PCCP

View Online Dynamic Article Links

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 6799–6807

www.rsc.org/pccp

PAPER

Tautomers of cytosine and their excited electronic states: a matrix isolation spectroscopic and quantum chemical study†‡

Gábor Bazsó,[a] György Tarczay,*[a] Géza Fogarasi[b] and Péter G. Szalay*[b]

Received 1st November 2010, Accepted 22nd December 2010
DOI: 10.1039/c0cp02354j

We have measured the IR and UV spectra of cytosine in a low-temperature argon matrix. An attempt was made to determine the tautomeric ratios existing in the matrix, making use of the matrix-isolation IR spectrum and computed IR intensities of the tautomers in a least squares fitting procedure. The mole fractions are about 0.22 for oxo(-amino) form, 0.26 and 0.44 for the two rotamers, respectively, of the hydroxy(-amino) form and 0.08 for the (oxo-)imino tautomer. These ratios were then used to simulate the matrix-isolation UV spectrum as a composite of the individual spectra, the latter calculated *ab initio* at high levels of electron correlation theory. The agreement between simulated and experimental UV spectra seems satisfactory. This indicates that, in contrast to the solid state and solution spectra described up to now by the oxo(-amino) form alone, the reproduction of the matrix-isolation UV spectrum needs at least the hydroxy(-amino) and oxo(-amino) forms, and probably also the (oxo-)imino form.

1. Introduction

The four nucleotide bases of DNA, including cytosine, are prototypical examples of molecules that are prone to tautomerism. This intramolecular proton transfer would lead to completely different hydrogen bond patterns in the double helix, thus—as recognized from the beginning of molecular genetics[1]—destroying the genetic code. One method of tautomerization may be through photochemical processes, the understanding of which necessitates detailed knowledge of the excited electronic states, including those of various tautomers.

The three lowest-energy tautomers of cytosine, two of which may exist in rotamer pairs, are shown in Fig. 1. There is general agreement that in the solid state as well as in aqueous solutions only the *canonical* oxo form (**1**) is present. In the vapor state, however, all spectroscopic and quantum chemistry (QC) studies prove that the hydroxy tautomer dominates. For the latter, very often only one rotamer (**2b**) was taken into consideration, although its partner **2a** is only 0.7–0.8 kcal mol^{-1} higher in energy.[2] Isomer **3b** is at $\Delta E^{\circ} \cong 3.0$ kcal mol^{-1} relative to **2b** and will be omitted. The other four isomers have relative Gibbs free energies within a narrow range of $\sim$2 kcal mol^{-1} only or even lower[2–7] and will be included in the present study. The isomers are well separated by high-energy barriers. All QC results in the literature put the barriers between the *tautomers* around 40 kcal mol^{-1}; to get an estimate for the *rotational* barriers we have run various DFT and MP2 calculations with results of 9–10 kcal mol^{-1} and 23–25 kcal mol^{-1} for **2b** → **2a** and **3a** → **3b**, respectively. In accordance with this, the existence of several tautomers in the gas phase or in low-temperature inert matrices was also proved by various spectroscopic methods.[8–20]

It is well known that the UV light absorption of nucleobases can lead to mutagenic and carcinogenic defects.[21] Therefore,

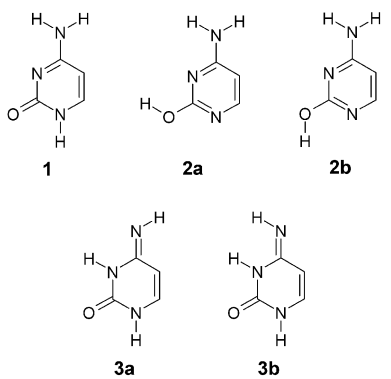

Fig. 1 The three most stable tautomers of cytosine: oxo(-amino) (**1**), hydroxy(-amino) (**2**) and (oxo-)imino (**3**). Both rotamers of **2** and **3** are shown.

[a] *Laboratory of Molecular Spectroscopy, Institute of Chemistry, Eötvös Loránd University, Pf. 32, Budapest, H-1518, Hungary. E-mail: tarczay@chem.elte.hu; Tel: +36 1 3722500*
[b] *Laboratory of Theoretical Chemistry, Institute of Chemistry, Eötvös Loránd University, Pf. 32, Budapest, H-1518, Hungary. E-mail: szalay@chem.elte.hu; Tel: +36 1 3723931*

† This article was submitted as part of an issue to coincide with Faraday Discussion 150: Frontiers in Spectroscopy.

‡ Electronic supplementary information (ESI) available: Experimental, and computational data used for calculating tautomer distribution, as well as equilibrium geometry and LVC parameters (ground state normal mode frequencies, and intrastate coupling parameters) for all three tautomers used in the spectrum simulation. See DOI: 10.1039/c0cp02354j

the understanding of the UV spectra of nucleotide bases is of outstanding importance. There have been numerous experimental[18,19,22–27] and theoretical[28–36] studies in this respect on cytosine. Among these, in a recent paper[35] some of us analyzed the UV spectrum of cytosine on the basis of high level QC calculations. This study—as with all other theoretical interpretations of the UV spectra in the literature—had two restrictions. First, even though QC results (normally referring to non-interacting molecules) should be compared to measurements carried out either in the gas phase or in an inert matrix, in each former study the computed spectra were compared to experimental spectra measured in the solid state or solutions. Second, only the oxo (**1**) form was taken into consideration in the analyses of the UV absorption.

The reason for using condensed phase UV spectra as experimental reference was that the only available matrix isolation UV (MI-UV)[25] spectrum was overlooked, and the more recent resonance enhanced multiphoton ionization (REMPI) spectrum[18,19] covers a very narrow range (only below 4.65 eV) of the spectral region only. Although not only MI-UV spectrum but also matrix isolation infrared (MI-IR)[11–14,37–39] spectra of cytosine are available in the literature, in the present study we have re-recorded these spectra in a low-temperature Ar matrix for several purposes. One of the most important goals was to obtain the MI-IR and MI-UV spectra under the same conditions as far as possible, in order to get spectra composed of the same tautomeric ratios. This made it possible to determine the ratios from the (well-resolved and easy-to-simulate) IR spectrum and use them for the simulation of the UV spectrum. The other goal was to obtain a better quality MI-UV spectrum, which is corrected much more carefully for scattering background than the previously published one.[25]

As for the theory, we have now included in the investigation of excited electronic states of all three tautomers, performing a series of coupled cluster (CC) calculations for the lowest lying singlet A' electronic states. Beside vertical excitation energies, the vibronic spectrum has been simulated by the Linear Vibronic Coupling (LVC) model. The IR spectra were calculated by Density Functional Theory (DFT) and by Møller–Plesset second-order perturbation theory, MP2.

The primary aim of the present study is to interpret the MI-UV spectrum of cytosine, with emphasis on the possibility of various tautomers coexisting. Thus, in the first step we try to determine the tautomer ratios from the IR spectrum, making use of the calculated IR intensities. The theoretical UV spectrum will then be constructed from the individual computed spectra of tautomers using the mole ratios obtained.

2. Methods

2.1 Spectroscopic measurements

In the course of the experiments we tried to ensure that, as far as possible, the conditions were the same for the IR and UV measurements. Some minor differences were unavoidable due to the different molar absorptions in the IR and UV, and due to the different thermal conductivity of window materials used for IR and UV spectroscopic measurements. Cytosine (Aldrich, 99%) was evaporated into the vacuum chamber by using a home built Knudsen effusion cell. The copper body was heated by a coil and its temperature was measured by a thermocouple. The sample was placed inside the cell in a glass sample holder. The evaporated sample first entered into a $\sim$0.5 cm^3 buffer chamber, and then left the cell through a variable-size pinhole. The evaporated sample was mixed with argon (Messer, 99.9997%) before deposition. The gas flow rate was regulated by an MKS flow controller system, kept at *ca.* 1.2 sccm and 0.3 sccm for IR and UV measurements, respectively. The temperature of the Knudsen cell was optimized to get the shortest possible deposition time and keep the concentration low enough to minimize the formation of dimers in the matrix. The applied evaporation temperature was $\sim$450 $\pm$ 5 K and $\sim$430 $\pm$ 5 K for IR and UV spectroscopic measurements, respectively.

The sample–noble gas mixture was deposited onto an 8 K CsI window for IR, and onto a 12 K quartz window for UV-Vis spectroscopic measurements. These windows were mounted on a Janis CCS-350R cold head cooled by a CTI Cryogenics 22 closed-cycle refrigerator unit coupled to a CTI 8200 compressor. The temperature of the copper cold head was controlled by a Lake Shore 321 thermostat equipped with a silicon diode thermometer.

IR spectra were recorded by a Bruker IFS 55 Fourier transform infrared (FT-IR) spectrometer equipped with a KBr beamsplitter and a DTGS detector. A total of 1000 scans were accumulated at 1 cm^{-1} resolution in the 400–4000 cm^{-1} spectral window. The Happ–Genzel apodization function, Mertz phase correction using phase resolution of 32 cm^{-1}, and zero filling factor of 4 were applied.

Absorption UV-Vis spectra were recorded by a Varian Cary3E spectrometer, using 5 nm min^{-1} scan rate, 0.333 nm step size, and 1 nm spectral band width. The data were collected in the 190–400 nm spectral region. In order to correct for the scattering background, the matrix was photolyzed with broadband UV light of a HPK 125 W high-pressure mercury vapor lamp. As a result of this, cytosine has completely decomposed in the matrix. Since, in our experimental conditions, the photolysis products (mainly HCN) has no observable UV absorption in this region, the UV spectrum of this photolized matrix was used to compensate for the scattering background in the corresponding UV spectrum of cytosine in Ar. In principle the matrix might anneal locally due to the photodecomposition of cytosine which could change the scattering, but according to our experience in similar UV decomposition experiments the scattering background of the MI-IR spectra did not change considerably.

Among the different conditions applied during the deposition of the matrix for UV and IR measurements and during acquisition of spectra only the vaporization temperature could affect directly the tautomeric ratios. The temperature difference in the two experiments is, however, so small on the absolute temperature scale that it changes the Boltzmann factors insignificantly.

2.2 Quantum chemistry

Vibrational spectra were calculated by standard techniques using the PQS program package.[40] To check the sensitivity of

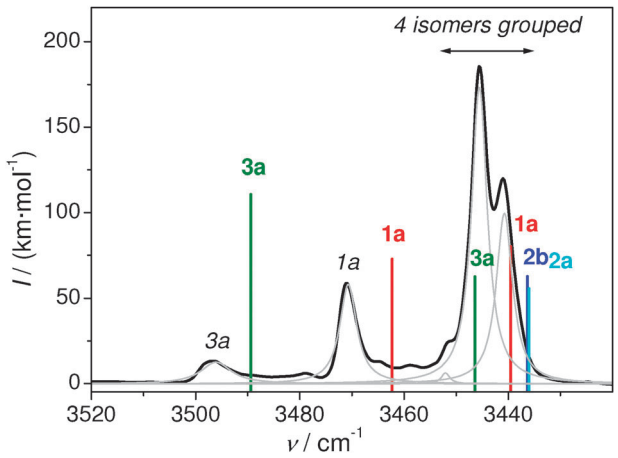

Fig. 2 A representative part of the MI-IR spectrum showing the decomposition of the spectrum to Lorentzian functions and also the computed positions and intensities (not weighted by mole ratios) of the individual tautomers. Symbols and text in italics show the individual or grouped assignments of the experimental spectral bands.

the results to the level of theory, both DFT (B3LYP)[41] and MP2 calculations[42] were performed, with basis sets ranging from 6-31++G(d,p)[43] up to aug-cc-pVTZ.[44]

An accurate treatment of excited electronic states requires more elaborate computations, involving less routine work. All of these calculations have been performed by the CFOUR suite of codes,[45] applying the methods CCSD (Coupled Cluster Singles Doubles),[46] CC3,[47] EOMEE-CCSD (Equation of Motion Excitation Energy CCSD)[48] and CC3-LR (CC3 Linear Response)[49] and using the correlation-consistent basis sets cc-pVDZ, aug-cc-pVDZ.[44] To stay consistent with our earlier work[35] the geometries have been optimized under the planarity constraint at the CCSD/cc-pVDZ level of theory for all three tautomers. These structures are given in the ESI.‡ Vertical excitation energies have been calculated at the EOMEE-CCSD and CC3-LR levels with aug-cc-pVDZ basis and using the frozen core (fc) approximation. It was found earlier (see also ref. 35) that the fc approximation has no significant effect on the excitation energies, *i.e.* the differences of the excitation energies from all-electron and frozen-core calculations are smaller than 0.02 eV. Therefore, it does not affect the final results while allows the inclusion of triple excitations in the electron correlation treatment (CC3) which have proven essential.

The UV spectrum has been simulated by the LVC (Linear Vibronic Coupling) model of Köppel *et al.*[50] using parameters from the CC calculations. Thus, the ground state vibrational frequencies were obtained at the CCSD/cc-pVDZ level, while the intrastate couplings were computed at the EOMEE-CCSD/cc-pVDZ level of theory. All parameters are included in the ESI.‡ Interstate couplings were not considered since, as shown in a previous paper,[35] these couplings between the A' and A'' states do not influence the spectrum. Since, in addition, all the transitions from the ground A' state to the A'' states have very low oscillator strengths, they need not be included in the spectrum simulations at all. Vertical excitation energies and transition moments have been taken from CC3-LR/aug-cc-pVDZ calculation using frozen core approximation. For all three tautomers, the first three A' excited states and the most important 10 vibrational modes have been included in the simulations. The program SIM developed by Stanton[51] was used to solve the LVC equations. The spectra were simulated by placing Lorentzian functions with 0.15 eV full width at half height around all eigenvalues. This linewidth was selected to approximately match the resolution of the experimental spectrum.

2.3 Determination of tautomeric ratios

In order to get semi-experimental tautomeric ratios, first the MI-IR spectrum was fitted by a sum of Lorentzian functions, using as many peaks as reasonably necessary to reproduce the experimental spectral band shapes. These bands were then assigned to calculated vibrational transitions and intensities. In the case of site splitting, a group of bands were assigned to a single vibrational mode, while in the case of incompletely resolved bands or ambiguous assignments a group of bands were assigned to several vibrational modes. Besides a comparison of our experimental and computational results, the assignments were based on former assignments available from the literature.[12–14]

The area under the curve was assumed to be proportional to the calculated intensity multiplied by a linear coefficient

Table 1 Selected quantum chemical results for relative electronic energies (ΔE°), Gibbs free energies (ΔG°_T) in kcal mol^{-1} and mole percentages of isomers of cytosine

	Trygubenko *et al.*[4] (2002)[a] CCSD(T)			Fogarasi[3] (2002)[b] CCSD(T)			Yang and Rodgers[5] (2004)[c] MP2			Wolken *et al.*[6] (2007)[d] CCSD(T)			Kosenkov *et al.*[7] (2009)[e] DFT			Kostko *et al.*[8] (2010)[f] CCSD			Present work (2010)[g] CCSD(T)		
Isomer	ΔE°	$\Delta G^\circ_{298\ K}$	%	ΔE°	$\Delta G^\circ_{298\ K}$	%	ΔE°	$\Delta G^\circ_{490\ K}$	%	ΔE°	$\Delta G^\circ_{473\ K}$	%	ΔE°	$\Delta G^\circ_{490\ K}$	%	ΔE°	$\Delta G^\circ_{582\ K}$	%	ΔE°	$\Delta G^\circ_{450\ K}$	%
1	1.44	0.63	21	1.66	0.86	13	1.4	0.9	20	0.93	0.33	31	—	0.28	42	1.35	—	24	0.76	0.21	29
2a	0.69	0.71	18	0.75	0.75	16	0.7	0.7	24	0.67	0.75	20	—	—	—	0.69	—	20	0.71	0.67	17
2b	0.00	0.00	60	0.00	0.00	57	0.0	0.0	51	0.00	0.00	43	—	0.00	57	0.00	—	35	0.00	0.00	37
3a	2.11	2.66	1	1.73	0.83	14	3.2	2.4	5	2.00	2.01	5	—	4.03	1	1.15	—	16	1.34	0.69	17
3b	3.68	4.10	0	—	—	—	4.9	4.0	1	—	—	—	—	—	—	2.75	—	5	—	—	—

[a] Energies extrapolated to "infinite" basis, geometries at RI-MP2/TZVPP, frequencies at HF/6-31G(d,p). [b] Energies with cc-pVTZ, geometries at CCSD/TZP, frequencies at MP2/TZP. [c] MP2 energies with 6-311++G(2d,2p), geometries and frequencies at MP2/6-31G(d). [d] Energies with aug-cc-pVTZ, including zero-point energy corrections; geometries at MP2/6-31G+(d,p), frequencies at B3LYP/6-31+G(d,p). [e] MPWB1K functional with aug-cc-pVDZ basis set used throughout. [f] Energies with CCSD/cc-pVTZ, geometries and frequencies at RI-MP2/6-311+G(d,p). [g] Energies with cc-pVQZ, geometries at CCSD/cc-pVTZ, frequencies at MP2/6-311++G(2d,2p), all-electron (ae) correlation in the CC calculations, frozen core (fc) in the MP2 frequency calculations.

proportional to the amount of the appropriate tautomer. Using as many peaks and groups of peaks as possible, an overdetermined system of linear equations was built and solved by the least squares method. This provided the unknown coefficients as the best approximation for the mole ratios of the tautomers. A representative region of the MI-IR spectra including computed IR intensities, fitted Lorentzian functions, assignment and groupings is shown in Fig. 2.

3. Results and discussion

3.1 Isomer ratios

In spite of much effort, the isomer ratios in gas-phase cytosine are still very uncertain. As we will see, the most intriguing question is the amount of the "rare" imino form (**3a**). In the following we critically review the available QC results, followed by a separate subsection on the comparison of experimental studies and on the derivation of experimental tautomeric ratios from the present MI-IR measurements.

3.1.1 Theoretical results. After many early results, QC calculations considered as high level by present standards have become available from the 2000s. In Table 1 we have compiled just a few representative results. In two 2002 papers, both Trygubenko *et al.*[4] and Fogarasi[3] used coupled cluster theory with fairly large basis sets, reporting energies at CCSD(T)/cc-pVTZ level. For the energies alone, the two sets represent well comparable results. There is, however, a drastic discrepancy in the calculated (relative) free energies: for **1** and **3a**, relative to **2b**, Fogarasi obtained $\Delta G^{\circ}_{298\ \mathrm{K}} = 0.86$ kcal mol^{-1} and 0.83 kcal mol^{-1} , respectively, which should be compared to Trygubenko's corresponding results of 0.63 and 2.66 kcal mol^{-1}, respectively. We have analysed the details of these calculations: the discrepancy appears already in the zero-point energies and is due to the differences in the vibrational frequencies, the latter taken from RHF calculations in ref. 4 and MP2 calculations in ref. 3. Given the narrow range of 1–2 kcal mol^{-1} between the tautomers, the final ΔG values are sensitive to small differences in details of the calculations. The consequencies are dramatic for the equilibrium concentration of the imino (**3a**) tautomer, as seen in Table 1.

Among other computed results listed in Table 1, Yang and Rodgers[5]—with interest primarily in the dimers—also reported electronic energies for the monomers: $\sim$1.5 kcal mol^{-1} for **1** and $\sim$3 kcal mol^{-1} for **3a**, values typical of MP2 results for these systems, see, *e.g.* ref. 3. Note that these authors were probably the first to bring up the important suggestion that a possible way of tautomerization may be through dimerization. The results of Wolken *et al.*,[6] obtained as part of their study on ionization of cytosine tautomers, are again based on coupled cluster calculations. (Note that their ΔE° values in Table 1 already include ZPEs.) The most striking difference to previous results is their extremely low ΔG° for tautomer **1**. (We read this value from their graph in Fig. 2, but checked it with the populations reported numerically.) Concerning the data from Kosenkov *et al.*[7] one should realize that density functional theory in general gives a qualitatively different stability order for cytosine's tautomers, preferring the oxo (**1**) form. It is therefore really interesting that the more recent functional used in ref. 7 may be the first one that reproduces the higher stability of the hydroxy form, although the oxo isomer is still very close to it with $\Delta G^{\circ} = 0.28$ kcal mol^{-1}. In conjunction with their experimental VUV photoionization study, Kostko *et al.*[8] reported recently on CCSD calculations. From our point of view, the most interesting point is that they get quite significant population for the imino (**3a**) form. Most probably this is again due to the better vibrational frequencies, obtained from MP2 calculations.

In the last column of Table 1, our present results represent the highest level of theory used for cytosine up to now. The electronic energy was obtained at CCSD(T)/cc-pVQZ level, in a geometry calculated by CCSD/cc-pVTZ. Note, however, that to make the CCSD(T)/cc-pVQZ calculations feasible we used the planarity constraint at this point. Full, nonplanar optimization would put the energies of the hydroxy and oxo forms a little lower, thus raising the relative energy of the planar imino form. This effect was estimated in ref. 3 as +0.15 and 0.24 kcal mol^{-1}, for **1** and **3a**, respectively, relative to **2b**. In the present study, we optimized only **2b** as nonplanar structure at CCSD/cc-pVTZ level, obtaining an energy lowering of 0.18 kcal mol^{-1}. Based on these data, a correction of $\delta\Delta E^{\circ} = 0.2$ kcal mol^{-1} could perhaps be added to the relative energy of **3a**, with $\delta\Delta E^{\circ} = 0.1$ kcal mol^{-1} added to tautomer **1**. However, this would change the mole fractions in Table 1 by 1–2 percentage points only.

In summary, the CC results in Table 1 may be collected into two groups: ref. 4 and 6 predict a very low population for the imino tautomer, while according to ref. 3 and 8 and the present work the imino (**3a**) form should have a significant mole fraction in an equilibrium mixture. The discrepancy comes mainly from the calculation of vibrational frequencies, for which HF or DFT theory was used in the first group, while MP2 theory in the second group.

3.1.2 Experimental results. Experimental results on tautomer ratios are also diverse, mainly of semiquantitative accuracy (see Table 2). The first elegant jet-cooled molecular-beam microwave study by Brown suggested roughly 1 : 1 : 0.25 ratios for **1**, **2b** and **3a**, respectively.[9] Although many papers considered this study to be the experimental gold standard, He droplet IR studies concluded that Brown *et al.* likely mixed up the **3a** with the **2a** form.[15] The He droplet IR studies identified the **1**, **2a** and **2b** forms. Very recently Alonso *et al.*[10] published a new laser-vaporization molecular-beam microwave study, in which they were able to identify all the five low-energy forms of cytosine.

Three tautomers were observed both by gas-phase valence-shell (UPS)[8] and core-level (XPS and Auger)[20] photoelectron studies. In the jet-cooled REMPI studies[18,19] the oxo (**1**) and one hydroxy (**2**) form were identified, but as was already suggested by some of us,[35] and as will be discussed in the Conclusions, the interpretation of REMPI spectra may need a revision.

Matrix isolation studies have the advantage that due to the fast freezing process the gas-phase mole ratios are conserved in the low-temperature matrix, unless the interconversion barrier between the different structures (isomers, conformers) is smaller than $\sim$1 kcal mol^{-1}.[52,53] This does not hold

Table 2 Experimental estimates for ratios of isomers of cytosine[a]

Method	**1**	**2a**	**2b**	**3a**
MW, jet-cooled[9]	1		1	0.25
MW, jet-cooled[10]	+	+	+	+[b]
IR, Ar matrix[11]	+		+	–
IR, Ar matrix[12]	0.45		1.00	≤0.05
IR, Ar matrix[13]	0.4–0.5	(+)	1.00	0.1–0.05
IR, Ar matrix[14]	+	+	+	+
IR, Ar matrix, present work[c]	0.50	0.59	1.00	0.18
IR, He-droplet[15–17]	+	+	+	–
REMPI, jet-cooled[18,19]	+		+	–
UPS, gas phase[8]	(+)		+	(+)
XPS, gas phase[20]	+		+	+

[a] Relative to **2b**. The + symbol indicates that the tautomer was identified but not quantified. It is given in brackets if the identification was ambiguous. The – symbol indicates unobserved tautomers. [b] They have also identified **3b**. [c] Obtained using B3LYP/6-31++G(d,p) transition intensities. The values given translate to mole fractions 0.22, 0.26, 0.44 and 0.08. Using MP2/6-31++G(d,p) intensities the latter are: 0.21, 0.24, 0.46 and 0.08.

for jet-cooling conditions, where certain isomers/conformers can be populated or depopulated by collisions. Thus, if thermal equilibrium exists in the gas phase, then the abundances observed in the matrix can directly be compared to QC ratios computed at the evaporation temperature of the sample.

Although only the oxo (**1**) and the hydroxy (**2**) forms were identified by the first MI-IR studies, later on it was proved that the imino (**3**) form is also present in the matrix. Furthermore, Nowak *et al.*[13] suggested that the observed splittings of the bands assigned to the hydroxy form are due to the presence of two rotamers. This was unambiguously proved by Nowak, Fausto and their co-workers,[14] by inducing rotamerization between the two hydroxy forms (**2a** and **2b**) by NIR laser irradiation. This experiment contributed also to the unambiguous assignment of MI-IR spectra of cytosine.

In the MI-IR studies by Nowak *et al.*[13] and Person *et al.*[12] the oxo (**1**) : hydroxy (**2**) ratio was estimated as the ratio of the integrated intensities of the OH and NH stretching bands appropriately weighted by the computed molar absorptions. Although the results from the two groups are consistent (see Table 2), they have some uncertain points. First, the rotamer pair **2a–2b** was not yet separated. Second, the computed IR intensities available at that time were rather uncertain.

The experimental determination of tautomeric ratios in the present study (see details in Section 2.3) differs from previous studies in that it used not only one or two characteristic bands but most of the experimental bands. Thus, the error of the computed IR transition intensities averages out for the entire range of the IR spectrum. Furthermore, present-day DFT and MP2 methods give better estimates for transition intensities than earlier computations.

As seen in Table 2 the present procedure using B3LYP/6-31++G(d,p) transition intensities yields 0.50 : 0.59 : 1.00 : 0.18 ratios for **1**, **2a**, **2b** and **3a**, respectively. This translates to mole fractions of 0.22, 0.26, 0.44 and 0.08. In order to check the sensitivity of the mole ratios to the theoretical level of the spectrum calculations, the mole ratios were also determined by using MP2/6-31++G(d,p) intensities. Reassuringly, this yielded quite similar mole fractions, *viz*. 0.21, 0.24, 0.46 and 0.08.

The obtained oxo (**1**) : hydroxy (**2**) ratio qualitatively agrees with that of Nowak *et al.*[13] and Person *et al.*,[12] while we obtained significantly larger mole fraction for the imino (**3a**) form. We have also determined the mole fraction of **2a**. The combined mole fraction of the hydroxy forms (**2a** + **2b**) at 450 K is 0.70, very large in the light of the fact that in solid or aqueous phase the oxo (**1**) form is considered to be the only form.

A comparison of the QC results in Table 1 and the present experimentally determined mole fractions shows essentially agreement for tautomers **1**, **2a** and **2b**. As noted above, the main question is the mole fraction of the imino form (**3a**). Our theoretical result of about 17% is significantly larger than most other studies suggested or assumed. Interestingly, however, our totally independent experimental result shows also relatively large, about 8% abundance. Obviously, given the difficulties of accurate determination both results may only be considered as semiquantitative. Apart from this, however, the following argument may suggest that the imino form (**3a**) is perhaps not in thermal equilibrium with the other forms. Intramolecular tautomerization of free cytosine in the gas phase is hindered by high barriers of about 40 kcal mol^{-1}. However, as suggested by Yang and Rodgers,[5] *bimolecular* tautomerization may be much more feasible, with barriers of 3–10 kcal mol^{-1} only. Moreover, preliminary computations in our laboratory indicate that bimolecular proton transfer leading from the oxo (**1**) to the hydroxy (**2**) form may have a lower barrier than to the imino (**3**) form. It is of interest in this respect that we tried to find signatures of dimers in the MI-IR spectra, but could not identify any of them.

Table 3 Vertical excitation energies (VEE in eV) from the ground electronic state to the lowest-lying singlet A' states with corresponding oscillator strengths (OS) for the oxo tautomer (**1**) of cytosine[a]

	EOMEE-CCSD/ cc-pVDZ (fc)		EOMEE-CCSD/ aug-cc-pVDZ (fc)		CC2-LR/ aug-TZVP[c]		CC3-LR/ cc-pVDZ (fc)		CC3-LR/ aug-cc-pVDZ (fc)	
State (transition)[b]	VEE	OS	VEE	OS	VEE	OS	VEE	OS	VEE	OS
$2A'$ ($\pi \rightarrow \pi^*$)	5.11	0.070	4.94	0.064	4.61	0.052	4.87	0.057	4.71	0.065
$3A'$ ($\pi(n_N) \rightarrow \pi^*$)	6.11	0.188	5.86	0.164	5.64	0.127	5.76	0.140	5.55	0.138
$4A'$ ($\pi \rightarrow 2\pi^*$)	7.08	0.677	6.50	0.508	6.28	0.388	6.88	0.553	6.30	0.426
$5A'$ ($n_N,n_O \rightarrow R$)	—	—	6.70	0.026	6.39	0.075			6.43	
$6A'$ ($\pi(n_N) \rightarrow 2\pi^*$)	7.47	0.208	6.88	0.181						

[a] Geometry from CCSD/cc-pVDZ calculations under the planarity constraint. [b] *R*: Rydberg state. [c] Barbatti *et al.*[36]

In the next section, for simulating the MI-UV spectrum we will apply the tautomer ratios experimentally determined by the help of B3LYP/6-31++G(d,p) IR intensities.

3.2 Vertical excitation energies

The computed vertical excitation energies from the ground electronic state to the first few singlet excited A' states of the three tautomers are compiled in Tables 3–5. A'' symmetry states are not listed because, as shown in ref. 35, they do not contribute to the spectrum, their transition moments being small and couplings with the A' states negligible. Of the two rotamers of the hydroxy tautomer, only the lower-energy form **2b** was considered, because test calculations at the EOMEE-CCSD/aug-cc-pVDZ(fc) level showed that the differences in the excitation energies are below 0.05 eV between rotamers **2a** and **2b** and the oscillator strengths are also practically the same.

First we compare the present results with our earlier ones for the oxo (**1**) form.[35] In that study we claimed that only two states would contribute to the spectrum below 6.5 eV. By contrast, Barbatti *et al.*[36] claimed recently on the basis of CC2-LR calculations that there are in fact three states contributing. The present results confirm their finding. As Table 3 shows, by adding diffuse functions to the basis the excitation energy of the third state decreases considerably, by about 0.5 eV. This is due to the effect of Rydberg states, which mix into the $4A'$ ($\pi \rightarrow 2\pi^*$) state to some extent. Note that the 4th A' excited state ($5A'$) is already a Rydberg state. As seen in Table 3, for the first three most intense transitions our best values (CC3-LR/aug-cc-pVDZ) are in good agreement with the values by Barbatti *et al.*[36] with deviations smaller than 0.1 eV. Note, however, that the separations of individual excitation energies differ by up to 0.2 eV, an effect which would change the shape of the spectrum by placing the bands further apart. The other discrepancy concerns the Rydberg states, *viz*. Barbatti *et al.*[36] predict Rydberg states below 6.5 eV which we do not see in our calculations. Since both works use similar diffuse functions in the calculations, presently one can not give a final conclusion on this issue; calculations with well designed Rydberg orbitals need to be performed in the future to answer this question. Nevertheless, these Rydberg states are of low intensity so would not influence the calculated spectrum, thus leaving the conclusion of the present study unchanged.

Some methodological conclusions from Table 3 are of interest: (a) diffuse basis functions and (b) triple excitations are necessary to get converged results. Therefore, for the other tautomers and the spectrum simulation the frozen-core CC3-LR/aug-cc-pVDZ method was used. Note that Schreiber *et al*. have shown (see Fig. 2 in ref. 33) that the aug-cc-pVXZ basis series, including the smallest DZ set, is particularly capable to describe the valence states of ethylene where the mixing in of Rydberg character is of concern. We face similar situation in the case of cytosine here.

A comparison of EOMEE-CCSD and CC3-LR results clearly shows that the inclusion of triple excitations in the electron correlation treatment has substantial effects also in the case of the other tautomers, lowering the excitation energies by about 0.2 eV on average. The deviation from this average value is, however, quite large, it is 0.3–0.4 eV for the $\pi(n_N) \rightarrow \pi^*$ excitations but only 0.13 eV for the $\pi \rightarrow 2\pi^*$ excitation of the hydroxy (**2b**) tautomer. As a consequence, the relative excitation energies are substantially influenced, thus triple excitations in the correlation treatment need to be considered to obtain reliable simulated spectra.

To the best of our knowledge no calculations exist on the excitation energies of the hydroxy (**2**) and the imino (**3**) tautomers, although, as obvious from simple Lewis-structures, tautomerization should drastically change the total electronic structure, including first of all the lone pairs and second the π-electrons. (For example, tautomer **2** is more aromatic than the others which may explain its stability.) The results presented in Tables 3–5 clearly support these expectations. As an example, the vertical excitation energy of the $\pi(n_N) \rightarrow \pi^*$ transition differs by about ~0.8 eV between the oxo (**1**) and hydroxy (**2**) tautomer, which results in a change of order of different electronic transitions. Beside excitation energies, the oscillator strengths are also strongly affected by tautomerization. While the three lowest-energy transitions of the hydroxy (**2**) tautomer have similar oscillator strengths, in the case of the oxo tautomer (**1**), the $\pi \rightarrow 2\pi^*$ transition is about eight and three times more intense than the $\pi \rightarrow \pi^*$ and the $\pi(n_N) \rightarrow \pi^*$ transition, respectively. In the case of the imino (**3**) tautomer, the $\pi(n_N) \rightarrow \pi^*$ transition is more than an order of magnitude weaker than the $\pi \rightarrow \pi^*$ and $\pi \rightarrow 2\pi^*$ transitions.

3.3 QC simulation and interpretation of the UV spectrum

In order to simulate the spectra up to 6.5 eV, the three lowest-energy singlet A'-symmetry transitions are considered for each tautomer. The higher-energy excitations are not expected to substantially modify this part of the spectrum, since their computed vertical excitation energies are higher than

Table 4 Vertical excitation energies (VEE in eV) from the ground electronic state to the lowest-lying singlet A' states with corresponding oscillator strengths (OS) for the hydroxy tautomer (**2b**) of cytosine[a]

	EOMEE-CCSD/aug-pVDZ (fc)		CC3-LR/aug-pVDZ (fc)	
State (transition)[a]	VEE	OS	VEE	OS
$2A'$ ($\pi \rightarrow \pi^*$)	5.04	0.104	4.88	0.113
$3A'$ ($\pi \rightarrow 2\pi^*$)	5.97	0.158	5.84	0.174
$4A'$ ($\pi(n_N) \rightarrow \pi^*$)	6.69	0.158	6.39	0.143
$5A'$ ($n_N \rightarrow R$)	6.83	0.069	6.59	0.026
$6A'$ ($\pi \rightarrow R$)[b]	7.17	0.310	7.02	

[a] Geometry from CCSD/cc-pVDZ calculations under the planarity constraint. *R*: Rydberg state. [b] Mixed with $\pi \rightarrow \pi^*$.

Table 5 Vertical excitation energies (VEE in eV) from the ground electronic state to the lowest-lying singlet A' states with corresponding oscillator strengths (OS) for the imino tautomer (**3a**) of cytosine[a]

	EOMEE-CCSD/aug-pVDZ (fc)		CC3-LR/aug-pVDZ (fc)	
State (transition)[a]	VEE	OS	VEE	OS
$2A'$ ($\pi \rightarrow \pi^*$)	5.34	0.228	5.07	0.209
$3A'$ ($\pi(n_N) \rightarrow \pi^*$)	6.29	0.017	5.89	0.003
$4A'$ ($\pi \rightarrow 2\pi^*$)	6.58	0.223	6.28	0.266
$5A'$ ($\pi \rightarrow R$)[b]	7.07	0.016	6.91	0.008
$6A'$ ($n_N \rightarrow R$)	7.20	0.000		

[a] Geometry from CCSD/cc-pVDZ calculations under the planarity constraint. *R*: Rydberg state. [b] Mixed with $\pi \rightarrow \pi^*$.

6.8 eV and they have considerably lower oscillator strengths. Nevertheless, the blue tail of the simulated spectra would certainly be affected if higher-energy excitations were also considered.

In Fig. 3 the calculated spectra of the three tautomers are plotted. The figure shows the peaks of the individual transitions as well as the complete spectra of each tautomer. As can be seen in the panels, in the case of the hydroxy (**2b**) tautomer all three A' states contribute to the spectrum with almost equal intensity, while they contribute with increasing intensity in the case of the oxo (**1**) tautomer. In contrast to these, the second A' excited state has negligible contribution in the case of the imino (**3a**) tautomer, while the transitions to the other two A' states give very similar contributions.

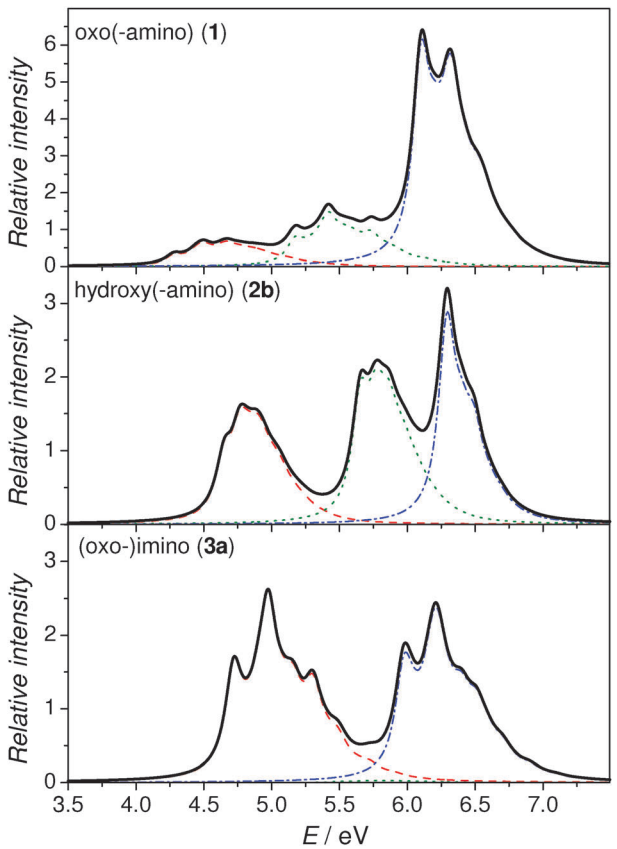

Fig. 3 Computed UV spectra (thick lines) of cytosine tautomers. The thin dashed, dotted and dash-dotted lines show the individual contribution of the transition from the ground A' to the first, second and third excited singlet A' state respectively.

In order to simulate the UV spectrum corresponding to experimental conditions, the appropriately weighted spectra of the tautomers have to be combined into a composite spectrum. It has been discussed in Section 3.2, that the vertical excitation energies and the corresponding oscillator strengths do not differ much for the rotamers; hence, in order to save computer time, the contribution of the higher-energy hydroxy rotamer (**2a**) to the UV spectrum was approximated by the spectrum of **2b**. The tautomer ratios obtained from the MI-IR measurements (see Section 3.2) were used as weighting factors, *i.e.* 0.22, 0.70 and 0.08 for **1**, **2** and **3a**, respectively.

The bottom panel of Fig. 4 shows how the different tautomers contribute to the spectrum. It is seen that the spectrum is dominated by the hydroxy (**2**) for low excitation energies,

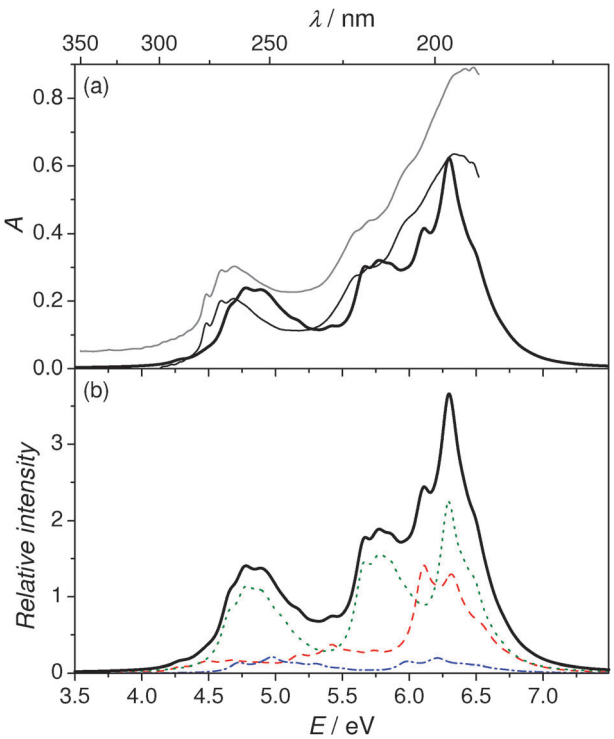

Fig. 4 Comparison of computed and experimental UV spectra of cytosine. On the top (a) panel the experimental MI-UV spectra are plotted by thin lines, (the upper gray line is the uncorrected spectrum, the lower black line is the one corrected for scattering background), and the composite computed spectrum is plotted by thick line. On the bottom (b) panel the tautomeric composition of the simulated composite spectrum (thick line) is shown; oxo (**1**) by the dashed line, hydroxy (**2b**) by the dotted line and imino (**3a**) by the dash-dotted line.

while above 6 eV both the hydroxy (**2**) form and the oxo (**1**) tautomers contribute significantly to the overall shape. One can conclude that, at this resolution (linewidth 0.15 eV), most of the bands should be assigned to the hydroxy (**2**) tautomer except the peak at 6.2 eV, which appears to be due to the oxo (**1**) tautomer. Furthermore, the oxo (**1**) tautomer gives a significant contribution to the intensity above 6 eV.

The top panel of Fig. 4 shows that the composite spectrum is in good agreement with the experimental (MI-UV) one. There is a small discrepancy in the position of the lower band, otherwise even minor features, including shoulders can be identified. A remaining small difference is that the integrated intensity of the spectra above 6 eV is slightly lower for the computed spectrum than for the experimental one. This cannot be explained by an eventual underestimation of abundance of the oxo (**1**) tautomer, since the increase of the weighting factor of the oxo (**1**) tautomer would result in a significant change in band shapes in the entire spectrum region. Hence, this small difference is most likely due to the neglect of the contribution of higher energy transitions to the blue tail of the spectrum.

4. Conclusions

We have recorded the MI-UV spectrum of cytosine and performed high level QC calculations for its interpretation. While earlier theoretical results could only be compared with UV experiments performed in aqueous solutions or solid films, the new matrix isolation spectrum refers to a situation simulating closely the vapor state of isolated molecules. This allows a more direct comparison of theory and experiment. Also, in contrast to the condensed phases where only the oxo (**1**) tautomer exists, several tautomers are expected in the matrix so we extended the computation of excited electronic states to the three most important (low-energy) tautomers. Thus, the theoretical UV spectrum has been constructed as a composite of the spectra of isomers.

To determine the tautomer ratios needed for the interpretation of the UV experiment, the MI-IR spectrum was also recorded at similar conditions, and complemented with theoretical calculation of the IR intensities. In the assignation, recent results of ref. 14 were also used. The mole ratios, that we consider semiquantitative—but more accurate than previous estimates—are as follows: 0.22 : 0.26 : 0.44 : 0.08 for **1** : **2a** : **2b** : **3a**.

Based on this, we draw first two methodological conclusions. First, when calculating relative free energies within a narrow range of 2–3 kcal mol^{-1}, the accuracy of the vibrational *frequencies* is as important as the accuracy of the *electronic* energies themselves; reliable frequencies seem to need MP2 calculations at least. About electronic excitations, early calculations on the oxo (**1**) tautomer's vertical excitation energies showed significant deviations (see *e.g.* ref. 35), but recent results seem to converge. For example, the excitation energy to the first excited singlet A' state is predicted to be 4.66, 4.63, and 4.68 eV by CC2-LR/aug-cc-pVTZ,[34] TDDFT-B3LYP/6-311++(d,p),[31] and CASPT2/TZVP[33] calculations, respectively, which compare well with our CC3-LR/aug-cc-pVDZ result of 4.71 eV. These methods also agree that the second singlet A' excited state is between 5.44 and 5.62 eV. Overall, among popular methods CC2-LR seems to be the closest to CC3-LR, but carefully designed CASPT2 wave function is also reliable.

Our results may also justify a discussion of some recent experimental results. Nir *et al.*[18,19] have measured the resonance enhanced multiphoton ionization (REMPI) spectrum of cytosine and found two regions of electronic excitations. A lower frequency region, 31 000–32 500 cm^{-1}, was assigned to the oxo (**1**) form and the 36 000–37 500 cm^{-1} region to the hydroxy (**2**) tautomer. Whatever the tautomeric ratios in their experiment, our results seriously challenge this assignment because the lowest-energy absorption bands of all three tautomers set up first at about 4.5 eV (36 000 cm^{-1}). This means that it is the second spectral region (36 000–37 500 cm^{-1}) where the band origin of the first excited A' state of each tautomer, if present, should appear. No sign of any singlet state below 4.5 eV is seen in our calculations. The discrepancy suggests that the REMPI spectrum may need a reinterpration.

Two recent electron-energy-loss (EEL) measurements on cytosine should also be discussed. In the first one Abouaf *et al.*[54] performed an experiment in the gas phase, while a very recent work by Bazin *et al.*[55] investigated cytosine as a film deposited on a cold surface. The assignments in both cases are based on the assumption that only the oxo (**1**) tautomer is present in the sample. While the experimental conditions in the latter experiment differ considerably form the present one, in the gas-phase experiment[54] both the evaporation and the measurement were performed at 150 °C, thus the conditions were not too far from ours. In our opinion it may well be that cytosine in the gas phase existed as a tautomeric mixture, possibly in a composition similar to that in our matrix sample. With regard to this, it would be worthwhile to investigate whether the slightly different assignments (see Table 1 in ref. 55) of the two EEL experiments are due to different tautomer abundances in the two measurements.

As a result of this study, the tautomer and rotamer ratios (although still semiquantitative) are considered to be more reliable and more complete than previous estimates: first, rather than using the intensities of a few characteristic bands, we made use of the information of the complete spectrum; second, it is a new feature that the two rotamers of the hydroxy tautomer were handled individually. In the electronic structure computations, excitation energies should be accurate to ~0.1 eV. The successful reproduction of the MI-UV spectrum as a composite of isomer vibronic spectra supports the correctness of the approach used in this study and may serve as an example of the combined use of spectroscopic and theoretical information.

Acknowledgements

Financial support by the Hungarian Scientific Research Foundation (OTKA, Grants No. K75877 and T68427) is gratefully acknowledged. The Project is supported by the European Union and co-financed by the European Social Fund (grant agreement no. TAMOP 4.2.1./B-09/1/KMR-2010-0003). The authors thank Prof. J. F. Stanton for providing access to and assistance with the SIM code and Dr Attila Tajti for continuous help with the calculations.

Notes and references

1 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 964–967; M. K. Shukla and L. Leszczynski, *J. Biomolecular Structure & Dynamics*, 2007, **25**, 93–117; C. E. Crespo-Hernandez, B. Cohen, P. M. Hare and B. Kohler, *Chem. Rev.*, 2004, **104**, 1977–2019.
2 G. Fogarasi, *J. Mol. Struct.*, 1997, **413–414**, 271–278.
3 G. Fogarasi, *J. Phys. Chem. A*, 2002, **106**, 1381–1390.
4 S. A. Trygubenko, T. V. Bogdan, M. Rueda, M. Orozco, F. J. Luque, J. Sponer, P. Slavicek and P. Hobza, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4192–4203.
5 Z. B. Yang and M. T. Rodgers, *Phys. Chem. Chem. Phys.*, 2004, **6**, 2749–2757.
6 J. K. Wolken, C. Yao, F. Turecek, M. J. Polce and C. Wesdemiotis, *Int. J. Mass Spectrom.*, 2007, **267**, 30–42.
7 D. Kosenkov, Y. Kholod, L. Gorb, O. Shishkin, D. M. Hovorun, M. Mons and J. Leszczynski, *J. Phys. Chem. B*, 2009, **113**, 6140–6150.
8 O. Kostko, K. Bravaya, A. Krylov and M. Ahmed, *Phys. Chem. Chem. Phys.*, 2010, **12**, 2860–2872.
9 R. D. Brown, P. D. Godfrey, D. McNaughton and A. P. Pierlot, *J. Am. Chem. Soc.*, 1989, **111**, 2308–2310.
10 I. Pena, V. Vaquero, J. L. Alonso and J. C. López, *64th International Symposium on Molecular Spectroscopy*, The Ohio State University, Columbus, Ohio (USA), June 22–26, 2009, Lecture TA02.
11 E. D. Radchenko, G. G. Sheina, N. A. Smorygo and Yu. P. Blagoi, *J. Mol. Struct.*, 1984, **116**, 387–396.
12 M. Szczesniak, K. Szczepaniak, J. S. Kwiatkowski, K. KuBulat and W. B. Person, *J. Am. Chem. Soc.*, 1988, **110**, 8319–8330.
13 M. J. Nowak, L. Lapinski and J. Fulara, *Spectrochimica Acta*, 1989, **45A**, 229–242.
14 L. Lapinski, M. J. Nowak, I. Reva, H. Rostkowska and R. Fausto, *Phys. Chem. Chem. Phys.*, 2010, **12**, 9615–9618.
15 F. Dong and R. E. Miller, *Science*, 2002, **298**, 1227–1230.
16 M. Y. Choi, F. Dong and R. E. Miller, *Philos. Trans. R. Soc. London, Ser. A*, 2005, **363**, 393–413.
17 A. Min, S. J. Lee, M. Y. Choi and R. E. Miller, *Bull. Korean Chem. Soc.*, 2009, **30**, 3039–3044.
18 E. Nir, M. Müller, L. I. Grace and M. S. de Vries, *Chem. Phys. Lett.*, 2002, **355**, 59–64.
19 E. Nir, Ch. Plützer, K. Kleinermanns and M. S. de Vries, *Eur. Phys. J. D*, 2002, **20**, 317–329.
20 V. Feyer, O. Plekan, R. Richter, M. Coreno, G. Vall-llosera, K. C. Prince, A. B. Trofimov, I. L. Zaytseva, T. E. Moskovskaya, E. V. Gromov and J. Schirmer, *J. Phys. Chem. A*, 2009, **113**, 5736–5742; V. Feyer, O. Plekan, R. Richter, M. Coreno, M. de Simone, K. C. Prince, A. B. Trofimov, I. L. Zaytseva and J. Schirmer, *J. Phys. Chem. A*, 2010, **114**, 10270–10276.
21 J. Cadet and M. Berger, *Int. J. Radiat. Biol.*, 1985, **47**, 127–143.
22 L. B. Clark and I. Tinoco Jr., *J. Am. Chem. Soc.*, 1965, **87**, 11–15.
23 T. Yamada and H. Fukutome, *Biopolymers*, 1968, **6**, 43–54.
24 K. Raksanyi, I. Foldvary, J. Fidy and L. Kittler, *Biopolymers*, 1978, **17**, 887–896.
25 G. G. Sheina, E. D. Radchenko, A. M. Plokhotnichenko and Yu. P. Blagoi, *Biofizika*, 1982, **27**, 983–986.
26 F. Zaloudek, J. S. Novros and L. B. Clark, *J. Am. Chem. Soc.*, 1985, **107**, 7344–7351.
27 D. W. Miles, M. J. Robins, R. K. Robins, M. W. Winkley and H. Eyring, *J. Am. Chem. Soc.*, 1969, **91**, 831–838.
28 M. P. Fülscher and B. O. Roos, *J. Am. Chem. Soc.*, 1995, **117**, 2089–2095.
29 N. Ismail, L. Blancafort, M. Olivucci, B. Kohler and M. A. Robb, *J. Am. Chem. Soc.*, 2002, **124**, 6818–6819.
30 M. Merchan and L. Serrano-Andres, *J. Am. Chem. Soc.*, 2003, **125**, 8108–8109.
31 M. K. Shukla and J. Leszczynski, *J. Comput. Chem.*, 2004, **25**, 768–778.
32 K. Tomic, J. Tatchen and C. M. Marian, *J. Phys. Chem. A*, 2005, **109**, 8410–8418.
33 M. Schreiber, M. R. Junior Silva, S. P. A Sauer and W. Thiel, *J. Chem. Phys.*, 2008, **128**, 134110.
34 T. Fleig, S. Knecht and C. Haettig, *J. Phys. Chem. A*, 2007, **111**, 5482–549.
35 A. Tajti, G. Fogarasi and P. G. Szalay, *ChemPhysChem*, 2009, **10**, 1603–1606.
36 M. Barbatti, A. J. A. Aquino and H. Lischka, *Phys. Chem. Chem. Phys.*, 2010, **12**, 4959–4967.
37 E. D. Radchenko, A. M. Plokhotnichenko, A. Yu. Ivanov, G. G. Sheina and Yu. P. Blagoi, *Biofizika*, 1986, **31**, 373.
38 Yu. P. Blagoi, E. D. Radchenko, S. G. Stepanian and G. G. Sheina, *J. Mol. Struct.*, 1990, **219**, 311–316, and references cited therein.
39 G. N. Ten, N. B. Zotov and V. I. Baranov, *Opt. Spectrosc.*, 2009, **107**, 235–243.
40 *PQS Version 3.3*, Parallel Quantum Solutions, 2013 Green Acres Road, Fayetteville, Arkansas 72703, USA, 2007.
41 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 564; C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785.
42 C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618.
43 R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.*, 1980, **72**, 650.
44 T. H. Dunning, Jr., *J. Chem. Phys.*, 1989, **90**, 1007; R. A. Kendall, T. H. Dunning, Jr. and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796.
45 *CFOUR, a quantum chemical program package* written by J. F. Stanton, J. Gauss, M. E. Harding, P. G. Szalay, with contributions from: A. A. Auer, R. J. Bartlett, U. Benedikt, C. Berger, D. E. Bernholdt, O. Christiansen, M. Heckert, O. Heun, C. Huber, D. Jonsson, J. Jusélius, K. Klein, W. J. Lauderdale, D. Matthews, T. Metzroth, D. P. O'Neill, D. R. Price, E. Prochnow, K. Ruud, F. Schiffmann, S. Stopkowicz, A. Tajti, M. E. Varner, J. Vázquez, F. Wang, J. D. Watts; and the integral packages MOLECULE (J. Almlöf, P. R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, J. Olsen), and ECP routines by A. V. Mitin, C. van Wüllen, 2010. For the current version, see http://www.cfour.de.
46 G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.*, 1982, **76**, 1910; J. F. Stanton, J. Gauss, J. D. Watts and R. J. Bartlett, *J. Chem. Phys.*, 1991, **94**, 4334; J. Gauss, J. F. Stanton and R. J. Bartlett, *J. Chem. Phys.*, 1991, **95**, 2623.
47 H. Koch, O. Christiansen, P. Jørgensen, A. M. Sanchez de Merás and T. Helgaker, *J. Chem. Phys.*, 1997, **106**, 1808–1818.
48 J. F. Stanton and R. J. Bartlett, *J. Chem. Phys.*, 1993, **98**, 7029–7039; J. F. Stanton and J. Gauss, *J. Chem. Phys.*, 1994, **100**, 4695.
49 O. Christiansen, H. Koch and P. Jørgensen, *J. Chem. Phys.*, 1995, **103**, 7429.
50 H. Köppel, W. Domcke and L. S. Cederbaum, *Adv. Chem. Phys.*, 1988, **57**, 59–246.
51 *SIM, a Vibronic Coupling program* by J. F. Stanton, University of Texas at Austin, USA, 2010.
52 A. J. Barnes, *J. Mol. Struct.*, 1984, **113**, 161–174.
53 E. Mátyus, G. Magyarfalvi and G. Tarczay, *J. Phys. Chem. A*, 2007, **111**, 450–459, and references cited therein.
54 R. Abouaf, J. Pommier, H. Dunet, P. Quan, P. C. Nam and M. T. Nguyen, *J. Chem. Phys.*, 2004, **121**, 11668–11674.
55 M. Bazin, M. Mivhaud and L. Sanche, *J. Chem. Phys.*, 2010, **133**, 155104.