Combined Matrix-Isolation FT-IR and Theoretical Study of the Intrinsic Tautomeric, Vibrational, and H-Bonding Properties of N4-Methoxycytosine

Riet Ramaekers,[†] Wim Dehaen,[†] Ludwik Adamowicz,[‡] and Guido Maes*[,†]

Department of Chemistry, University of Leuven, Celestijnenlaan 200 F, 3001 Heverlee, Belgium, and Department of Chemistry, University of Arizona, Tucson, Arizona 85721

Received: July 17, 2001; In Final Form: November 19, 2002

The tautomeric, vibrational, and H-bonding characteristics of N4-methoxycytosine (N4MC) are investigated by a combined experimental, matrix-isolation Fourier transform infrared spectroscopic and computational study. For the theoretical part of the study, the methods RHF, MP2, and DFT(B3LYP) are used with the 6-31++G** basis set. N4MC occurs in a low-temperature Ar matrix exclusively in the oxo-imino form. Two isomers of this tautomer exist. Using IR intensities of characteristic bands, an approximate value of 2.3 $\times 10^{-2}$ is obtained for the isomerization constant K_I (anti/syn). Theoretically, two closed H-bonded complexes with water are found for both isomers, i.e., N1−H...O−H...O═C2 and N3−H...O−H...O═C2, as well as one less stable open complex, N7...H−O and O8...H−O, for the syn and anti isomer, respectively. In the experimental FT-IR spectra, both closed H-bonded complexes are clearly identified based on characteristic H-bonding absorptions predicted by the calculations. Trace amounts of the open complexes are detected as well. A correlation between the frequency shift of the bonded O−H stretching mode and the elongation of the O−H distance is presented for N−H...O−H...X type complexes of N4MC and some comparable compounds.

Introduction

Change in the genetic material in the form of point mutations may result from the modification of a single base. The occurrence of minor quantities of tautomeric forms of the naturally occurring bases in DNA or the presence of base analogues may be at the origin of these alterations. Mutations can either be spontaneous, i.e., they result from natural processes in cells, or they can be induced by a chemical mutagen, radiation, etc.

N^4-Methoxycytosine (N4MC) is the reaction product of the known chemical mutagen methoxyamine (NH_2OCH_3) with cytosine. N4MC is able to form stable base-pairs with both adenine and guanine during replication. As a consequence, this compound may be responsible for point mutations of the C → T type. A detailed knowledge of the intrinsic tautomeric, vibrational, and H-bonding characteristics of this base analogue may contribute to a better understanding of the mechanism of point mutations.

The compound N4MC has been the subject of intensive experimental research. Brown et al.[1] have studied the tautomeric equilibria of N4MC and some of its analogues by IR, UV, and NMR spectroscopy. These studies have indicated that N4MC occurs in the imino form. UV experiments performed by Janion[2] have confirmed this result. The orientation of the methoxy group has not been investigated in the former studies. An infrared spectroscopic study of the structure of 1,5-dimethyl-N4-methoxycytosine in several polar solvents by Kulinska et al.[3] has also demonstrated that this compound occurs predominantly in the imino form. However, this study has also revealed that an intramolecular H-bond exists between the ring N3−H group and the O atom of the methoxy group, which implies a syn orientation of the methoxy group with respect to the ring N3 atom. UV absorption spectra of N4-methoxycytidine[4] have also demonstrated the dominating occurrence of the imino tautomer, e.g., for the N3-methyl derivative of this compound the syn isomer occurs for 95%. Analysis of the high-resolution ^{15}N NMR spectra of 1-methyl-N4-methoxycytosine in chloroform by Kierdaszuk et al.[5] has indicated that a closed H-bonded complex is formed between N4MC and adenine, with the proton donor N3−H and the proton acceptor C2═O groups of N4MC being involved. Van Meervelt has determined the crystal structure of 3′5′-protected N4MC.[6] This free-base analogue occurs as the imino tautomeric form with the methoxy group in the syn configuration. The base pair geometry of N4MC with guanine has been performed by the same research group.[7] They have determined the structure of d(CGCGmo^{4}CG) by single-crystal X-ray diffraction techniques (mo4C = N4MC). In the left-handed double helical structure d(CGCGmo^{4}CG)$_2$ of type Z, N4MC forms a wobble base-pair with guanine. From this study, it appeared that N4MC occurs in the imino form, with the methoxy group oriented syn with respect to the ring N3 atom.

The combination of high-level theoretical methods with experimental data from matrix-isolation FT-IR spectroscopy is an extremely informative method for the study of tautomeric and H-bonding properties of polyfunctional bases such as N4MC. This combination of matrix-isolation FT-IR experiments with theory has a mutual advantage. On one hand, accurate theoretical methods are essential for the interpretation of the highly complicated experimental FT-IR spectra for H-bonded polyfunctional bases such as the nucleic acid bases. On the other hand, such a comparison also allows the different theoretical methods to be evaluated for their capability to predict tautomeric and H-bonding properties. In this paper, we present a combined theoretical and experimental matrix-isolation FT-IR study of the tautomeric, vibrational, and H-bonding characteristics of N4-methoxycytosine.

Methodology

Experimental Method. The cryogenic and FT-IR (Bruker IFS-88) equipment used here have been described in detail in

* To whom correspondence should be addressed. E-mail: guido.maes@chem.kuleuven.ac.be. Fax: 016/32 79 92.
† Department of Chemistry, University of Leuven.
‡ Department of Chemistry, University of Arizona, Tucson.

10.1021/jp012744k CCC: $25.00
Published on Web 02/22/2003

earlier papers.[8,9] The solid compound N4-methoxycytosine was evaporated in a homemade mini-furnace positioned close to the cold CsI window (18 K) in the cryostat. The optimal sublimation temperature was found to be 306 K. For the study of the H-bonded complexes, water-doped matrixes were deposited. Fairly high water-to-base ratios were used to ensure an excess amount of 1:1 H-bonded complexes N4MC/water. Former studies of H-bonded complexes have demonstrated that such ratios ensure excess amounts of the 1:1 complex to be formed without disturbing effects of higher-stoichiometry complexes.[8]

Because the solid compound N4MC is not commercially available, it was synthesized in the Laboratory of Organic Synthesis (W.D.). This synthesis has been performed according to the procedure described by Brown et al.[10] The product was purified by chromatographic separation, and product identification was performed by mass spectroscopy and ^{1}H and ^{13}C NMR spectroscopy. High purity argon gas (N60 = 99.99990%) was obtained from Air Liquide. Twice-distilled water was used for the experiments with water-doped matrixes.

Theoretical Method. Molecular properties such as geometries, energies, and vibrational frequencies of the different tautomers, isomers, and complexes were calculated using the density functional theory (DFT) and the RHF methods. The DFT method is used with the combined Becke's nonlocal three parameter exchange functional[11] and the gradient-corrected functional of Lee, Yang, and Parr (B3LYP functional).[12] It has been demonstrated by former studies that the latter method produces quite accurate results for isolated molecules modeling nucleic acid bases.[13−17] The standard 6-31++G** basis set was used in all these calculations. Because the extension of the basis set with diffuse functions (++) allows a considerably better description of the long-range interaction of H-bondings,[18] it is essential to employ sets of orbitals that possess sufficient diffuseness and angular flexibility. For consistency with former studies, we have also included diffuse orbitals for the calculations on the monomeric compounds.

The structure optimizations of the different forms at the RHF level was followed by single-point MP2 calculations with the frozen core electron option. This was necessary to account for electron correlation. The MP2 energy obtained for the RHF optimized molecular structures will be further denoted as MP2//RHF.

The computed total energy for each system includes the zero-point vibrational energy correction calculated either at the RHF level with a uniform scaling factor 0.900 or at the DFT level with a uniform scaling factor 0.970.

The H-bond interaction energy of each complex was computed as the difference between the energy of the complex and the energies of the monomer base and water. These results were corrected for the basis set superposition error (BSSE) by recalculating the monomer energies in the basis set of the heterodimers using the Boys−Bernardi counterpoise correction.[19]

Finally, potential energy distributions (PED) were calculated, and the computed IR frequencies were scaled to account for various systematic errors in the theoretical approach, i.e., the use of a finite basis set, the neglect of the vibrational anharmonicity, and the incomplete account for electron correlation. Either a single scaling factor was used (0.900 for RHF and 0.970 for DFT) or a set of different scaling factors was used (DFT), reflecting the difference in the anharmonicity of the different types of vibrational modes. The latter scaling procedure has been proposed by several authors in the past.[20−22] All calculations were performed with the Gaussian 94 and Gaussian 98 programs.[23,24]

Results

Monomeric N4-Methoxycytosine. Because of the amino⇌imino, the oxo⇌hydroxy, and the N1H⇌N3H prototropy, N4MC can occur in six different tautomeric forms. Because of the orientation of the methoxy group with respect to the N3 atom, two isomers are possible for each tautomer. Scheme 1 shows all the 12 possible structures for N4-methoxycytosine in decreasing order of stability. Table 1 summarizes the total and relative energies of all of these forms calculated on the different levels of theory. According to all of the theoretical methods used, the two isomers (syn and anti) of the oxo-imino tautomer are by far the most stable forms, with all other structures being at least 32 kJ mol^{-1} less stable. The attractive interaction between the N3−H group and the O atom of the methoxy group in the oxo-imino syn structure (intramolecular H-bond) is most probably at the origin of the energy difference with the anti form. This stability order is different from the stability order for the nonsubstituted cytosine, where the hydroxy-amino form is the most stable structure.[25] This implies that the methoxy group on the amino N atom of cytosine clearly modifies the order of stability. The ΔE values in Table 1 suggest that the oxo-imino syn form as well as the oxo-imino anti form (by its most intense vibrational modes) will be observable in the experimental FT-IR spectrum of N4MC. Identification of all of the other forms in the spectrum is extremely improbable. Figure 1 illustrates the FT-IR spectrum of N4MC in Ar. As usual, very small amounts of water are observable in the spectrum, but because it is essentially free water, this has no further influence on the product spectrum. The analysis of the experimental spectrum was done by comparison with the theoretical predictions.

Tables with the experimental and theoretically predicted spectral parameters for the two most abundant forms, oxo-imino syn (Table 1-S) and oxo-imino anti (Table 2-S), are available as Supporting Information. N4MC has 45 fundamental vibrational modes. When the methyl group is considered as a point mass, these can be classified under C_s symmetry into 31 in-plane and 14 out-of-plane vibrations. There is no large difference between the predicted spectra of the two isomers of the oxo-imino tautomer, except for some modes of or connected to the methoxy group, such as ν(N3H), ν(C5H), ν(CH$_3$), δ(N3H), etc. Almost all fundamental modes of the oxo-imino-syn isomer are found in the experimental spectrum, with the exception of a few very weakly predicted modes. For the anti isomer, only the most intense vibrational modes are assigned tentatively. For some vibrational bands of the latter isomer, it is not possible to measure the intensities, because they overlap with more intense bands of the syn isomer.

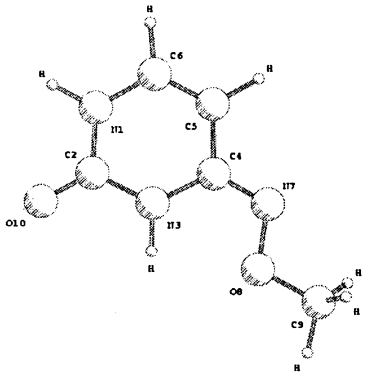
Combination of experimental and theoretical intensities of some characteristic bands of the anti form with some bands of the syn form allows the estimation of the position of the isomeric equilibrium syn⇌anti. The value of the isomerization constant K_I (denoted as "experimental" constant) is calculated using the procedure earlier described by Nowak et al.[26−28]:

$$K_I = \frac{[\text{anti}]}{[\text{syn}]} = \frac{I_{\text{anti}}/A_{\text{anti}}}{I_{\text{syn}}/A_{\text{syn}}}$$

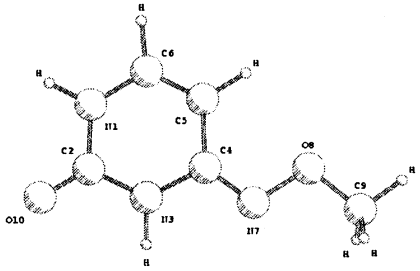
with I_i the experimental intensity and A_i the theoretically predicted intensity.

Because only the bands at 2895 and 1290 cm^{-1} of the anti isomer can be integrated accurately, these bands can be used in combination with some bands of the syn form to estimate the

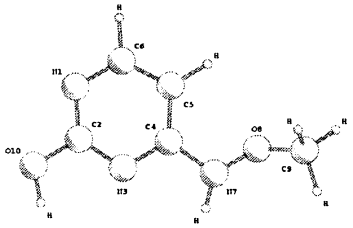
SCHEME 1: Optimal Geometries of the Tautomers and Isomers of N4-Methoxycytosine



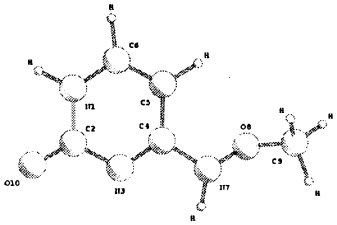
oxo-imino-syn



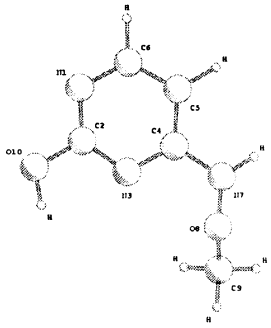
oxo-imino-anti



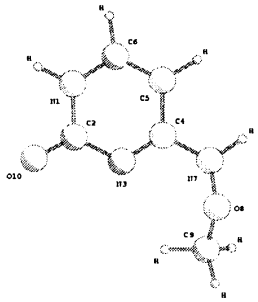
hydroxy-amino-anti



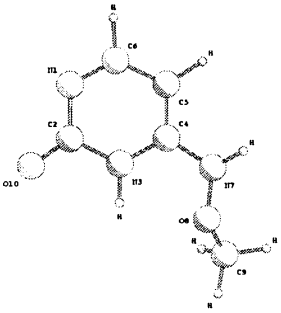
oxo-amino-N1H-anti



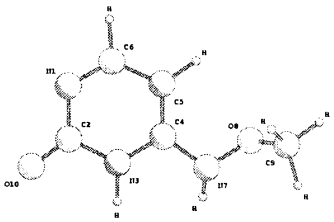
hydroxy-amino-syn



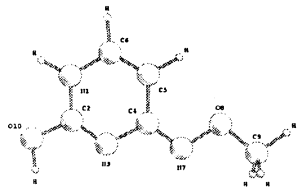
oxo-amino-N1H-syn



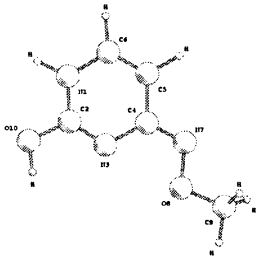
oxo-amino-N3H-syn



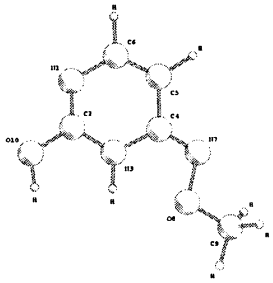
oxo-amino-N3H-anti



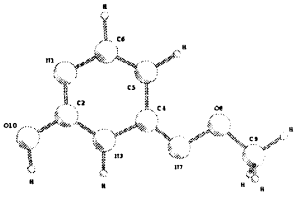
hydroxy-imino-N1H-anti



hydroxy-imino-N1H-syn



hydroxy-imino-N3H-syn



hydroxy-imino-N3H-anti

isomerization constant K_I. The mean experimental value for K_I obtained in this way is 2.2×10^{-2}, when RHF intensities are used. For DFT intensities, this value is close, i.e., 2.4×10^{-2}.

The isomerization constant K_I can also be estimated from the theoretical energy difference between the two isomeric forms:

$$\Delta E_I(0\ \mathrm{K}) \approx \Delta G_I^\circ(T) = \Delta H_I^\circ - T\Delta S_I^\circ = -RT \ln K_I$$

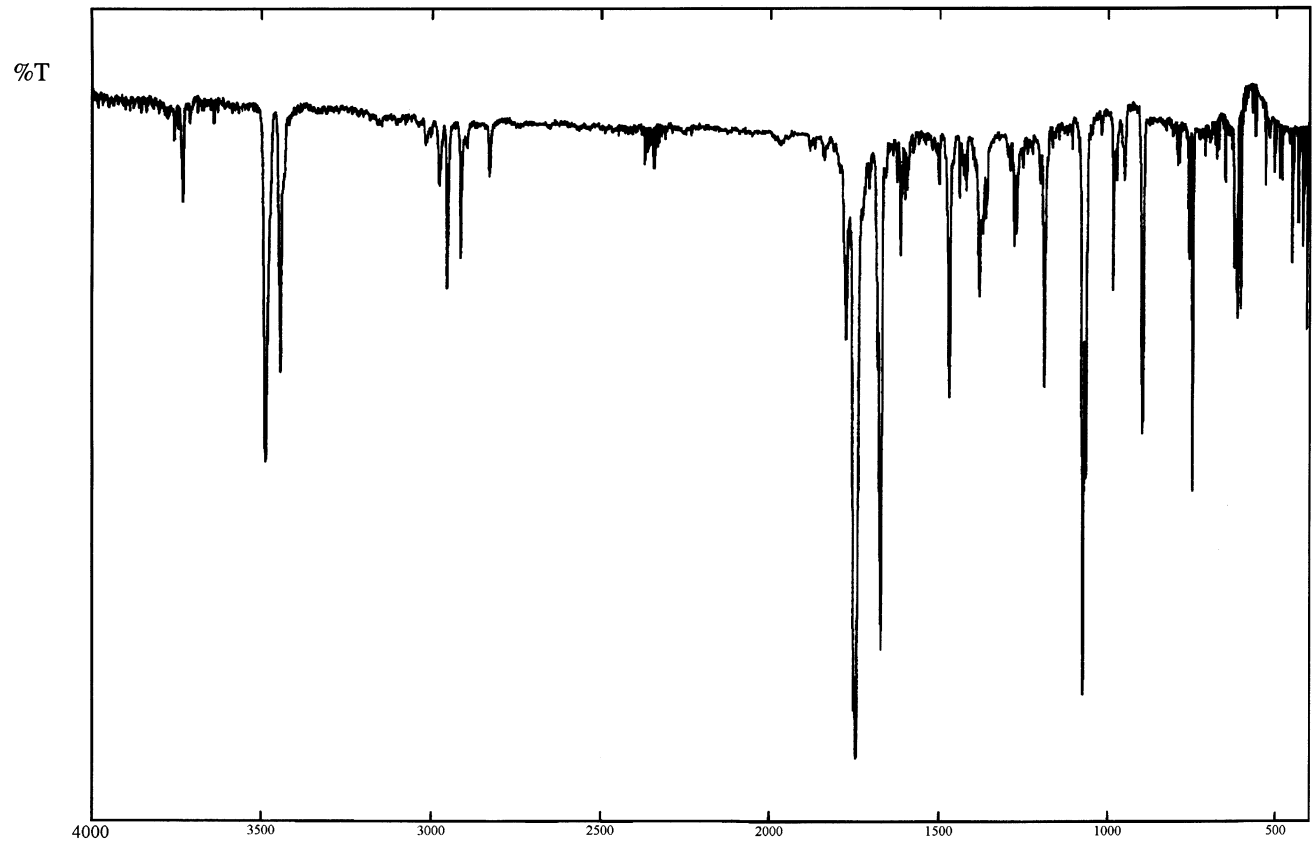


Figure 1. FT-IR spectrum of N4-methoxycytosine in Ar at 18 K (furnace temperature: 306 K).

TABLE 1: Total (a.u.) and Relative Energies (kJ/mol) of the Tautomers of N⁴-Methoxycytosine Calculated at the RHF/MP2 and DFT Levels of Theory with the 6-31++G Basis Set[a]**

	RHF[b]	ΔE^c (RHF)	μ (D) (RHF)	MP2[b]	ΔE^c (MP2)	DFT[d]	ΔE^c (DFT)	μ (D) (DFT)
oxo-imino-syn	−506.479831/	0.00/	3.30	−507.996532/	0.00/	−509.447889/	0.00/	2.87
	−506.352655	0.00		−507.869356	0.00	−509.321061	0.00	
oxo-imino-anti	−506.475354/	11.75/	3.73	−507.991093/	14.28/	−509.442975//	12.90/	3.41
	−506.348370	11.25		−507.864108	13.78	−509.316332	12.42	
hydroxy-amino-anti	−506.467538	32.28/	4.34	−507.983317/	34.70/	−509.430056/	46.82/	4.47
	−506.340156	32.82		−507.855935	35.24	−509.303208	46.87	
oxo-amino-N1H-anti	−506.465770/	36.92/	6.90	−507.981027/	40.71/	−509.431700/	42.50/	6.48
	−506.338478	37.22		−507.853735	41.01	−509.305044	42.05	
hydroxy-amino-syn	−506.462910/	44.43/	4.05	−507.978714/	46.78/	−509.425855/	57.85/	4.07
	−506.335535	44.95		−507.851339	47.30	−509.298998	57.93	
oxo-amino-N1H-syn	−506.461456/	48.24/	6.83	−507.976741/	51.96/	−509.427522/	53.47/	6.30
	−506.334167	48.54		−507.849453	52.26	−509.300865	53.02	
oxo-amino-N3H-syn	−506.455983/	62.61/	8.65	−507.971334/	66.16/	−509.422719/	66.08/	8.52
	−506.328905	62.36		−507.844256	65.90	−509.296062[f]	65.63	
oxo-amino-N3H-anti	−506.454005/	67.81/	8.02	−507.968794/	72.83/	−509.419713/	73.98/	7.98
	−506.327007	67.34		−507.841796	72.36	−509.293058[f]	73.52	
hydroxy-imino-N1H-anti	−506.446917/	86.42/	4.37	−507.966489	78.88/	−509.416115/	83.42/	4.05
	−506.320262	85.05		−507.839834	77.51	−509.289472	82.94	
hydroxy-imino-N1H-syn	−506.446684/	87.03/	4.22	−507.965662/	81.05/	−509.416110/	83.44/	3.81
	−506.320088	85.50		−507.839067	79.53	−509.289283[g]	83.43	
hydroxy-imino-N3H-syn	−506.444605/	92.49/	3.16	−507.963693/	86.22/	−509.414075	88.78/	3.18
	−506.318009[e]	90.96		−507.837097[e]	84.70	−509.287248[g]	88.78	
hydroxy-imino-N3H-anti	−506.439628/	105.55/	2.90	−507.957284/	103.05/	−509.408437/	103.58/	2.82
	−506.312973[e]	104.19		−507.830629[e]	101.68	−509.281794[g]	103.10	

[a] syn: syn position of methoxy group with respect to the N3 atom. anti: anti position of methoxy group witht respect to the N3 atom. [b] First row: not ZPE corrected; second row: ZPE-corrected, ZPE-values scaled with uniform scaling factor 0.900 (RHF). [c] Energy difference between the tautomers relative to the most stable form. [d] First row: not ZPE corrected; second row: ZPE-corrected, ZPE-values scaled with uniform scaling factor 0.970 (DFT). [e] ZPE-contribution not calculated, ZPE of hydroxy-imino-N1H is used. [f] ZPE-contribution not calculated; ZPE of oxo-amino-N1H is used. [g] ZPE-contribution not calculated; ZPE of oxo-imino is used.

The $\approx$ sign indicates two assumptions, i.e., the neglected difference between $\Delta H°(0\ \mathrm{K})$ and $\Delta H°(T)$ and the neglect of the $T\Delta S°$ term. These assumptions are acceptable according to the calculations performed by Kwiatkowski et al.,[29] at least if the internal energy $\Delta E(0\ \mathrm{K})$ is relatively large. The theoretical K_I values obtained at the different theoretical levels are 1.2×10^{-2} (RHF), 4.4×10^{-3} (MP2), and 7.6×10^{-3} (DFT). These values differ from the experimental values with factors varying

TABLE 2: Total (a.u.) and Relative Energies (kJ/mol) of the Possible H-Bond Complexes of the Two Isomers of the oxo-imino Form of N⁴-Methoxycytosine with Water, Calculated at the RHF/MP2 and DFT Levels of Theory with the 6-31++G Basis Set**

	RHF[a]	ΔE^b (RHF)	μ (D) (RHF)	MP2[a]	ΔE^b (MP2)	DFT[c]	ΔE^b (DFT)	μ (D) (DFT)
				syn				
N1−H...O−H...O═C2	−582.525484 /	0.00 /	1.85	−584.247297 /	0.00 /	−585.899089 /	0.00 /	1.75
	−582.374430	0.00		−584.096243	0.00	−585.748055	0.00	
N3−H...O−H...O═C2	−582.521698 /	9.94 /	3.25	−584.243850 /	9.05 /	−585.895196 /	10.22 /	2.96
	−582.370936	9.17		−584.093089	8.28	−585.744494	9.35	
N7...H−O...(H−C5)	−582.519345 /	16.12 /	1.93	−584.241186 /	16.04 /	−585.892535 /	17.21 /	1.59
	−582.368794	14.80		−584.090635	14.72	−585.742133	15.55	
				anti				
N1−H...O−H...O═C2	−582.521172 /	11.32 /	2.24	−584.242022 /	13.85 /	−585.894337 /	12.48 /	2.20
	−582.370300	10.84		−584.091146	13.38	−585.743420	12.17	
N3−H...O−H...O═C2	−582.518298 /	18.87 /	3.49	−584.239431 /	20.65 /	−585.891424 /	20.12 /	3.41
	−582.367583	17.98		−584.088716	19.76	−585.740801	19.05	
O8...H−O(...H−C5)	−582.513095 /	32.53 /	1.58	−584.234391 /	33.88 /	−585.884912 /	37.22 /	1.34
	−582.362925	30.21		−584.084222	31.56	−585.734971	34.35	

[a] First row: not ZPE-corrected; second row: ZPE-corrected RHF and MP2 energies, ZPE values scaled with the uniform scaling factor 0.900 (RHF). [b] Energy difference between the possible complexes relative to the most stable complex N1−H...O−H...O═C2 of the oxo-imino syn form. [c] First row: not ZPE-corrected; second row: ZPE-corrected DFT energy, ZPE values scaled with the uniform scaling factor 0.970 (DFT).

from 3 (DFT) to 5 (MP2). For tautomerization processes, much larger differences between experiment and theory are usually found.[30] The detailed mechanism of the tautomerization process is still not known, but there are strong indications for a bimolecular process, where H-bond-assistance is needed to pass the energy barrier between the two tautomers. This implies that the tautomeric equilibrium is not solely governed by the free enthalpy difference between the monomeric tautomers.[31] On the other hand, in the case of rotamerization or isomerization processes, the experimental and the theoretical equilibrium constants appear to be comparable,[30] as can be expected for a monomolecular process. One still needs more values of experimental and theoretical constants for rotamerization and isomerization in order to draw definite conclusions. In this case, the minor difference observed between theoretical and experimental K_I values is probably due to the imperfection of the theoretical calculations in combination with the different simplifications in the equation used for the calculation of the isomerization constant K_I.

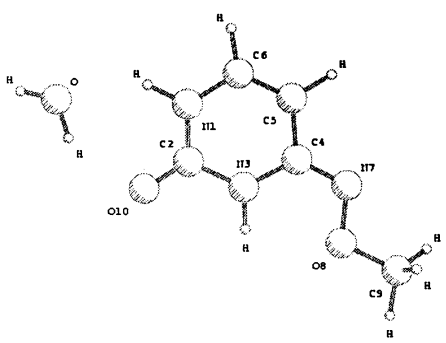
The different scaling procedures described in the theoretical section can be evaluated by comparing the observed and theoretical IR frequencies. When a single scaling factor is used (0.900 for RHF and 0.970 for DFT), the mean frequency deviations $|\Delta\nu| = \overline{|\nu_{exp}-\nu_{theor}|}$ are 19.2 and 19.0 cm^{-1}, for RHF and DFT, respectively. The use of variable scaling factors, i.e., 0.950 for ν(X−H), 0.980 for the out of plane modes γ and τ, and 0.975 for all other modes, decreases the mean deviation to 10.4 cm^{-1}. These values are similar to the mean deviations obtained in earlier studies, e.g., for imidazole,[13] benzimidazole,[14] and hypoxanthine.[16]

H-Bonding Properties of N4-Methoxycytosine. Because water can act simultaneously as proton donor and as proton acceptor, it is a suitable model molecule to study the H-bonding properties of nucleic acid bases and related polyfunctional compounds. We have limited the investigation of the H-bonding properties of N4MC to the H-bonding capability of the two isomers of the abundant oxo-imino tautomer. The dipole moments of the two isomers are not significantly different (syn, 3.30 D; anti, 3.73 D, Table 1), which implies that a shift in the conformational equilibrium is not to be expected in a more polar environment. Both isomers of the oxo-imino tautomer have different H-bonding interaction sites available: two H-donor sites (N1−H and N3−H) and three H-acceptor sites (C2═O and the N and O atoms of the N-methoxy group). The vicinal position of two H-bonding sites of opposite nature allows the formation of a closed complex with water, linked by two H-bonds. Two similar closed complexes are possible for both isomers: N1−H...O−H...O═C2 and N3−H...O−H...O═C2. The larger acidity of the N1−H group, which is easily understood from resonance structures, suggests that the N1−H...O−H...O═C2 complex would be stronger than the N3−H...O−H...O═C2 complex. A third complex, N7...H−O (...H−C5), is possible for the syn isomer. In this complex, a rather strong interaction takes place with the N7 atom, whereas an additional weak H bond appears with the H atom on C5. For the anti isomer, a third water complex involving the O atom of the methoxy group and the H atom on C5, i.e., O8...H−O (H−C5), is possible. Because the H-bond interaction with C5−H is very weak, both the latter complexes may be considered as "open" complexes.

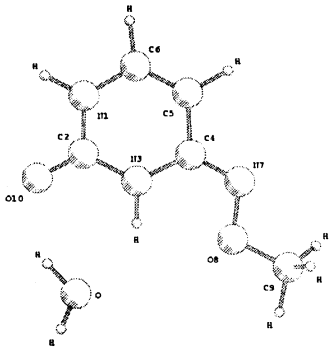
Table 2 lists the total and relative energies of the six complexes considered in this work, all calculated at the three theoretical levels RHF, MP2, and DFT. The optimal geometries of these six structures are shown in Schemes 2 and 3. Both closed complexes of the syn isomer are more stable than the two closed complexes of the anti isomer. Nevertheless, the H-bond interaction energies of the N1−H...O−H...O═C2 complexes of both isomers are very similar and much stronger than for the N3−H...O−H...O═C2 complexes (Table 3), as can be expected from the difference in acidity between the two N−H groups. Comparison with the results in Table 1 allows us also to conclude that the energy difference between the two isomers of the oxo-imino tautomer does not change by the H bonding with water.

From the above theoretical results, it can be expected that it will be extremely difficult to discriminate between the experimental spectra of the two closed complexes of the syn isomer and the closed complexes of the anti isomer. We have therefore only analyzed the complex spectra in a first stage in terms of the presence of two syn complexes. Tables with the assignments of the complex bands in terms of these two complexes, N1−H...O−H...O═C2 (Table 3-S) and N3−H...O−H...O═C2 (Table 4-S), are available as Supporting Information. Only the DFT predicted frequencies (with the procedure of different scaling factors) are used to analyze the experimental spectra. Furthermore, the spectral analysis is limited to the vibrational modes

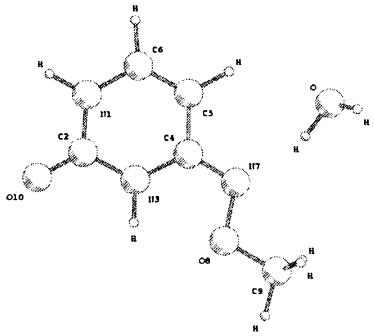
SCHEME 2: Optimal Geometries of the Complexes of the oxo-imino syn Form of N4-Methoxycytosine with Water



N1-H…O-H…O=C2

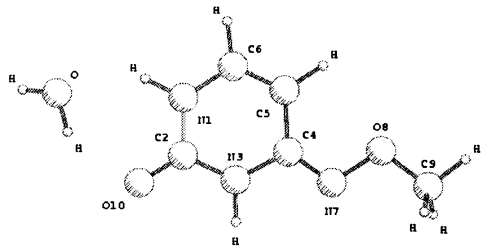


N3-H…O-H…O=C2

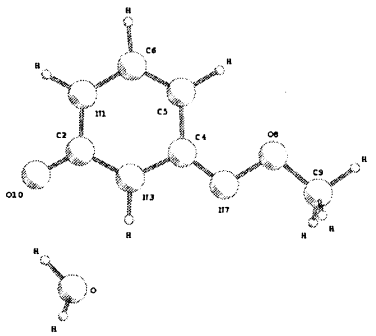


N7…H-O(…H-C5)

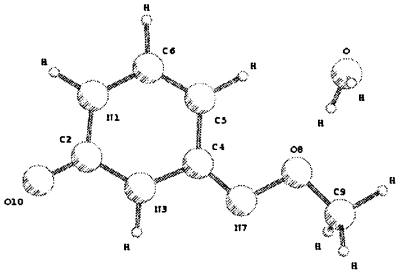
SCHEME 3: Optimal Geometries of the Complexes of the oxo-imino anti Form of N4-Methoxycytosine with Water



N1-H…O-H…O=C2



N3-H…O-H…O=C2



O8…H-O(…H-C5)

directly involved in the H-bond interaction. The important spectral regions of the IR spectra are discussed separately.

***ν*^{f}(OH) Region.** As in the case of hypoxanthine·H_2O complexes in Ar,[17] this spectral region is not suitable to make any distinction between the two complexes. As a matter of fact, the ν^{f}(OH) frequency and shift $\Delta\nu^{f}$(OH) for both complexes are predicted to be nearly the same. Only around 3700 cm^{-1}, different intensities are observable between the complex spectrum N4MC/H_2O/Ar and the water spectrum H_2O/Ar. This implies that a very weak complex band is present in this region, but it is largely obscured by the spectral noise. The location of the stretching mode of the free O−H bond in bonded water is obviously the reason for the higher intensity observed around 3700 cm^{-1} in the complex spectrum, superimposed on the trimer

TABLE 3: Interaction Energies (kJ mol^{-1}) of the Most Stable H-Bonded Complexes of N4-Methoxycytosine with Water

	RHF//RHF	MP2//RHF	DFT//DFT
Isomer 1: oxo-imino SYN Form (3 Complexes)			
N1−H...O−H...O═C2			
ΔE	−37.66	−47.84	−44.83
$\Delta E + \Delta$ZPE	−29.58	−39.76	−35.47
$\Delta E + \Delta$ZPE − BSSE	−27.29	−31.79	−32.89
N3−H...O−H...O═C2			
ΔE	−27.72	−38.79	−34.61
$\Delta E + \Delta$ZPE	−20.40	−31.48	−26.13
$\Delta E + \Delta$ZPE − BSSE	−17.67	−23.26	−22.67
N7...H−O(...H−C5)			
ΔE	−21.54	−31.80	−27.63
$\Delta E + \Delta$ZPE	−14.78	−25.03	−19.92
$\Delta E + \Delta$ZPE − BSSE	−12.66	−18.13	−17.14
Isomer 2: oxo-imino ANTI Form (2 Complexes)			
N1−H...O−H...O═C2			
ΔE	−38.09	−48.27	−45.26
$\Delta E + \Delta$ZPE	−29.97	−40.15	−35.72
$\Delta E + \Delta$ZPE − BSSE	−27.35	−32.43	−32.40
N3−H...O−H...O═C2			
ΔE	−30.55	−41.47	−37.62
$\Delta E + \Delta$ZPE	−22.85	−33.77	−28.84
$\Delta E + \Delta$ZPE − BSSE	−20.18	−25.84	−25.57
O8...H−O(...H−C5)			
ΔE	−16.89	−28.24	−20.51
$\Delta E + \Delta$ZPE	−10.62	−21.98	−13.54
$\Delta E + \Delta$ZPE − BSSE	−8.21	−13.84	−10.67

absorption in free water.[34] The experimental frequency shift $\Delta\nu^f$(OH), calculated as the difference of the complex band and the monomer band, is around −34 cm^{-1}, in accordance with the theoretical prediction.

***ν*^{b}(OH) and *ν*(NH) Region.** As is usually the case for H-bonded complexes with water, it is much easier to discriminate between different complexes using the ν^b(OH) frequencies of H-bonded water. As indicated in Table 3 (H-bond interaction energies), the strength of the H-bonding is not the same for both "similar" complexes, and this is also reflected by the $\Delta\nu^b$(OH) values which differ by about 60 cm^{-1}. In the experimental spectrum (Figure 2), two new bands appear clearly in the ν^b(OH) region, although they are relatively weak. The rather broad band at 3405 cm^{-1} can be assigned to the ν^b(OH) band of the N1−H...O−H...O═C2 complex, whereas the new band at 3470 cm^{-1}, appearing as a shoulder on the ν(N1H) monomer mode, can be reasonably assigned to ν^b(OH) of the complex N3−H...O−H...O═C2. As a matter of fact, the experimental frequency shifts $\Delta\nu^b$(OH) of −233 and −168 cm^{-1} for the N1−H...O−H...O═C2 and N3−H...O−H...O═C2 complexes agree rather well with the predicted shifts of −212 and −176 cm^{-1}, respectively. For the N−H...O−H...O═C complexes in the system hypoxanthine·H_2O, $\Delta\nu^b_{OH}$ shifts of −198 → −292 cm^{-1} have been observed.[17]

The bonded N−H stretching region of the H-bonded N4-methoxycytosine is situated somewhat lower in the spectrum. From a predicted shift of −191 cm^{-1} for the N1−H stretching mode, an absorption band around 3290 cm^{-1} is expected in the experimental spectrum. The weak band at 3277 cm^{-1} can tentatively be assigned to this ν(N1H) mode, whereas the similar band at 3246 cm^{-1} is probably caused by the bonded N3−H stretching mode in the second complex which is characterized by a predicted shift of −183 cm^{-1}.

The $\nu(CH_3)$ region (around 2950 cm^{-1}) remains almost unchanged after addition of water to the matrix.

***ν*(C═O) and *δ*(H_2O) Region.** The C═O stretching vibration is predicted to be very intense for both H-bonded complexes, but the predicted shift is rather small in both cases. Comparison of the experimental monomer spectrum with different spectra of the complexes (different amounts of water) in Figure 3 reveals that the C═O stretching mode is shifted from 1743 to 1726 cm^{-1}, with a shoulder at 1717 cm^{-1}, because of the H-bonding with water. These bands are assigned to the bonded ν(C═O)

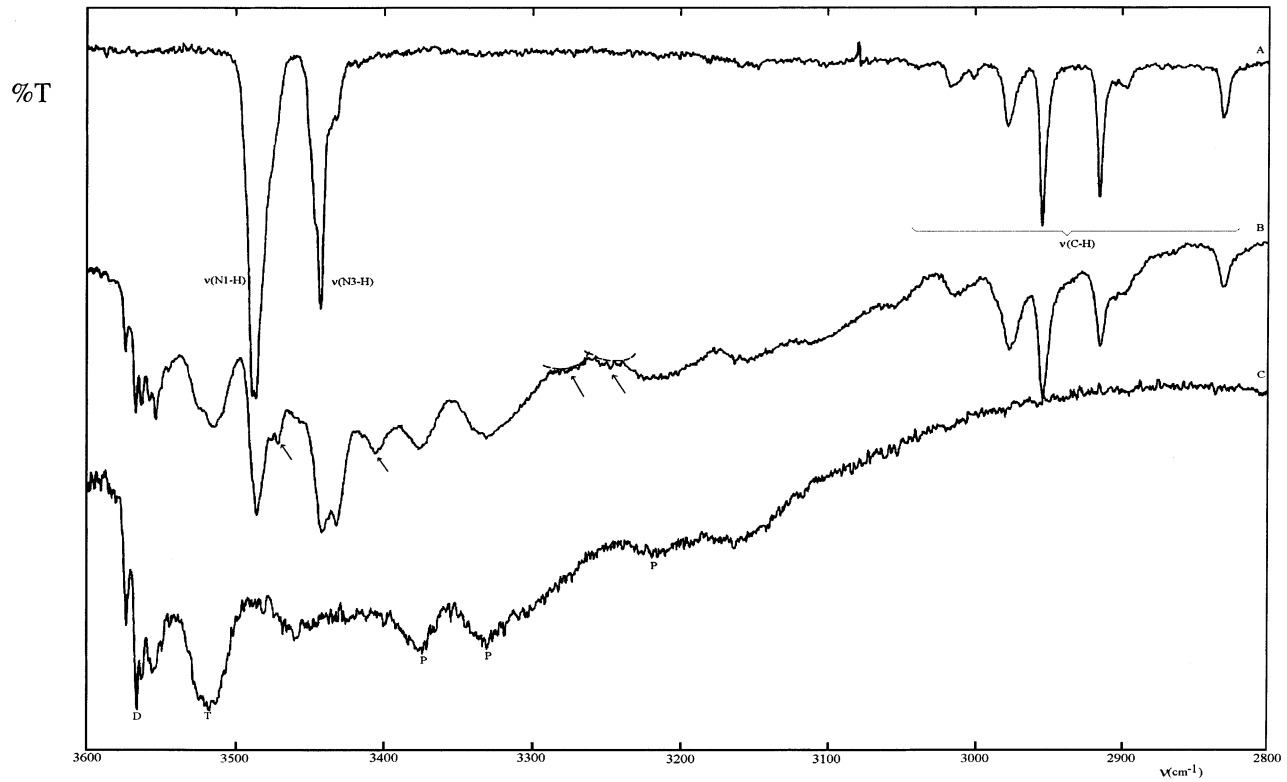


Figure 2. The 3600−2800 cm^{-1} region of the FT-IR spectrum of N4-methoxycytosine/Ar (A), N4-methoxycytosine/H_2O/Ar (B), and H_2O/Ar (C). ↑, Complex band; D, water dimer; T, water trimer; P, water polymer.[34]

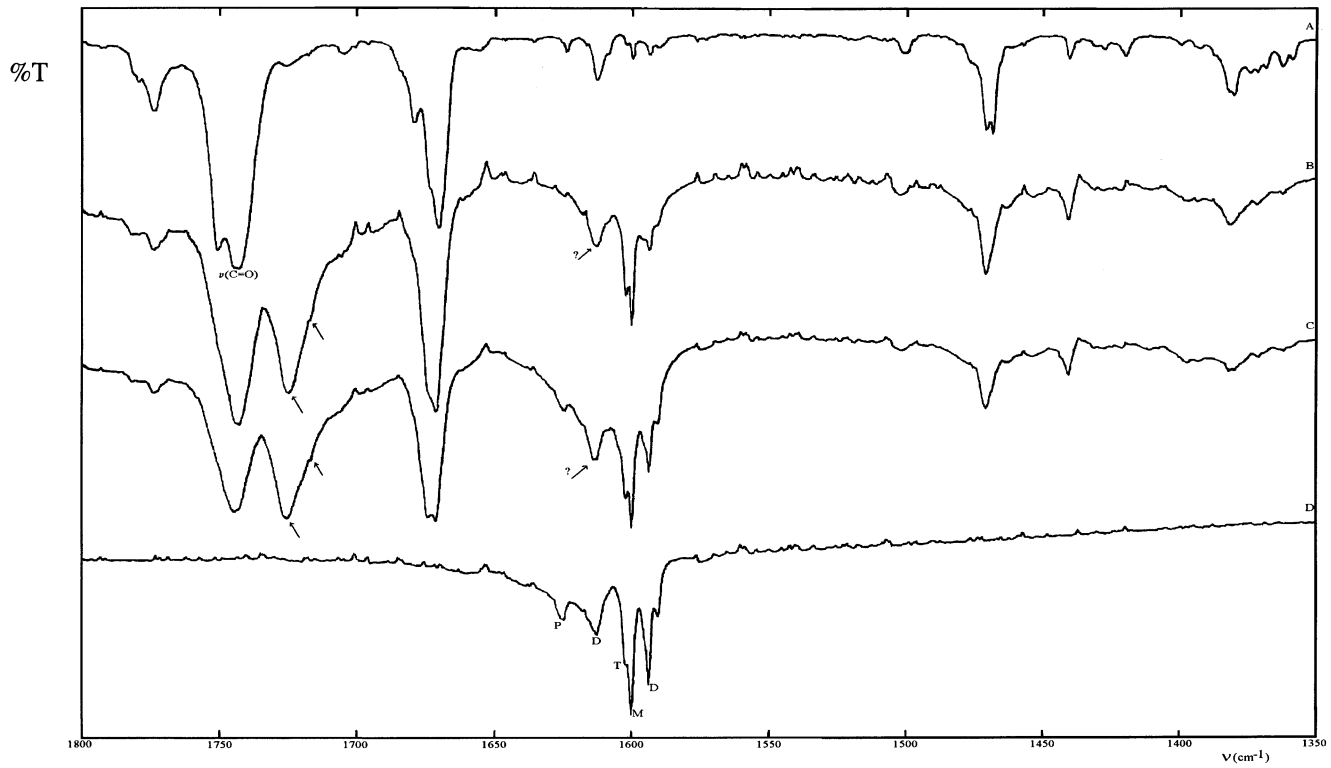


Figure 3. The 1800−1350 cm^{-1} region of the FT-IR spectrum of N4-methoxycytosine/Ar (A), N4-methoxycytosine/H_2O/Ar (H_2O/Ar: 1/800) (B), N4-methoxycytosine/H_2O/Ar (H2O/Ar:1/400) (C), and H_2O/Ar (D). ↑, Complex band; M, water monomer; D, water dimer; T, water trimer; P, water polymer.[34]

mode of the N1−H...O−H...O═C2 complex and the N3−H...O−H...O═C2 complex, respectively. As expected, the intensities of the bands due to the free and bonded C═O stretchings are reversed on addition of more water to the matrix.

The water bending vibration $\delta(H_2O)$ has about the same predicted frequency for both complexes. This vibration must be situated in the region of 1610 cm^{-1} in the experimental spectrum (Figure 3). It is hard to distinguish this band from the dimer band of pure water and from the ring stretching of the monomer situated in the same spectral region.

The δ(N−H) vibrations have a very small intensity in the complexes and these modes cannot be identified in the experimental spectrum.

***τ*(...H_2O...), *γ*(N−H), and *γ*(C═O) Region.** In the low-frequency region (850−500 cm^{-1}) in the complex spectrum (Figure 4), some typical, broad new bands appear. The band at ±777 cm^{-1} is most probably due to the out-of-plane γ(N1H) vibration of the N1−H...O−H...O═C2 complex, because the experimental and theoretical shifts for this band agree well. The band at about 800 cm^{-1} can tentatively be assigned to the γ(N3H) mode of the second complex, N3−H...O−H...O═C2.

The out-of-plane vibrations of the C2═O group are too weak to be identified with confidence in the experimental spectrum.

Two broad bands also appear around 650 cm^{-1}. These bands are due to the new, intermolecular torsion vibration of H-bonded water τ(...OHH...). In the case of hypoxanthine·H_2O, similar bands have been observed at 680 and 630 cm^{-1} for the same type of complexes N−H...O−H...O═C.[17]

Although the third complex N7...H−O (...H−C5) is 15.55 kJ mol^{-1} less stable than the N1−H...O−H...O═C2 complex, some very weak new bands in the complex spectrum might be due to this complex. The ν^b(O−H) vibration overlaps presumably with the ν(N3H) mode of the monomer, whereas the ν^f(O−H) and the $\delta(H_2O)$ overlap with the same modes of the other complexes. The two new bands at 1084 and 916 cm^{-1} might be due to the δ(C4N7) and ν(N7O8) modes, respectively, whereas the broad band at 846 cm^{-1} (Figure 4) probably reflects the γ(C4N7...) mode of this complex.

Correlation

The results obtained in this work allow us to further develop the correlation between the $-\Delta\nu^b$(O−H) values (experimental or theoretical) and the increase of the O−H distance (theoretical) in the N−H...O−H...X complexes. Figure 5 illustrates this correlation obtained for these types of complexes found for water with N4-methoxycytosine (this work), hypoxanthine,[17] and 2-oxo-pyridine.[32] As can be expected, the O−H distance increases linearly with increasing O−H frequency shifts. Both parameters are directly related to the extent of the shift of electrons from the X (O or N) atom lone pair to the antibonded σ^* orbital of the O−H bond. A similar correlation has been found for open N...H−O H-bonded systems.[33]

Conclusions

The combined theoretical and experimental matrix FT-IR study of N4-methoxycytosine and its H-bonded complexes with water leads to the following important conclusions:

1. The two isomers (syn and anti) of the oxo-imino tautomer of N4-methoxycytosine are by far the most stable structures according to the RHF, MP2, and DFT calculations with the 6-31++G** basis set.

2. The experimental FT-IR matrix spectral data support the theoretical predictions.

3. The methoxy group on the N4 atom of cytosine modifies the stability order of the tautomers completely, compared to cytosine.

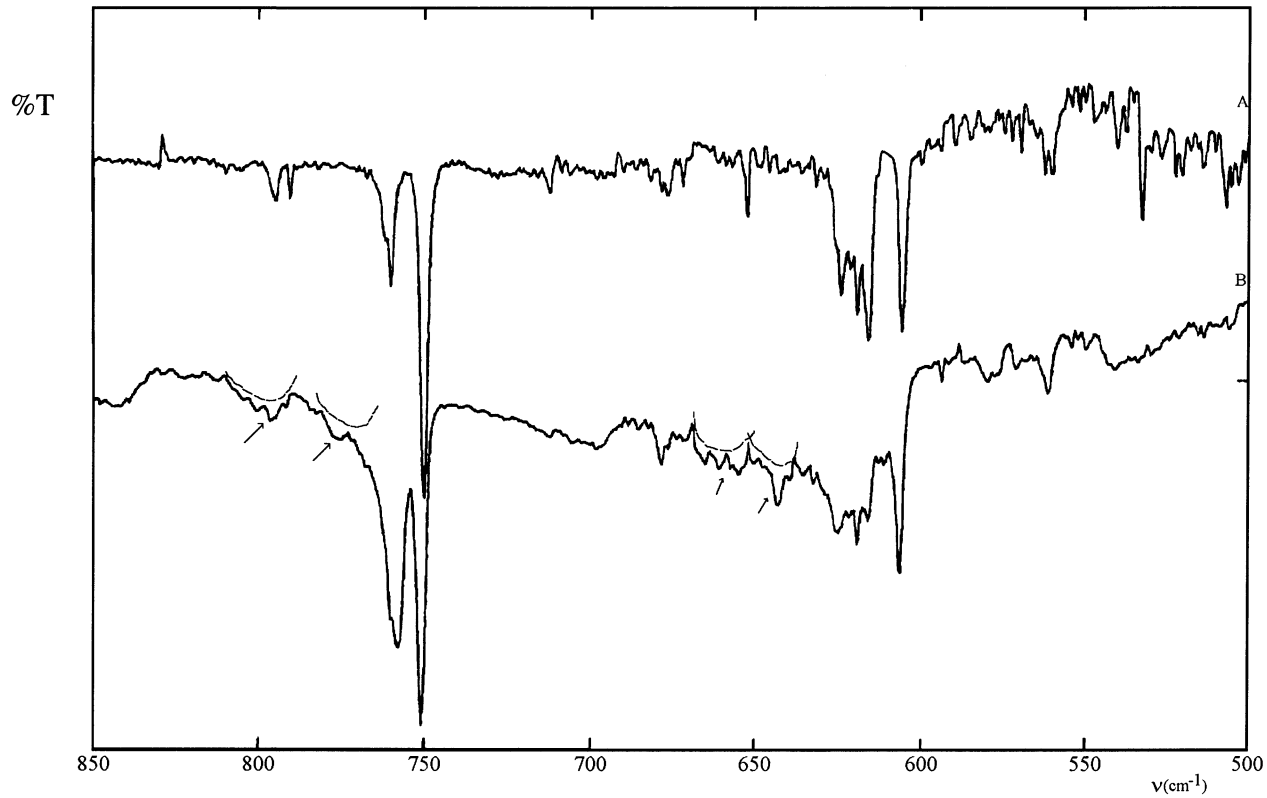


Figure 4. The 850−500 cm^{-1} region of the FT-IR spectrum of N4-methoxycytosine/Ar (A) and N4-methoxycytosine/H_2O/Ar (B). ↑, Complex band.

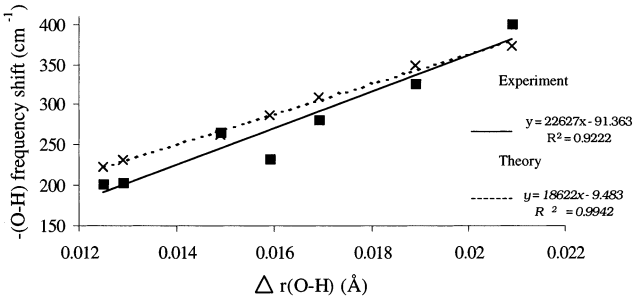


Figure 5. Correlation between the experimental or the theoretical frequency shift $-\Delta\nu^b$(OH) and the predicted increase of the O−H distance in N−H...O−H...X H-bonded systems for N4-methoxycytosine, hypoxanthine, and 2-oxopyridine with water.

4. The agreement between experimental and theoretical frequencies is satisfactory at the RHF and DFT level, when a single scaling factor is used. However, a clear improvement is observed when different scaling factors are used at the DFT level.

5. Two closed complexes of N4-methoxycytosine with water are formed for both oxo-imino isomers (N1−H...O−H...O═C2 and N3−H...O−H...O═C2) and one less stable open complex (N7...H−O and O8...H−O, respectively) for both isomers. No tautomeric shift is induced by the addition of water.

6. The experimental spectra of the water complexes agree well with the predictions at the DFT level of theory. Especially the modes directly involved in the H-bond interactions such as ν^b(O−H), ν(N−H), ν(C═O), γ(N−H), and the new water mode τ(...OHH...) are clearly identified in the experimental spectra. From these spectra, it is obvious that both closed complexes are present. Some very weak complex bands might be due to the weaker complex N7...H−O(...H−C5) of the syn isomer.

Calculation of the optimal scaling factor for the DFT predicted frequencies confirms earlier calculated scaling factor values.

The rather classic but important correlation between the shift of the bonded O−H stretching vibration and the elongation of the O−H bond can be extended with these results.

Acknowledgment. R.R. acknowledges financial support from the Flemish Institute for the Promotion of the Scientific-Technological Research in Industry (IWT).

Supporting Information Available: Tables with the experimental and theoretical predicted spectral parameters for the most abundant forms, i.e., oxo-imino syn (Table 1-S) and oxo-imino anti (Table 2-S). Tables with the assignments of the complex bands in terms of the two most stable H-bonding complexes, i.e., N1−H...O−H...O═C2 (Table 3-S) and N3−H...O−H...O═C2 (Table 4-S). This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Brown, D. M.; Hewlins, M. J. E.; Schell, P. *J. Chem. Soc.* **1968**, 1925.

(2) Janion, C. *Acta Biochim. Polon.* **1972**, *19*, 261.

(3) Kulinska, K.; Psoda, A.; Shugar, D. *Acta Biochim. Polon.* **1980**, *27*, 57.

(4) Morozov, Yu. V.; Savin, F. A.; Chekhov, V. O.; Budowsky, E. I.; Yakovlev, D. Yu. *J. Photochem.* **1982**, *20*, 229.

(5) Kierdaszuk, B.; Stolarski, R.; Shugar, D. *Eur. J. Biochem.* **1983**, *130*, 559.

(6) Van Meervelt, L. *Acta Cryst.* **1991**, *C47*, 2635.

(7) Van Meervelt, L.; Moore, M. H.; Lin, P. K. T.; Brown, D. M.; Kennard, O. *J. Mol. Biol.* **1990**, *216*, 773.

(8) Maes, G. *Bull. Soc. Chim. Belg.* **1981**, *90*, 1093. Destexhe, A.; Smets, J.; Adamowicz, L.; Maes, G. *J. Phys. Chem.* **1994**, *98*, 1506.

(9) Graindourze, M.; Smets, J.; Zeegers-Huyskens, Th.; Maes, G. *J. Mol. Struct.* **1990**, *222*, 345.

(10) Brown, D. M.; Schell, P. *J. Chem. Soc.* **1965**, 208.

(11) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(12) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev.* **1988**, *B37*, 785.

(13) Van Bael, M. K.; Schoone, K.; Houben, L.; Smets, J.; McCarthy, W.; Adamowicz, L.; Nowak, M. J.; Maes, G. *J. Phys. Chem. A* **1997**, *101*, 2397.

(14) Schoone, K.; Smets, J.; Houben, L.; Van Bael, M. K.; Adamowicz, L.; Maes, G. *J. Phys. Chem. A* **1998**, *102*, 4863.

(15) Smets, J.; Schoone, K.; Ramaekers, R.; Adamowicz, L.; Maes, G. *J. Mol. Struct.* **1998**, *442*, 201.

(16) Ramaekers, R.; Maes, G.; Adamowicz, L.; Dkhissi, A. *J. Mol. Struct*. **2001**, *560*, 205.

(17) Ramaekers, R.; Dkhissi, A.; Adamowicz, L.; Maes, G. *J. Phys. Chem. A* **2002**, *106*, 4502.

(18) Hobza, P.; Sponer, J.; Reschel, T. *J. Comput. Chem*. **1995**, *1611*, 1315.

(19) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(20) Florian, J.; Johnson, B. G. *J. Phys. Chem*. **1994**, *98*, 3681.

(21) Rauhut, G.; Pulay, P. *J. Phys. Chem.* **1995**, *99*, 3093.

(22) Florian, J.; Leszczynski, J. *J. Phys. Chem*. **1996**, *100*, 5578.

(23) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, revision E.2; Gaussian, Inc.: Pittsburgh, PA, 1995.

(24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.5; Gaussian, Inc.: Pittsburgh, PA, 1998.

(25) Les, A.; Adamowicz, L.; Bartlett, R. J. *J. Phys. Chem*. **1989**, *93*, 4001.

(26) Lapinski, L.; Fulara, J.; Nowak, M. J. *Spectrochim. Acta* **1990**, *46A*, 61.

(27) Lapinski, L.; Fulara, J.; Czerminsky, R.; Nowak, M. J. *Spectrochim. Acta* **1990**, *46A*, 1087.

(28) Nowak, M. J.; Lapinski, L.; Fulara, J.; Les A.; Adamowicz, L. *J. Phys. Chem*. **1992**, *96*, 1562.

(29) Kwiatkowski, J. S.; Leszczynski, J. *Chem. Phys. Lett*. **1993**, *204*, 5−6, 430.

(30) Houben, L.; Schoone, K.; Smets, J.; Adamowicz, L.; Maes, G. *J. Mol. Struct.* **1997**, *410*−*411*, 397.

(31) Nowak, M. J.; Lapinski, L.; Fulara, J.; Les, A.; Adamowicz, A. *J. Phys. Chem*. **1992**, *96*, 1562.

(32) Dkhissi, A.; Houben, L.; Adamowicz, L.; Maes, G. *J. Phys. Chem. A* **2000**, *104*, 9785.

(33) Schoone, K. MS Thesis, KULeuven 1995.

(34) Bentwood, R. M.; Barnes, A. J.; Orville-Thomas, W. J. *J. Mol. Spect.* **1980**, *84*, 391.