonation, nucleobase orientation, and sugar puckering are also indicated for each conformer. The measured IRMPD spectrum is overlaid with each theoretical spectrum and scaled to match the intensities of the computed bands in the FELIX and OPO regions separately to facilitate comparisons

substituent, and thus provides greater opportunities for the low-energy conformers to exhibit various favorable orientations of the 2'- and 3'-hydroxyl substituents (i.e., the **Tii**, **O4i**, **Tiii**, **O4iii**, **O4iv**, **O4vi**, **O2ii**, and **O2iii** conformers. In addition, the O2H···O2'H···O3' and O3'H···O2'H···O2 dual hydrogen-bonding interactions are enabled by the presence of the 2'-hydroxyl substituent. Both dual hydrogen-bonding interactions stabilize the anti-oriented nucleobase and C2'-endo sugar puckering (i.e., the **Ti**, **O2i**, **O4ii**, and **O4v** conformers. Similar effects of the 2'-hydroxyl substituent on the conformation are also found for the protonated forms of the DNA and RNA nucleosides of guanine, adenine, and cytosine [29–31]. Conversely, our previous study of the protonated forms of uridine and 2'-deoxyuridine [22] reveals the effects of the 2'-deoxy modification. In particular, this modification reduces the

number of favorable orientations of the free 3'-hydroxyl substituent and eliminates the formation of the dual O2H···O2'H···O3' and O3'H···O2'H···O2 hydrogen-bonding interactions.

## Discussion

### Comparison of the Measured IRMPD and Computed IR Spectra of [dThd + H]$^+$.

The experimental IRMPD and theoretical IR spectra of the low-energy conformers, **TA**, **TB**, **TC**, **O2A**, and **O4A** of [dThd + H]$^+$ in the FELIX and OPO regions are compared in Figure 3. The calculated IR spectra of these **T** and **O2** conformers complement each other and exhibit very good agreement with the measured IRMPD spectra. In

particular, the bands observed at ~1785 and ~3395 cm$^{-1}$ are contributed only by the **O2A** conformer, whereas the band at ~3580 cm$^{-1}$ and its shoulder to the red at ~3565 cm$^{-1}$ are contributed only by the **T** conformers, and all other bands arise from both the **T** and **O2** conformers. The calculated band at ~1565 cm$^{-1}$ of the **TC** conformer may contribute to the nonzero valley between the bands measured at ~1605 and ~1520 cm$^{-1}$. In contrast, the calculated IR spectrum of **O4A** exhibits obvious discrepancies with the measured spectrum. These discrepancies are highlighted in the comparison. In particular, the strikingly strong band predicted at ~1575 cm$^{-1}$ for **O4A** would be expected to broaden the bands observed at ~1605 and ~1520 cm$^{-1}$ if they were populated in measurable abundance. The bands predicted at ~3605 cm$^{-1}$ and ~1290 cm$^{-1}$ for **O4A** are not observed in the measured spectrum. Also, the band calculated at ~1630 cm$^{-1}$ for **O4A** is shifted to a lower frequency relative to the measured band at ~1650 cm$^{-1}$. Therefore, it can be concluded that **O4A** is not populated by ESI. Comparisons between the measured IRMPD and calculated IR spectra for the other low-energy conformers computed are shown in Figures S3 and S4 and are discussed in the Supporting Information. In particular, similar to **O4A**, the disagreement between the calculated IR spectra of the other O4 protonated conformers and the measured IR spectrum eliminates their presence in the experimental population. The calculated IR spectrum of **TD** exhibits good agreement with the measured spectrum, but its calculated relative stability suggests that if it is present, its population is likely negligible.

In summary, O4 protonated conformers of [dThd + H]$^+$ are clearly not populated in the experiments. The 2,4-dihydroxy tautomers, **TA**, **TB**, and **TC** are populated in the experiments. **TD** may be present, but likely in negligible abundance. **O2A** also contributes to the experimental population.

## Resonant Vibrational Modes of [dThd + H]$^+$

Comparison of the IRMPD spectrum with the calculated IR spectra of the conformers populated allows vibrational assignments to be made. In the FELIX region, the small IR absorption measured at ~1785 cm$^{-1}$, contributed by **O2A**, represents O4 carbonyl stretching. The two bands measured at ~1650 and ~1605 cm$^{-1}$ primarily represent coupled C2–N3 and C5=C6 stretching, and coupled N1–C2 and C4–C5 stretching, respectively. The bands measured at ~1520 and ~1495 cm$^{-1}$ represent C2–O2 and C4–O4 stretching, respectively. The weak band measured at ~1375 cm$^{-1}$ arises from bending of the hydrogen atoms of the sugar moiety. The measured band at ~1210 cm$^{-1}$, primarily arising from the 2,4-dihydroxy tautomers, represents coupled O2−H and O4−H bending. The band measured at ~1095 cm$^{-1}$ represents stretching of the sugar ring. In the OPO region, the measured band at ~3665 cm$^{-1}$ represents coupled O3'−H and O5'−H stretching. The intense IR

absorption measured at ~3580 cm$^{-1}$ and its shoulder to the red at ~3565 cm$^{-1}$, arising from the 2,4-dihydroxy tautomers, represent strong O2−H and O4−H stretching, respectively. The weak band measured at ~3395 cm$^{-1}$, contributed by **O2A**, represents N3−H stretching.

## Comparison of the Measured IRMPD and Computed IR Spectra of [Thd + H]$^+$

The experimental IRMPD and theoretical IR spectra of the low-energy conformers, **TA**, **TC**, **TB**, **O2A**, and **O4A** of [Thd + H]$^+$, in the FELIX and OPO regions are compared in Figure 4. Similar to [dThd + H]$^+$, the calculated IR spectra of these **T** and **O2** conformers complement each other and exhibit very good agreement with the measured IRMPD spectrum. In particular, the calculated IR bands of the ground-state conformer, **TA**, best reproduce the measured IR bands in the regions of ~1400–1700 cm$^{-1}$ and ~1000–1300 cm$^{-1}$. The calculated bands at ~3580 cm$^{-1}$ for **TA** and **TB** and the calculated band at ~3560 cm$^{-1}$ for **TC** may contribute to the band observed at ~3580 cm$^{-1}$ and its strong shoulder to the red at ~3560 cm$^{-1}$, respectively. In contrast, the bands observed at ~1775 and ~3395 cm$^{-1}$ can only be attributed to the presence of the **O2A** conformer. The calculated band at ~1565 cm$^{-1}$ for **TC** may contribute to the broadening between the measured bands at ~1595 and ~1520 cm$^{-1}$. In contrast, the highlighted bands predicted for **O4A** exhibit disagreement with the measured spectrum. In particular, the bands predicted at ~3600 and ~1800 cm$^{-1}$ are shifted to higher frequencies relative to the measured bands at ~3580 and ~1775 cm$^{-1}$, respectively. The band predicted at ~1630 cm$^{-1}$ is shifted to a lower frequency relative to the measured band at ~1650 cm$^{-1}$. The bands predicted at ~1570, ~1275, ~1040, and ~990 cm$^{-1}$ are too intense relative to the respective measured bands. Therefore, **O4A** is not populated by ESI. The experimental IRMPD and theoretical IR spectra of the low-energy conformers, **Ti**, **Tii**, **O2i**, and **Tiii** of [Thd + H]$^+$, in the FELIX and OPO regions are compared in Figure 5. In particular, the weak band observed at ~3610 cm$^{-1}$ is contributed by **Tii**, and the weak band measured at ~3510 cm$^{-1}$ is contributed by **Ti** and **O2i**. The bands observed at ~1775 and ~3395 cm$^{-1}$ are also contributed by **O2i**. The calculated IR spectra of **Tii** and **Tiii** are highly parallel to those of **TA** and **TB**, respectively, indicating that the changes in the orientations of the sugar hydroxyls do not exert a significant impact on the spectral features. Therefore, **TA**, **TC**, **TB**, and **O2A** of [Thd + H]$^+$, the conformations analogous to those of [dThd + H]$^+$ populated by ESI, may also be present in the experiments. In addition, the minor 2,4-dihydroxy tautomers, **Ti**, **Tii**, and **Tiii**, as well as the **O2i** conformers of [Thd + H]$^+$ may also contribute to the experimental population. Comparisons between the measured IRMPD and calculated IR spectra for the other low-energy conformers computed are shown in
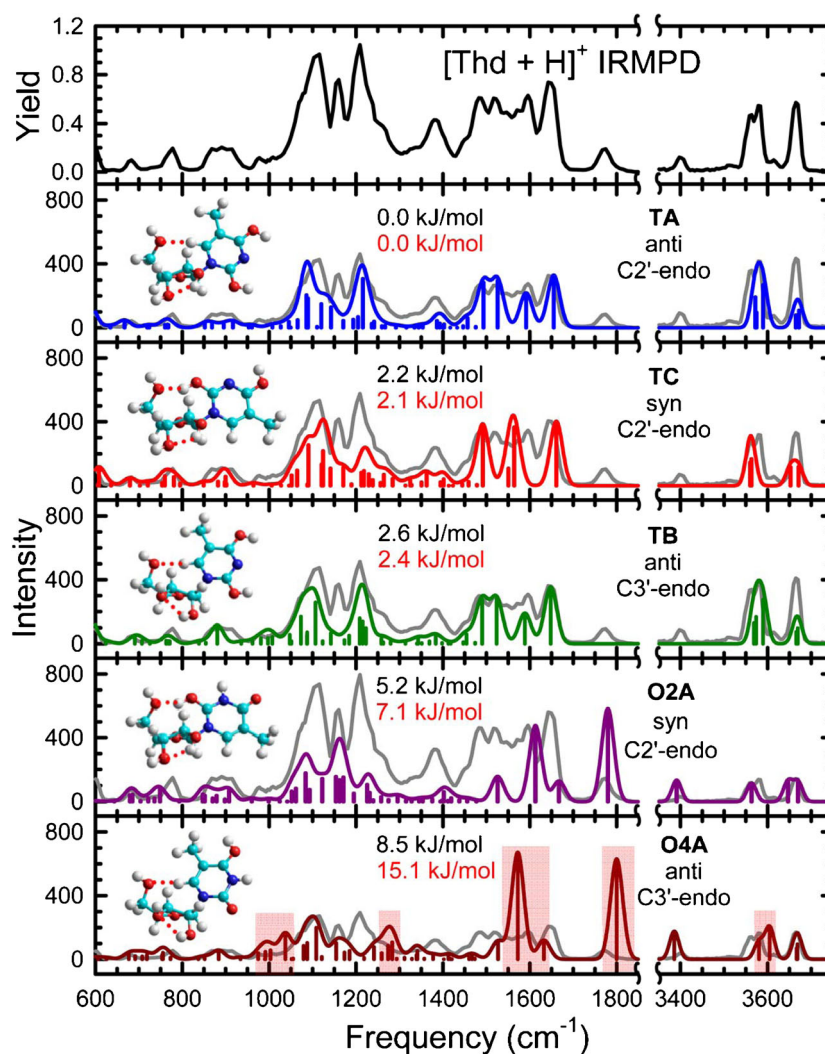
**Figure 4.** Comparison of the measured IRMPD action spectrum of [Thd + H]$^+$ with the theoretical linear IR spectra for the select low-energy conformers of [Thd + H]$^+$ that are analogous to the conformations shown in Figure 3 for [dThd + H]$^+$ and the corresponding optimized structures calculated at the B3LYP/6-311+G(d,p) level of theory. Also shown are the B3LYP/6-311+G(2d,2p) (in black) and MP2(full)/6-311+G(2d,2p) (in red) relative Gibbs free energies at 298 K. The state of protonation, nucleobase orientation, and sugar puckering are also indicated for each conformer. The measured IRMPD spectrum is overlaid with each theoretical spectrum and scaled to match the intensities of the computed bands in the FELIX and OPO regions separately to facilitate comparisons

Figures S5 and S6 and are discussed in the Supporting Information. In particular, various bands predicted for the O4 protonated conformers in the regions of ~1500–1700 cm$^{-1}$ and ~3550–3650 cm$^{-1}$ exhibit discrepancies with the measured IR bands and provide evidence that excludes the O4 protonated conformers from the experimental population. Similar to [dThd + H]$^+$, the **TD** conformer of [Thd + H]$^+$ cannot be ruled out based on its calculated IR spectrum. However, its calculated relative stability suggests that if it is populated, it is likely present in negligible abundance.

In summary, O4 protonated conformers of [Thd + H]$^+$ are not populated in the experiments. Instead, a diverse mixture of minor 2,4-dihydroxy tautomers, **TA**, **Ti**, **TC**, **TB**, **Tii**, and **Tiii**, are present in the experiments. They exhibit various combinations of nucleobase orientations and sugar puckering. **TD** may also be populated, but is likely of very low abundance. The O2 protonated conformers, **O2A** and **O2i**, are also present in the experiments. They adopt different nucleobase orientations, but both exhibit C2'-endo sugar puckering.

*Vibrational Assignments of [Thd + H]$^+$*

Comparison of the IRMPD spectrum with the calculated IR spectra of the conformers of [Thd + H]$^+$ that may be populated in the experiments allows vibrational assignments to be made. In the FELIX region, the small IR band measured at ~1775 cm$^{-1}$ reflects O4-carbonyl stretching from the **O2A** and **O2i** conformers. The measured IR features at ~1650, ~1595, ~1520, and ~1490 cm$^{-1}$ represent

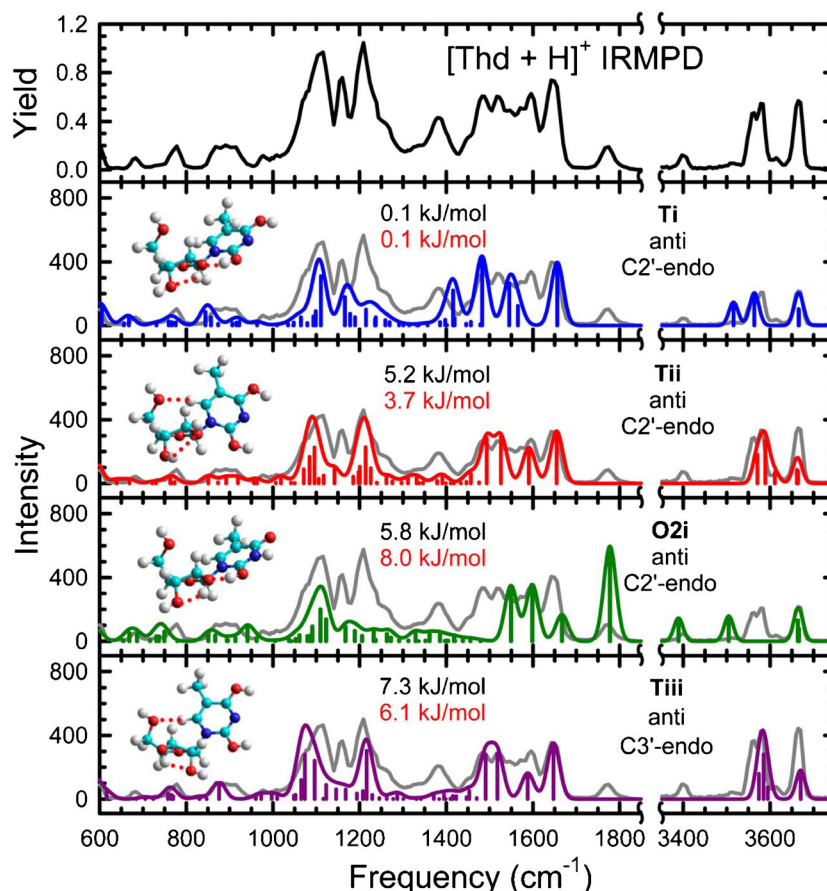**Figure 5.** Comparison of the measured IRMPD action spectrum of [Thd + H]⁺ with the theoretical linear IR spectra for additional low-energy conformers of [Thd + H]⁺ that may also be populated in the experiments and the corresponding optimized structures calculated at the B3LYP/6-311+G(d,p) level of theory. Also shown are the B3LYP/6-311+G(2d,2p) (in black) and MP2(full)/6-311+G(2d,2p) (in red) relative Gibbs free energies at 298 K. The state of protonation, nucleobase orientation, and sugar puckering are also indicated for each conformer. The measured IRMPD spectrum is overlaid with each theoretical spectrum and scaled to match the intensities of the computed bands in the FELIX and OPO regions separately to facilitate comparisons

coupled C2–N3 and C5=C6 stretching, coupled N1–C2 and C4–C5 stretching, C2−O2 stretching, and C4–O4 stretching respectively. The IR band measured at ~1390 cm⁻¹ represents bending motions of the ring hydrogen atoms. The IR band measured at ~1210 cm⁻¹ represents coupled O2–H and O4–H bending. The IR band measured at ~1160 cm⁻¹ represents bending of the sugar hydrogen atoms coupled with O2–H and O4–H bending. The IR band measured at ~1115 cm⁻¹ arises from ring breathing of the sugar. In the OPO region, the band measured at ~3665 cm⁻¹ represents coupled O3'–H and O5'–H stretches, arising from **TA**, **Ti**, **TC**, **TB**, **O2A**, and **O2i**, or coupled O2'–H and O5'–H stretches, arising from **Tii** and **Tiii**. The weak band measured at ~3610 cm⁻¹, primarily contributed by **Tii** and **Tiii**, represents O3'–H stretches. The IR band measured at ~3580 cm⁻¹, and its shoulder at ~3560 cm⁻¹, reflect O2–H and O4–H stretches, respectively. The weak feature observed at ~3510 cm⁻¹ is primarily contributed by the O2'–H stretch of **Ti** and **O2i**, in which the 2'-hydroxyl is involved in dual hydrogen-bonding interactions. The weak IR band measured

at ~3395 cm⁻¹, contributed only by the O2 protonated conformers, **O2A** and **O2i**, reflects the N3–H stretch.

## Maxwell-Boltzmann Weighted and Least Squares Fitted Populations

Visual comparisons between the measured IRMPD and calculated IR spectra for the low-energy conformers of [dThd + H]⁺ and [Thd + H]⁺ suggest that minor 2,4-dihydroxy tautomers as well as O2 protonated conformers are populated in the experiments. In particular, the **TA**, **TB**, **TC**, **O2A**, and **TD** conformers of [dThd + H]⁺, and the **TA**, **Ti**, **TC**, **TB**, **Tii**, **O2A**, **O2i**, **Tiii**, and **TD** conformers of [Thd + H]⁺ may be populated in the experiments. To enhance the interpretation, least squares fitting (LSF) of the measured spectra based on the calculated IR spectra of all low-energy conformers within 10 kJ/mol (B3LYP), including **T**, **O2**, and **O4** conformers, as well as the **TD** conformers of both [dThd + H]⁺ and [Thd + H]⁺ was performed. The best least squares fits found for [dThd + H]⁺ suggest that **TB**, **O2A**, **TC** and **TD** or **TA**, **O2A**, **TC** and **TD**

are present, and represent 71.6%, 13.8%, 13.7%, and 0.9%, or 70.9%, 13.5%, 14.3%, and 1.3% of the ESI population. As found from our visual comparison, the O4 protonated conformers of $[dTh + H]^+$ are absent in the population. **TA** and **TB** were not readily differentiated by LSF because the calculated IR features of **TA** and **TB** are highly parallel and exchangeable. The best least squares fit found for $[Thd + H]^+$ suggests that **Ti**, **TC**, **TB**, **Tii**, **TD**, **O2A**, and **O4ii** are present, and represent 16.4%, 12.6%, 31.4%, 11.6%, 7.3%, 18.3%, and 2.4% of the ESI population. The LSF chose **TB** and **Tii** over **Tiii** and **TA** because the calculated IR spectra of **TB** and **Tiii** as well as **Tii** and **TA**, respectively, are highly parallel, but **TB** and **Tii** reproduce the measured spectrum slightly better. The best LSF suggests that the **O4ii** conformer is populated at 2.4% of the ESI population, but when the O4 protonated conformers are not included in the analyses, the LSF fits are still able to fit the spectrum and give rise to residuals that are only slightly larger, i.e., they suggest that **Ti**, **TC**, **TB**, **Tii**, **TD**, and **O2A** or **Ti**, **TC**, **TB**, **Tii**, **TD**, and **O2i** are present, and represent approximately 9.7%, 18.5%, 33.6%, 10.0%, 2.2%, and 26.0% or 5.7%, 11.0%, 40.7%, 8.0%, 12.3%, and 22.3%, respectively. This suggests that when a different O2 protonated conformer, **O2A** or **O2i** is present, the percentage of each **T** tautomer in the population varies, but the **T** versus **O2** ratios are similar. The various fits that give rise to small residuals suggest that the populations of the 2,4-dihydroxy tautomers versus the O2 protonated conformers for $[dThd + H]^+$ are 87% ± 2% and 13% ± 2%, whereas the populations of the 2,4-dihydroxy tautomers versus the O2 protonated versus O4 protonated conformers for $[Thd + H]^+$ are 81% ± 6%, 18% ± 6%, and 1% ± 1%.

The measured IRMPD spectra of $[dThd + H]^+$ and $[Thd + H]^+$ are compared with the linear IR spectra predicted for Maxwell-Boltzmann weighted (MBW) distributions at room temperature based on the B3LYP and MP2 relative stabilities as well as the best LSF fits. The B3LYP and MP2 MBW spectra include all of the low-energy conformers that are concluded to be populated based on visual comparison between the measured IRMPD and calculated IR spectra. These comparisons are shown in Figures S7 and S8. Comparisons in Figure S7 suggest that overall the LSF, B3LYP, and MP2 MBW spectra exhibit good agreement with the measured spectrum. Different ratios of the 2,4-dihydroxy tautomers versus the O2 protonated conformers of $[dThd + H]^+$ predicted by LSF, B3LYP, and MP2 lead to the modest differences among the spectra. B3LYP and MP2 predict ratios of 2,4-dihydroxy tautomers versus O2 protonated conformers to be 92%:8% and 94%:6%, respectively, whereas the best LSF fit suggests the ratio to be 86%:14%. Therefore, both B3LYP and MP2 appear to underestimate the relative stabilities of the O2 protonated conformers of $[dThd + H]^+$. In Figure S8, the intensities of the bands at ~1775 and ~3395 cm$^{-1}$ that arise from O2 protonated conformers are underestimated by both B3LYP and MP2. The best LSF fit suggests the O2 protonated conformers account for ~18% of the ESI population, whereas B3LYP and MP2 predict their population to be ~7% and ~3%, respectively. Again, B3LYP and MP2 appear to underestimate the relative stabilities of the

O2 protonated conformers of $[Thd + H]^+$. The calculated IR spectra of $[Thd + H]^+$ based on the ratio of the tautomeric conformations, **T**:**O2**:**O4** = 79.3%:18.3%:2.4%, suggested by the best LSF fit, and without the O4 protonated conformers, the ratio of **T**:**O2** = 79.3%:20.7% are compared in Figure S9. These two calculated IR spectra are nearly identical, suggesting that the 2.4% population of the O4 protonated conformers does not exert a significant impact on the averaged IR spectra. Therefore, the O4 protonated conformers of $[Thd + H]^+$ are likely not important contributors to the ESI population.

## Comparison to Protonated Thymine

A previous IRMPD and theoretical study of protonated thymine [9] suggests that the lowest-energy conformer, the minor 2,4-dihydroxy tautomer with both the 2- and 4-hydroxyl hydrogen atoms pointed toward N3, dominates the population. The second most stable conformer, the O4 protonated conformer is also a minor contributor to the population. O2 protonated conformers are absent. In contrast, the current study of $[dThd + H]^+$ and $[Thd + H]^+$ suggests that the presence of the sugar moieties allows a larger variety of minor 2,4-dihydroxy tautomers to exist and to be populated in the experiments. Moreover, comparison between the measured IRMPD and calculated IR spectra indicates the absence of O4 protonated conformers, whereas the O2 protonated conformers are present in the experiments. Thus, hydrogen-bonding interactions with the sugar moieties alter the tautomeric conformations populated.

## Comparison to Protonated Uridine and 2'-Deoxyuridine

Our previous IRMPD and theoretical study of the protonated forms of uridine and 2'-deoxyuridine, $[Urd + H]^+$ and $[dUrd + H]^+$, suggests that a variety of 2,4-dihydroxy tautomers and the O4 protonated conformers are populated in the experiments [22]. In contrast, the current study suggests that the O4 protonated conformers of $[dThd + H]^+$ and $[Thd + H]^+$ are not populated or are populated in negligible abundance in the experiments, whereas O2 protonated conformers are clearly present. This suggests that the presence of the 5-methyl substituent in $[dThd + H]^+$ and $[Thd + H]^+$ alters the relative stabilities of the O4 protonated conformers and their populations in the experiments. Moreover, our previous IRMPD and theoretical study of $[Urd + H]^+$ and $[dUrd + H]^+$ suggests that the minor 2,4-dihydroxy tautomers may account for 75% ± 2% and 68% ± 7% of the ESI population [22], respectively, whereas the current study suggests that 87% ± 2% and 81% ± 6% of the ESI population are contributed by the minor 2,4-dihydroxy tautomers of $[dThd + H]^+$ and $[Thd + H]^+$, respectively. Thus, the minor 2,4-dihydroxy tautomers are somewhat more important for $[dThd + H]^+$ and $[Thd + H]^+$. Interestingly, the minor 2,4-dihydroxy tautomers of both the naturally occurring $[Urd + H]^+$ and $[dThd + H]^+$ account for a larger portion of the ESI population than those of their modified counterparts, $[dUrd + H]^+$ and $[Thd + H]^+$,

respectively, and thus both modifications slightly reduce the propensity for tautomerization.

## Comparison to Salpin's IRMPD study of [dThd + H]⁺

Essentially simultaneous with this work, Salpin et al. [23] used a quadrupole ion trap coupled to the CLIO free electron laser to measure the IRMPD spectrum of $[dThd + H]^+$ in the IR fingerprint region between 1000 and 2000 cm$^{-1}$. Their measured IRMPD spectrum for $[dThd + H]^+$ is highly parallel to that reported here, but is somewhat less resolved, particularly in the region of ~1400 to 1550 cm$^{-1}$. Differences in the relative IRMPD yields of various bands are also observed. The ground-state conformation of $[dThd + H]^+$ and its linear IR spectrum computed at B3LYP/6-311++G(3df,2p)// B3LYP/6-31++G(d,p) level of theory determined in that work are consistent with those found here. We also agree that 2,4-dihydroxy tautomers and O2 protonated conformers coexist in the experimental populations. Based solely on comparison of the calculated spectra with their measured spectrum in the IR fingerprint region, O4 protonated conformers cannot be ruled out. However, differences in the spectral features observed in the hydrogen-stretching region in the present work and those predicted for O4 protonated conformers eliminate them from the experimental population. The LSF procedure employed here also provides an estimate of the relative populations of the T versus O2 conformations populated by ESI.

## Conclusions

The IRMPD action spectra of $[dThd + H]^+$ and $[Thd + H]^+$ in the region of ~600–1900 cm$^{-1}$ and ~3300–3800 cm$^{-1}$ have been measured and compared with the linear IR spectra calculated at the B3LYP/6-311+G(d,p) level of theory to determine the favorable state(s) of protonation/tautomerization and the conformations populated by ESI. Both B3LYP and MP2 levels of theory predict the minor 2,4-dihydroxy tautomer with the nucleobase in the anti-orientation, both 2- and 4-hydroxyl groups pointing toward N3, and C2'-endo sugar puckering as the most stable conformation for both $[dThd + H]^+$ and $[Thd + H]^+$. The 2'-hydroxyl substituent does not lead to significant conformational differences between $[dThd + H]^+$ versus $[Thd + H]^+$. However, the 2'-hydroxyl substituent increases the opportunities for hydrogen-bonding interactions with the nucleobase as well as the adjacent 3'-hydroxyl substituent. Therefore, more low-energy conformations are found for $[Thd + H]^+$. The comparisons between experiment and theory suggest that the minor 2,4-dihydroxy tautomers account for over 80% of the ESI population for both $[dThd + H]^+$ and $[Thd + H]^+$. The 2'-hydroxyl substituent allows a larger variety of the 2,4-dihydroxy tautomers of $[Thd + H]^+$ to be present in the experiments. However,

the 2,4-dihydroxy tautomers of $[Thd + H]^+$ contribute 81% ± 6% to the ESI population, less than those of $[dThd + H]^+$, 87% ± 2%. **O2A** of both $[dThd + H]^+$ and $[Thd + H]^+$ as well as **O2i** of $[Thd + H]^+$ also contribute to the ESI population. Electronic effects of the 5-methyl substituent lead to the absence of O4 protonated conformers in the experiments. In contrast, for both $[Urd + H]^+$ and $[dUrd + H]^+$, the O2 protonated conformers are absent in the experiments [22] and the O4 protonated conformers of $[Urd + H]^+$ and $[dUrd + H]^+$ account for a larger portion of the ESI population than the O2 protonated conformers of $[dThd + H]^+$ and $[Thd + H]^+$. Thus, the 5-methyl substituent clearly alters the protonation preferences of $[dThd + H]^+$ and $[Thd + H]^+$ versus those of $[dUrd + H]^+$ and $[Urd + H]^+$. This work along with parallel IRMPD and theoretical studies of the protonated forms of the other common DNA and RNA nucleosides [22, 29-31] have enabled the preferred states of protonation/tautomerization and the conformations populated by ESI to be established. In general, the measured nonlinear IRMPD spectra are well reproduced by the room-temperature MBW linear IR spectra based on the low-energy conformers populated by ESI in spite of the anharmonicity effects associated with the multiple photon dissociation process employed. These studies clearly establish IRMPD action spectroscopy as a powerful tool for characterizing intrinsic structural properties of the fundamental building blocks of nucleic acids.

## Acknowledgments

## References

1. Kang, H., Lee, K.T., Jung, B., Ko, Y.J., Kim, S.K.: Intrinsic lifetimes of the excited state of DNA and RNA bases. J. Am. Chem. Soc. **124**, 12958–12959 (2002)

2. Cadet, J., Sage, E., Douki, T.: Ultraviolet radiation-mediated damage to cellular DNA. Mutat. Res. Fundam. Mol. Mech. Mutagen. **571**, 3–17 (2005)

3. Kim, S.-T., Sancar, A.: Effect of base, pentose, and phosphodiester backbone structures on binding and repair of pyrimidine dimers by *Escherichia coli* DNA photolyase. Biochemistry **30**, 8623–8630 (1991)

4. Nosenko, Y., Kunitski, M., Brutschy, B.: Specific photodynamics in thymine clusters: the role of hydrogen bonding. J. Phys. Chem. A **115**, 9429–9439 (2011)

5. Harris, V.H., Smith, C.L., Cummins, W.J., Hamilton, A.L., Adams, H., Dickman, M., Hornby, D.P., Williams, D.M.: The effect of tautomeric constant on the specificity of nucleotide incorporation during DNA replication: support for the rare tautomer hypothesis of substitution mutagenesis. J. Mol. Biol. **326**, 1389–1401 (2003)

6. Alyoubi, A.O., Hilal, R.H.: A theoretical and experimental investigation of the electronic spectra and tautomerization of nucleobases. Biophys. Chem. **55**, 231–237 (1995)

7. Colarusso, P., Zhang, K.Q., Guo, B.J., Bernath, P.F.: The infrared spectra of uracil, thymine, and adenine in the gas phase. Chem. Phys. Lett. **269**, 39–48 (1997)

8. Zhang, S.L., Michaelian, K.H., Loppnow, G.R.: Vibrational spectra and experimental assignments of thymine and nine of its isotopomers. J. Phys. Chem. A **102**, 461–470 (1998)

9. Salpin, J.-Y., Guillaumont, S., Tortajada, J., MacAleese, L., Lemaire, J., Maitre, P.: Infrared spectra of protonated uracil, thymine and cytosine. ChemPhysChem **8**, 2235–2244 (2007)

10. Nonin, S., Leroy, J.-L.: Structure and conversion kinetics of a bi-stable DNA i-motif: broken symmetry in the [d(5mCCTCC)]4 tetramer. J. Mol. Biol. **261**, 399–414 (1996)

11. Ramreddy, T., Kombrabail, M., Krishnamoorthy, G., Rao, B.J.: Site-specific dynamics in TAT triplex DNA as revealed by time-domain fluorescence of 2-aminopurine. J. Phys. Chem. B **113**, 6840–6846 (2009)

12. Jochims, H.-W., Schwell, M., Baumgärtel, H., Leach, S.: Photo-ion mass spectrometry of adenine, thymine, and uracil in the 6-22 eV photon energy range. Chem. Phys. **314**, 263–282 (2005)

13. Imhoff, M., Deng, Z., Huels, M.A.: Identification of ion fragments produced from thymine and deuterated thymine by low energy ion impact in films and electron impact in the gas phase. Int. J. Mass Spectrom. **245**, 68–77 (2005)

14. Itälä, E., Ha, D.T., Kooser, K., Rachlew, E., Huels, M.A., Kukk, E.: Fragmentation patterns of core-ionized thymine and 5-bromouracil. J. Chem. Phys. **133**, 154316-1–154316-9 (2010)

15. Bell, A.F., Hecht, L., Barron, L.D.: Vibrational RAMAN optical activity of pyrimidine nucleosides. J. Chem. Soc. Faraday Trans. **93**, 553–562 (1997)

16. Razvazhnaya, O.V., Burmykin, D.A., Revelsky, A.I.: Study of the influence of derivatization conditions on the structure and stability of nucleoside derivatives using gas chromatography-mass spectrometry. J. Anal. Chem. **69**, 1307–1312 (2014)

17. Ptasińska, S., Candori, P., Denifl, S., Yoon, S., Grill, V., Scheier, P., Märk, T.D.: Dissociative ionization of the nucleosides thymidine and uridine by electron impact. Chem. Phys. Lett. **409**, 270–276 (2005)

18. Itälä, E., Huels, M.A., Rachlew, E., Kooser, K., Hägerth, T., Kukk, E.: A comparative study of dissociation of thymidine molecules following valence or core photoionization. J. Phys. B Atomic Mol. Phys. **46**, 215102-1–215102-9 (2013)

19. Palafox, M.A., Iza, N., de la Fuente, M., Navarro, R.: Simulation of the first hydration shell of nucleosides D4T and thymidine: structures obtained using MP2 and DFT methods. J. Phys. Chem. B **113**, 2458–2476 (2009)

20. Sinden, R.R.: DNA structure and function, pp. 1−57. Academic Press, Waltham, MA (1994)

21. Szyf, M.: The DNA Methylation machinery as a target for anticancer therapy. Pharmacol. Ther. **70**, 1–37 (1996)

22. 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)

23. 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)

24. Valle, J.J., Eyler, J.R., Oomens, J., Moore, D.T., van der Meer, A.F.G., von Heldon, G., Meijer, G., Hendrickson, C.L., Marshall, A.G., Blakney, G.T.: Free electron laser-Fourier transform ion cyclotron resonance mass spectrometry facility for obtaining infrared multiphoton dissociation spectra of gaseous ions. Rev. Sci. Instrum. **76**, 023103-1–023103-7 (2005)

25. 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)

26. Polfer, N.C., Oomens, J.: Reaction products in mass spectrometry elucidated with infrared spectroscopy. Phys. Chem. Chem. Phys. **9**, 3804–3817 (2007)

27. 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)

28. HyperChem computational chemistry software package, version 5.0. Hypercube, Inc., Gainesville (1997)

29. 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)

30. 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)

31. 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)

32. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Montgomery Jr., J.A., Vreven, T., Kudin, K.N., Burant, J.C., Millam, J.M., Iyengar, S.S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G.A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J.E., Hratchian, H.P., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomeli, C., Ochterski, J.W., Ayala, P.Y., Morokuma, K., Voth, G.A., Salvador, P., Dannenberg, J.J., Zakrzewski, V.G., Dapprich, S., Daniels, A.D., Strain, M.C., Farkas, O., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Ortiz, J.V., Cui, Q., Baboul, A.G., Clifford, S., Cioslowski, J., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R.L., Fox, D.J., Keith, T., Al-Laham, M.A., Peng, C.Y., Nanayakkara, A., Challacombe, M., Gill, P.M.W., Johnson, B., Chen, W., Wong, M.W., Gonzalez, C., Pople, J.A.: Gaussian 09, revision C.01. Gaussian, Inc, Wallingford (2009)

33. Wong, M.W.: Vibrational frequency prediction using density functional theory. Chem. Phys. Lett. **256**, 391–399 (1996)

34. Maejima, K., Kawabata, H., Fujita, S.: Density functional theory study on beta-hydride elimination as thermal decomposition process of diethylzinc. Jpn. J. Appl. Phys. **47**, 1098–1100 (2008)