Neutron Inelastic Scattering, Optical Spectroscopies and Scaled Quantum Mechanical Force Fields for Analyzing the Vibrational Dynamics of Pyrimidine Nucleic Acid Bases: 3. Cytosine

A. Aamouche and M. Ghomi*

Laboratoire de Physicochimie Biomoléculaire et Cellulaire, CNRS URA 2056, Case courrier 138, Université Pierre et Marie Curie, 4 Place Jussieu, 75252 Paris Cedex 05, France

L. Grajcar, M. H. Baron, and F. Romain

Laboratoire de Spectrochimie Infrarouge et Raman, CNRS UPR 2631, 2 rue H. Dunant, 94320 Thiais, France

V. Baumruk and J. Stepanek

Institute of Physics, Charles University, Ke Karlovu 5, 12116 Prague 2, Czech Republic

C. Coulombeau

Laboratoire d'Etudes Dynamiques et Structurales de la Sélectivité, CNRS URA 332, Université J. Fourier, BP 53X, 38041 Grenoble, France

H. Jobic

Institut de Recherche sur la Catalyse, 2 avenue A. Einstein, 69626 Villeurbanne, France

G. Berthier

Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, F-75231 Paris Cedex 05, and Laboratoire de Radioastronomie Millimétrique, Ecole Normale Supérieure, 75231 Paris Cedex 05, France

Received: June 23, 1997; In Final Form: October 3, 1997[⊗]

Neutron inelastic scattering (NIS), Raman scattering (R), and infrared absorption (IR) spectra of cytosine and its N-deuterated species in solid phase were recorded at 15 K. These spectra were completed by the room temperature Raman and IR spectra from both solid samples and aqueous solutions. NIS spectra are sensitive to the modes in which the hydrogen displacements are involved. The evolution of the vibrational spectra in going from solid phase to solution is discussed. To analyze the structural and vibrational features of an isolated cytosine, we resorted to the quantum mechanical calculations performed at SCF or MP2 levels of theory, with double-ζ form basis sets enlarged or not by adequate polarization functions. The energetically most favorable tautomers, e.g. amino−oxo (a−o) and amino−hydroxy (a−h) forms have been analyzed in this framework. MP2 calculations have shown that the a−h tautomeric form has a lower energy than the well-known a−o form. However, taking into account the small energy difference between the above tautomers, all of them have been considered in assigning the vibrational modes observed in gas phase or in Ar matrix. It turns out that the conformation of the amino group in cytosine deduced from the calculations depends strongly on the extension of the basis functions used. At MP2 level, as soon as the d-polarization functions have been added to the basis sets the NH_2 group prefers a pyramidal conformation. It has been verified that the inversion of the nitrogen located at the top of this pyramid does not correspond to a double minima and that the planar geometry of this chemical group, leading to an overall C_s symmetry for cytosine, corresponds to a transition state. However, as the calculations show, the preferred amino group pyramid orientation (upward or downward with respect to the cytosine ring) also depends on the extension of the basis sets used. The addition of p-polarization functions on hydrogen orbitals leads to the inversion of the NH_2 orientation. In condensed phase, experimental evidence is in favor of the predominance of the a−o form. In order to assign the vibrational modes observed in the condensed phase, where intermolecular interactions are important, the ab initio force field of the a−o tautomer has been scaled. Although, in many cases this scaling improves the agreement between the experimental and calculated results (wavenumbers and NIS intensities), some crucial problems still exist. Indeed, as has been confirmed by the comparison between the experimental and calculated NIS intensities, the scaled theoretical force field overestimates the coupling between N1−H (or NH_2) wagging and torsional motions, whereas the coupling between C−H waggings and skeletal torsional motions is largely underestimated. Similar effects have also been observed in the other two pyrimidine bases (Aamouche, A. et al. *J. Phys. Chem*. **1996**, *100*, 5224; *J. Phys. Chem. A* **1997**, *101*, 1808). Recently, it has been shown that an implicit introduction of intermolecular hydrogen bonds by taking into consideration supermolecular models allows to solve partly the above-mentioned problems without a crucial need to scale the molecular force field (Ghomi, M. et al. *J. Mol. Struct.* **1997**, *411*, 323).

I. Introduction

Following our recent investigations on the vibrational analysis of the two pyrimidine bases, i.e., uracil[1] and thymine,[2] we report here our results concerning cytosine (DNA and RNA base). The aim of this paper is to bring a new insight to the structural and vibrational features of cytosine (C) and its N-deuterated (C-d_3) species. Both neutron inelastic scattering (NIS) and optical (Raman and IR) techniques have been jointly used in order to collect a complete set of experimental data in solid and in aqueous phases. Ab initio quantum mechanical calculations

* To whom correspondence should be addressed. Fax: 33 1 44277560. E-mail: ghomi@lpbc.jussieu.fr.

⊗ Abstract published in *Advance ACS Abstracts,* November 15, 1997.

S1089-5639(97)02037-9 CCC: $14.00

considering electronic correlation have been undertaken in order to estimate as accurately as possible the structural parameters and harmonic vibrations of the compounds under study.

The utility of the present results can be understood in the context of the previous experimental and theoretical investigations on cytosine and its N-deuterated species. Experimentally, only the vibrational spectra analyzed by Raman and IR techniques were available in the literature. One can mention here the spectra analyzed in gas phase[3] (for C), in rare gas matrixes[4−9] at low temperature (for C and C-d_3), in aqueous solutions,[10] and finally in solid state at room temperature by Raman and IR techniques[11−14] or by far-IR spectroscopy (25−1000 cm^{-1}).[15] Theoretically, one can recall all normal-mode calculations using empirical[12,13] or quantum mechanical force fields, e.g. those obtained at SCF level[5,8,9,14,16−18] or with correlation corrections performed by density functional theory (DFT).[14,18−20]

In light of the above-mentioned data, the most important problem concerns the tautomerism of cytosine. In fact, since the earliest study of the X-ray diffraction patterns of anhydrous cytosine,[21] a possible coexistence of different tautomers, with a predominance of the amino−oxo (a−o) form in solid phase, has emerged. As evidenced by T-jump[22] and ultraviolet (UV) electronic absorption[23] the a−o tautomer appeared as the most stable form in neutral aqueous solutions. IR spectra recorded in rare gas matrixes at low temperature[4−9] showed the coexistence of both the a−o and the amino−hydroxy (a−h) forms, with an [a−h]/[a−o] equilibrium ratio close to 0.5 in normal conditions. UV irradiation of matrix-isolated cytosine leads to a considerable decrease of the a−o form characteristic IR bands along with an enhancement of those assigned to the a−h tautomer.[7] Other possible tautomers of cytosine as the imino−oxo (i−o) and the imino−hydroxy (i−h) forms have been shown to form minor populations.[5,7] Only in particular cases, have NMR data shown that the formation of C−A base pairs in double-helical nucleic acids is favored by the i−h form of cytosine.[24]

The second attractive problem concerning the structural properties of the major tautomers (a−o and a−h) is the fact that all quantum mechanical calculations using extended basis sets, i.e., double-ζ enlarged by at least d-polarization orbitals for heavy atoms, especially on nitrogens, gave rise to an amino group which is not coplanar with the cytosine ring.[18−20,25−29] The pyramidalization of the amino group appears to be a common property of all nucleic acid bases containing this chemical group.[25,28−30] The stabilization of the nonplanar form of cytosine has been attributed to the interplay between the delocalization of the $N(H_2)$ lone pair into the ring (see ref 26 and references therein) and the valence state of the nitrogen atom. This effect is reinforced by the effective hybridization of the nitrogen s and p orbitals in the presence of d orbital functions.

Thus, here in order to interpret our new experimental results, we have resorted to ab initio calculations including correlation effect. To our knowledge, no attempt has been made to obtain a MP2 force field for cytosine after full geometry optimization at this level. All the previous investigations at MP2 level[25−28,31−35] have only been devoted to geometrical optimization. The combination of the used experimental and theoretical results allows the crucial problems related to cytosine, e.g. tautomerism and amino group nonplanarity, to be discussed.

II. Materials and Methods

Cytosine (C) powder was purchased from Sigma-Aldrich and used as supplied. The experimental method of the labile hydrogen−deuterium exchange to obtain C-d_3 samples, was similar to that used for the other two pyrimidine bases.[1,2] Aqueous solutions of C and C-d_3 were used at approximately 6×10^{-2} M.

NIS spectra of C and C-d_3 powder samples were obtained at $T = 15$ K on the time-focused crystal analyzer spectrometer TFXA located at the ISIS pulsed neutron source of the Rutherford Appleton Laboratory (RAL), United Kingdom. The whole experimental setup, the data acquisition system, and data treatment for Raman and IR spectra have been described elsewhere.[1]

Ab initio quantum mechanical calculations using GAUSSIAN92 package[36] were carried out to estimate the structural parameters of the cytosine tautomers and the fundamental vibrations of the isolated molecule. The level of calculation was self-consistent field (SCF) plus second-order Møller−Plesset perturbation treatment (MP2). All geometry optimizations were performed by using *very tight* option. To estimate the harmonic wavenumbers and atomic displacements, only the second-order derivatives of the electronic energy with respect to the atomic Cartesian coordinates were computed. Double-ζ Gaussian basis sets, i.e., 6-31G and D95V, alone or enlarged by nonstandard d-polarization functions,[1,2,37] hereafter referred to as 6-31G(*) and D95V(*), were used. Moreover, in order to check the dependence of the calculated results (molecular geometry, wavenumbers and NIS intensities) on the addition of p-polarization functions on hydrogens, we have also used the standard basis functions 6-31G** and D95V**.

Postprocessing of the vibrational modes was done on the basis of the Wilson GF-method[38] by using a homemade program (BORNS). First, F_X matrix (interatomic Cartesian force constants) was converted to F_R matrix (interatomic force constants expressed in terms of internal coordinates). Second, the redundant internal coordinates were treated. Harmonic wavenumbers as well as potential energy distribution (PED) associated to them (by taking into account the whole F matrix elements) were calculated for assigning the normal modes. Ab initio force field as output from quantum mechanical calculations was used for assigning the experimental modes observed in gas and in rare gas matrixes. Then, this force field was scaled to improve the agreement between the calculated wavenumbers and those observed in condensed phase. NIS intensities were calculated by the scaled force field, following the method described in our previous papers.[1,2]

III. Results and Discussion

Tautomerism and NH_2 Group Nonplanarity of Cytosine. During the geometry optimization of different cytosine tautomers (a−o, a−h cis and a−h trans, Figure 1) several interesting problems emerged which can be summarized as below:

Due to the presence of the amino group in the 4 position of these tautomers, it appears that the optimized geometry depends strongly on the basis sets used. Whatever the starting conformation was (planar or nonplanar), the full geometry optimization either at SCF or MP2 level without d-polarization functions (6-31G or D95V sets) leads to a perfect planar cytosine. No imaginary vibrational wavenumber has been found in these calculations. The geometry parameters of this step have not been reported here, because of the insufficiency of these basis sets.

With the basis sets containing d-polarization functions (6-31G(*), D95V(*), 6-31G** or D95V**) the calculations either at SCF or MP2 level converge to different geometries depending on the starting point. With an initial planar molecule, the optimization converges to a final planar one. The vibrational calculations with this planar geometry contains one imaginary wavenumber (transition state). On the contrary, with a starting

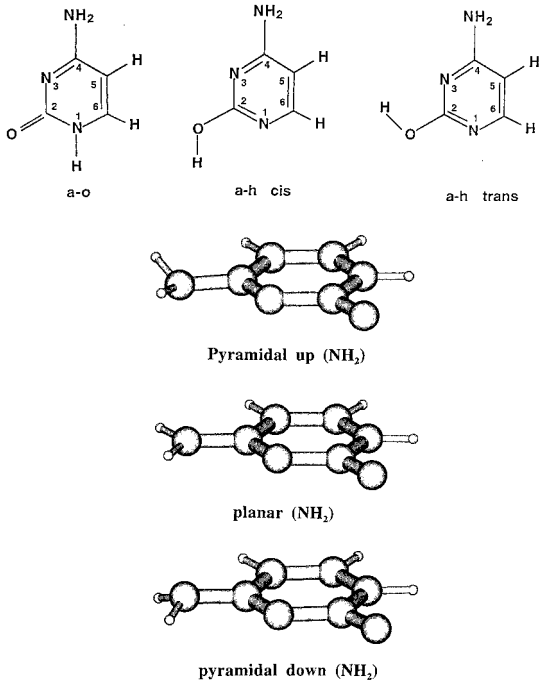

Figure 1. Top: atom numbering and chemical structure of the three different and energetically most stable cytosine tautomers considered in this work: a−o (amino−oxo), a−h (amino−hydroxy). Cis and trans refer to two different orientations of the O−H group in the 2-position of the a−h cytosine tautomer. Bottom: 3D representations of the cytosine a−o tautomer with different amino group conformations (pyramidal up, planar and pyramidal down conformations).

geometry containing a slightly distorted NH_2, the optimization leads to a distinct pyramidalization of this group (Figure 1). In this case, vibrational calculations give no imaginary wavenumber. We have checked by further calculations that the inversion of the NH_2 pyramid does not correspond to a double minima. This inversion of the pyramid creates a local minimum which differs from the lowest one by about 12 kcal/mol (Table 1). It is interesting to mention that the potential barrier separating the transition state (planar) from the lowest minimum (pyramidal) is very low (about 0.3 kcal/mol). On the other hand, the orientation of the NH_2 group relative to the lowest minimum depends on the extension of the basis sets used. Indeed, the lowest energy geometry obtained at MP2 level has a NH_2 group orientation which points upward the cytosine ring (pyramidal up geometry), or downward (pyramidal down geometry) with either 6-31G(*) (D95V(*)) or with 6-31G** (D95V**) basis sets, respectively (Table 1, Figure 1).

As the energy values calculated at MP2/6-31G(*) level show (Table 1), the a−h cis form has the lowest energy among the three tautomers considered in our theoretical work. The a−o tautomer with an energy difference of $\Delta E = 0.04$ kcal/mol with respect to the lowest energy is located just above. Finally, the a−h trans is the highest energy tautomer, with a $\Delta E = 0.73$ kcal/mol, with respect to that of the a−h cis form. Thus, the energy range of these tautomers in vacuo is less than 0.75 kcal/mol. Similar results have been obtained previously by single-point MP2/6-31G** calculations.[33] Full geometry optimization with DFT/6-31G** is also consistent with this small amount of energy separating these tautomers.[19,20]

The experimental rotational constants (*A*, *B*, and *C*) as derived from the microwave experiments[39] for the a−o and a−h tautomers are compared with those calculated at MP2 level in Table 1. Addition of d-polarization functions to the basis sets, leading to the pyramidal NH_2 group, improves considerably the calculated values. However, the small difference between the 6-31G(*) and the 6-31G** rotational constants does not allow to precisely orient the NH_2 group (upward or downward) with respect to the cytosine ring. As far as the a−h tautomer is concerned, the only experimental rotational constants available have been evaluated for the cis form.[39] The value of the constant *A* differs considerably in going from the a−o to a−h tautomers, and this confirms the reliability of the theoretical level used for calculating the rotational constants of different cytosine tautomers.

For the sake of brevity, only the lowest energy geometrical parameters calculated at MP2/6-31G(*) and MP2/6-31G** are reported in Table 2. In both cases, disregarding the amino group of the molecule, the cytosine ring is only slightly distorted, as can be verified by the values of the calculated torsional angles defined about the ring skeletal bonds. The calculated bond lengths and valence angles have been compared to those deduced from X-ray diffraction patterns of anhydrous cytosine.[21]

Calculated results (geometry, vibrational wavenumbers, and NIS intensities) obtained with D95V(*) and D95V** basis sets (not reported here) are similar to those related to the 6-31G(*) and 6-31G** functions, respectively.

Finally, it should be mentioned here that no clear experimental evidence on the pyramidalization of the NH2 group can be directly extracted from the observed vibrational spectra. We have verified that on the basis of the simple comparison between the theoretical isotopic shifts of the NH_2 group characteristic modes obtained for planar and pyramidal geometries, and those observed in Ar-matrix-isolated cytosine, one cannot promote the pyramidal or the planar conformation of this chemical group in cytosine. However, this particularity of amino group has also been encountered in the calculated results performed by extended basis functions on other nucleic acid bases as guanine and adenine.[25,28−30]

TABLE 1: Electronic (E_e), Zero Point Vibrational (ZPV) Energies and Rotational Constants of Different Cytosine Tautomers as Calculated at the MP2 Level[a]

				rotational constants/MHz								
	NH2 group	E_e(ZPV)[c]/au		MP2/6-31G(*)			MP2/6-31G**			exp.[d]		
tautomer	orientation[e]	MP2/6-31G(*)	MP2/6-31G**	A	B	C	A	B	C	A	B	C
a−o	pyramidal (up)	−393.77153 (0.09981)	−393.79318	3854.98	2008.67	1321.58	3856.65	2008.24	1321.53	3871.55	2024.93	1330.31
	planar[b]	−393.77098	−393.81265	3858.53	2011.37	1322.16	3866.71	2013.30	1323.95			
	pyramidal (down)	−393.75259	−393.81307 (0.10040)	3848.61	2006.11	1319.76	3863.38	2010.93	1323.45			
a−h cis	pyramidal (up)	−393.67179 (0.09992)		3928.60	1997.33	1325.28				3951.72	2008.89	1332.46
a−h trans	pyramidal (up)	−393.67063 (0.09988)		3867.90	2013.89	1325.54						

[a] See Figure 1 for the definition of the a−o and a−h cis and trans tautomers. [b] Transition state. One imaginary wavenumber is calculated for this geometry. [c] ZPV is only indicated for the calculated geometry corresponding to the lowest energy minima. [d] Experimental values obtained from the microwave experiments.[39] [e] Pyramidal and planar geometries are displayed in Figure 1.

TABLE 2: Calculated Geometry Parameters of the Cytosine Tautomers (calc)[a]

		calc (a−o)		calc (a−h)	
	exp	MP2/6-31G(*) pyramidal up	MP2/6-31G** pyramidal down	cis MP2/6-31G(*) pyramidal up	trans MP2/6-31G(*) pyramidal up
bonds (Å)					
N1−H	0.88	1.014	1.009		
O2−H				0.977	0.976
C2═O2		1.226	1.227		
	1.234				
C2−O2				1.352	1.352
C4−N4	1.330	1.371	1.368	1.374	1.376
N4−H4′	0.87	1.011	1.005	1.011	1.012
N4−H4″	0.86	1.014	1.008	1.013	1.013
C5−H	0.87	1.084	1.078	1.086	1.086
C6−H	1.01	1.087	1.081	1.089	1.089
N1−C2		1.418	1.417		
	1.374				
N1═C2				1.337	1.333
C2−N3	1.364	1.381	1.381	1.335	1.340
N3═C4	1.337	1.318	1.318	1.341	1.343
C4−C5	1.424	1.437	1.436	1.409	1.405
C5═C6	1.342	1.360	1.359	1.383	1.386
C6−N1	1.357	1.358	1.357	1.348	1.346
angles (deg)					
C2−N1−H	120.0	114.656	114.706		
C6−N1−H	117.0	121.339	121.246		
C2−O2−H				105.049	105.476
N1−C2═O2		118.834	118.859		
	119.8				
N1═C2−O2				116.829	115.348
N3−C2═O2		125.244	125.265		
	122.2				
N3−C2−O2				114.720	116.255
N3═C4−N4	118.2	116.755	116.817	116.068	116.303
C5−C4−N4	119.9	118.642	118.602	121.982	122.413
C4−N4−H4′	123.0	118.501	118.704	117.517	117.103
C4−N4−H4″	124.0	114.409	114.661	114.163	114.277
H4′-N4−H4″	114.0	116.023	116.517	115.211	114.817
C4−C5−H	123.0	122.458	122.409	121.994	121.955
C6═C5−H	119.0	121.563	121.573	121.677	121.702
C5═C6−H	122.0	123.536	123.564	120.901	120.463
N1−C6−H	118.0	116.870	116.862	115.866	115.608
N1C2N3	118.1	115.920	115.875	128.450	128.396
C2N3C4	119.9	119.956	119.959	115.582	116.158
N3═C4C5	122.0	124.548	124.527	121.886	121.217
C4C5═C6	117.3	115.973	116.015	116.329	116.342
C5═C6N1	120.1	119.594	119.574	123.233	123.930
torsions (deg)					
H−C6−N1−H		0.107	−0.110		
C5═C6−N1−H		179.905	−179.895		
N3−C2−N1−H		179.602	−179.628		
O2═C2−N1−H		0.028	−0.028		
N1═C2−O2−H				−0.161	−179.665
N3−C2−O2−H				−179.820	0.647
C6−N1═C2−O2				−179.931	179.952
C4═N3−C2−O2				−179.923	−179.808
C6−N1−C2═O2		179.888	−179.929		
C4═N3−C2═O2		−179.737	179.780		
N3═C4−N4−H′		156.541	−157.67	156.245	156.460
N3═C4−N4−H″		13.920	−13.427	16.770	18.030
C5−C4−N4−H′		−26.028	24.868	−26.612	−26.487
C5−C4−N4−H″		−168.649	169.111	−166.087	−164.912
C2−N3═C4−N4		176.780	−176.892	176.998	176.939
C6═C5−C4−N4		−177.240	177.338	−177.216	−177.209
N4−C4−C5−H		1.909	−2.008	2.484	2.575
N3═C4−C5−H		179.123	−179.259	179.461	179.485
N1−C6═C5−H		−178.928	179.074	−179.309	−179.330
C4−C5═C6−H		180.000	179.960	−179.761	−179.679
H−C5═C6−H		0.850	−0.690	0.539	0.536
C2−N1−C6−H		−179.729	179.777		
C2═N1−C6−H				180.000	−179.665
C6−N1−C2−N3		−0.549	0.477		
C6−N1═C2−N3				−0.325	−0.410
N1−C2−N3═C4		0.731	−0.652		
N1═C2−N3═C4				0.465	0.552
C2−N3═C4−C5		−0.482	0.403	−0.147	−0.152
N3═C4−C5═C6		−0.020	0.086	−0.238	−0.298
C4−C5═C6−N1		0.229	−0.278	0.392	0.455
C5═C6−N1−C2		0.059	0.000		
C5═C6−N1═C2				−0.136	−0.130

[a] The experimental results (exp) are from the X-ray diffraction patterns of anhydrous[21] cytosine. For the definition of a−o, a−h cis and trans tautomers as well as the pyramidal and planar conformations of the amino group, see Figure 1.

TABLE 3: Vibrational Wavenumbers Observed in the Gas Phase, in Ar Matrix Isolated Cytosine, and Calculated at the MP2/6-31G(*) Level[a] (Unscaled Force Field)

exp				calc		
gas Raman (ref 3)	IR (ref 3)	matrix IR (ref 5)	IR (ref 7)	assignment		tautomer
				3731 (*7*)		a−o
3365	3360	3565 (*12*)	3564 (*18*)	3726 (*6*)	NH2 asym st	a−h cis
				3724 (*7*)		a−h trans
				3720 (*13*)		a−h trans
		3591 (*21*)	3600 (*18*)		O−H	
				3714 (*16*)		a−h cis
3253	3249	3471 (*16*)	3471 (*12*)	3632 (*14*)	N1−H	a−o
		3441 (*24*)		3600 (*13*)		a−o
3206	3203	3445 (*15*)	3445 (*33*)	3599 (*11*)	NH2 sym st	a−h cis
				3599 (*10*)		a−h trans
				3281 (*0*)	C5−H; C6−H	a−o
3112	3107			3262 (*1*)	C5−H	a−h cis
				3262 (*0*)	C5−H	a−h trans
				3253 (*0*)	C6−H; C5−H	a−o
3100	3096			3224 (*3*)	C6−H	a−h cis
				3221 (*3*)	C6−H	a−h trans
1690	1684	1720 (*100*)	1718 (*100*)	1837 (*100*)	C2═O2	a−o
1666	1661	1656 (*49*)	1656 (*63*)	1743 (*55*)	C5═C6; N3═C4	a−o
1621	1619	1623 (*69*)	1622 (*90*)	1713 (*71*)	C4−N4; NH2-sciss; C5═C6	a−h cis
				1710 (*67*)	NH2-sciss; C4−N4; C4−C5	a−h trans
		1595 (*8*)	1599 (*68*)	1680 (*14*)	NH2-sciss; C4−N4; CNH asym bend	a−o
				1680 (*14*)	NH2-sciss; C4−N4; N1═C2	a−h trans
		1600 (*7*)	1589 (*25*)	1678 (*2*)	NH2-sciss; C5═C6; C2−N3	a−h cis
				1668 (*23*)	C5═C6; C6−N1; N3═C4	a−h trans
		1569 (*21*)	1570 (*14*)			
1537	1540			1664 (*33*)	C4−C5; N3═C4; C6−N1	a−h cis
		1539 (*11*)	1539 (*21*)			
				1625 (*23*)	N3═C4; C4−C5; C5═C6	a−o
				1562 (*4*)	C4−N4; C2−N3; N3═C4	a−h cis
1498	1500	1496 (*3*)	1496 (*14*)			
				1558 (*8*)	C4−N4; N3═C4; N1═C2	a−h trans
		1475 (*22*)	1475 (*38*)	1553 (*16*)	C4−N4; C6−N1; N3═C4	a−o
				1513 (*58*)	C2−O2; C2−N3; N1C6H	a−h cis
		1439 (*67*)	1439 (*52*)			
				1504 (*61*)	C2−O2; C2−N3; N1═C2−N3	a−h trans
1460	1458	1422 (*6*)	1423 (*26*)	1475 (*11*)	C2N1H; C6N1H; C6−N1	a−o
		1379 (*9*)	1380 (*17*)	1440 (*11*)	C5═C6; C2−O2; C4−N4	a−h cis
				1437 (*6*)	C5═C6; C4−N4; N3═C4	a−h trans
			1374 (*5*)	1392 (*2*)	C5C6H; N1C6H; C4−N4	a−h trans
1360	1351	1337 (*5*)	1338 (*11*)	1387 (*7*)	C4−N4; C4C5H; C5C6H	a−o
		1320 (*13*)	1333 (*15*)	1384 (*20*)	C4−N4; CNH asym bend; N1C6H	a−h cis
				1357 (*0*)	C6−N1; N3═C4; N1═C2	a−h trans
			1304 (*3*)			
				1349 (*0*)	C6−N1; C5C6H; N1C6H	a−h cis
1287	1290			1293 (*3*)	C2−N3; C4−N4; N1−C2	a−o
				1281 (*11*)	C2O2H; C2−N3	a−h cis
		1257 (*2*)	1244 (*4*)			
				1272 (*31*)	C2O2H; C2−N3; C2−O2	a−h trans
1218	1220	1244 (*4*)	1224 (9)	1249 (*6*)	N1C6H; C6−N1; C6N1H	a−o
		1193 (*14*)	1211 (*6*)	1160 (*5*)	CNH asym bend; C4C5H	a−h cis
		1192 (*8*)		1158 (*2*)	C4C5H; CNH asym bend	a−h trans
		1124 (*4*)	1125 (*11*)	1152 (*0*)	C6C5H; C6−N1; CNH asym bend	a−o
1119	1118	1110 (*3*)	1108 (*3*)	1133 (*8*)	CNH asym bend; N1−C2	a−o
		1090 (*6*)	1091 (*6*)	1129 (*6*)	CNH asym bend; C2O2H; C6−N1	a−h cis
		1087 (*1*)				
		1084 (*6*)	1083 (*12*)	1127 (*4*)	C5═C6; N3═C4; C6−N1	a−h trans
				1032 (*0*)	C4−C5; C2−O2	a−h trans
1042	1040					
				1029 (*3*)	C4−C5; CNH asym bend; C2−O2	a−h cis
				1007 (*1*)		a−h cis
1004	1000	980 (*1*)	988 (*1*)		N1═C2; C2−N3; C2N1C6	
				1006 (*0*)		a−h trans
			981 (*6*)	1001 (*0*)	C4−C5; C4C5C6	a−o
				968 (*0*)		a−h trans
968	970	955 (*1*)	955 (*1*)		ω(C6H); τ(C5C6)	
				961 (*1*)		a−h cis
				950 (*0*)	N1−C2; C2N1C6	a−o
823	827	818 (*2*)	818 (*8*)	918 (*0*)	ω(C6H); τ(C5C6)	a−o
				801 (*3*)		a−h cis
		807 (*7*)	807 (16)		ω(C5H); C5C6N1; C2−O2	
				800 (*3*)		a−h trans
				793 (*2*)		a−h cis
799	797	782 (*2*)	794 (*1*)		ω(C5H); τ(C4C5)	
				793 (*2*)		a−h trans

TABLE 3 (Continued)

exp				calc		
gas Raman (ref 3)	IR (ref 3)	matrix IR (ref 5)	IR (ref 7)	assignment		tautomer
777				782 (*0*)	N1−C2; C4−C5; C4−N4	a−o
				773 (*6*)		a−h cis
	780	781 (*5*)	782 (*15*)	ω(C2O2); ω(C4N4)		
				772 (*6*)		a−h trans
760	758	751 (*0*)	767 (*2*)	755 (*2*)	ω(C2O2)	a−o
			747 (*4*)	738 (*11*)	ω(C5H); ω(C6H); τ(C4−C5)	a−o
		747 (*1*)	717 (*8*)	710 (*2*)	ω(C4N4); ω(C5H)	a−o
692	690			696 (*2*)		a−h cis
		717 (*2*)	710 (*3*)	ω(C4N4); ω(C2O2)		
				692 (*2*)		a−h trans
626	630		636 (*12*)	647 (*17*)	ω(N1H); τ(C6N1); ω(C4N4)	a−o
				609 (*2*)		a−h trans
		601 (*1*)	600 (*3*)	C2N3C4; N1C2N3, C2−O2		
				603 (*1*)		a−h cis
		614 (*5*)	613 (*18*)	575 (*1*)	C2N3C4; N1C2N3; N3C2O2	a−o
590	588		576 (*12*)	572 (*31*)	τ(C2O2)	a−h cis
				568 (*30*)	τ(CNH2); C2N1C6; C2−O2	a−h trans
571	570	569 (*1*)	568 (*8*)	568 (*6*)	τ(C2O2); N3C4C5; C2N1C6	a−h cis
560	561	537 (*1*)	558 (*4*)	563 (*14*)	τ(CNH2); τ(C4N4)	a−o
				553 (*19*)	τ(C2O2)	a−h trans
546	544	532 (*1*)	535 (*16*)	537 (*5*)	N3C4C5; τ(C4N4); C4−N4	a−o
519	522	520 (*8*)	520 (*57*)	537 (*29*)	τ(CNH2); C4−N4; NH2-sciss	a−h cis
				533 (*25*)	τ(CNH2); CNH asym bend; C4−N4	a−h trans
		507 (*4*)	507 (*22*)	530 (*4*)		a−o
				504 (*2*)	N3C2O2; N3C4N4; C5C4N4	a−h trans
			498 (*11*)	503 (*3*)		a−h cis
				456 (*6*)		a−h trans
453	450	451 (*1*)	443 (*3*)	τ(C6N1); ω(C2O); ω(C4N4)		
				454 (*6*)		a−h cis
				414 (*13*)		a−h cis
436	431			τ(C4N4); τ(CNH2)		
				411 (*9*)		a−h trans
392	390	409 (*2*)	400 (*4*)	408 (*36*)	τ(CNH2); τ(C4N4)	a−o
				381 (*2*)	ω(C5H); ω(N1H); τ(C4N4);	a−o
363	360	360 (*3*)		357 (*3*)	C5C4N4; N3C4N4; CNH asym bend	a−o
				342 (*2*)		a−h cis
		350 (*1*)	343 (*5*)	C5C4N4; N3C4N4; N3C2O2		
				340 (*0*)		a−h trans
				226 (*0*)		a−h trans
			236 (*3*)	222 (*2*)	ring tors	a−h cis
				204 (*2*)		a−o
				189 (*0*)		a−h cis
				ring tors; ω(C5H)		
				187 (*0*)		a−h trans
				131 (*0*)	ω(N1H); ring tors	a−o

[a] See also Figure 1 for the definition of the tautomers. In parentheses are reported the experimental and calculated IR intensities in relative values as normalized to 100 with respect to the most intense peak in all cases. The most intense calculated peak at 1837 cm^{-1} (a−o tautomer) has an absolute IR intensity value of 653.8 km/mol.

Thus, other spectroscopic methods as NMR or others (for review see the paper published by J. M. Lehn in 1970)[40] might be more adequate to elucidate this particular behavior of the amino group in nucleic acid bases.

Choice of Internal Coordinates for Vibrational Mode Assignments. Full geometry optimization gives an overall C_1 symmetry for all the three tautomers studied here (Table 2). Thus, no *F* matrix partitioning into in-plane and out-of-plane submatrixes is justified. In order to analyze accurately the vibrational modes in each tautomer, *natural* internal coordinates have been used. We have considered all bond stretching coordinates (ν), angular bending coordinates (δ), wagging coordinates (ω) for N1−H, C2O, and C4−N4 groups, average torsional coordinates (τ) about the ring skeletal and the C4−N4 extracyclic bond, and finally one additional *improper torsion* for describing completely the amino group vibrational motion. It should be mentioned that for N1−H, C2O, and C4−N4 bond, one can use the term *wagging*, because each of these groups along with its two nearest neighbors in the ring lie approximately in the same plane (see the values of interatomic torsions in Table 2), whereas this becomes impossible for the NH_2 group, due to its important distortion predicted by MP2 calculations using d-polarization functions (see above). For this reason, we use here the term *improper torsion* which reflects a particular torsional angle defined as τ(C4−N4−H4′H4″), referred to as $\tau(CNH_2)$ in the text. This notation should not be confused with τ(C4N4) which designates the internal torsional angle defined about C4−N4 bond.

Comparison of the Unscaled Vibrational Wavenumbers and IR Intensities from the a−o and a−h Tautomers with Those Experimentally Observed in Gas Phase and Rare Gas Matrixes. Taking into account the small energy differences separating the different tautomeric forms (see above), one should conclude that in gas or even in rare gas matrixes all the three forms (a−o, a−h cis, a−h trans) should coexist and must be considered together in assigning the observed vibrational modes.

As shown in Table 2, the addition of the p-polarization functions to the hydrogen atoms (6-31G** or D95V** basis sets) leads to the shortening of the N−H and C−H bond lengths and consequently to the increasing of their bond-stretching wavenumbers. Thus, in the following we focus our discussion on the vibrational wavenumbers calculated at MP2/6-31G(*)

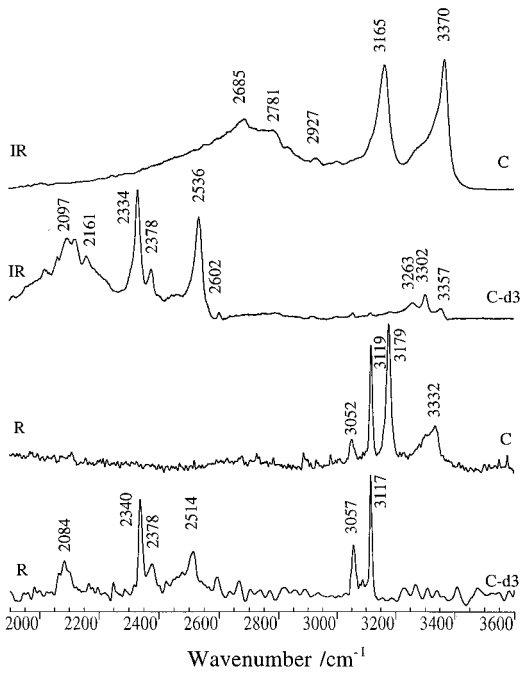

Figure 2. Low-tempertaure ($T = 15$ K) infrared (IR, up) and Raman (R, down) spectra of polycrystalline samples of cytosine (C) and its N-deuterated species (C-d_3) in the 1900−3600 cm^{-1} spectral region.

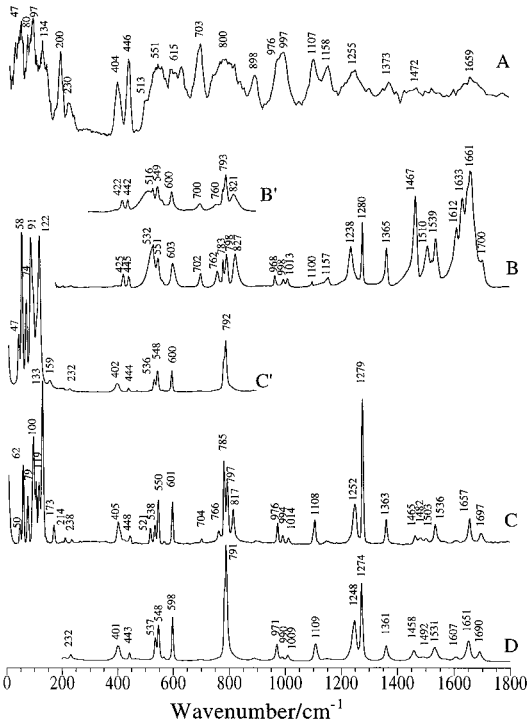

Figure 3. Vibrational spectra of polycrystalline samples of cytosine (C) in the spectral region below 1800 cm^{-1}. (A) NIS spectrum recorded at $T = 15$ K, (B) IR spectrum at $T = 15$ K, (B′) low-wavenumber IR spectrum at room temperature, (C) Raman spectrum ($\lambda_{exc} = 514.5$ nm) at $T = 15$ K, (C′) low-wavenumber Raman spectrum ($\lambda_{exc} = 514.5$ nm) at room temperature, (D) FT-Raman spectrum ($\lambda_{exc} = 1.06$ μm) at room temperature.

level with the geometrical parameters corresponding to the lowest energy minimum (Tables 1 and 2).

In Table 3, the whole set of calculated vibrational modes and their assignments have been used in order to analyze the gas-phase Raman and IR spectra,[3] as well as the low-temperature IR spectra of argon matrixes.[5,7] To propose more reliable assignments, both experimental (Ar matrix) and calculated IR

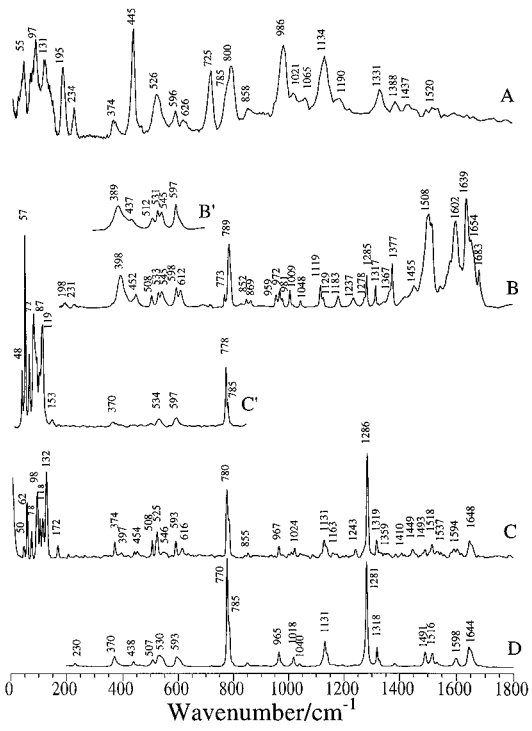

Figure 4. Same as Figure 3 but for the vibrational spectra of polycrystalline samples of N-deuterated cytosine (C-d_3) in the spectral region below 1800 cm^{-1}. See also the caption of Figure 3.

intensities have been used in relative values (normalized to 100 for the most intense peak).

Vibrational Spectra of C and C-d_3 Species in the Condensed Phase Obtained by NIS and Optical Spectroscopies. The geometry parameters obtained by MP2 calculations for the a−o tautomer are in good agreement with those observed on the X-ray diffraction patterns of anhydrous cytosine crystal. See for instance the comparison between the experimental and calculated results for the C2═O2 bond length and the N1−C2═O2, N3−C2═O2, and C5═C6−N1 angles (Table 2). Moreover, a careful examination of the various spectra recorded in the solid phase allows to conclude that these spectra are not constituted by the superposition of several spectra arising from different cytosine tautomers. Therefore, we assume that the a−o tautomer is the predominant form in solid phase.

Low-temperature Raman and IR spectra recorded in solid phase are shown in Figure 2 in the spectral region above 2000 cm^{-1} for both C and C-d_3 species.

Figures 3A and 4A show the NIS spectra of C and C-d_3 species, respectively, in the spectral region below 1800 cm^{-1}. In the same figures are also reported the IR absorption (Figures 3B and 4B) and Raman spectra ($\lambda_{exc} = 514.5$ nm, Figures 3C and 4C) recorded at 15 K as well as the FT-Raman spectra obtained with $\lambda_{exc} = 1.06$ μm (Figures 3D, 4D) at room temperature. In order to follow the evolution of the vibrational modes as a function of temperature, the spectral regions showing considerable changes as the low wavenumber spectral ranges recorded at room temperature have been shown in traces B′ and C′ of Figures 3 and 4, for IR and Raman ($\lambda_{exc} = 514.5$ nm) spectra, respectively. It should be mentioned that the variation of the high wavenumber region spectra (above 2000 cm^{-1}) in going from 20 K to room temperature is negligible (spectra not shown).

Room temperature Raman spectra obtained with three different excitations ($\lambda_{exc} = 257$ nm, 514.5 nm, and 1.06 μm) as well as FT-IR spectra of aqueous solutions in the spectral region below 1800 cm^{-1} are shown in Figures 5 and 6 for C and C-d_3 species, respectively.

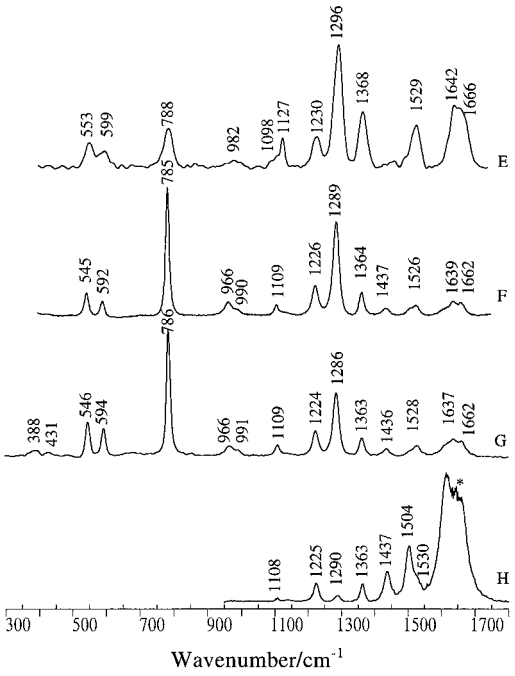

Figure 5. Vibrational spectra of aqueous solutions of cytosine at room temperature in the spectral region below 1800 cm^{-1}. (E, F, G) Raman spectra obtained by the excitation wavelengths $\lambda_{exc} = 257$ nm, 514.5 nm, and 1.06 μm, respectively. (H) FT-IR spectrum after subtraction of the FT-IR spectrum of water. In the domain indicated by an asterisk, the peaks resulting from the subtraction of the water absorption band are not reliable.

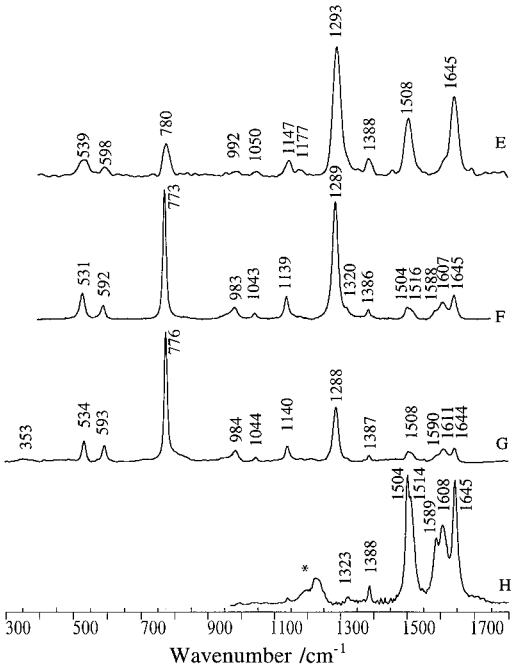

Figure 6. Same as Figure 5 but for the vibrational spectra in D_2O buffer (C-d_3 species).

The whole set of the observed vibrational wavenumbers in solid phase and solution are reported in Tables 4 and 5 for C and C-d_3 species, respectively.

Comparison between the Scaled Vibrational Wavenumbers and Those Experimentally Observed in the Condensed Phase. Calculation of NIS Intensities. In order to improve the agreement between the calculated wavenumbers and the observed ones in the solid state and aqueous solutions as well as to get a good fit with the observed NIS intensities, the molecular force field of the a−o tautomer (dominant form in condensed phase) has been scaled.

Our scaling procedure of the ab initio force field has been described in detail in the two previously published papers on the pyrimidine nucleic acid bases.[1,2] Bond-stretching and angular-bending force constants have been scaled by using the Pulay method.[41] In order to decrease the number of scaling factors, only one scaling factor has been assigned to a group of similar internal coordinates (Table 6). As far as the ring internal coordinates are concerned, the values of the scaling factors are comparable with those used in the case of uracil and thymine for similar internal coordinates. The majority of these values are in the range 0.82−1.22 (Table 6), which is acceptable for MP2 harmonic force fields. The only exceptions concern the N1−H and C2═O stretching coordinates associated to rather low value scaling factors, respectively, 0.62 and 0.79. This is due to the fact that N1−H and C2═O groups are involved in the intermolecular hydrogen bond network in solid phase. This fact is partially revealed by the N1−H bond stretching motion which gives rise to a broad band located around 2700 cm^{-1} (Figure 2, Table 4). Again the NH_2 symmetric and antisymmetric bond stretchings (sensitive to H bondings) have been scaled with a factor of 0.81 which was necessary to reproduce the observed vibrational wavenumbers of these modes. The other scaling factors related to the NH_2 group internal coordinates are in the 0.95−1.25 range, comparable with the above-mentioned ring scaling factors. Torsional and wagging motions have been scaled separately by using the previously reported method.[1,2] In this case the diagonal force constants have been left unchanged with respect to the unscaled theoretical values and only the off-diagonal terms have been modified by a least-squares method in order to improve both vibrational wavenumbers and NIS intensities corresponding to the torsional and wagging modes.

Figures 7 and 8 show the comparison between the calculated NIS intensities (up to the third order)[1,2] and the experimental ones in the region below 2000 cm^{-1} for C and C-d_3 species, respectively.

Figure 9 shows the picture of the scaled harmonic normal modes located in the spectral region below 2000 cm^{-1}.

Overview of the Vibrational Modes Observed in the Condensed Phase and Their Assignments. In the 1600−1750 cm^{-1} spectral region, the cytosine NIS spectrum shows only a broad and nonresolved band peaking around 1659 cm^{-1} which should mainly correspond to NH_2 scissoring mode (Figure 3A). This band disappears upon selective deuteration (Figure 4A). IR spectra provide more details on the above spectral range (Figures 3 and 4). The shift upon deuteration of several intense IR bands in the 1720−1600 cm^{-1} range allows us to confirm the calculated results (Table 4) showing that NH motions are involved in the modes calculated in this region (Figure 9). The intense IR band observed at 1467 cm^{-1} in the solid phase and at 1437 cm^{-1} in solution, both vanishing upon deuteration, corresponds to the calculated mode at 1467 cm^{-1} in which N1 atom takes mainly part. The 30 cm^{-1} downshift of this mode from the solid state to solution reveals the sensitivity of the imino group at the N1 site to the molecular environment. In NIS (Figure 3A) a broad and intense band is observed around 1255 cm^{-1} which also vanishes upon deuteration. This mode corresponds to that calculated at 1232 cm^{-1} involving mainly N1C6 bond stretch and δ(N1H) bending motions (Figure 9). Raman spectra also confirm this assignment because the bands of medium intensity around 1252 cm^{-1} in solid (Figure 3) and 1224 cm^{-1} in solution disappear by selective deuteration. In contrast, the intense Raman and IR bands located around 1280 and 1286 cm^{-1} in solid and aqueous phases, respectively, are less sensitive to N-deuteration. They are satisfactorily assigned by the calculated mode at 1277 cm^{-1} involving mainly ring

TABLE 4: Comparison between Experimental (Condensed Phase) and Scaled (Calculated with MP2/6-31G(*) Force Field) Vibrational Wavenumbers (cm^{-1}) for Cytosine a−o Tautomer (Figure 1)[a]

exp (NIS)	exp (Raman)		exp (IR)			
A	C	G	B	H	calc	potential energy distribution (%)
	3332(br,m)		3370(s)		3358	NH2 asym st (99)
	3179(s)		3165(s)		3240	NH2 sym st (99)
	3119(s)				3113	C5−H (89); C6−H (10)
	3052(w)				3087	C6−H (89); C5−H (10)
	2700(br,w)		2700(br)		2868	N1−H (99)
	1697(w)		1700(sh)		1710	N3═C4 (19); C2═O2 (17); C6−N1 (12); C2−N3 (10); C5═C6 (8)
		1662(sh)	1661(vs)	1660(sh)	1696	NH2-sciss (30); C2═O2 (19); CNH asym bend (12); C4−N4 (10)
1659(br, m)						
	1657(m)	1637(w)	1633(s)	1616(vs)	1651	NH2-sciss (28); C2═O2 (15); CNH asym bend (13); C5═C6 (9)
			1612(s)		1608	C4−C5 (29); C5═C6 (15); N3═C4 (9); CNH asym bend (9)
	1536(w)	1528(w)	1539(m)	1530(sh)	1528	C4−N4 (15); N3═C4 (14); C6−N1 (14); C5═C6H (11); C6−N1 (9); N3═C4C5 (7)
			1510(m)	1504(s)		
1472(br, w)	1465(w)	1436(w)	1467(s)	1437(m)	1467	C6−N1 (17); C2N1H (15); C6N1H (15); C2═O2 (14)
1373(br, m)	1363(m)	1363(m)	1365(m)	1363(m)	1358	C5═C6H (13); C4C5H (11); C5═C6 (11); C6═C5H (8); N1C6H (8); C4−N4 (7)
	1279(vs)	1286(s)	1280(s)	1290(w)	1277	C2−N3 (30); C4−N4 (20); N3═C4 (10); C2═O2 (8); N1−C2 (7)
1255(br, s)						
	1252(m)	1224(m)	1238(m)	1225(m)	1232	C6−N1 (20); C2N1H (15); C6N1H (15); N1C6H (15); N1−C2 (9)
1158(s)			1157(vw)		1170	CNH asym bend (44); N1−C2 (8); C2N3═C4 (6)
1107(s)	1108(m)	1109(w)	1100(vw)	1108(vw)	1113	C6═C5H (31); C5═C6 (18); C4C5H (14); C6−N1 (11)
	1014(w)		1013(vw)		1028	C4−C5 (21); C4C5═C6 (19); C4C5H (11); N1C2N3 (7); C5═C6N1 (7)
997(vs)	994(vw)	991(sh)	998(w)		1003	τ(C6N1) (21); τ(C5C6) (17); τ(N1C2) (16); τ(C4C5) (9); ω(C4N4) (8)
976(sh, s)	976(m)	966(w)	968(w)		960	N1−C2 (14); C2N1C6 (14); C4−N4 (9); C2−N3 (8); C5═C6N1 (8); C5═C6H (7)
898(s)					890	ω(C5H) (27); τ(C5C6) (12); τ(N1C2) (12); τ(C4C5) (11); ω(C6H) (10); ω(N1H) (9); ω(C4−N4) (8)
	817(m)		827(m)		812	ω(C6H) (65); ω(C2O2) (15)
	797(s)	785(vs)	798(m)		778	N1−C2 (25); C4−C5 (16); C4−N4 (8); C2−N3 (7)
800(br, s)	785(vs)		783(w)			
	766(w)		762(w)		760	ω(C2O2) (30); ω(C4N4) (16); ω(C6H) (12); ω(N1H) (10); ω(C5H) (10); τ(N1C2) (10)
703(vs)	704(vw)		702(w)		688	ω(C4N4) (36); ω(C2O2) (35); τ(C6N1) (12)
	601(m)	594(m)			584	N1C2N3 (18); N3C2═O2 (18); C2N3═C4 (12); C5═C6N1 (10)
615(s, br)						
			603(m)		576	N3═C4N4 (15); N3C2═O2 (14); N1C2═O2 (11); τ(CNH2) (11); C2N1C6 (9)
	550(s)	545(m)	551(m)		558	N1C2═O2 (21); N3═C4C5 (13);C5C4N4 (12); N1C2N3 (10); τ(CNH2) (10)
551(br, s)						
	538(m)		532(s)		537	τ(C4N4) (29); τ(CNH2) (19); C4−N4 (8)
	521(m)					
446(vs)	448(vw)	431(vw)	445(w)		435	τ(CNH2) (30); τ(C4N4) (17); ω(N1H) (9); τ(C5═C6) (9)
404(s)	405(w)		425(w)		392	C5C4N4 (14); ω(N1H) (12); ω(C5H) (12); N3═C4N4 (11); τ(C4N4) (8)
		388(w)			381	τ(C4N4) (22); C5C4N4 (13); N3═C4N4 (10); ω(C5H) (7)
230(m)	238(w)	243(w)			230	τ(N3C4) (28); τ(C2N3) (13); τ(C4C5) (12); ω(C4N4) (11); ω(C5H) (11)
200(vs)	214(vw)				181	ω(N1H) (42); τ(C2N3) (31); ω(C5H) (11); τ(N1C2) (8); τ(C4C5) (8)

[a] A, B, and C refer to the labels given to the solid phase spectra displayed in Figure 3, whereas G and H refer to those recorded in solution and shown in Figure 5. vs, very strong; s, strong; m, medium; w, weak; vw, very weak; sh, shoulder; br, broad. Assignments are given in terms of internal coordinates.

bond-stretch motions (Figure 9). The NIS band observed at 1158 cm^{-1} is also very sensitive to N-deuteration and is assigned by the present calculations to the C4N4H bending (Figure 3, Table 4). The ring breathing mode (observed in Raman spectra around 795 cm^{-1} in solid and 785 cm^{-1} in solution) (Figures 3 and 5) shows a 10 cm^{-1} downshift upon deuteration. The wavenumber as well as the isotopic shift of this mode is very well interpreted by the calculations. The change in the relative intensities of the intense Raman bands observed around 1280 and 795 cm^{-1} in solid (1286 and 785 cm^{-1} in solution) versus exciting wavelength is worthy to be emphasized (Figures 3 and 5).

The wagging and torsional motions in the cytosine ring appear in the region below 1020 cm^{-1} (Table 4, Figures 3 and 9) as confirmed by NIS showing intense bands in this spectral region. For the calculated mode at 1003 cm^{-1}, the PED (Table 4) reveals a concerted torsional motion which is responsible for the large-amplitude out-of-plane motions of the hydrogens connected to N1, C5, and C6 atoms (Figure 9). This mode should be partly responsible for the intense NIS band observed at 997 cm^{-1} with a shoulder at 976 cm^{-1} (Figure 3). In the C-d_3 NIS spectrum, the intensity of this band decreases and a new and narrower band appears at 986 cm^{-1} which is mainly assigned to the C−H wagging coupled with their adjacent torsional motions (calculated at 964 cm^{-1}, Table 5). On the basis of the calculated results, the intense NIS band at 898 cm^{-1} (Figure 3) which is canceled in the C-d_3 spectrum (Figure 4), should arise mainly from the C5−H wagging coupled with a considerable out-of-plane motion of the hydrogen connected to the N1 atom (calculated mode at 890 cm^{-1}, Table 5, Figure 9). Obviously a component of the broad NIS band centered at 800 cm^{-1}, canceled in C-d_3, should involve a strong N1H contribution. This mode is calculated at 760 cm^{-1} with a weak N1−H wagging contribution. The calculated mode at 812 cm^{-1} involving mainly C6−H wagging contributes as expected to this band.

The main problem encountered in the calculated assignments is that N1−H wagging does not contribute as it should to the

TABLE 5: Comparison between Experimental (Condensed Phase) and Scaled (Calculated with MP2/6-31G(*) Force Field) Vibrational Wavenumbers (cm^{-1}) for N-Deuterated Cytosine (C-d_3) a−o Tautomer (Figure 1)[a]

exp (NIS)	exp (Raman)		exp (IR)			
A	C	G	B	H	calc	potential energy distribution (%)
	3117(s)				3113	C5−H (89); C6−H (10)
	3057(s)				3087	C6−H (89); C5−H (10)
	2514(m)		2536(s)		2487	ND2 asym st (99)
	2340(s)		2334(s)		2344	ND2 sym st (99)
	2084(br, m)		2100(br, s)		2111	N1-D (98)
			1683(m)		1696	C2═O2 (20); N3═C4 (18); C6−N1 (12); C2−N3 (11); C5═C6 (11)
	1648(m)	1644(w)	1654(sh)	1645(vs)	1657	C2═O2 (39); C5═C6 (10); C4−N4 (8)
			1639(vs)			
	1594(w)	1611(w)	1602(vs)	1608(s)	1591	C4−C5 (33); N3═C4 (14); C5═C6 (10); C2═O2 (7)
	1570(w)	1590(sh)		1589(s)		
1520(br, w)	1518(w, br)	1508(m)	1519(sh)	1504(vs)	1533	C4−N4 (20); N3═C4 (14); C5═C6 (10); C5═C6H (10); C6−N1 (8); N1C6H (8)
	1493(w)		1508(vs)			
1437(br, w)	1449(w)	1387(vw)	1455(m)	1388(m)	1415	C6−N1 (21); C4−N4 (14); C5═C6N1 (8); C2N3═C4 (7)
1388(w)	1359(vw)		1377(w)			
					1343	N1C6H (21); C5═C6H (18); C6−N1 (13)
1331(m)	1319(m)		1317(m)	1323(w)		
	1286(s)	1288(s)	1285(s)		1293	C2−N3 (31); N1−C2 (13); ND2-sciss (10); N3═C4 (9)
			1278(sh)			
1190(w)	1243(w)		1237(w)	1193(sh)	1189	ND2-sciss (39); CND sym bend (18); C2−N3 (8)
			1183(w)			
1134(vs)	1163(vw)	1140(m)	1129(sh)		1139	C6═C5H (34); C4C5H (19); C5═C6 (13)
	1131(m)		1119(m)			
1065(w)		1044(vw)	1048(w)		1059	N3C2═O2 (12); CND asym bend (11); C4C5═C6 (10); N1−C2 (9)
1021(w)	1024(w)		1009(m)		985	CND asym bend (11); C2N1C6 (8)
986(s)		984(w)	972(w)		964	τ(C5C6) (28); τ(C4C5) (16); ω(C6H) (11); ω(C5H) (9)
	967(m)		959(w)		953	C6N1D (34); C2N1D (18); CND asym bend (12); C6−N1 (9)
	855(w)		869(vw)		862	CND asym bend (35); C2N1D (14); C4−C5 (9)
858(br, w)						
			852(vw)		841	ω(C4N4) (35); ω(C5H) (24); τ(C2N3) (7)
800(vs)			789(s)		808	ω(C6H) (68); ω(C2O2) (11); ω(C5H) (8); τ(C4C5) (8)
785(sh)	780(vs)	776(vs)	773(w)		766	N1−C2 (26); C4−C5 (14); C4−N4 (9); C2−N3 (7)
725(vs)			725(vw)		718	ω(C2O2) (64); ω(C6H) (8)
626(w)	616(w)		612(m)		592	ω(C4N4) (30); τ(C6N1) (29); τ(N1C2) (27); ω(N1D) (11)
596(m)	593(m)	593(m)	598(m)		580	N1C2N3 (21); N3C2═O2 (19); C2N3═C4 (10); C5═C6N1 (9)
	546(m)	531(m)	545(m)		537	N1C2═O2 (30); N3C2═O2 (15); C2N1D (10), N3═C4N4 (7); C5C4N4 (7)
526(s)	525(m)		533(m)		534	N3═C4C5 (24); C4−N4 (12); C5C4N4 (7)
	508(m)		508(m)			
445(vs)	454(w, br)		452(m)		442	τ(C4N4) (24); ω(N1D) (17); τ(CND2) (15); ω(C5H) (13); τ(C5C6) (11)
	397(w)		398(s)		389	τ(CND2) (51); ω(N1D) (9); τ(C6N1) (7)
374(m)						
	374(m)	353(vw)			359	C5C4N4 (28); N3═C4N4 (22); C2N3═C4 (9); CND asym bend (8)
					312	τ(C4N4) (54); ω(C5H) (9); ω(C4N4) (7)
234(m)			231(vw)		205	τ(N3C4) (19); τ(C2N3) (15); τ(C4N4) (13); τ(CND2) (9); τ(C4C5) (8); ω(N1D) (7); ω(C4N4) (7); τ(C6N1) (7)
195(s)	172(m)		198(vw)		179	ω(N1D) (37); τ(C2N3) (27); ω(C5H) (13); τ(C4C5) (11); τ(N1C2) (7)

[a] A, B, and C refer to the labels given to the solid phase spectra displayed in Figure 4, whereas G and H refer to those recorded in solution and shown in Figure 6. vs: very strong, s: strong, m: medium, w: weak, vw: very weak, sh: shoulder, br: broad. Assignments are given in terms of internal coordinates.

PED of the vibrational motions above 750 cm^{-1} (Table 4). In fact, the contribution of this internal coordinate does not exceed 10% in the PED of the calculated modes at 890, 760, and 435 cm^{-1} (Table 4). Inversely, a very important contribution of N1−H wagging is found in the calculated mode at 181 cm^{-1} (42% of PED, Table 4). Surprisingly, this mode gives rise to a low-intensity calculated NIS band (Figure 7 down). Indeed, in this vibrational motion (Figure 9) the N1 atom moves with a large amplitude out of the average ring plane, whereas the hydrogen connected to this atom remains motionless, leading to a low-intensity calculated NIS band. A similar behavior was found for N1−H wagging in uracil,[1] where the graphical representation of the low-wavenumber vibrational modes shows notable out-of-plane motion of N1 atom.

The intense band at 703 cm^{-1} (calc. 688 cm^{-1}) is upshifted to 725 cm^{-1} (calc. 718 cm^{-1}) upon deuteration. A large ω(C4N4) contribution and an insufficient ω(CH) one, found in the PED of the calculated mode, is responsible for its deficiency in reproducing the experimental NIS intensity. The comparison between the C and C-d_3 NIS spectra (Figures 3 and 4) allows the assignments of the NH and NH2 group vibrations. NIS bands at 615 (broad), 551, and 404 cm^{-1} are altered by the selective deuteration. The calculated results (Tables 4 and 5) for the amino group vibrational modes below 700 cm^{-1} show a strong coupling between ω(C4N4), τ(C4N4), and τ(CNH2) coordinates. The calculated band at 576 cm^{-1} could correspond to the 615 cm^{-1} strong NIS band which contains a very weak $\tau(CNH_2)$ contribution (Table 4, Figures 7 and 9). The 537 cm^{-1} calculated mode with important τ(C4N4) and τ(CNH2) contributions is in good agreement with that observed in NIS at 551 cm^{-1}. The 404 cm^{-1} mode calculated at 392 cm^{-1} contains ω(N1H) rather than τ(CNH2).

The intense NIS band at 446 cm^{-1} and also those at 230 and 200 cm^{-1} remain rather unchanged (wavenumber and intensities) upon deuteration. These facts are not reproduced by the present calculations which show considerable isotopic shifts for these modes and notable decrease in their intensities in the theoretical NIS spectrum of C-d_3 (Figure 8, down).

TABLE 6: Scaling Factors Associated with the Cytosine Internal Coordinates[a]

coord specificn	uracil	thymine	cytosine	coord specificn (amino group)	cytosine
ν(C−N)	0.942	0.903	0.976	ν(C4−N4)	0.950
ν(C═N)			0.941	δ(N3═C4−N4)	1.250
ν(C−C)	0.952	1.180	1.044	δ(C5−C4−N4)	1.250
ν(C═C)	0.834	0.890	0.816	NH2 sciss = δ(H′−N4−H″)	0.951
ν(C═O)	0.792	0.831	0.787	NH2 asym st = $(1/\sqrt{2})[\nu(\text{N4−H}') - \nu(\text{N4−H}'')]$	0.810
ν(C−H)	0.897	0.889	0.900		
ν(N1−H)	0.905	0.905	0.620	NH2 sym st = $(1/\sqrt{2})[\nu(\text{N4−H}') - \nu(\text{N4−H}'')]$	0.810
δ(C−N−H)	1.032	1.058	0.973		
δ(N−C−H)	1.069	1.198	1.003	CNH asym bend = $(1/\sqrt{2})[\delta(\text{C4N4H}') + \delta(\text{C4N4H}'')]$	1.111
δ(C−N−C)	0.996	0.996	0.996		
δ(N−C−N)	0.992	0.992	0.928	CNH sym bend = $(1/\sqrt{2})[\delta(\text{C4N4H}') + \delta(\text{C4N4H}'')]$	1.111
δ(C−C═C)	1.144	1.032	1.222		
δ(N−C═C)	1.144	1.043	1.222		
δ(N═C−C)			1.000		
δ(N−C═O)	1.092	1.048	1.168		
δ(C−C−H), δ(C═C−H)	0.992	1.058	0.930		

[a] For the amino group, internal coordinates of local symmetry (C_{2v}) have been used, whose expressions are given in front of each scaling factor. The values of the scaling factors for similar coordinates used in uracil (ref 1) and thymine (ref 2) are also reported for comparison.

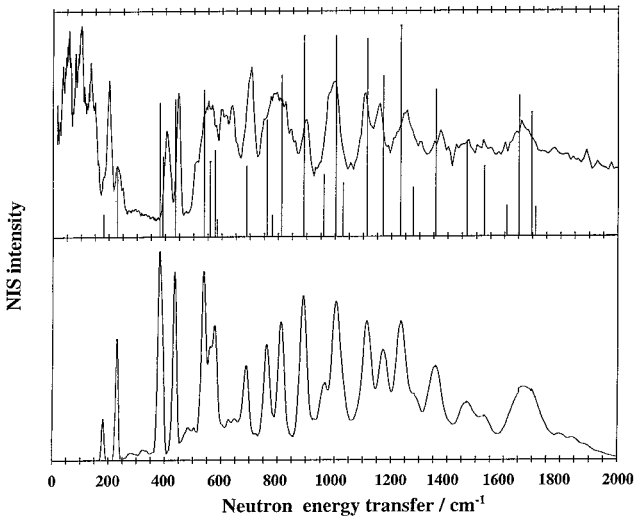

Figure 7. Comparison between experimental (top) and calculated NIS (obtained with scaled MP2/6-31G(*) force field of the a−o tautomer) spectra for cytosine (C) in the spectral region below 1900 cm^{-1}. Both spectra are shown in the region above the lattice mode vibrations. The vertical lines (top) represent the calculated first-order intensities. The calculated NIS spectrum (down) has been reconstituted by adding up to the third-order spectra.

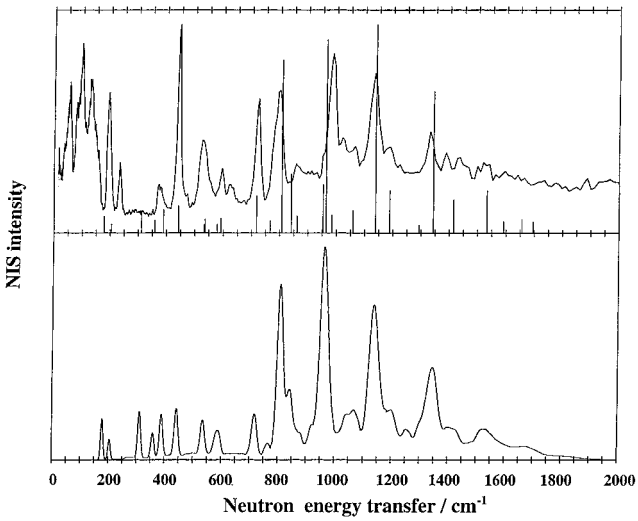

Figure 8. Comparison between experimental (top) and calculated NIS (obtained with scaled MP2/6-31G(*) force field of the a−o tautomer) spectra for N-deuterated cytosine (C-d_3) in the spectral region below 1900 cm^{-1}. For details see also the caption of Figure 7.

At last, all the NIS and Raman intense bands observed in solid phase, below 150 cm^{-1} and barely affected by N-deuteration, are assigned to the lattice modes (Figures 3 and 4). The effect of temperature on these external modes, as evidenced by the Raman spectra, should be mentioned.

IV. Concluding Remarks

The present paper and the two previous ones[1,2] have been devoted to the structural features and to the vibrational mode analysis of the nucleic acid pyrimidine bases. In all these investigations, it has been shown that the joint use of NIS and optical (Raman and IR) spectroscopies, constitutes a powerful methodology to analyze the molecular vibrational dynamics. We have also shown that quantum mechanical calculations performed at a reasonable level of correlation, self-consistent field (SCF) plus second-order Møller−Plesset, can give an excellent agreement with the spectra observed in the gas phase and in the rare gas matrixes. However, it has been demonstrated that in all cases the quantum mechanical force field should be scaled in order to improve the agreement between the theoretical and experimental results in condensed phases. As far as the modes involving the wagging and torsional motions are concerned, the various scaling procedures based on the Pulay or other methods cannot correct completely all the discrepancies related to intermolecular H bonds and probably to molecular packing in solid phase. For instance, the problem arising from the comparison between the calculated and experimental results concerning N1−H wagging and its coupling with the torsional motions has been fully discussed here and in our previous paper on uracil.[1] Upon N-deuteration, in contrast with the experimental results, the calculated NIS intensities decrease considerably in the spectral region below 450 cm^{-1}. This fact leads us to conclude that generally the scaled (or even unscaled) force field overestimates the coupling between the N1−H (or NH_2) wagging and the torsional motions. Although the C−H wagging modes calculated by the scaled force field are located in the adequate spectral range (900−1050 cm^{-1}), their coupling with the ring torsional motions is largely underestimated. This seems to be the main reason of the abnormally weak calculated NIS intensity below 450 cm^{-1} in the N-deuterated spectra. In fact the experimental constancy of these low-wavenumber NIS bands

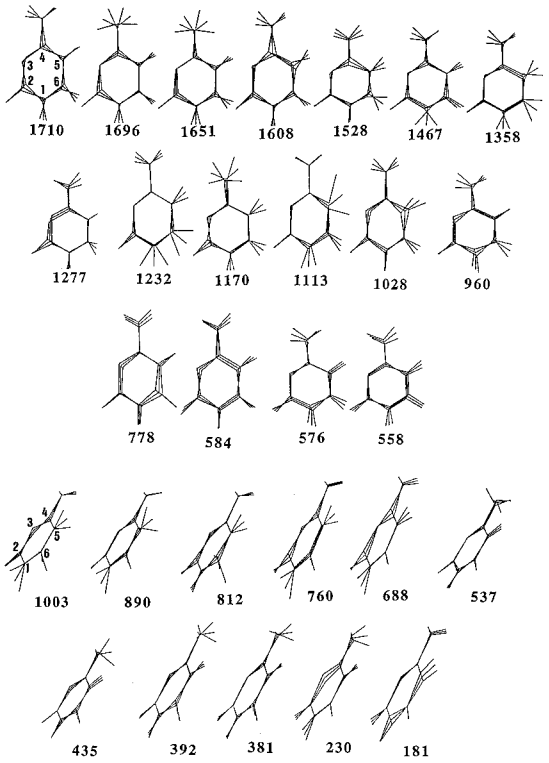

Figure 9. Graphic representation of the calculated vibrational modes of the cytosine a−o tautomers as obtained with scaled MP2/6-31G(*) force field. For atom numbering see Figure 1. Only the vibrational modes below 2000 cm^{-1} have been drawn. For each mode the equilibrium and the most distorted configurations are presented. The three first rows correspond to the modes involving mainly bond-stretch and angle bends, whereas the last two rows show the modes corresponding mainly to the wagging and torsional motions. For assignment of these modes in terms of internal coordinates, see Table 5.

should result from the strong coupling between C−H waggings and ring skeletal motions.

To conclude this series of investigations on pyrimidine bases, we should mention that the force field obtained by scaling a quantum mechanical force field determined on the isolated base is not a sufficient tool to assign all vibrational features observed in the condensed phase. Intermolecular H bonds occurring on N−H and C═O groups are in part responsible for the encountered discrepancies upon the comparison between the experimental and calculated results. Recently, in a series of DFT calculations based on a *supermolecule* constituted by an uracil surrounded by two water molecules,[42,43] we have shown that the above-mentioned problems may partially be solved by introducing explicitly the intermolecular H bonds in the theoretical model. Another problem concerning the coupling of the C−H waggings with the ring torsional motions has also emerged upon the analysis of the NIS spectra observed in the N-deuterated solid samples. It may be related to the molecular stacking in solid phase which should influence the C−H group out-of-plane motions.

Full unscaled and scaled force constant matrixes can be provided upon request.

Acknowledgment. The authors thank the Rutherford Appleton Laboratory staff for their technical assistance in using TFXA to obtain the NIS spectra. All the quantum chemical computations reported in this paper were carried out on Cray C98 computers. The authors thank IDRIS (Institut du Développement et des Ressources en Informatique Scientifique, CNRS) for access to computational facilities. A.A. was partially supported by a fellowship from the Government of Morocco for his PhD study.

References and Notes

(1) Aamouche, A.; Ghomi, M.; Coulombeau, C.; Jobic, H.; Grajcar, L.; Baron, M. H.; Baumruk, V.; Turpin, P. Y.; Henriet, C.; Berthier, G. *J. Phys. Chem.* **1996**, *100*, 5224.

(2) Aamouche, A.; Ghomi, M.; Coulombeau, C.; Grajcar, L.; Baron, M. H.; Jobic, H.; Berthier, G. *J. Phys. Chem. A* **1997**, 101, 1808.

(3) Ferro, D.; Bencivenni, L.; Teghil, R.; Mastromarino, R. *Thermochim. Acta* **1980**, *42*, 75.

(4) Radchenko, E. D.; Sheina, G. G.; Smorygo, N. A.; Blagoi, Yu. P. *J. Mol. Struct.* **1984**, *116*, 387.

(5) Szczesniak, M.; Szczepaniak, K.; Kwiatkowski, J. S.; KuBulat, K.; Person; W. B. *J. Am. Chem. Soc.* **1988**, *110*, 8319.

(6) Person; W. B; Szczepaniak, K.; Szczesniak, M.; Kwiatkowski, J. S.; Hernandez, L.; Czerminski, R. *J. Mol. Struct.* **1989**, *194*, 239.

(7) Nowak, M. J.; Lapinski, L.; Fulara, J. *Spectrochim. Acta* **1989**, *45A*, 229.

(8) Gould, I. R.; Vincent, M. A.; Hillier, I. H.; Lapinski, L.; Nowak, M. J. *Spectrochim. Acta* **1992**, *48A*, 811.

(9) Person; W. B.; Szczepaniak, K., In *Vibrational Spectra and Structure*; During, J. R., Ed.; Elsevier: Amsterdam, 1993; Vol. 20, p 239.

(10) Lord, R. C.; Thomas Jr., G. J. *Spectrochim. Acta* **1967**, *23A*, 2551.

(11) Angell, C. L. *J. Chem. Soc.* **1961**, 504.

(12) Susi, H.; Ard, J. S., Purcell, J. M. *Spectrochim. Acta* **1973**, *29A*, 725.

(13) Susi, H.; Ard, J. S. *Spectrochim. Acta* **1974**, *30A*, 1843.

(14) Florián, J.; Baumruk, V.; Leszczynski, J. *J. Phys. Chem.* **1996**, *100*, 5578.

(15) Beetz Jr., C. P.; Ascarelli, G. *Spectrochim. Acta* **1980**, *36A*, 299.

(16) Nishimura, Y.; Tsuboi, M. *Chem. Phys.* **1985**, *98*, 71.

(17) Tsuboi, M.; Nishimura, Y.; Hirakawa, A. Y.; Peticolas, W. L. In *Biological Applications of Raman Spectroscopy*; Spiro, T. G., Ed.; J. Wiley & Sons: New York, 1987; Vol. 2, p 109.

(18) Kwiatkowski, J. S.; Leszczynski, J. *J. Phys. Chem.* **1996**, *100*, 941.

(19) Estrin, D. A.; Paglieri, L.; Corongiu, G. *J. Phys. Chem.* **1994**, *98*, 5653.

(20) Paglieri, L.; Corongiu, G.; Estrin, D. A. *Int. J. Quantum Chem.* **1995**, *56*, 615.

(21) Barker, D. L.; Marsh, R. E. *Acta Crystallogr.* **1964**, *17*, 1581.

(22) Dreyfus, M.; Bensaude, O.; Dodin, G.; Dubois, J. E. *J. Am. Chem. Soc.* **1976**, *98*, 6338.

(23) Katritzky, A. R.; Waring, A. J. *J. Am. Chem. Soc.* **1963**, *85*, 3046.

(24) Fazakerley, G. V.; Gdaniec, Z.; Sowers, L. C. *J. Mol. Biol.* **1993**, *230*, 10.

(25) Leszczynski, J. *Int. J. Quantum Chem* **1992**, *QBS19*, 43.

(26) Sponer, J.; Hobza, P. *J. Phys. Chem.* **1994**, *98*, 3161.

(27) Stewart, E. L.; Foley, C. K.; Allinger, N. L.; Bowen, J. P. *J. Am. Chem. Soc.* **1994**, *116*, 7282.

(28) Bludsky, O.; Sponer, J., Leszczynski, J., Spirko, V.; Hobza, P. *J. Chem. Phys.* **1996**, *105*, 11042.

(29) Sponer, J., Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *14*, 117.

(30) Riggs, N. V. *Chem. Phys. Lett.* **1991**, *177*, 447.

(31) Gould, I. R.; Hillier, I. H. *Chem. Phys. Lett.*. **1989**, *161*, 185.

(32) Ha, T.-K.; Gunthard, H. H. *J. Mol. Struct.*(*THEOCHEM*) **1992**, *276*, 209.

(33) Ha, T.-K.; Gunthard, H. H. *J. Mol. Struct* **1993**, *300*, 619.

(34) Hall, R. J.; Burton, N. A.; Hillier, I. H.; Young, P. E. *Chem. Phys. Lett.* **1994**, *220*, 129.

(35) Gould, I. R; Burton, N. A.; Hall, R. J.; Hillier, I. H. *J. Mol. Struct.*(*THEOCHEM*) **1995**, *331*, 147.

(36) Frisch, M. J. G.; Trucks, W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. G.; Petersson, A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J. B.; Stefanov, 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 92;* Gaussian, Inc.: Pittsburgh, PA, 1993.

(37) Gallinella, E.; Cadioli, B.; Flament, J. P.; Berthier, G. *J. Mol. Struct.* (*THEOCHEM*) **1994**, *315*, 137−148.

(38) Wilson Jr., E. B.; Decius, J. C.; Cross, P. C., In *Molecular Vibrations*; McGraw-Hill: New York, 1955.

(39) Brown, R. D.; Godfrey, P. D., McNaughton, D., Pierlot, A. P. *J. Am. Chem. Soc.* **1989**, *111*, 2308.

(40) Lehn, J. M. In *Nitrogen Inversion—Experiment and Theory*; Topics in Current Chemistry, Band 15, Heft 3; Springer-Verlag, Berlin, 1970; p 319.

(41) Fogarassy, G.; Pulay P. In *Vibrational spectra and structure*; Durig, J. R., Ed.; Elsevier: Amsterdam, 1985; Vol. 14, p 125.

(42) Ghomi, M.; Aamouche, A.; Cadioli, B.; Berthier, G.; Grajcar, L.; Baron, M. H. *J. Mol. Struct.* **1997**, *411*, 323.

(43) Aamouche, A.; Berthier, G.; Cadioli, B.; Gallinella, E.; Ghomi, M. *J. Mol. Struct.* (*THEOCHEM*), in press.