



HAL
open science

Structure of protonated thymidine characterized by InfraRed Multiple Photon Dissociation and quantum calculations

Jean-Yves Salpin, D. Scuderi

► **To cite this version:**

Jean-Yves Salpin, D. Scuderi. Structure of protonated thymidine characterized by InfraRed Multiple Photon Dissociation and quantum calculations. *Rapid Communications in Mass Spectrometry*, 2015, 29 (20), pp.1898-1904. 10.1002/rcm.7296 . hal-01221301

HAL Id: hal-01221301

<https://hal.science/hal-01221301>

Submitted on 5 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Structure of protonated thymidine characterized by InfraRed Multiple

Photon Dissociation and quantum calculations

Jean-Yves Salpin^{1,2*}

1) Université d'Evry Val d'Essonne – Laboratoire Analyse et Modélisation pour la Biologie et l'Environnement – Boulevard François Mitterrand – 91025 Evry – France

2) CNRS- UMR 8587.

Debora Scuderi^{3,4}

3) Université Paris Sud Orsay - Laboratoire de Chimie Physique – Bâtiment 350 —91405 ORSAY; FRANCE

4) CNRS – UMR 8000

Number of pages (including Tables, Figures, legends and schemes) : 22

Corresponding author:

Jean-Yves Salpin

Tel: 33 1 69 47 76 47 Fax: 33 1 69 47 76 55

e-mail : jean-yves.salpin@univ-evry.fr

Abstract

Rationale.

Many fundamental studies are motivated by the probable relationship between the presence of rare enol tautomers of nucleobases and point mutation developing during nucleic acid replication. The evaluation of the tautomeric behaviour of nucleobases is therefore of fundamental importance. This can be probed in the gas phase by combining action spectroscopy and mass spectrometry.

Methods

Experimental Infrared Multiple Photon Dissociation spectra in the fingerprint region of electrospray-generated and subsequently selected ions were recorded at the CLIO free electron laser (FEL) facility, by coupling the FEL to a quadrupole ion trap, and compared to calculated harmonic vibrational infrared spectra of the different low-lying isomers computed at the B3LYP/6-31++G(d,p) level. Relative energies were refined using the extended basis set 6-311++G(3df,2p).

Results

The Density Functional Theory (DFT) study shows that, like for protonated thymine, the global energy minimum of protonated thymidine corresponds to an enol tautomer, whose infrared absorption spectrum is found to be in very good agreement with the experimental IRMPD spectrum. A very weak IRMPD signal observed at $\sim 1780\text{ cm}^{-1}$ is very likely the signature of the second energy lying oxo tautomer. Consequently, like for thymine, protonated thymidine generated by electrospray corresponds to a mixture of at least two tautomeric forms.

Conclusion

Tautomerization can be characterized by IRMPD spectroscopy. Interestingly, the dominant enolic tautomeric form(s) presently observed cannot be directly generated from the most stable neutral tautomer of the thymine residue.

Keywords : infrared spectroscopy, multiple photon absorption, photodissociation, protonated thymidine, mass spectrometry, electrospray ionization.

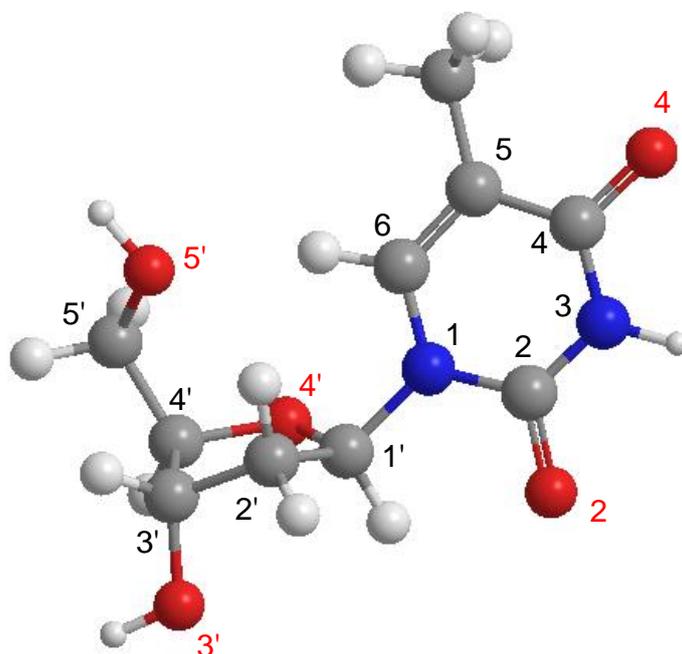
Running title : IRMPD spectroscopy in the mid-IR region of protonated thymidine

1. Introduction

During the last decade, InfraRed Multiple Photon Dissociation (IRMPD) spectroscopy has been established as a powerful approach for probing the structure of gaseous ions.^[1-5] In this context, this technique has been notably widely applied to study the structure of either bare^[6-19] or metal-cationized DNA building blocks.^[20-33] Many studies were motivated by the probable relationship between the presence of rare enol tautomers of nucleobases within nucleic acid building blocks and point mutation developing during nucleic acid replication^[34-39] The evaluation of the tautomeric behaviour of nucleobases is therefore of fundamental importance. In previous papers,^[7, 8] we showed that for protonated uracil and thymine, under electrospray conditions, two tautomeric forms were formed in the gas phase, the most abundant tautomer being an enolic structure that could not be directly generated from the most stable neutral tautomer. More recently, structure of protonated nucleosides has also been probed by IRMPD. Several studies were indeed dedicated to protonated cytidine and deoxycytidine.^[16, 17, 19] Filipini and co-workers showed that IRMPD spectra were consistent with the presence of both the O2 and N3 protonated nucleosides.^[16] This was further confirmed by the study carried out by Rodgers and co-workers^[19], and IRMPD experiments carried out on protonated cytosine lead to the same conclusion.^[7] On the other hand, a very recent report by Morton and co-workers,^[17] involving partial deuteration experiments, suggest that for 2'-deoxycytidine, the experimental bands in the XH stretch region, initially attributed to the O2-protonated form would in fact correspond to overtones and combination bands, thereby discarding the presence of two distinct forms. The structure of protonated adenosine has also been recently examined,^[18] and IRMPD spectra in both fingerprint and XH stretch regions points to protonation onto the N3 position of the adenine residue. Such protonation results in a structure stabilized by a very strong intramolecular hydrogen bond. However,

protonation onto N1, although less favorable energetically, cannot be rigorously excluded because IRMPD spectra of N3 and N1 protonated forms are very similar.

To complement these studies, we present in this short paper the IRMPD spectroscopy in the fingerprint region between 1000 and 2000 cm^{-1} of protonated thymidine (scheme 1), generated by electrospray ionization and trapped in a quadrupole ion trap. Comparison with DFT-computed IR absorption spectra of different protonated forms indicates the presence of at least two distinct tautomeric forms.



Scheme 1: structure of thymidine

2. Experimental and theoretical methods

2.1 Materials

The thymidine used in this work is a research grade product from Sigma-Aldrich (Saint-Quentin Fallavier, France) and was used without further purification. Stock solutions were prepared at 10^{-2} mol/L in purified water (purified with a Milli-Q water purification system).

2.2 IRMPD experiments.

Thymidine was diluted at 10^{-4} M in a water/methanol mixture (50/50 v/v), prior to its introduction into electrospray source. Typical ESI conditions were a flow rate of 3 $\mu\text{L}/\text{min}$, capillary spray voltage at -4.2 kV, nebulizer at 14 PSI, drying gas flow at 5 L/min, and drying gas temperature at 200°C .

The present IRMPD spectroscopic investigation has been performed using an experimental platform based a modified Bruker Esquire 3000+ quadrupole ion trap (Bruker), which has been described in details previously.^[40-42] Tunable mid-infrared radiation is produced by the Free Electron Laser (FEL) of CLIO (Centre Laser Infrarouge d'Orsay).^[43] This laser is based on a 16-48 MeV linear electron accelerator, and bunches of electrons are injected in an undulator, a periodic magnetic field, which is placed in the optical cavity. Using a fixed electron energy, the energy of the emitted photons can be continuously scanned by adjusting the undulator gap. The mean average power was monitored and found to be quite stable (700-800 mW) over the $1000\text{-}2000\text{ cm}^{-1}$ energy range scanned using electrons at 45 MeV. Protonated thymidine ions were accumulated for 3 ms and mass selected prior to IR irradiation. After mass selection, ions were then irradiated for 200 ms with the IR FEL light, after which the resulting ions were mass-analyzed.

IRMPD spectra are obtained by plotting the photofragmentation efficiency ($-\ln[I_{\text{parent}}/(I_{\text{parent}} + \sum I_{\text{fragment}})]$), where I_{parent} and I_{fragment} are the integrated intensities of the precursor and of the fragment ions, respectively, as a function of the frequency of the laser wavelength.^{[44], [45]}

2.3 theoretical calculations.

DFT calculations were carried out using the B3LYP density functional^[46, 47], as implemented in the Gaussian 09 set of programs.^[48]

The different protonated forms considered have first been optimized with the 6-31++G(d,p) basis set, without any symmetry constraint. Harmonic vibrational frequencies were estimated at this level to characterize the stationary points as local minima or saddle points, and to estimate the zero-point vibrational energy (ZPE) corrections. Finally, relative energies were refined using the extended basis set 6-311++G(3df,2p). The infrared absorption spectra were calculated within the harmonic approximation. Ghomi et al. ^[49, 50] demonstrated in detail that DFT and MP2 levels of theory gave similar results for the geometrical and vibrational features of the nucleic acid bases. Furthermore, a recent study about protonated cytosine by Tureček and co-workers demonstrates that use of diffuse functions and extended basis sets was essential to accurately evaluate the relative energies between the various protonated forms.^[51] Furthermore, provided the use of an appropriate scaling factor, hybrid DFT methods such as B3LYP have been shown to outperform other DFT methods as well as traditional *ab initio* approaches to describe both position^[52] and relative intensities^[53, 54] of IR bands. As far as the positions are concerned, a scaling factor value of 0.98 has been chosen.

3. Results, discussion

3.1 Low-energy isomers of protonated thymidine.

Mononucleosides are flexible molecules, and consequently performing a complete conformational survey is already challenging. Based upon our previous study on protonated thymine^[7], we only considered the oxygen atoms of the thymine moiety (T) as possible protonation sites and examined both oxo and enol tautomeric forms, the thymine residue adopting either Syn or Anti orientation. Two different sugar-ring puckerings, namely C2'endo and C3'endo, have been considered. The name of various forms is made of two parts. The first part describes the ring puckering and the nucleobase orientation (for example a structure starting with C2A means a C2'endo ring puckering and an Anti orientation of T). The various

structures are named according the following nomenclature. The second part indicates which oxygen(s) of T are indeed protonated (O2 and/or O4). The six most stable protonated forms of thymidine obtained at the B3LYP/6-311++G(3df,2p)//B3LYP/6-31++G(d,p)+ZPE level of theory, are given in Figure 1. Energetics and structure of all the protonated forms considered in this study are provided in sections 1S of the Supporting Information.

Six forms are lying within 7 kJ/mol in relative Gibbs free energies. Our theoretical survey indicates that C2'endo forms are systematically more stable than C3'endo structures for a given protonation state and T orientation. Several detailed theoretical studies^[6, 7, 55, 56] showed that the most stable protonated form of uracil and thymine is an enol structure, regardless of the method of calculation (B3LYP, MP2 or QCISD(T)). A similar finding is presently obtained, the **C2A_O2O4a** structure being the global minimum. The oxo form (**C2A_O4b**) appears slightly less stable (+ 3.2 kJ/mol). These two structures turned to be degenerate when considering the C3'endo ring puckering. Several **C2S** forms appears particularly stable (**C2S_O2a**) due to the establishment of a very strong intramolecular hydrogen bond (HB) between the protonated C2=O2 carbonyl and the hydroxymethyl group the deoxyribose unit. Comparison with the **C2S_O2b** structure shows that the stabilization energy provided by this HB is about 37 kJ/mol. (see Supporting Information) On the other hand, when T is in Anti configuration, the O4 oxygen is the most basic site, as already reported for thymine in the literature.^[7, 57] We also considered different hydroxymethyl orientations. The most stable forms correspond to structures for which the O5'C5'C4'O4' torsional angle is about -60° (G-).

It should be also recalled that the orientation of the hydroxyl group of T in the protonated forms plays an important role in the stabilization of the structures. For instance **C2A_O4a** and **C2A_O4b** structures only differ by the rotation of 180° of the C4-O4H hydroxyl group but the Gibbs free energy difference between these two forms is significant (13.3 kJ/mol). This

illustrates the effects of the repulsive interaction between parallel N3-H and O4-H bonds (**C2A_O4a**) and of the attractive interaction between the O4-H bond and the lone pair of adjacent N3 (**C2A_O4b**). A similar behaviour is observed for enol forms. Finally, it is worth noting that the enol forms and notably global minimum **C2A_O2O4a** cannot be generated directly from protonation of the dioxo form of the thymine residue.

3.2 IRMPD spectra of protonated thymidine

On resonance with an infrared active mode of the mass-selected ions, photo-fragmentation leading to formation of protonated thymine (m/z 127) is observed. High fragmentation yield (~20%) was readily achieved with only two IR-FEL macropulses. The IRMPD spectrum presently reported corresponds to the fragmentation efficiency, defined as $-\ln[I_{\text{parent}}/(I_{\text{parent}} + \sum I_{\text{fragment}})]$, as a function of the photon wavenumber.^{[44],[45]} This spectrum is given in Figure 2a. The assignment of the IRMPD spectrum is based on its comparison with the spectra computed for the various low energy lying isomers. To be consistent with the experimental spectral resolution, the calculated spectra have been convoluted by a 15 cm^{-1} large Lorentzian function.

As it will be discussed in details later, there is a quite good agreement between the IRMPD spectrum and the calculated IR absorption spectrum of the lowest energy isomer of protonated thymidine (**C2A_O2O4a**), which corresponds to an enol tautomeric form, with the exception of two bands: a relatively weak IRMPD signal observed at about 1304 cm^{-1} , and a band detected at c.a. 1780 cm^{-1} . Such a behavior has been also encountered for protonated thymine^[7] and was interpreted by the presence in the gas phase of at least two different structures (*vide infra*).

The experimental IRMPD spectrum of protonated thymidine (Figure 2a) is characterized by five very intense features, dominated by a band centered at 1215 cm^{-1} . Four slightly less

intense bands were also detected at c.a. 1112, 1384, 1600 and 1640 cm^{-1} . Another intense signal whose maximum is at about 1490 cm^{-1} , displays an asymmetric profile. Its bandwidth (fwhm~65 cm^{-1}) clearly suggests that this band probably corresponds to the sum of the contributions of two strongly infrared active modes. Indeed, using the same experimental setup, isolated IR active vibrational modes give rise to narrower IRMPD band, with a width of typically 25 cm^{-1} or less.^[41] Finally, a weak signal could also be considered around 1165 cm^{-1} . Assignment of most of the IRMPD bands can be made on the basis of the analysis of the infrared active modes of **C2A_O2OA4a** given in Table 1. All the experimental features but the two bands at 1304 and 1780 cm^{-1} can indeed be interpreted by the calculated IR absorption spectrum of **C2A_O2O4a** (Figure 2b). For the sake of clarity, only the positions and intensities (km/mol) of the main IR active vibrational modes of this form are given in Table 1. The strong IRMPD band centered at 1112 cm^{-1} could be assigned to the contribution of three IR active modes corresponding to sugar ring breathing and the stretch of different C-O bonds of the deoxyribose moiety (Table 1). The most intense experimental band (~1215 cm^{-1}) is consistent with three strongly IR active modes of **C2A_O2OA4a**, which correspond to the twisting of the C2' methylene group and to different bending modes of the enolic groups of the thymine residue. Another combined bending modes of these two particular enolic groups may also be responsible for the experimental signal detected at ~1165 cm^{-1} . The calculated infrared absorption spectrum of this structure also displays four strongly active IR modes corresponding between 1450 and 1650 wavenumbers, which nicely reproduce experimental spectrum in this energy region. The asymmetric profile of the large signal detected between 1460 and 1560 cm^{-1} can be ascribed to two intense computed modes associated to the C4-O4 (1502 cm^{-1}) and C4-O4 (1533 cm^{-1}) stretches. A shoulder can also be observed at the blue side of this band that can be ascribed to bending modes of the thymine methyl group. The two strong bands detected at 1600 and 1641 cm^{-1} can be attributed to the

coupling of the C4C5 and N1C2 (1598 cm^{-1}), and C5C6 and C2N3 (1660 cm^{-1}) stretches, respectively. Finally the experimental band measured at 1384 cm^{-1} may correspond to C-H bending modes of the ribose moiety.

As already mentioned previously, the computed IR spectrum of the enol tautomer **C2A_O2O4a** cannot account for the two bands experimentally observed at 1304 and 1780 cm^{-1} . This particular situation has been already encountered for protonated thymine^[7], and based on these previous results, it is likely that different tautomeric forms are simultaneously present under our experimental conditions. The IRMPD band observed at $\sim 1780\text{ cm}^{-1}$ could indeed be the result of a resonant absorption through a very strongly infrared active C=O stretch. The C=O stretching mode indeed constitutes an excellent infrared diagnostic of the presence of an oxo form of the protonated thymine moiety, and the experimental band measured at 1780 cm^{-1} is in good agreement with the C=O stretch computed for the second most stable form, namely the **C2S_O2a** oxo tautomer. Given the multiple photon character of the IRMPD process, the relative IRMPD band intensities should be taken with caution. However, the calculated infrared cross section of the **C2S_O2a** isomer at 1780 cm^{-1} is almost as large as at 1600 cm^{-1} . As a result, if the **C2S_O2a** oxo isomer would be selectively formed, one would have expected a stronger IRMPD signal at 1780 cm^{-1} , i.e. of the same order of magnitude as the one observed at $\sim 1600\text{ cm}^{-1}$. Consequently, the present IRMPD spectrum could therefore result from the superposition of a minor contribution of the oxo isomer (**C2S_O2a**) and a major contribution of the lowest energy structure, an enolic isomer (**C2A_O2O4a**). In addition, a dominant enol form is consistent with the calculated relative Gibbs free energies of these two competing structures. In this case, the band measured at 1304 cm^{-1} may correspond to the combined $\delta\text{O2H}+\delta\text{C6H}$ bending mode of **C2S_O2a**. Note that the IR spectrum computed for **C2S_O2a** could also account for the tailing of the strong band centered at 1215 cm^{-1} . It is worth noting that another oxo forms such as **C2A_O4b** (Figure 2e)

cannot be rigorously ruled out, in spite of a computed C=O stretch for this particular form slightly blue-shifted with respect to the experimental trace. Other oxo forms such as **C2A_O2a**, **C2A_O2b** or **C3A_O2a** (see Supporting Information), although sensibly less stable, may also account for the experimental trace around 1780 wavenumbers. Finally, one has to comment the fact that the IRMPD spectrum is also in very good agreement with the computed spectrum of other enol forms. This is the case for example of the **C3A_O2O4a** enol form (Figure 2c), which only differs from the global minimum by its sugar-ring puckering (C3'endo). Note also that **C2S_O2O4a** and **C3S_O2O4a** present computed spectra in good agreement with the IRMPD spectrum (Supporting Information). These two forms differ from **C2A_O2O4a** and **C3A_O2O4a**, respectively, by the nucleobase orientation (Syn instead of Anti). Although higher in relative free energies, such structures cannot be rigorously discarded as interconversion between Anti and Syn structures may occur during the ESI process. As a matter of fact, the rotational barrier associated with such a process may be small in solution (less than 10 kJ/mol), as suggested by a study made on cytidine.^[58] To summarize, for this system, IRMPD spectrum in the mid-IR region cannot be characteristic of a given type of sugar ring puckering. Speranza and co-workers lead to the same finding for protonated 2'-deoxycytidine.^[16] A similar conclusion can be drawn as regards to Syn/Anti nucleobase orientation. On the other hand, such distinctions might be achieved in the far-IR region. A recent study have indeed demonstrated that IRMPD experiments in the far-IR energy can allow discrimination among conformers of peptides that could not be distinguished in the mid-IR region.^[59]

Conclusion

The structure of gaseous protonated thymidine was probed using mid-infrared multiple photon dissociation (IRMPD) spectroscopy. Overall, there is a good agreement between experimental

IRMPD spectrum recorded between 1000 and 2000 cm^{-1} and the calculated infrared absorption spectrum of the lowest energy lying tautomer (enol form). Three other enol forms having different sugar-ring puckering and/or nucleobase orientation are in agreement with the experimental spectrum. A fraction of oxo tautomer(s) is also evidenced by the presence of a IRMPD signals at 1304 and 1780 cm^{-1} . Two types of protomers of protonated thymidine are therefore generated under our experimental conditions, the most abundant corresponding to enol structure(s).

Acknowledgements

The authors wish to thank the CLIO team (J. M. Ortega, C. Six, G. Perilhous, J. P. Berthet) as well as P. Maître and V. Steinmetz for their support during the experiments. A generous allocation of computing time at the Centro de Computación Científica of the UAM is also acknowledged. J. Musso is also warmly acknowledged for fruitful discussion.

Supporting information

Structures, energetics and computed IR absorption spectre of the different optimized structures are provided as a supporting information.

Bibliography

- [1]P. M. Luke MacAleese. Infrared spectroscopy of organometallic ions in the gas phase: From model to real world complexes. *Mass Spectrom. Rev.* **2007**, 26, 583.
- [2]T. D. Fridgen. Infrared Consequence Spectroscopy Of Gaseous Protonated And Metal Ion Cationized Complexes. *Mass Spectrom. Rev.* **2009**, 28, 586.

- [3]N. C. Polfer, J. Oomens. VIBRATIONAL SPECTROSCOPY OF BARE AND SOLVATED IONIC COMPLEXES OF BIOLOGICAL RELEVANCE. *Mass Spectrom. Rev.* **2009**, 28, 468.
- [4]J. R. Eyler. INFRARED MULTIPLE PHOTON DISSOCIATION SPECTROSCOPY OF IONS IN PENNING TRAPS. *Mass Spectrom. Rev.* **2009**, 28, 448.
- [5]J. S. Brodbelt, J. J. Wilson. Infrared Multiphoton Dissociation In Quadrupole Ion Traps. *Mass Spectrom. Rev.* **2009**, 28, 390.
- [6]R. Wu, T. B. McMahon. Investigation of Proton Transport Tautomerism in Clusters of Protonated Nucleic Acid Bases (Cytosine, Uracil, Thymine, and Adenine) and Ammonia by High-Pressure Mass Spectrometry and Ab Initio Calculations. *J. Am. Chem. Soc.* **2007**, 129, 569.
- [7]J. Y. Salpin, S. Guillaumont, J. Tortajada, L. MacAleese, J. Lemaire, P. Maitre. Infrared spectra of protonated uracil, thymine and cytosine. *ChemPhysChem* **2007**, 8, 2235.
- [8]J. M. Bakker, R. K. Sinha, T. Besson, M. Brugnara, P. Tosi, J.-Y. Salpin, P. Maitre. Tautomerism of Uracil Probed via Infrared Spectroscopy of Singly Hydrated Protonated Uracil. *J. Phys. Chem. A* **2008**, 112, 12393.
- [9]B. Chiavarino, M. E. Crestoni, S. Fornarini, F. Lanucara, J. Lemaire, P. Maitre, D. Scuderi. Infrared spectroscopy of isolated nucleotides 1. The cyclic 3',5'-adenosine monophosphate anion. *Int. J. Mass Spectrom.* **2008**, 270, 111.
- [10] J. M. Bakker, J.-Y. Salpin, P. Maitre. Tautomerism of cytosine probed by gas phase IR spectroscopy. *Int. J. Mass Spectrom.* **2009**, 283, 214.
- [11] F. Lanucara, M. E. Crestoni, B. Chiavarino, S. Fornarini, O. Hernandez, D. Scuderi, P. Maitre. Infrared spectroscopy of nucleotides in the gas phase 2. The protonated cyclic 3',5'-adenosine monophosphate. *Rsc Advances* **2013**, 3, 12711.
- [12] Y. W. Nei, K. T. Crampton, G. Berden, J. Oomens, M. T. Rodgers. Infrared Multiple Photon Dissociation Action Spectroscopy of Deprotonated RNA Mononucleotides: Gas-Phase Conformations and Energetics. *J. Phys. Chem. A* **2013**, 117, 10634.
- [13] Y. W. Nei, N. Hallowita, J. D. Steill, J. Oomens, M. T. Rodgers. Infrared Multiple Photon Dissociation Action Spectroscopy of Deprotonated DNA Mononucleotides: Gas-Phase Conformations and Energetics. *J. Phys. Chem. A* **2013**, 117, 1319.
- [14] H. U. Ung, A. R. Moehlig, R. A. Kudla, L. J. Mueller, J. Oomens, G. Berden, T. H. Morton. Proton-bound dimers of 1-methylcytosine and its derivatives: vibrational and NMR spectroscopy. *Phys. Chem. Chem. Phys.* **2013**, 15, 19001.
- [15] B. Yang, R. R. Wu, G. Berden, J. Oomens, M. T. Rodgers. Infrared Multiple Photon Dissociation Action Spectroscopy of Proton-Bound Dimers of Cytosine and Modified Cytosines: Effects of Modifications on Gas-Phase Conformations. *J. Phys. Chem. B* **2013**, 117, 14191.

- [16] A. Filippi, C. Fraschetti, F. Rondino, S. Piccirillo, V. Steinmetz, L. Guidoni, M. Speranza. Protonated pyrimidine nucleosides probed by IRMPD spectroscopy. *Int. J. Mass Spectrom.* **2013**, 354, 54.
- [17] H. U. Ung, K. T. Huynh, J. C. Poutsma, J. Oomens, G. Berden, T. H. Morton. Investigation of proton affinities and gas phase vibrational spectra of protonated nucleosides, deoxynucleosides, and their analogs. *Int. J. Mass Spectrom.* **2015**, 378, 294.
- [18] R. R. Wu, B. Yang, G. Berden, J. Oomens, M. T. Rodgers. Gas-Phase Conformations and Energetics of Protonated 2'-Deoxyadenosine and Adenosine: IRMPD Action Spectroscopy and Theoretical Studies. *J. Phys. Chem. B* **2015**, 119, 2795.
- [19] R. R. Wu, B. Yang, C. E. Frieler, G. Berden, J. Oomens, M. T. Rodgers. N3 and o2 protonated tautomeric conformations of 2'-deoxycytidine and cytidine coexist in the gas phase. *J. Phys. Chem. B* **2015**, 119, 5773.
- [20] E. A. L. Gillis, K. Rajabi, T. D. Fridgen. Structures of Hydrated Li⁺-Thymine and Li⁺-Uracil Complexes by IRMPD Spectroscopy in the N-H/O-H Stretching Region. *J. Phys. Chem. A* **2009**, 113, 824.
- [21] E. A. L. Gillis, T. D. Fridgen. The hydrated Li⁺-adenine-thymine complex by IRMPD spectroscopy in the N-H/O-H stretching region. *Int. J. Mass Spectrom.* **2010**, 297, 2.
- [22] O. Y. Ali, T. D. Fridgen. Structures of electrosprayed Pb(Uracil-H)(⁺) complexes by infrared multiple photon dissociation spectroscopy. *Int. J. Mass Spectrom.* **2011**, 308, 167.
- [23] J.-Y. Salpin, L. Gamiette, J. Tortajada, T. Besson, P. Maitre. Structure of Pb²⁺/dCMP and Pb²⁺/CMP complexes as characterized by tandem mass spectrometry and IRMPD spectroscopy. *Int. J. Mass Spectrom.* **2011**, 304, 154.
- [24] J.-Y. Salpin, S. Guillaumont, D. Ortiz, J. Tortajada, P. Maitre. Direct Evidence for Tautomerization of the Uracil Moiety within the Pb²⁺/Uridine-5'-monophosphate Complex: A Combined Tandem Mass Spectrometry and IRMPD study. *Inorg. Chem.* **2011**, 50, 7769.
- [25] Y. W. Nei, T. E. Akinyemi, C. M. Kaczan, J. D. Steill, G. Berden, J. Oomens, M. T. Rodgers. Infrared multiple photon dissociation action spectroscopy of sodiated uracil and thiouracils: Effects of thioketo-substitution on gas-phase conformation. *Int. J. Mass Spectrom.* **2011**, 308, 191.
- [26] O. Y. Ali, T. D. Fridgen. Structures and Fragmentation of Cu(Uracil-H)(Uracil) plus in the Gas Phase. *ChemPhysChem* **2012**, 13, 588.
- [27] E. A. L. Gillis, M. Demireva, K. Nanda, G. Beran, E. R. Williams, T. D. Fridgen. Structures and energetics of electrosprayed uracil(n)Ca(2⁺) clusters (n=14-4) in the gas phase. *Phys. Chem. Chem. Phys.* **2012**, 14, 3304.
- [28] B. Chiavarino, M. E. Crestoni, S. Fornarini, D. Scuderi, J.-Y. Salpin. Interaction of Cisplatin with Adenine and Guanine: A Combined IRMPD, MS/MS, and Theoretical Study. *J. Am. Chem. Soc.* **2013**, 135, 1445.

- [29] B. Yang, R. R. Wu, N. C. Polfer, G. Berden, J. Oomens, M. T. Rodgers. IRMPD Action Spectroscopy of Alkali Metal Cation-Cytosine Complexes: Effects of Alkali Metal Cation Size on Gas Phase Conformation. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1523.
- [30] J.-Y. Salpin, V. Haldys, S. Guillaumont, J. Tortajada, M. Hurtado, A. M. Lamsabhi. Gas-Phase Interactions between Lead(II) Ions and Cytosine: Tandem Mass Spectrometry and Infrared Multiple-Photon Dissociation Spectroscopy Study. *ChemPhysChem* **2014**, *15*, 2959.
- [31] J.-Y. Salpin, L. MacAleese, F. Chiro, P. Dugourd. Structure of the Pb²⁺-deprotonated dGMP complex in the gas phase: a combined MS-MS/IRMPD spectroscopy/ion mobility study. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14127.
- [32] B. Chiavarino, M. E. Crestoni, S. Fornarini, D. Scuderi, J.-Y. Salpin. Interaction of Cisplatin with 5'-dGMP: A Combined IRMPD and Theoretical Study. *Inorg. Chem.* **2015**, *54*, 3513.
- [33] B. Power, V. Haldys, J.-Y. Salpin, T. D. Fridgen. Structures of bare and singly hydrated M(Ura-H)(Ura) (+) (M = Mg, Ca, Sr, Ba) complexes in the gas phase by IRMPD spectroscopy in the fingerprint region. *Int. J. Mass Spectrom.* **2015**, *378*, 328.
- [34] P. O. Lowdin. Proton tunneling in dna and its biological implications. *Reviews of Modern Physics* **1963**, *35*, 724.
- [35] M. D. Topal, J. R. Fresco. Complementary Base-Pairing And Origin Of Substitution Mutations. *Nature* **1976**, *263*, 285.
- [36] L. C. Sowers, G. V. Fazakerley, R. Eritja, B. E. Kaplan, M. F. Goodman. Base-Pairing And Mutagenesis - Observation Of A Protonated Base Pair Between 2-Aminopurine And Cytosine In An Oligonucleotide By Proton Nmr. *Proc. Nat. Acad. Sci. USA* **1986**, *83*, 5434.
- [37] J. Florian, J. Leszczynski. Spontaneous DNA mutations induced by proton transfer in the guanine cytosine base pairs: An energetic perspective. *J. Am. Chem. Soc.* **1996**, *118*, 3010.
- [38] E. S. Kryachko. The origin of spontaneous point mutations in DNA via Lowdin mechanism of proton tunneling in DNA base pairs: Cure with covalent base pairing. *Int. J. Quantum Chem.* **2002**, *90*, 910.
- [39] E. S. Kryachko, J. R. Sabin. Quantum chemical study of the hydrogen-bonded patterns in A center dot T base pair of DNA: Origins of tautomeric mispairs, base flipping, and Watson-Crick double right arrow Hoogsteen conversion. *Int. J. Quantum Chem.* **2003**, *91*, 695.
- [40] B. Chiavarino, M. E. Crestoni, S. Fornarini, J. Lemaire, P. Maitre, L. MacAleese. pi-complex structure of gaseous benzene-NO cations assayed by IR multiple photon dissociation spectroscopy. *J. Am. Chem. Soc.* **2006**, *128*, 12553.
- [41] L. MacAleese, A. Simon, T. B. McMahon, J. M. Ortega, D. Scuderi, J. Lemaire, P. Maitre. Mid-IR spectroscopy of protonated leucine methyl ester performed with an FTICR or a Paul type ion-trap. *Int. J. Mass Spectrom.* **2006**, *249*, 14.

- [42] A. Simon, L. MacAleese, P. Maitre, J. Lemaire, T. B. McMahon. Fingerprint Vibrational Spectra of Protonated Methyl Esters of Amino Acids in the Gas Phase. *J. Am. Chem. Soc.* **2007**, *129*, 2829.
- [43] R. Prazeres, F. Glotin, C. Insa, D. A. Jaroszynski, J. M. Ortega. Two-colour operation of a Free-Electron Laser and applications in the mid-infrared. *Eur. Phys. J. D* **1998**, *3*, 87.
- [44] J. Lemaire, P. Boissel, M. Heninger, G. Mauclaire, G. Bellec, H. Mestdagh, A. Simon, S. Le Caer, J. M. Ortega, F. Glotin, P. Maitre. Gas Phase Infrared Spectroscopy of Selectively Prepared Ions. *Phys. Rev. Lett.* **2002**, *89*, 273002.
- [45] J. S. Prell, J. T. O'Brien, E. R. Williams. IRPD Spectroscopy and Ensemble Measurements: Effects of Different Data Acquisition and Analysis Methods. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 800.
- [46] C. Lee, W. Yang, R. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter* **1988**, *37*, 785.
- [47] A. D. Becke. Density-functional Thermochemistry .3. the Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648.
- [48] M. J. Frisch, e. al., Gaussian09, Revision C.01. See Supporting Information for complete citation. ed.
- [49] M. P. Gaigeot, N. Leulliot, M. Ghomi, H. Jobic, C. Coulombeau, O. Bouloussa. Analysis of the structural and vibrational properties of RNA building blocks by means of neutron inelastic scattering and density functional theory calculations. *Chem. Phys.* **2000**, *261*, 217.
- [50] A. Hocquet, N. Leulliot, M. Ghomi. Ground-state properties of nucleic acid constituents studied by density functional calculations. 3. Role of sugar puckering and base orientation on the energetics and geometry of 2'-deoxyribonucleosides and ribonucleosides. *J. Phys. Chem. B* **2000**, *104*, 4560.
- [51] J. K. Wolken, C. Yao, F. Turecek, M. J. Polce, C. Wesdemiotis. Cytosine neutral molecules and cation-radicals in the gas-phase structures, energetics, ion chemistry, and neutralization-reionization mass spectrometry. *Int. J. Mass Spectrom.* **2007**, *267*, 30.
- [52] M. D. Halls, J. Velkovski, H. B. Schlegel. Harmonic frequency scaling factors for Hartree-Fock, S-VWN, B-LYP, B3-LYP, B3-PW91 and MP2 with the Sadlej pVTZ electric property basis set. *Theor. Chem. Acc.* **2001**, *105*, 413.
- [53] M. D. Halls, H. B. Schlegel. Comparison of the performance of local, gradient-corrected, and hybrid density functional models in predicting infrared intensities. *J. Chem. Phys.* **1998**, *109*, 10587.
- [54] A. Sen, K. Le Barbu-Debus, D. Scuderi, A. Zehnacker-Rentien. Mass Spectrometry Study and Infrared Spectroscopy of the Complex Between Camphor and the Two Enantiomers of Protonated Alanine: The Role of Higher-Energy Conformers in the Enantioselectivity of the Dissociation Rate Constants. *Chirality* **2013**, *25*, 436.

- [55] E. S. Kryachko, M. T. Nguyen, T. Zeegers-Huyskens. Theoretical study of tautomeric forms of uracil. 1. Relative order of stabilities and their relation to proton affinities and deprotonation enthalpies. *J. Phys. Chem. A* **2001**, *105*, 1288.
- [56] J. K. Wolken, F. Turecek. Proton affinity of uracil. A computational study of protonation sites. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1065.
- [57] A. K. Chandra, M. T. Nguyen, T. Zeegers-Huyskens. Theoretical study of the interaction between thymine and water. Protonation and deprotonation enthalpies and comparison with uracil. *J. Phys. Chem. A* **1998**, *102*, 6010.
- [58] B. Knobloch, H. Sigel. A quantitative appraisal of the ambivalent metal ion binding properties of cytidine in aqueous solution and an estimation of the anti-syn energy barrier of cytidine derivatives. *J. Biol. Inorg. Chem.* **2004**, *9*, 365.
- [59] S. Jaqx, J. Oomens, A. Cimas, M.-P. Gageot, A. M. Rijs. Back Cover: Gas-Phase Peptide Structures Unraveled by Far-IR Spectroscopy: Combining IR-UV Ion-Dip Experiments with Born–Oppenheimer Molecular Dynamics Simulations (Angew. Chem. Int. Ed. 14/2014). *Angew. Chem. Int. Ed.* **2014**, *53*, 3750.

Table and Figure caption

Figure 1: Structure and B3LYP/6-311++G(3df,2p)//B3LYP/6-31+G(d,p)+ZPE relative Gibbs free energies (kJ/mol) of the lowest energy structures obtained for protonated thymidine. See section 1S of the supporting information for a complete survey of the various structures considered.

Figure 2: IR spectrum of protonated thymidine. DFT calculated IR absorption spectrum of the four lowest energy isomers **C2A_O2O4a** (b), **C3A_O2O4a** (c), **C2S_O2a** (d) and **C2A_O4b** (e), compared to the experimental IRMPD spectrum (a) recorded with two IR-FEL macropulses

Table 1: Experimental IRMPD spectrum of protonated thymidine and computed vibrational modes for the **C2A_O2O4a** structure

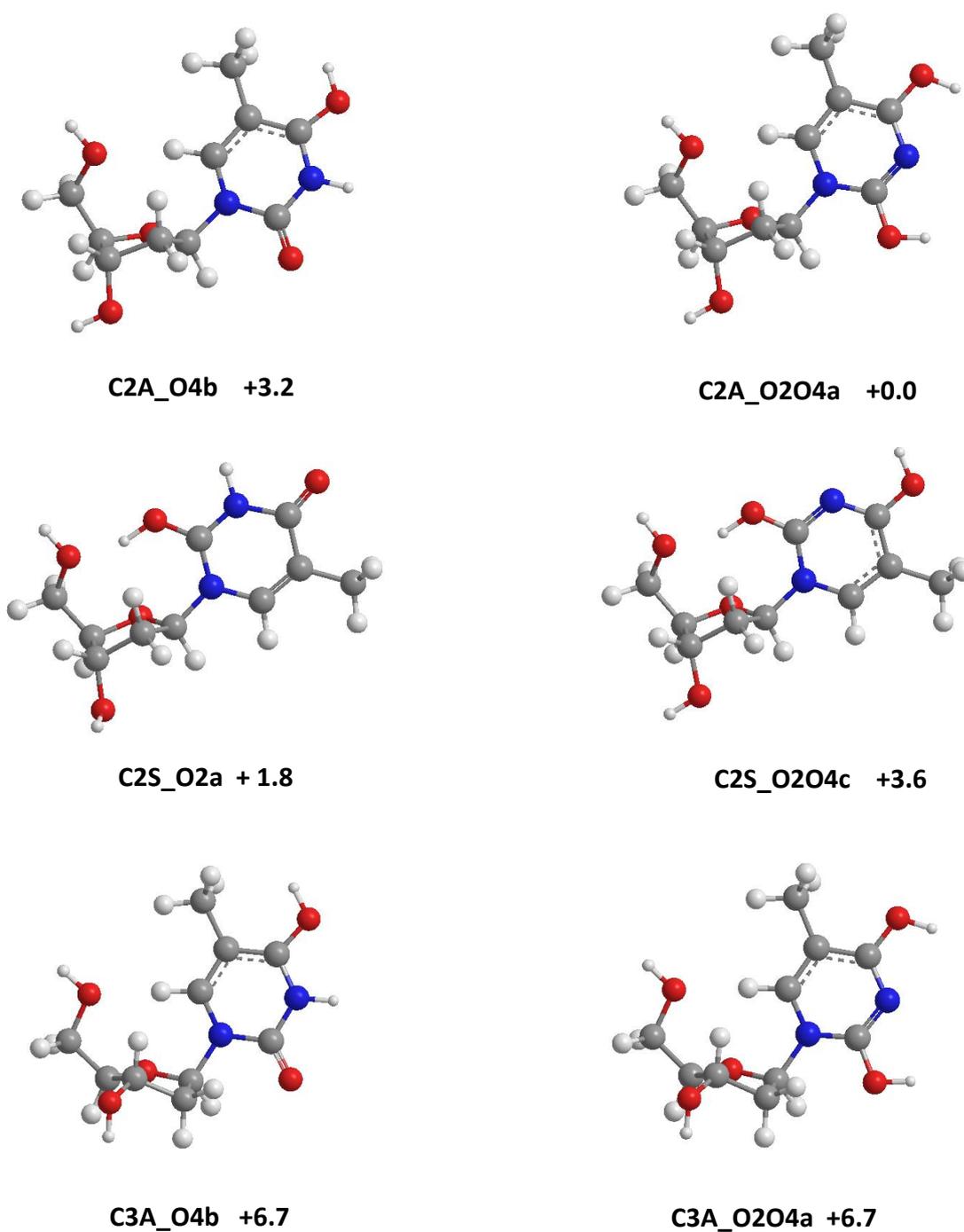


Figure 1: Structure and B3LYP/6-311++G(3df,2p)//B3LYP/6-31+G(d,p)+ZPE relative Gibbs free energies (kJ/mol) of the lowest energy structures obtained for protonated thymidine. See section 1S of the supporting information for a complete survey of the various structures considered.

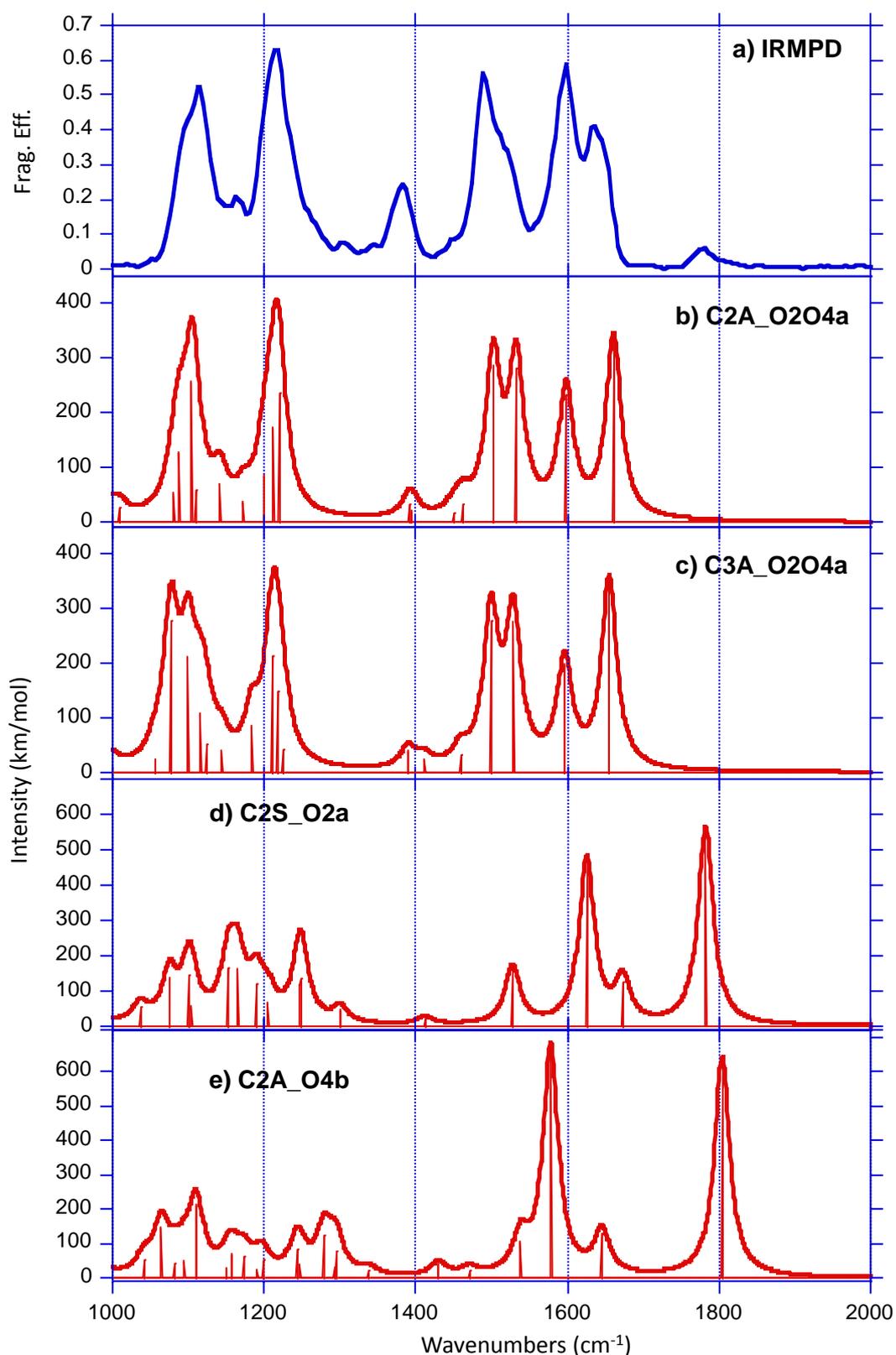


Figure 2: IR spectrum of protonated thymidine. DFT calculated IR absorption spectrum of the four lowest energy isomers **C2A_O2O4a** (b), **C3A_O2O4a** (c), **C2S_O2a** (d) and **C2A_O4b** (e), compared to the experimental IRMPD spectrum (a) recorded with two IR-FEL macropulses.

Table 1: Experimental IRMPD spectrum of protonated thymidine and computed vibrational modes for the **C2A_O2O4a** structure

Wavenumber (cm ⁻¹)		Vibrational mode
IRMPD	Calculated (km/mol) ^a	
1112	1081 (52)	Sugar ring breathing
	1088 (126)	ν C5'O5'
	1004 (256)	ν C3'O3' + ν C1'O4'
	1109 (58)	ν C'1N1
1165	1142 (69)	combined δ O4-H+ δ O2-H sciss.
	1172 (35)	δ O3'H
1215	1200 (85)	δ C2'H ₂ twist.
	1212 (170)	combined δ O4-H+ δ O2-H rock.
	1220 (237)	δ O4-H
1304		
1384	1392 (32)	δ C3'H
	1395 (19)	δ C'H
1450-1560	1450 (16)	δ CH ₂ rock. of CH ₃
	1462 (33)	δ CH ₂ sciss. of CH ₃
	1502 (286)	ν C4O4
	1533 (283)	ν C2O2
1600	1598 (234)	ν C4C5 + ν N1C2
1640	1660 (334)	ν C5C6 + ν C2N3
1780		

^a scaled by a 0.98 scaling factor.

Graphical abstract

