JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Identifying Cytosine-Specific Isomers via High-Accuracy Single Photon Ionization

Ziyong Chen,[†] Kai-Chung Lau,[*,†] Gustavo A. Garcia,[‡] Laurent Nahon,[‡] Dušan K. Božanić,[‡] Lionel Poisson,[§] Muneerah Mogren Al-Mogren,[∥] Martin Schwell,[⊥] Joseph S. Francisco,[#] Ayad Bellili,[▽] and Majdi Hochlaf[*,▽]

[†]Department of Biology and Chemistry, City University of Hong Kong, Kowloon, Hong Kong
[‡]Synchrotron SOLEIL, L'orme des Merisiers, Saint-Aubin - BP 48, 91192 Gif-sur-Yvette Cedex, France
[§]LIDYL, CEA, CNRS, Université Paris-Saclay, CEA Saclay, F-91191 Gif-sur-Yvette, France
[∥]Chemistry Department, Faculty of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Kingdom of Saudi Arabia
[⊥]Laboratoire Interuniversitaire des Systèmes Atmosphériques (LISA), UMR 7583 CNRS, Universités Paris-Est Créteil et Paris Diderot, 61 Avenue du Général de Gaulle, 94010 Créteil, France
[#]Department of Chemistry, University of Nebraska-Lincoln. 433 Hamilton Hall, Lincoln, Nebraska 68588-0304, United States
[▽]Laboratoire Modélisation et Simulation Multi Echelle, MSME UMR 8208 CNRS, Université Paris-Est. 5 bd Descartes, 77454 Marne-la-Vallée, France

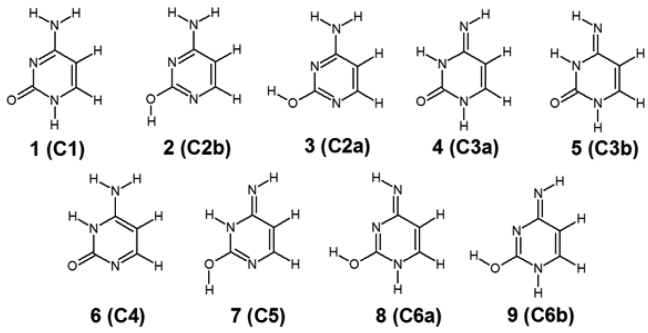
Supporting Information

ABSTRACT: Biological entities, such as DNA bases or proteins, possess numerous tautomers and isomers that lie close in energy, making the experimental characterization of a unique tautomer challenging. We apply VUV synchrotron-based experiments combined with state-of-the-art ab initio methodology to determine the adiabatic ionization energies (AIEs) of specific gas-phase cytosine tautomers produced in a molecular beam. The structures and energetics of neutral and cationic cytosine tautomers were determined using explicitly correlated methods. The experimental spectra correspond to well-resolved bands that are attributable to the specific contributions of five neutral tautomers of cytosine prior to ionization. Their AIEs are experimentally determined for the first time with an accuracy of 0.003 eV. This study also serves as an important showcase for other biological entities presenting a dense pattern of isomeric and tautomeric forms in their spectra that can be investigated to understand the charge redistribution in these species upon ionization.

Ionization of large biomolecules, such as DNA or proteins, may cause structural damage and hence make their structural identification difficult. This is closely connected with biomolecules with hazardous genetic mutations with, for instance, enhanced risk for cancer.[1−3] The underlying mechanisms when energetic particles (e.g., photons, electrons) interact with biological species are not well understood. In particular, complete knowledge of the effects of ionizing radiation on biomolecules requires a deep understanding of the makeup and pattern of the molecular building blocks (i.e., nucleobases and amino acids) and of their spectroscopic and thermochemical properties. Despite a large number of experimental and theoretical studies,[4] fundamental properties such as adiabatic ionization energies (AIEs) of specific DNA bases are still not well established or not known. The experimental complexities of this problem arise from the large number of tautomers and isomers that lie close in energy and possess relatively low potential barriers for interconversion.[5] This results in mixtures of unknown composition of several tautomers and isomers prior to experimental interrogation. In this case, the assignment of species-specific spectra is difficult and not straightforward and hence does not allow the unequivocal determination of their spectroscopic characterization.

Cytosine, of interest in this contribution, falls into this category. Previous theoretical studies[6−8] showed that cytosine possesses several tautomers that lie close in energy (cf. Scheme 1). The most stable form is C2b, followed by C1, C2a, C3a, and C3b. However, all are located within energy differences of less than 0.15 eV. The four other tautomers (C4, C5, C6a, and C6b) are predicted to be less stable and higher in energy (at 0.362, 0.537, 0.725, and 0.858 eV). Therefore, prior to any

Scheme 1. Nine Low-Lying Tautomer and Rotamer Isomers of Cytosine



Received: October 8, 2016
Published: December 6, 2016

DOI: 10.1021/jacs.6b10413
J. Am. Chem. Soc. 2016, 138, 16596−16599

experimental interrogation, there will most likely be a mixture of cytosine tautomers. Strictly speaking, upon ionization all forms are expected to contribute to the ionization spectra with overlapping bands, which would complicate the analysis of cytosine photoionization spectra. Knowledge of the AIEs of cytosine tautomers would provide important fundamental data useful for the analysis. However, such basic data are not well established despite the large number of experimental studies that have examined the photoionization of cytosine (see ref 5). For instance, the measured values for photoionization of cytosine range from 9.0 to 8.45 eV.[9−13] It should be noted that the up-to-date AIE determinations correspond to average values (AIE = 8.60 ± 0.05 eV[7] and 8.66 ± 0.01 eV[14]), i.e., without specificity to any tautomer. Previous studies assumed that the prevalent contribution to the photoionization of cytosine arises from the three lowest-energy isomers. Such an assumption is questionable and biases the spectral assignment and the deduction of the corresponding ionization energies for specific cytosine tautomers. Moreover, the photoionization signal associated with this DNA base might be too low (low density of cytosine in the gas phase upon vaporization, thermal decomposition, and formation of clusters instead of monomers). This highlights the difficulties and challenges associated with the analysis of the photoionization spectra of such monomeric species.

In this combined theoretical and experimental work, we investigated the VUV single photon ionization of cytosine. The experiments were performed at the DESIRS beamline[15] of the French synchrotron SOLEIL coupled to the DELICIOUS3 double-imaging electron/ion coincidence spectrometer (i2PEPICO) of the double-skimmer molecular beam setup SAPHIRS[16,17] combining a velocity map imaging (VMI) electron analyzer with a Wiley−McLaren ion time of flight (TOF) mass spectrometer to detect and image the photoions. Cytosine tautomers were vaporized using an in-vacuum temperature-controlled oven. The resulting vapor was expanded through a micrometric nozzle/skimmer assembly comprising a 70 μm nozzle using Ar as the carrier gas to form a molecular beam, which was skimmed twice before arriving at the ionization chamber. Two Ar backing pressures were used, 0.5 and 3 bar, leading to different translational temperatures. These were measured by ion imaging, as explained in the Supporting Information (SI), and found to be markedly different (390 and 150 K, respectively). Representative ion imaging spectra are shown in Figure 1. The molecular beam conditions (oven temperature, carrier gas pressure, etc.) were optimized to get enough cytosine monomers into the gas phase prior to ionization (cf. Figure 2). Moreover, we were able to get a "cold" cytosine molecular beam possessing several low-energy isomers.

The single photon ionization of cytosine isomers was performed by electron/ion imaging coincidence spectroscopy. After photoionization, the spectra of close to zero kinetic energy photoelectrons (threshold photoelectron photoion coincidence (TPEPICO) spectra) associated with cytosine cations were recorded as a function of the wavelength close to the corresponding adiabatic ionization thresholds. We also collected the ion images and TOFs of these species in order to deconvolve the direct and dissociative ionization processes by their 3D momenta. Hence, we were able to certify that the cytosine did not come from decomposition of cytosine clusters (mainly dimers).

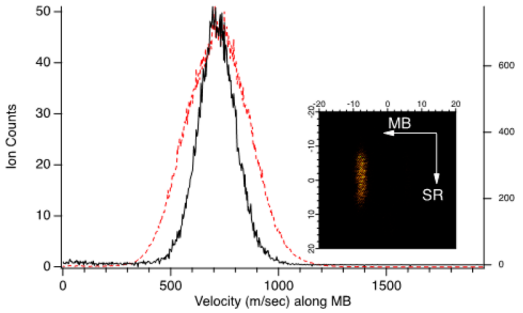


Figure 1. Velocity distributions along the molecular beam (MB) direction extracted from the ion images of the parent cytosine$^+$ for two different Ar backing pressures, 0.5 bar (dashed red) and 3 bar (solid black). The Boltzmann translational temperatures were determined from the widths of the distributions. The inset shows the mass-selected ion image for m/z = 111 for the case of Ar at 3 bar.

Figure 2 displays the TPEPICO spectra of cytosine recorded in the 7.5−9.7 eV photon energy range using different molecular beam conditions. This range covers the thresholds of the low-lying isomers of this nucleobase. As established by

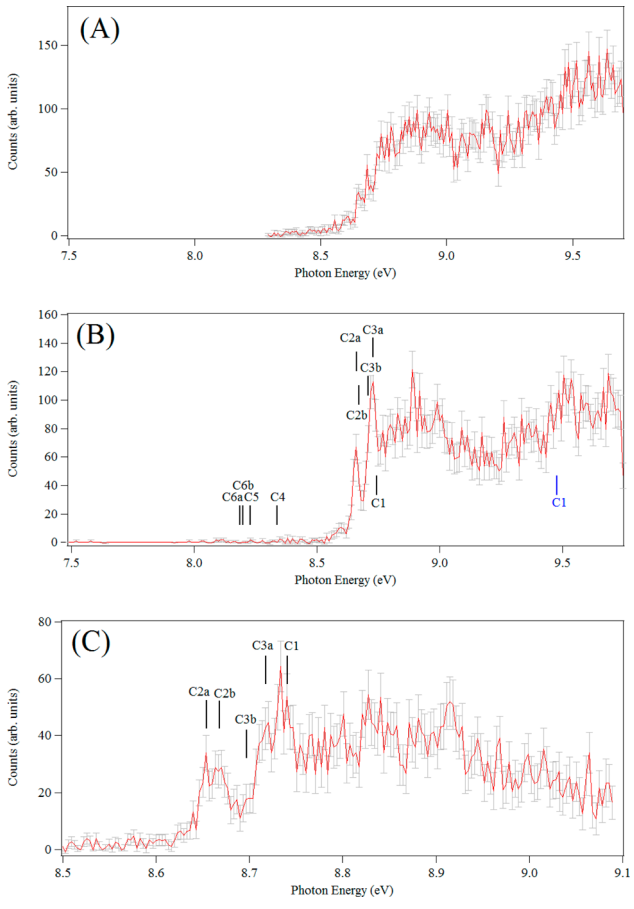


Figure 2. TPEPICO spectra of the cytosine parent (m/z = 111 amu) recorded using (A) an Ar backing pressure of 0.5 bar, (B) Ar at 3 bar and low resolution, and (C) Ar at 3 bar and high resolution. The error bars are shown in gray. In (B), the vertical combs correspond to the computed adiabatic ionization energies (AIEs) of cytosine tautomers as given in Tables 1 and 2. The black (blue) ones are for the population of the cationic ground (first excited) states. In (C), the vertical combs are for the AIEs derived from the fitting of the spectrum (cf. Table 1 and Figure S2).

DOI: 10.1021/jacs.6b10413
J. Am. Chem. Soc. 2016, 138, 16596−16599

Table 1. Adiabatic Ionization Energies (in eV) of Cytosine Isomers As Deduced from the TPEPICO Spectrum and As Computed at the CCSD(T)-F12/cc-pVTZ-F12(+CV+SR+ZPVE) Level of Theory Using the Equilibrium Structures Optimized at the PBE0/aug-cc-pVDZ Level

	C1 → C1$^+$	C2a → C2a$^+$	C2b → C2b$^+$	C3a → C3a$^+$	C3b → C3b$^+$	C4 → C4$^+$	C5 → C5$^+$	C6a → C6a$^+$	C6b → C6b$^+$
				Theory					
this work	8.741	8.661	8.671	8.727	8.700	8.341	8.223	8.188	8.193
ref 6[a]	8.71	8.62	8.64	8.58	8.64	8.31			
ref 8[b]	8.64	8.51	8.53	8.67	8.64				
				Experiment					
this work	8.738 ± 0.003	8.652 ± 0.003	8.669 ± 0.003	8.717 ± 0.003	8.695 (not resolved)	no peak	no peak	no peak	no peak
VUV-SPI[c]				8.68 ± 0.05					
EI[d]				9.0 ± 0.1					
PES[e]				8.45					
VUV-SPI[b]				8.60 ± 0.05					
SPES[f]				8.66 ± 0.01					

[a]CCSD(T), ZPE-corrected [B3LYP/6-31+G(d,p) frequencies]. [b]EOM-IP-CCSD/cc-pVTZ//IP-CISD/6-31+G(d), ZPE-corrected [ωB97X-D/6-31+G(d,p) frequencies]. [c]From ref 10. [d]From ref 9. [e]From ref 11. [f]From ref 14.

Touboul et al.,[14] single photoionization of cytosine occurs mainly by a direct process, so the TPEPICO spectra correspond to the populations of the cationic levels rather than to autoionization lines.

Although the TPEPICO spectra in Figure 2 have similar shapes, there exist noticeable differences due to the different neutral internal energies of cytosine prior to ionization, i.e., a variation of the tautomer population, rotational broadening, and Franck−Condon factors. Indeed, this figure shows that cooling the cytosine molecular beam (by increasing the Ar backing pressure from 0.5 to 3 bar) results in well-resolved photoelectron spectra for the first time for this DNA base, in contrast to the previous large, structureless photoelectronic bands available in the literature for cytosine. The peaks in Figure 2 correspond to the populations of the ground states of the cationic cytosine isomers from their corresponding neutrals. Therefore, the peaks in the cold cytosine spectrum can be associated with the respective AIEs.

In order to deduce the AIEs, we performed a multipeak fit of the high-resolution spectrum (Figure S2). The peak positions are given in Table 1. The resolved peaks were fully assigned after considering the following adiabatic transitions: C1 → C1$^+$, C2a → C2a$^+$, C2b → C2b$^+$, C3a → C3a$^+$, and C3b → C3b$^+$. The adiabatic peak for C3b → C3b$^+$ could not be experimentally resolved, but we nevertheless took it into account for the multipeak fit following the theoretical prediction and found that would be compatible with a location between the C2b and C3a peaks. However, the contribution of the less-stable neutral tautomers (C4, C5, C6a, and C6b) is not evidenced since no peaks were recorded at their expected AIEs (in the 8−8.35 eV range). We confirmed the contributions of the five lowest-energy tautomers of cytosine to the spectra recorded by the group at the Advanced Light Source.[8] It should be noted that the assignment of the band at 8.695 eV to the AIE transition for C3b may not be obvious because of the low intensity and relatively large error bars of this band (see Table 1).

The present work and that of Bravaya et al.[8] raise questions about the assumptions of Trofimov et al.,[18] who assigned their cytosine photoelectron spectrum by assuming that the only contribution to the spectrum arises from the C2b tautomer. For $h\nu$ > 8.75 eV, the spectra are composed of several overlapped bands corresponding to the populations of the vibronic bands of cytosine tautomer cations.

State-of-the art computations were carried out as described in the SI to test and further validate the spectral assignments. They correspond to the determination of the equilibrium structures, energetics, and frequencies of both the neutral and ionic species. We list in Table 1 the computed AIEs together with their comparison to the present experimental ones. The agreement between the measured and computed AIEs of the cytosine tautomers is quite good, allowing the individual identification of at least four cytosine conformers. Indeed, the differences between the observed and computed AIEs are less than 10 meV (absolute, without prior scaling). The current value of the uncertainty is significantly smaller than the error limit of ±0.035 eV after comparing pulsed field ionization−photoelectron (PFI−PE) experimental results with standard CCSD(T) predictions of the AIEs of hydrocarbon radicals.[19] In addition, our calculations show that only slight geometrical changes occur upon ionization of the observed cytosine isomers (cf. Figure S4). This suggests that the Franck−Condon factors for the 0 ← 0 transitions are highly favorable.

We also computed the $D_0 \rightarrow D_1$ and $D_0 \rightarrow D_2$ excitation energies at the MRCI-F12/aug-cc-pVDZ level (see the SI). Our data are listed in Table 2. Combining the AIEs of Table 1 and the excitation energies of Table 2 shows that the computed ionization energies for the D_1 transition from their corresponding neutral states are higher than 9.9 eV for all of the isomers except for C1. Therefore, the second band at the right-hand side of the upper traces of Figure 1 is assigned to the population of the D_1 state of the C1 tautomer solely.

Table 2. MRCI-F12/aug-cc-pVDZ Vertical Ionization Energies (in eV) of the Two Lowest Doublet States (D_1 and D_2) of C1$^+$, C2a$^+$, C2b$^+$, C3a$^+$, and C3b$^+$, Computed at the Equilibrium Geometries of the Corresponding Ions and Given with Respect to the Energies of the Cations' Ground States (D_0) at Equilibrium

	C1$^+$	C2a$^+$	C2b$^+$	C3a$^+$	C3b$^+$
$D_0 \rightarrow D_1$	0.736	1.244	1.456	1.350	1.573
$D_0 \rightarrow D_2$	1.583	2.358	2.294	1.937	2.148

To date, only the AIE of thymine has been determined accurately by Choi et al.[20] after analysis of their mass-analyzed

DOI: 10.1021/jacs.6b10413
J. Am. Chem. Soc. 2016, 138, 16596−16599

threshold ionization spectrum for thymine obtained with a very high resolution (~0.1 meV). It should be noted that thymine is less challenging than cytosine since only a unique tautomer of thymine contributes to the experimental spectra. At present we have identified without ambiguity the specific contributions of the five lowest-lying isomers of cytosine to the photoionization spectra. This was possible after monitoring of the physical conditions of the neutral molecular beam together with the outstanding performance of the imaging detectors. Moreover, the first-principles ab initio computed energetics are in excellent agreement with those deduced experimentally (within the error bars of the experiments).

We show in the SI that the charge distribution upon ionization depends on the considered tautomer. Accordingly, the charge dynamics after ejection of an electron from the outermost molecular orbitals of cytosine tautomers should be quite different and tautomer-dependent. Such effects can be probed using the recently developed ultrafast XUV pump−probe excitation light source-based experiments.[21,22] Therefore, our work adds a new dimension that leads to a much better definition of the electronic structures of the cations and grants access to their state-selected photochemistry and to an understanding of the effects of the interaction of the ionizing matter on DNA double-strand formation and RNA folding. More generally, this work not only highlights the importance of a combined experimental and theoretical approach to provide accurate spectral characterization of a complex biomolecule such as cytosine but also establishes the capabilities of this approach for future studies of other biological entities and their oligomers and clusters.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b10413.

Experimental and theoretical methods, density differences, and optimized structures and frequencies of the neutral and ionic forms of the cytosine tautomers (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*kaichung@cityu.edu.hk
*hochlaf@univ-mlv.fr

ORCID

Kai-Chung Lau: 0000-0003-2125-6841
Gustavo A. Garcia: 0000-0003-2915-2553
Majdi Hochlaf: 0000-0002-4737-7978

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors extend their appreciation to the International Scientific Partnership Program ISPP at King Saud University for funding this research work through ISPP Grant 0045. We are indebted to the general technical staff of Synchrotron Soleil for running the facility under Proposal 20150956. We also thank Jean-François Gil for his technical help on the SAPHIRS molecular beam chamber.

■ REFERENCES

(1) Hagen, U. *Radiat. Environ. Biophys.* **1986**, *25*, 261.
(2) Kumar, A.; Sevilla, M. D. *Chem. Rev.* **2010**, *110*, 7002.
(3) Shikazono, N.; Noguchi, M.; Fujii, K.; Urushibara, A.; Yokoya, A. *J. Radiat. Res.* **2009**, *50*, 27.
(4) Schwell, M.; Hochlaf, M. *Top. Curr. Chem.* **2014**, *355*, 155.
(5) Das, T.; Ghosh, D. *J. Phys. Chem. A* **2014**, *118*, 5323.
(6) Wolken, J. K.; Yao, C.; Turecek, F.; Polce, M. J.; Wesdemiotis, C. *Int. J. Mass Spectrom.* **2007**, *267*, 30.
(7) Kostko, O.; Bravaya, K. B.; Krylov, A. I.; Ahmed, M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 2860.
(8) Bravaya, K. B.; Kostko, O.; Dolgikh, S.; Landau, A.; Ahmed, M.; Krylov, A. I. *J. Phys. Chem. A* **2010**, *114*, 12305.
(9) Verkin, B. I.; Sukodub, L. F.; Yanson, I. K. *Dokl. Akad. Nauk SSSR* **1976**, *228*, 1452.
(10) Orlov, V. M.; Smirnov, A. N.; Varshavsky, Y. M. *Tetrahedron Lett.* **1976**, *17*, 4377.
(11) Dougherty, D.; Younathan, E. S.; Voll, S.; Abdulnur, S.; McGlynn, S. P. *J. Electron Spectrosc. Relat. Phenom.* **1978**, *13*, 379.
(12) Hush, N. S.; Cheung, A. S. *Chem. Phys. Lett.* **1975**, *34*, 11.
(13) Fulfer, K. D.; Hardy, D.; Aguilar, A. A.; Poliakoff, E. D. *J. Chem. Phys.* **2015**, *142*, 224310.
(14) Touboul, D.; Gaie-Levrel, F.; Garcia, G. A.; Nahon, L.; Poisson, L.; Schwell, M.; Hochlaf, M. *J. Chem. Phys.* **2013**, *138*, 094203.
(15) Nahon, L.; de Oliveira, N.; Garcia, G. A.; Gil, J.-F.; Pilette, B.; Marcouille, O.; Lagarde, B.; Polack, F. *J. Synchrotron Radiat.* **2012**, *19*, 508.
(16) Tang, X.; Garcia, G. A.; Gil, J.-F.; Nahon, L. *Rev. Sci. Instrum.* **2015**, *86*, 123108.
(17) Garcia, G. A.; Cunha de Miranda, B. K.; Tia, M.; Daly, S.; Nahon, L. *Rev. Sci. Instrum.* **2013**, *84*, 053112.
(18) Trofimov, A. B.; Schirmer, J.; Kobychev, V. B.; Potts, A. W.; Holland, D. M. P.; Karlsson, L. *J. Phys. B: At., Mol. Opt. Phys.* **2006**, *39*, 305.
(19) Lau, K.-C.; Ng, C. Y. *J. Chem. Phys.* **2006**, *124*, 044323.
(20) Choi, K. W.; Lee, J. H.; Kim, S. K. *J. Am. Chem. Soc.* **2005**, *127*, 15674.
(21) Marciniak, A.; Despré, V.; Barillot, T.; Rouzée, A.; Galbraith, M. C. E.; Klei, J.; Yang, C.-H.; Smeenk, C. T. L.; Loriot, V.; Reddy, S. N.; Tielens, A. G. G. M.; Mahapatra, S.; Kuleff, A. I.; Vrakking, M. J. J.; Lépine, F. *Nat. Commun.* **2015**, *6*, 7909.
(22) Calegari, F.; Ayuso, D.; Trabattoni, A.; Belshaw, L.; De Camillis, S.; Anumula, S.; Frassetto, F.; Poletto, L.; Palacios, A.; Decleva, P.; Greenwood, J. B.; Martín, F.; Nisoli, M. *Science* **2014**, *346*, 336.

DOI: 10.1021/jacs.6b10413
J. Am. Chem. Soc. 2016, 138, 16596−16599