International Journal of Mass Spectrometry 283 (2009) 214–221

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

Tautomerism of cytosine probed by gas phase IR spectroscopy

Joost M. Bakker[a], Jean-Yves Salpin[b], Philippe Maître[c,*]

[a] FOM Institute for Plasma Physics Rijnhuizen, Edisonbaan 14, 3439 MN Nieuwegein, The Netherlands
[b] Laboratoire Analyse et Modélisation pour la Biologie et l'Environnement (LAMBE), Université d'Evry Val d'Essonne – CNRS – Bâtiment Maupertuis – Boulevard François Mitterrand, 91025 Evry, France
[c] Laboratoire de Chimie Physique, Université Paris-Sud – CNRS – Faculty of Sciences, Bâtiment 350, 91405 Orsay, France

ARTICLE INFO

Article history:
Received 4 December 2008
Received in revised form 27 March 2009
Accepted 30 March 2009
Available online 10 April 2009

Keywords:
Mass spectrometry
Gas-phase ion structure
IR Photodissociation spectroscopy
Electrospray ionization
Cytosine

ABSTRACT

The tautomerism of the gaseous protonated cytosine is studied using infrared multiple photon dissociation (IRMPD) spectroscopy of singly hydrated complexes of protonated cytosine in the 2700–3750 cm^{-1} wavenumber range. The hydrated complexes are formed through argon-mediated collisions between bare electrosprayed cytosine and low-pressure water vapor. In the spectra, where X–H (X=C, O, and N) stretching vibrations are probed, evidence is found for the coexistence in the gas phase of hydrated complexes of two different cytosine tautomers. As the addition of a water molecule to either tautomer of protonated cytosine is energetically highly unlikely to induce interconversion, it is deduced that both C(2)=O and N(3) protonated tautomers of cytosine are formed under electrospray conditions.

1. Introduction

The study of the conformational space of DNA and its building blocks is of fundamental interest for a better understanding of its physical and chemical properties relevant for replication mechanisms, DNA–protein and –drug interactions. Although most studies of the structures of DNA subunits have been done in solute and crystalline phases, gas-phase studies are also of particular interest. Not only is it possible to obtain information on their intrinsic properties by eliminating any influence from solvent and conformational ensemble averaging effects, studies in the gas phase also allow for a direct comparison between experimental data and quantum chemical calculations.

While it has been difficult to transfer the notoriously labile DNA fragments intact into the gas phase, the advent of soft ionization methods such as electrospray [1] and matrix-assisted laser desorption [2] that allow for facile formation of ions in the gas phase, has made gas-phase studies of these species feasible. The use of tandem mass spectrometric (MS) methods with structure sensitive activation techniques has enabled to obtain direct structural information on these species.

One such activation technique is ion mobility (IM) which allows for the measurement of the collision cross-sections of molecules and can separate different conformations of flexible molecules by making use of their different collision cross-sections. Prof. Bowers, who has been a pioneer in the development of this technique, has realized the possibilities it offered for structural characterizations of DNA building blocks and has studied as one of the first such aspects as the conformational shapes of gas-phase DNA complexes [3], G-quadruplexes [4,5], oligonucleotides [6,7] and hydrated mononucleotides [8].

A second strongly structure sensitive activation technique is infrared (IR) multiple photon dissociation (IRMPD). Where IM is the method of choice for larger species of a flexible nature where there exists multiple conformers or families of conformers for which the collision cross-sections exhibit clear differences, for smaller species IRMPD spectroscopy is the more relevant technique. The use of free-electron lasers (FELs) as a source of intense, tunable IR radiation has made it possible to perform IRMPD spectroscopy over a wide spectral range in the mid-IR. The combination of powerful Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometers with FELs has opened up the possibility of multi-step tandem mass spectrometry with resonant IRMPD as a highly structure-sensitive activation technique [9–11].

Despite the continuously growing capabilities to transfer larger and larger species into the gas phase, some very fundamental questions at the scale of the nucleobases remain. Such questions range from the possibility of tautomerism of the nucleobases and its implications on point mutations which develop in the nucleic acid replication [12,13], to the location of the protonation site of the nucleobases, which plays a role in stabilization of triplex structures and in mutagenic processes.

In their canonical form the three pyrimidic nucleobases, cytosine, thymine and uracil, adopt keto or diketo forms to complement

* Corresponding author. Fax: +33 1 69 15 61 88.
E-mail address: philippe.maitre@u-psud.fr (P. Maître).

1387-3806/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2009.03.014

their partner nucleobases guanine and adenine. However, through internal proton transfer, different tautomeric forms of the pyrimidic bases can be formed. The thermochemistry of gas-phase ions derived from biologically important molecules has been studied extensively by tandem mass spectrometry (MS). Although several MS [14–17] studies have characterized the gas-phase basicity and proton affinity of the nucleobases, structural information from MS can only be derived using theory [18–27]. In the case of protonated cytosine, for instance, the transition state associated with tautomerization lies well below the fragmentation threshold, and it is thus difficult to distinguish the tautomers through collision induced dissociation [19]. It is therefore clear that direct information on the protonation sites of nucleic bases and on the presence of tautomers in the gas phase is of fundamental importance.

In a previous study in this group the IRMPD spectra of bare protonated uracil, thymine and cytosine, formed under ESI conditions, were reported, studied using the IR-FEL at CLIO in Orsay, France [27]. All experimental spectra, recorded in the fingerprint region 1000–1900 cm^{-1}, compared favorably to the calculated spectrum of the lowest energy structure, all enolic forms. From this finding it was deduced that the predominant tautomer of the protonated nucleobases under electrospray conditions is enolic. However, the presence in each experimental spectrum of a clear resonance around 1800 cm^{-1} hinted at the presence of another tautomeric form. As in this spectral region only the characteristic frequency for stretching vibrations of a C=O bond are found, it was deduced that a second tautomer must be present. To verify this it would be of interest to probe the spectral region where N–H, O–H and C–H stretching vibrations are found, the range between roughly 2500 and 3500 cm^{-1}. Unfortunately, the FEL at CLIO does not cover this wavenumber range. Tabletop IR sources are typically not intense enough to drive the IRMPD process in such strongly bound systems as the protonated nucleobases. In a recent publication, we have found an elegant way of avoiding these problems [28]. After generating the bare protonated uracil by ESI, and complexing this with one water molecule in a low-pressure collision cell, the IRMPD spectrum for the protonated uracil–water complex was recorded between 2500 and 3500 cm^{-1}. The IRMPD spectrum of this system revealed the presence of two different structures, where a quantum chemical analysis convincingly demonstrated the origin of the two structures in two tautomeric forms of uracil, an enolic and a ketonic one [28]. Here we present an IR photodissociation spectroscopic study of protonated cytosine–water clusters, where the loss of a water molecule is used as the signature of IR absorption. Quantum chemical calculations will provide evidence that the water molecule does not alter the tautomeric form, and thus purely acts as a messenger for the absorption. It will be shown below that the earlier evidence for the presence of multiple tautomers of protonated cytosine under ESI conditions is reinforced.

2. Experimental and computational methods

IRMPD spectroscopy of the protonated cytosine–water complex is performed employing a commercial 7 T FT-ICR mass spectrometer (Bruker, Apex Qe) into which pulsed IR light is coupled. The detailed layout of this experimental apparatus is described elsewhere [11] and only a brief overview is given here. Protonated cytosine ions are prepared in an ESI source by introducing a 10^{-4} M aqueous solution of the nucleobase prepared in pure water (purified with a Milli-Q water purification system) into the source using direct infusion with a syringe pump. Typical ESI parameters are a flow rate of 3 μL/min, a spray voltage of 3500 V and a capillary temperature of 200 °C. The most important feature of the mass spectrometer for the present study is its quadrupole–hexapole interface between the electrospray source and the ICR cell. The bias voltage and RF amplitude of the quadrupole are adjusted to selectively transmit protonated cytosine. Mass-selected ions are then trapped in a ~5 cm long hexapole ion-trap contained within a collision cell where ions normally are collisionally cooled using a flow of high-purity argon buffer gas. In the present study, water vapor is seeded in the argon gas flow. Ions are typically stored in the ion-trap over periods of ~50 ms, leading to solvation of some ~60% of the protonated cytosine ions. Ions are then pulse-extracted towards the ICR cell where mass-selection of mono-solvated protonated cytosine is performed. They are then irradiated with IR light, after which the resulting ions are mass-analyzed. If the IR light is in resonance with an IR active vibrational mode of singly hydrated protonated cytosine ions stored in the ICR cell, IR photons can be absorbed by the ions. The absorption of one or a few IR photons can lead to fragmentation of the mass-selected hydrated ions yielding bare protonated cytosine. By monitoring the number of bare protonated cytosine ions and that of the protonated cytosine water complex while varying the frequency of the IR light, an IRMPD spectrum is obtained where the IRMPD yield is plotted as a function of the IR wavenumber. The IRMPD yield R is expressed as a function of the abundances of the parent and fragments ions as $R = -\log(I_{\text{parent}}/(I_{\text{parent}} + \sum I_{\text{fragment}}))$. For each wavelength, the mass spectrum is the Fourier transform of the average of five time-domain transients.

In the current study, the 2700–3750 cm^{-1} wavenumber range is explored using an IR OPO/OPA laser (LaserVision), pumped by a 25 Hz Nd:YAG laser (Innolas Spitlight 600, 550 mJ per pulse, 1064 nm, 4–6 ns pulse duration). In the presently covered wavenumber range, the typical output energy slowly decreases from ~12 mJ per pulse at 3600 cm^{-1} to ~4 mJ at 2800 cm^{-1} with a spectral bandwidth of ~5 cm^{-1}. The IR laser beam is steered into the FT-ICR along the axis of the magnetic trapping field and loosely focused using a spherical mirror with a 2 m focal distance. The light beam has a nearly constant diameter throughout the ICR cell region.

To aid in interpretation of the experimental data *ab initio* calculations were performed. Structure characterizations were performed using density functional theory (DFT) calculations with the Gaussian03 program suite [29]; relative energies were also determined at the MP2 and CCSD(T) level using the MOLPRO program [30]. Geometries were first optimized at the B3LYP/6-31G++(d,p) level, and then re-optimized at the B3LYP/6-311++G(3df,2p) level. Enthalpies and entropies were determined using standard formulas of statistical thermodynamics based on calculated moments of inertia and harmonic frequencies as implemented in Gaussian03 [29].

The characterization of the energy profile associated with the water-assisted tautomerization was performed at the B3LYP/6-31G++(d,p) level. Using the corresponding optimized geometries, single point energy calculations at the MP2 level were carried out using both 6-311G++(d,p) and 6-311++G(3df,2p) Gaussian basis set. CCSD(T) energies were only calculated in the smaller basis set, and the CCSD(T) energy values in the larger basis set were estimated using a standard formula for the basis set expansion, i.e., CCSD(T)/6-311++G(3df,2p) ≈ CCSD(T)/6-311G++(d,p) + MP2/6-311++G(3df,2p) − MP2/6-311G++(d,p).

IR absorption spectra of the three lowest energy structures of monohydrated protonated cytosine presented here are derived using B3LYP/6-311++G(3df,2p) harmonic vibrational frequencies. In order to facilitate the comparison of IRMPD spectra with calculated IR absorption spectra, each calculated band is convoluted assuming a lorentzian profile with an associated width (FWHM) of 10 cm^{-1}.

3. Results

3.1. IRMPD spectrum

The IRMPD spectrum for the protonated cytosine–water clusters is presented in the top panel of Fig. 1 as the logarithm of the ratio of

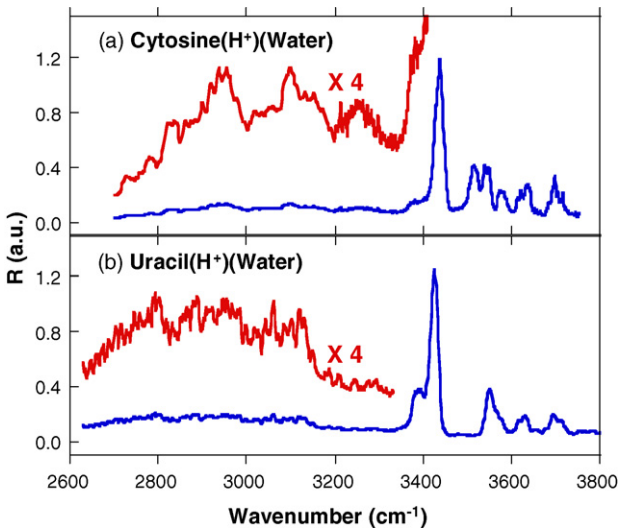


Fig. 1. IR photodissociation spectrum recorded with the OPO/OPA laser system of singly hydrated protonated cytosine (a). Below is the spectrum for hydrated protonated uracil obtained using the same experimental conditions [28] (b). The photodissociation efficiency *R* is given in arbitrary units.

parent ions to the sum of parent and fragment ions. For comparison it is accompanied by the spectrum of protonated uracil–water.

The IR spectra for cytosine and uracil show great similarities. Both are dominated by sharp resonances in the 3400–3750 cm^{-1} wavenumber range and exhibit a broad structured resonance in the lower wavenumber range (2700–3300 cm^{-1}). Both spectra exhibit a continuous background in the observed IRMPD signal. This background is observed even when the IR photodissociation laser is switched off and is attributed to spontaneous dissociation under influence of blackbody radiation produced by the cell wall. As the spontaneous dissociation rate is low as compared to the laser-induced dissociation it will not affect the discussion below.

Some general remarks can be made before a detailed discussion of the IRMPD spectrum. First, it is a well-known feature of IRMPD spectra of complex ions involving hydrogen bonds that the stretching vibrations of NH or OH groups that are donor in a hydrogen bond exhibit a strong red-shift accompanied by a dramatic broadening of the peak [28,31,32]. The broad resonance structure observed in the 2700–3300 cm^{-1} region could well be a signature of these effects.

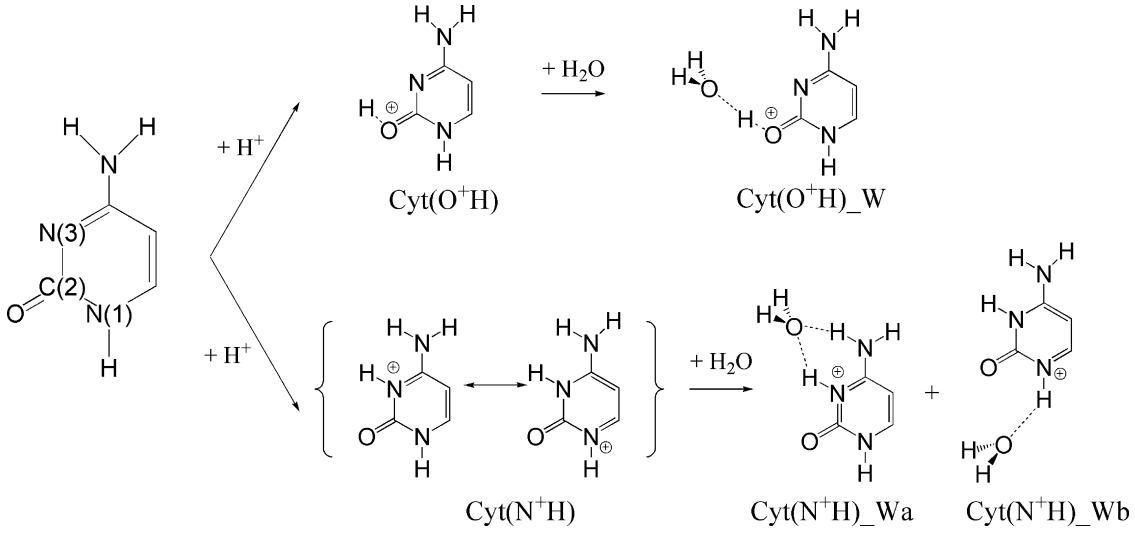
Second, each cytosine–water structure under consideration would produce two resonances for OH stretching vibrations of the water molecule plus four due to N–H and/or O–H stretching vibrations of the cytosine moiety to form a total of six resonances. It can therefore immediately be deduced that multiple isomers are responsible for the observed IRMPD spectrum as one single conformer cannot account for the at least seven observed resonances.

When comparing the spectra for cytosine and uracil one distinguishes in both spectral regions more resonances for cytosine than for uracil. This could indicate either the presence of more isomers for cytosine–water or that the structural differences between the various conformers of cytosine–water are more pronounced than for uracil–water.

3.2. Calculated structures

To aid in the interpretation of the observed spectra DFT calculations were performed. As in our previous communication on hydrated uracil, the lowest energy structures of protonated cytosine form the starting point. Two structures that are formed upon protonation of the most stable tautomer of neutral cytosine compete as the lowest energy structure. In previous high level *ab initio* calculations it was found that protonation of the oxygen C(2)=O oxo group is slightly more favorable than protonation of the ring N(3) and a calculated value for ΔG at 298 K at the CCSD(T)/aug-cc-pVTZ level of 6.4 kJ/mol was found [19]. Also, it was reported that a correct energy ordering can be obtained using the B3LYP method provided that a sufficiently large basis set, including diffuse function is used, but that the energy difference is significantly underestimated [19]. This is indeed found in the present calculations at the B3LYP/6-311++G(3df,2p) level, where the Cyt(O^+H) is only slightly lower in energy than the Cyt(N^+H) tautomer and the ΔG value at 298 K is only 0.6 kJ/mol. Nevertheless, the near degeneracy of these two tautomers is consistent with our recently reported IRMPD spectrum of protonated cytosine ions generated by electrospray since the IRMPD spectrum in the 1000–2000 cm^{-1} spectral region indicate the presence of both tautomers [27].

Addition of a water molecule leads to structures where the water molecule will coordinate towards the proton, as is schematically depicted in Scheme 1. Hydration of Cyt(O^+H) leads to structure Cyt(O^+H)_W where the water molecule shares the proton with the keto group. In both Cyt(N^+H)_Wa and Cyt(N^+H)_Wb structures, the water molecule shares the proton with a ring nitrogen. In the Cyt(N^+H)_Wa structure, however, the water molecule acts as a dou-



Scheme 1. The formation of different tautomers of protonated cytosine and the possible structures of its hydrated complexes.

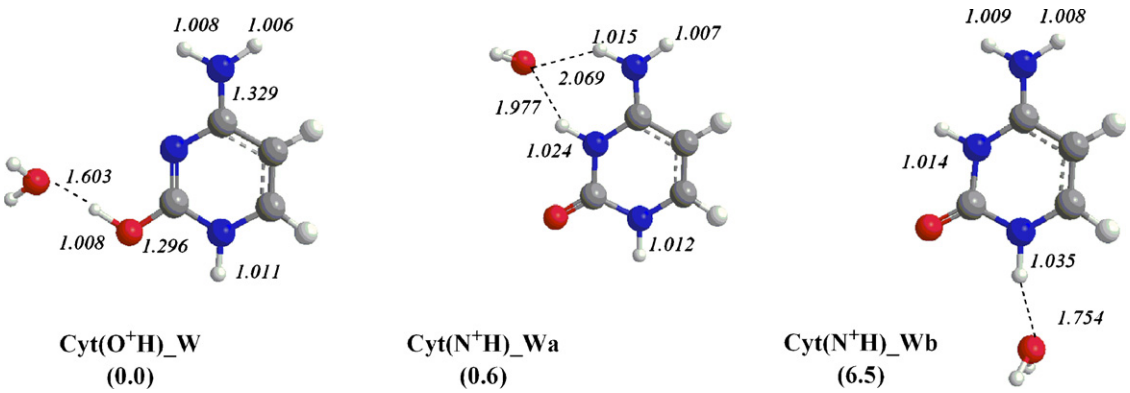


Fig. 2. Optimized structures of monohydrated protonated cytosine at the B3LYP/6-311++G(3df,2p) level of theory. Relative Gibbs free enthalpies at 298 K (within parentheses) are given in kJ/mol, bond lengths in Å.

Table 1
Calculated (B3LYP/ B3LYP/6-311++G(3df,2p)) enthalpies (ΔH) at 0 and 298 K and Gibbs free enthalpies (ΔG) at 298 K of the optimized structures of the protonated cytosine–water complexes, and the binding energies of water at 298 K. All values are given in kJ/mol.

Structure	Relative energies			Water binding energy
	ΔH(0 K)	ΔH(298 K)	ΔG(298 K)	ΔG(298 K)
Cyt(O+H)_W	3.5	4.4	0	26.3
Cyt(N+H)_Wa	0	0	0.6	25.2
Cyt(N+H)_Wb	9.4	10.9	6.5	19.2

ble hydrogen bond acceptor, and forms a second hydrogen bond with an NH of the amino group. As can be seen in Table 1, the formation of this cyclic hydrogen bond network is favorable from an enthalpy point of view. If the enthalpies are considered at 298 K, structure Cyt(N+H)_Wa is more favorable than Cyt(O+H)_W and Cyt(N+H)_Wb by 3.5 and 9.4 kJ/mol, respectively. However, if the free energies are considered at 298 K structures Cyt(O+H)_W and Cyt(N+H)_Wa are nearly degenerate (Fig. 2).

The binding energy of a water molecule to protonated cytosine is on the same order of magnitude as for uracil. At the B3LYP/6-311++G(3df,2p) level of theory, the 0 K binding energy of structure Cyt(O+H)_W is 60.3 kJ/mol, and the difference in free energy at 298 K is 26.3 kJ/mol (25.0 kJ/mol for uracil [28]). One can thus estimate that the absorption of one or two IR photons is in principle sufficient for inducing the fragmentation of monohydrated protonated cytosine in the frequency range explored here.

4. Discussion

The IR photodissociation spectrum of singly hydrated protonated cytosine is given in Fig. 3 together with the IR absorption spectra calculated for the three structures discussed above, where the calculated vibrational frequencies are scaled by a factor 0.965. The stick spectra are accompanied by a convolution of the stick spectrum with a 10 cm^{-1} width (fwhm) lorentzian line shape function. The discussion of the IR photodissociation spectrum will be divided in two parts. First, the 3300–3750 cm^{-1} range, where sharp resonances (bands B–G in Fig. 3) are observed that are typically due to free NH and OH stretching vibrations. Second, we discuss the 2700–3300 cm^{-1} spectral region, where broader features (band A in Fig. 3) are observed.

The observed frequencies in the IRMPD spectra for both cytosine are reported in Table 2, along with those observed in the case of uracil [28]. In the following, we will discuss the band assignment proposed in Table 2 for singly hydrated protonated cytosine ions.

4.1. Free NH and OH stretches in the 3300–3750 cm^{-1} spectral range

From our knowledge on the IR photodissociation spectrum of protonated uracil–water, for all calculated structures, IR active normal modes in the 3300–3750 cm^{-1} spectral region are free NH and OH stretching vibrations. On the blue side of the IR photodissocia-

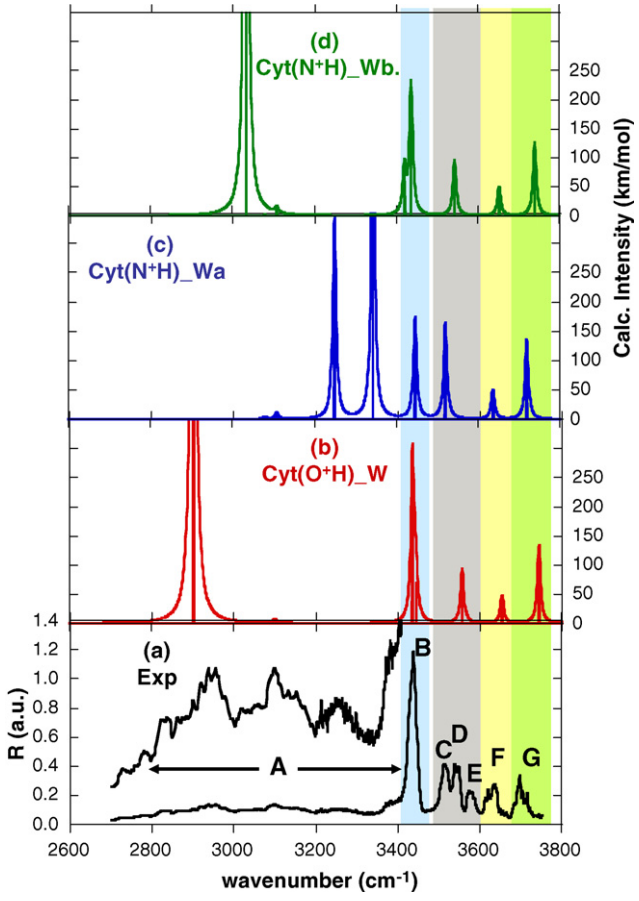


Fig. 3. IR spectra of monohydrated protonated cytosine in the 2700–3750 cm^{-1} spectral range. (a) Experimental spectrum: the photodissociation efficiency R (left vertical scale) is given in arbitrary units. (b–d) Calculated IR absorption spectra of structures Cyt(O+H)_W, Cyt(N+H)_Wa and Cyt(N+H)_Wb, respectively; calculated intensities (right vertical scale) are given in km/mol. Additionally, each IR stick spectrum is accompanied by a convoluted spectrum (lorentzian function, fwhm = 10 cm^{-1}). Scaling factor is 0.965.

Table 2
Observed IRMPD resonances (in cm^{-1}), their labeling in Fig. 3 and their assignments for singly hydrated protonated cytosine (left) and uracil (right). Where no structure is indicated the resonance is assigned to a common vibration in all structures.

Singly hydrated protonated cytosine				Singly hydrated protonated uracil		
Observed frequency	Label	Structure	Mode description	Observed frequency		Mode description
				2920		H bonded C(2)O–H···OH_2(A)
2948	A	Cyt(O^+H)_W	C(2)O–H stretch			
3120		Cyt(N^+H)_Wb	N(1)–H stretch			
3253		Cyt(N^+H)_Wa	N(3)–H stretch			
3396		Cyt(N^+H)_Wa	H bonded NH_2 sym NH···OH_2(AA)			
				3390	[b]	Free N(3)–H
				3425		Free N(1)–H
3436	B		Free NH			
3516	C	[a]	NH_2			
3544	D	[a]	asym			
			NH	3556	[c]	Free O–H
3577	E	[a]				
3628	F		Water sym OH	3625		Water sym OH
3700	G		Water asym OH	3702		Water asym OH

[a] No structure assigned.
[b] Shoulder on red side of peak B (Ref. [28], Fig. 4).
[c] Assigned to structures U(KE)H^+_Wa and U(KE)H^+_Wb, Ref. [28]).

tion spectrum, two features are observed at 3628 and 3700 cm^{-1}, corresponding to bands labeled as F and G in Fig. 3, respectively. As can be seen in Fig. 1, these bands are very similar to those observed in the high frequency part of the IR photodissociation spectrum of the singly hydrated protonated uracil, where they were assigned to the symmetric and asymmetric water OH stretching vibrations. The calculated spectra offer a favorable comparison which suggests the same assignment here. One can notice that the observed photodissociation efficiencies for these two vibrational modes of water are approximately equal. This might be somewhat surprising since the calculated IR intensity for the asymmetric combination is predicted to be approximately three times larger than that for the symmetric combination of the water OH stretches. Previously similar discrepancies between the ratio of calculated IR intensities and the ratio of observed photodissociation yields for the symmetric and asymmetric combinations of water stretches have been observed for singly hydrated protonated uracil, and for other water solvated clusters. It was recently proposed that this low photodissociation efficiency is in turn the result of the inefficiency of the intramolecular vibrational redistribution (IVR) of the energy from the excited asymmetric water stretch to the bath of the vibrational modes of the ion [33].

At 3426 cm^{-1}, the most intense feature in the spectrum is observed. In the case of Cyt(O^+H)_W and Cyt(N^+H)_Wb, there are two IR active modes: the ring NH stretch and the symmetric combination of the amino NH stretches. The latter is the most intense and is predicted at the same frequency (scaled value of 3455 cm^{-1}) for the two structures. In the case of Cyt(N^+H)_Wa, there is only one IR active mode in this region, and it corresponds to the free ring amide N(1)H stretch which is predicted to be at a slightly higher frequency (3464 cm^{-1}). Considering the small energy differences between these calculated IR active modes for the various isomers, the observation of the band at 3426 cm^{-1} is not structurally informative either since it could result from the IR photodissociation of any of the three potentially competing isomers. The same conclusion was reached for singly hydrated protonated uracil, where the strongest signal in the IR photodissociation spectrum was observed at 3425 cm^{-1}, which was assigned to a free NH stretching mode in multiple isomers [28].

It turns out that the spectral range between 3500 and 3600 cm^{-1} is the most diagnostically useful part of the IR photodissociation spectrum. Three bands, labeled C, D and E in Fig. 3, are observed at 3516, 3544, and 3577 cm^{-1}, respectively. Comparison to the calculated IR absorption spectra of the three structures (Fig. 3b–d) suggests that each of these bands could be the signature of one of the three structures, each being characterized by a single IR active mode in this spectral region. Structures Cyt(N^+H)_Wb and Cyt(O^+H)_W are characterized by the asymmetric combination of the free amino NH stretches calculated at 3563 and 3575 cm^{-1}, respectively. The analysis of the normal modes of structure Cyt(N^+H)_Wa suggests that the absorption at 3534 cm^{-1} is associated with the free amino NH stretch.

The accuracy of the theoretical method is likely to prevent the assignment of the bands C–E to a specific structure. Nevertheless, the observation of these three well resolved IR photodissociation bands clearly points to the formation of the three isomers of singly hydrated protonated cytosine displayed in Fig. 2.

As illustrated in other hydrogen bonded systems [34], the position of the free OH stretching vibration of the water molecules can be highly structurally diagnostic. On the basis of these experiments, both the symmetric and the asymmetric OH combinations for a double acceptors water molecule are expected to be red-shifted with respect to those for a single acceptor water molecule. While the water OH stretching frequencies for Cyt(O^+H)_W and Cyt(N^+H)_Wb are predicted to be the same, symmetric and asymmetric OH stretches for structure Cyt(N^+H)_Wa are predicted to be red-shifted by 15 (20) cm^{-1}. Unfortunately, it seems that a better resolution would be needed for observing this red-shift effect.

4.2. Red-shifted NH and OH stretches in the 2700–3300 cm^{-1} spectral range

It is well-known that the absorption band associated with the stretching vibration of NH or OH groups involved in a hydrogen bond as a donor are strongly red-shifted, and dramatically broadened [28,31,32]. Upon complexation with H_2O, the NH_4^+ ion, for example, exhibits marked frequency red-shifting in its NH stretches accompanied by a large broadening [31]. This tendency for the ammonium NH stretches to shift to lower frequency and broaden upon forming hydrogen bond with water molecules has also been observed for water solvated protonated valine [32], for example. For singly hydrated protonated uracil [28], a broad band was observed from 2650 through 3150 cm^{-1}, and had been assigned to the contribution of the red-shifted OH stretching vibrations of the two types of structures that could be formed experimentally by the addition of water on the two lowest-energy tautomers of protonated uracil.

In the present case, the IR photodissociation spectrum is more structured than in the case of uracil, and three broad features can

be distinguished in the low frequency range (2700–3300 cm^{-1}) of the spectrum (Fig. 3a). Comparison with the calculated IR absorption spectrum of the three isomers may further confirm that the three structures are formed under our experimental conditions. Indeed, each of the three structures of singly hydrated protonated cytosine has a characteristic band in the lower frequency part of the spectrum. The most red-shifted band, with a maximum at ~2950 cm^{-1}, is likely associated to the hydrogen bonded OH stretch of the Cyt(O^+H)_W structure for which the calculated frequency (scaled) is 2916 cm^{-1}.

On the opposite, the two broad bands observed at 3253 and 3396 cm^{-1} could be the signatures of two relatively weakly hydrogen bonded NH stretches as in structure Cyt(N^+H)_Wa. Due to the cyclic constraints of the hydrogen bond network, which is characteristic of this isomer, the two hydrogen bonds are relatively weak and the shift in the corresponding NH stretching frequencies is consequently not as large as those for structures Cyt(O^+H)_W and Cyt(N^+H)_Wb: the two NH stretching vibrations are predicted at 3267 and 3362 cm^{-1}. When comparing these values to the maximum found around 3250 cm^{-1} and to the shoulder of band B just below 3400 cm^{-1} it seems not unlikely that these could be the signature for structure Cyt(N^+H)_Wa. Finally, the band maximum observed around 3100 cm^{-1} can be compared to the calculated frequency for the N(1)H stretching vibration (3048 cm^{-1}) for Cyt(N^+H)_Wb and could provide a signature of this third isomer.

4.3. Tautomerization of protonated cytosine in the gas phase

From the above discussion it appears that the recorded IR photodissociation spectrum of singly hydrated protonated cytosine shown in Fig. 3a provides clear evidence for the formation of the three isomers displayed in Fig. 2. These three structures are formed upon addition of water on mass-selected protonated cytosine ions after their formation in the electrospray process. For this purpose, ions are stored for about 50 ms in a pressurized linear hexapole ion trap where water vapor is seeded in the argon gas flow. The remaining question to be answered is whether the relative abundance of the two tautomers of protonated cytosine that result from the protonation of C(2)=O and N(3), respectively, may change due to water catalyzed tautomerization that could also take place in the experiment.

In a recent study in our group, it was shown that the IRMPD spectrum of bare protonated cytosine suggested that the two Cyt(N^+H) and Cyt(O^+H) tautomers are already formed by electrospray [27]. It was suggested that the C(2)=O protonated tautomer is the most abundant, since the IR signature of the N(3) protonated tautomer, the normally strong IR active carbonyl stretch, was only observed as a relatively weak signal. Although these earlier experiments were conducted in another experimental apparatus consisting of a 3D Paul ion trap equipped with an ESI source, the current apparatus has a very similar source and similar ESI parameters are used. It is therefore reasonable to assume that similar populations of the two tautomers are created in the source of the current experiment. Nevertheless, it is important to evaluate whether further tautomerization can occur in the gas phase during the microsolvation process.

The activation energy corresponding to the 1,3 proton transfer from Cyt(O^+H) to Cyt(N^+H) for the isolated protonated cytosine is associated with a large isomerization barrier. An activation energy value of 157 kJ/mol has been calculated at a high *ab initio* level (DH(0K) at the CCSD(T)/6-311++G(3df,2p) level) [19]. Upon addition of a base, however, such 1,3 proton transfer can be efficiently catalyzed. This has been observed for protonated thymine where the corresponding activation barrier is significantly reduced upon addition of an ammonia molecule [35]. Since the proton affinity of water is significantly lower than that of ammonia (691 and 854 kJ/mol, respectively), one would expect that the decrease in the tautomerization activation energy is lower in the case of water. This is supported by our recent study of the water-assisted tautomerization of protonated uracil where it has been shown that the free energy barrier at 298 K is reduced from 164.7 to 69.5 kJ/mol by the addition of water. Nevertheless, the corresponding rate-limiting transition state is energetically well above (ΔG(298 K) = 25.0 kJ/mol) the fragmentation channels associated with the loss of water [28].

The energy profile associated with the water-assisted tautomerization of protonated cytosine has been studied, and the results are summarized in Fig. 4. The geometries of the stationary points were determined at the B3LYP/6-31++G(d,p) level and the corresponding ΔG(298 K) values are also given in Fig. 4. Single point energy calculations were also performed using the MP2 and CCSD(T) approaches, and the relative energies of the most important stationary points are given in Table 3. As in the case of uracil [28], the rate-limiting transition state (TS2 in Fig. 4) is found to correspond to the transfer of H_3O^+ from N(3) to C(2)=O. The activation energy is indeed significantly reduced as compared to the case of the bare

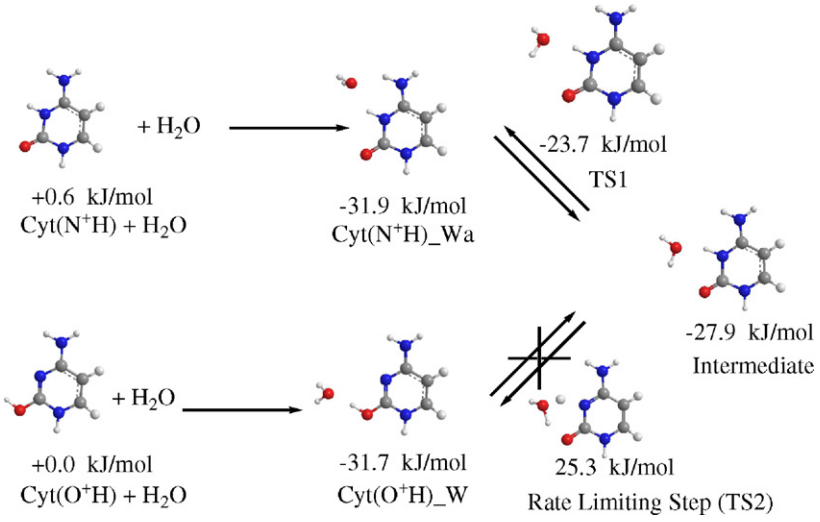


Fig. 4. Optimized structures and corresponding Gibbs free enthalpies at 298 K (B3LYP/6-31++G(d,p)) associated with the water catalyzed 1,3 proton transfer reaction converting Cyt(O^+H)_W into Cyt(N^+H)_Wa. The Gibbs free enthalpies are given relative to the Cyt(O^+H) plus water asymptote.

Table 3
Relative energies in units of kJ/mol of stationary points along the energy profile associated with the tautomerization of singly hydrated protonated cytosine calculated with respect to the lowest fragmentation channels corresponding to O(2) protonated cytosine plus water. Geometries were optimized at the B3LYP/6-31++G(d,p) level, and thermal and entropy corrections were determined at the same level of theory.

	Method[a]	Relative energies (kJ/mol)[b]						
		Cyt(N^+H)_Wa	TS1	Intermediate	TS2	Cyt(O^+H)_W	Cyt(N^+H) + H_2O	Cyt(O^+H) + H_2O
ΔH(0 K)	B3LYP/B1	−67.0	−59.9	−60.0	−15.5	−63.5	1.1	0.0
ΔH(298 K)	B3LYP/B1	−68.7	−62.5	−60.3	−21.5	−64.3	1.5	0.0
ΔG(298 K)	B3LYP/B1	−31.9	−23.7	−27.9	25.3	−31.7	0.6	0.0
ΔG(298 K)	MP2/B1	−30.0			33.5	−27.9	3.9	0.0
ΔG(298 K)	MP2/B2	−28.1			27.0	−29.0	4.3	0.0
ΔG(298 K)	RCCSD(T)/B1	−34.4			34.2	−32.4	5.0	0.0
ΔG(298 K)	RCCSD(T)/B2	−32.5			27.7	−33.5	5.4	0.0

[a] Single point MP2 and CCSD(T) energies were calculated in two basis sets, B1 and B2 refers to 6-31++G(d,p) and 6-311++G(3df,2p) basis set respectively.
[b] The CCSD(T) electronic energy in the B2 basis set was estimated as E[CCSD(T)/B2] = E[CCSD(T)/B1] = E[MP2/B2] − E[MP2/B1].

ions. At the B3LYP/6-31++G(d,p) level, the calculated 0 K energy value for the activation energy is 51.5 kJ/mol for the water catalyzed process, as compared to 158 kJ/mol for the isolated protonated cytosine. Nevertheless, as can be seen in Fig. 4, the corresponding transition state is energetically well above the fragmentation channels corresponding to the loss of water from either of the two protonated cytosine tautomers. As can be seen in Table 3, the relative energy of TS2 with respect to the two lowest energy of naked and singly hydrated tautomers of cytosine calculated at the *ab initio* MP2 and CCSD(T) levels seems to confirm that the transition state associated with the rate limiting step lies well above (∼27.7 kJ/mol at our highest level of theory) the lowest fragmentation threshold of singly hydrated protonated cytosine.

Based on these results, it is very likely that loss of water from the singly hydrated protonated cytosine complex is strongly preferred over water-assisted tautomerization for each tautomeric form. As the observed dissociation in the ICR cell over retention times exceeding 250 ms is below 10%, it appears reasonable to neglect water-assisted tautomerization. These results are consistent with our previous IR spectroscopic results on bare protonated pyrimidic bases which led us to the conclusion that both C(2)=O and N(3) protonated cytosine ions are formed by electrospray ionization. Therefore, as in the case of protonated uracil, the microsolvation occurring in the pressurized hexapole ion-trap leaves the tautomer population formed in the ESI source unchanged while forming clusters by the addition of one water molecule to each of the two tautomers.

5. Concluding remarks

The gas phase IR photodissociation spectrum of singly hydrated protonated cytosine has been recorded in the 2700–3750 cm^{-1} spectral range. The number of well resolved and sharp features observed above 3300 cm^{-1} readily shows that more than one structure is formed under our experimental conditions. Since in the present case the hydrated complex ions are generated by water attachment to bare protonated cytosine it is proposed that the structures involved reflect the tautomeric abundances of protonated cytosine, as formed under ESI conditions. An analysis of possible water-assisted tautomerization between calculated structures suggests that the solvation process does not affect the relative abundances of the tautomeric forms. *Ab initio* calculations suggest three possible structures for the hydrated protonated cytosine resulting from the addition of water on C(2)=O and N(3) protonated cytosine tautomers. Comparison of the experimental IR photodissociation spectrum with the calculated linear IR absorption spectra for the three structures allows for a clear assignment of most of the observed features, while others can tentatively be attributed to one of the three proposed structures. These conclusions are further supported by the analysis of the spectrum below 3300 cm^{-1} where broad features characteristic of hydrogen bonded NH and OH stretching vibrations are observed.

Acknowledgment

We gratefully acknowledge the European Commission for a generous grant associated to the EPITOPES (Electron Plus Infrared TO Probe and Elucidate Structures, EC project #15637) funded through the NEST (New and Emerging Science and Technology) program of the 6th framework program.

References

[1] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Science 246 (1989) 64.
[2] M. Karas, F. Hillenkamp, Anal. Chem. 60 (1988) 2299.
[3] J. Gidden, E.S. Baker, A. Ferzoco, M.T. Bowers, Int. J. Mass Spectrom. 240 (2005) 183.
[4] V. Gabelica, E.S. Baker, M.P. Teulade-Fichou, E. De Pauw, M.T. Bowers, J. Am. Chem. Soc. 129 (2007) 895.
[5] E.S. Baker, S.L. Bernstein, V. Gabelica, E. De Pauw, M.T. Bowers, Int. J. Mass Spectrom. 253 (2006) 225.
[6] J. Gidden, J.E. Bushnell, M.T. Bowers, J. Am. Chem. Soc. 123 (2001) 5610.
[7] J. Gidden, M.T. Bowers, Eur. Phys. J. D 20 (2002) 409.
[8] D.F. Liu, T. Wyttenbach, M.T. Bowers, J. Am. Chem. Soc. 128 (2006) 15155.
[9] J. Lemaire, P. Boissel, M. Heninger, G. Mauclaire, G. Bellec, H. Mestdagh, A. Simon, S. Le Caer, J.M. Ortega, F. Glotin, P. Maitre, Phys. Rev. Lett. 89 (2002) 273002.
[10] J.J. Valle, J.R. Eyler, J. Oomens, D.T. Moore, A.F.G. van der Meer, G. von Helden, G. Meijer, C.L. Hendrickson, A.G. Marshall, G.T. Blakney, Rev. Sci. Instrum. 76 (2005) 023103.
[11] J.M. Bakker, T. Besson, J. Lemaire, D. Scuderi, P. Maitre, J. Phys. Chem. A 111 (2007) 13415.
[12] M.D. Topal, J.R. Fresco, Nature 263 (1976) 285.
[13] L.C. Sowers, G.V. Fazakerley, R. Eritja, B.E. Kaplan, M.F. Goodman, Proc. Natl. Acad. Sci. U.S.A. 83 (1986) 5434.
[14] M.S. Wilson, J.A. McCloskey, J. Am. Chem. Soc. 97 (1975) 3436.
[15] M. Meot-Ner, J. Am. Chem. Soc. 101 (1979) 2396.
[16] F. Greco, A. Liguori, G. Sindona, N. Uccella, J. Am. Chem. Soc. 112 (1990) 9092.
[17] C.C. Nelson, J.A. McCloskey, J. Am. Soc. Mass Spectrom. 5 (1994) 339.
[18] C.X. Yao, M.L. Cuadrado-Peinado, M. Polasek, F. Tureček, J. Mass Spectrom. 40 (2005) 1417.
[19] C.X. Yao, F. Tureček, M.J. Polce, C. Wesdemiotis, Int. J. Mass Spectrom. 265 (2007) 106.
[20] A.K. Chandra, M.T. Nguyen, T. Uchimaru, T. Zeegers-Huyskens, J. Phys. Chem. A 103 (1999) 8853.
[21] A.K. Chandra, M.T. Nguyen, T. Zeegers-Huyskens, J. Phys. Chem. A 102 (1998) 6010.
[22] Y. Podolyan, L. Gorb, J. Leszczynski, J. Phys. Chem. A 104 (2000) 7346.
[23] A.K. Chandra, D. Michalska, R. Wysokinsky, T. Zeegers-Huyskens, J. Phys. Chem. A 108 (2004) 9593.
[24] C. Colominas, F.J. Luque, M. Orozco, J. Am. Chem. Soc. 118 (1996) 6811.
[25] M.T. Nguyen, A.K. Chandra, T. Zeegers-Huyskens, J. Chem. Soc., Faraday Trans. 94 (1998) 1277.
[26] J.K. Wolken, F. Tureček, J. Am. Soc. Mass Spectrom. 11 (2000) 1065.
[27] J.Y. Salpin, S. Guillaumont, J. Tortajada, L. MacAleese, J. Lemaire, P. Maitre, Chemphyschem 8 (2007) 2235.
[28] J.M. Bakker, R.K. Sinha, T. Besson, M. Brugnara, P. Tosi, J.Y. Salpin, P. Maitre, J. Phys. Chem. A 112 (2008) 12393.
[29] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, N.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski,

G.A.A. Petersson, P.Y.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople, Gaussian 98, Gaussian, Inc., Pittsburgh, PA, 1998.

[30] H.-J. Werner, P.J. Knowles, R. Lindh, F.R. Manby, M. Schütz, P. Celani, T. Korona, G. Rauhut, R.D. Amos, A. Bernhardsson, A. Berning, D.L. Cooper, M.J.O. Deegan, A.J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, A.W. Lloyd, S.J. McNicholas, W. Meyer, M.E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A.J. Stone, R. Tarroni, T. Thorsteinsson, MOLPRO, version 2006.1, A package of ab initio programs, see http://www.molpro.net., 2006.

[31] Y.S. Wang, H.C. Chang, J.C. Jiang, S.H. Lin, Y.T. Lee, H.C. Chang, J. Am. Chem. Soc. 120 (1998) 8777.

[32] A. Kamariotis, O.V. Boyarkin, S.R. Mercier, R.D. Beck, M.F. Bush, E.R. Williams, T.R. Rizzo, J. Am. Chem. Soc. 128 (2006) 905.

[33] T. Pankewitz, A. Lagutschenkov, G. Niedner-Schatteburg, S.S. Xantheas, Y.T. Lee, J. Chem. Phys. 126 (2007).

[34] M.F. Bush, R.J. Saykally, E.R. Williams, J. Am. Chem. Soc. 130 (2008) 15482.

[35] R.H. Wu, T.B. McMahon, J. Am. Chem. Soc. 129 (2007) 569.