IR and Raman Spectra, Tautomeric Stabilities, and Scaled Quantum Mechanical Force Fields of Protonated Cytosine†

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY · MARCH 1996

Impact Factor: 2.78 · DOI: 10.1021/jp953284w

3 AUTHORS, INCLUDING:

Jan Florián
Loyola University Chicago
95 PUBLICATIONS **4,144** CITATIONS
SEE PROFILE

Vladimir Baumruk
Charles University in Prague
88 PUBLICATIONS **1,901** CITATIONS
SEE PROFILE

IR and Raman Spectra, Tautomeric Stabilities, and Scaled Quantum Mechanical Force Fields of Protonated Cytosine†

Jan Florián,*,‡,§,|| Vladimír Baumruk,§ and Jerzy Leszczyński‡

Department of Chemistry, Jackson State University, Jackson, Mississippi 39217, Institute of Physics, Charles University, Ke Karlovu 5, 12116 Prague, Czech Republic, and Department of Chemistry, University of Southern California, 3620 S. McClintock Ave., Los Angeles, California 90089-1062

Received: November 7, 1995; In Final Form: January 4, 1996⊗

We present an experimental and theoretical investigation of the vibrational spectra of cytosine and protonated cytosine. In addition, we study spectra of cytosine and protonated cytosine with the mass of N1−H hydrogen changed to the mass of the methyl group. In this way the spectra of the heterocyclic part of cytidine are simulated. Spectral interpretation is based on the double harmonic approximation and scaled quantum mechanical (SQM) methodology. The scale factors of the *ab initio* HF/6-31G* and density functional Becke3-LYP/6-31G* force fields are adjusted using the frequencies and intensities measured for crystalline samples of anhydrous cytosine, cytosine monohydrate, and cytosine hydrochloride of known crystal structures. The same sets of scale factors are used for interpretation of vibrational spectra of both the neutral and protonated cytosine. These scale factors are also recommended for the spectral calculations of other nucleic acid bases and oligonucleotides. The consistency in transferring scale factors among different molecules is improved by using ring stretching scale factors that are linearly decreasing with the magnitude of the respective force constant. The original IR and Raman spectra of neutral and acidic aqueous solutions of cytosine are compared to the spectra of crystalline samples and to the calculated spectra. It is shown that frequency deviations caused by molecular environment and deviations between calculated spectra and spectra of aqueous solutions are of the same magnitude. Larger uncertainties that remain for wagging and torsional vibrations of the cytosine amino group, due to its largely anharmonic character, are significantly decreased upon cytosine protonation, which makes the cytosine amino group planar and more stiff in its out-of-plane degrees of freedom. The relative stabilities of tautomers of protonated cytosine in the gas phase and in a polar solvent are investigated at the *ab initio* level of theory with methods involving Hartree−Fock, MP2, MP4, and polarizable continuum approximations. In addition, discrete description of aqueous solvation was employed by using an empirical Langevine dipole model. According to these calculations, O2-protonated (enol) cytosine should slightly prevail in the gas phase, whereas in aqueous solution the N3-protonated form of cytosine is 3 ± 1 kcal/mol more stable than the enol tautomer. The energetic grounds, as well as our analysis of the vibrational spectra, support the recent finding of Purrello et al. (*J. Am. Chem. Soc.* **1993**, *115*, 760) that the enol form of protonated cytosine, in the form of cytidine monophosphate, is present in small amounts in acidic aqueous solution.

Introduction

Among nucleic acid bases, cytosine is the most alkaline in aqueous solution (p*K*a = 4.6).[1] This property plays an important role in many biochemical processes. According to the crystal diffraction data[2] and *ab initio* Hartree−Fock calculations, carried out with the minimal STO-3G basis set (HF/STO-3G),[3] cytosine protonation should occur in aqueous solution or the gas phase at the N3 position (Figure 1). More specifically, a large (7.5 kcal/mol) energy difference between N3- and O2-protonated cytosine was predicted.[3] Consequently, like other nucleic acid bases, protonated cytosine was generally assumed to occur only in its keto−amino tautomeric form. However, the exact proton position has been determined indirectly in most crystal studies, either from residual electron density maps or from the magnitude of the ring angle at the protonated nitrogen atom.[2] Also, the use of only a minimal basis set for the calculation of tautomeric equilibria may lead to misleading results. Thus, a more sophisticated theoretical examination of tautomeric equilibria of protonated cytosine is clearly needed.

In several cases—such as the formation of the C·C$^+$ DNA duplex,[4−6] C·G·C$^+$ DNA triplex,[7,8] the C·C$^+$ four-stranded DNA,[9,10] and the base pair between cytosine and 2-aminopurine (a mutagenic base analog)[11,12]—cytosine protonation at the N3 position has been unambiguously identified. On the other hand, cytosine protonation at the O2 oxygen could be responsible for the stabilization of the adenine−cytosine mispair, which was observed in single crystals of oligonucleotide duplexes,[13] via the formation of the O2−H···N1(ade) hydrogen bond. However, recent NMR[14] and *ab initio*[15] studies indicated that adenine rather than cytosine is protonated in this complex.

Interestingly, the presence of a considerable amount of the enol tautomer of protonated cytidine monophosphate in its 5 mM aqueous solution at pH values below 4.8 was revealed by a resonance Raman spectroscopic study.[16] The Raman spectra have been interpreted by comparison with the spectra of structurally related molecules and by examining band excitation profiles. Obviously, a theoretical prediction of the vibrational spectra of different tautomers of protonated cytosine would be helpful to support this finding.

* Corresponding author, at Department of Chemistry, University of Southern California, E-mail: florian@usc.edu.

† This paper is the eighth in the series "Scaled quantum mechanical force fields and vibrational spectra of solid state nucleic acid constituents".

‡ Jackson State University.

§ Charles University.

|| University of Southern California.

⊗ Abstract published in *Advance ACS Abstracts,* March 1, 1996.

0022-3654/96/20100-5578$12.00/0

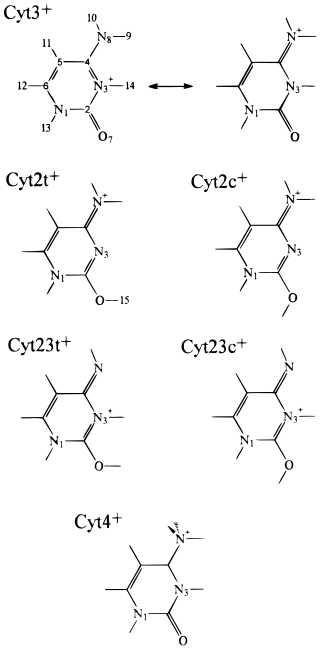

Figure 1. Studied tautomers of protonated cytosine.

Nucleobases, as the main chromophores in DNA, dominate the Raman spectra of nucleotides. Thus, the experimental data can be plausibly modeled with the spectra calculated for isolated cytosine. The inaccuracies of *ab initio* force fields, which originate from the use of truncated basis sets of atomic orbitals, from improper treatment of electron correlation and anharmonicity, and from neglected intermolecular interactions, can be partly corrected by multiplying elements of the force-constant matrix by empirical scale factors. Usually, force constants of a similar type are scaled by a common scale factor. The main virtue of this methodology, developed by Fogarasi and Pulay,[17,18] stems from the transferability of scale factors between chemically related molecules. This feature, which results from the systematic nature of the errors of *ab initio* force fields, has been proven for a wide range of chemical compounds,[19,20] including protonated adenine and guanine.[21,22] The resulting scaled quantum mechanical (SQM) force fields thus represent an efficient tool for investigation of spectral variations caused by cytosine protonation and tautomerization.

Previously, vibrational spectra of neutral cytosine have been studied using empirical[23−26] and *ab initio*[27−31] force fields. The frequencies and IR intensities calculated for tautomers of cytosine[29] and 1-methylcytosine[30] at the Hartree−Fock 6-31G(d,p) level of theory were shown to be in a good agreement with the low-temperature matrix-isolation IR spectra of these compounds.[28,32,33] However, for interpretation of aqueous solution spectra, empirical force fields determined by fitting force constants to the spectra of crystalline nucleobases are usually employed. This is because solid samples provide spectra similar to the aqueous solution spectra,[34] while having the advantage of better resolved band structure.

In this study, we have measured the IR and Raman spectra of aqueous solutions of cytosine at neutral and acidic pH, as well as the spectra of crystalline anhydrous cytosine, cytosine monohydrate, and cytosine hydrochloride. Subsequently, we have used these experimental data to obtain a plausible interpretation of the vibrational spectra of neutral and protonated cytosine via the scaled quantum mechanical (SQM) methodology. In addition, we have evaluated the energetics of the tautomeric transitions in protonated cytosine. All tautomers depicted in Figure 1 have been treated at the *ab initio* Hartree−Fock (HF) level using the split-valence 3-21G basis set. Subsequently, we have studied the relative stability of the two most stable tautomeric forms, the $N3^+$ oxo−amino ($Cyt3^+$) and the $O2^+$(trans) hydroxy−amino ($Cyt2t^+$), using extended basis sets and Møller−Pleset perturbation theory, up to the fourth order (MP4). A large part of the electron correlation effects, which have been repeatedly shown to be important for tautomeric equilibria,[35−38] is taken into account by these computational methods. Finally, by using polarized continuum and Langevin dipole models we have examined the effect of a polar environment on the relative stabilities of the keto and enol tautomers of protonated cytosine.

Experimental Section

Samples. Anhydrous cytosine was purchased from Calbiochem and used without further purification. Single crystals of cytosine monohydrate and cytosine hydrochloride were powdered to obtain polycrystalline samples. Single crystals were grown using techniques described by Jeffrey[39] and by Zachová.[40] Aqueous solutions were prepared by dissolving cytosine (Calbiochem) in double-distilled water at a 0.05 M concentration. The final pH value of each sample was set by adding a proper amount of 1 M HCl or NaOH.

Raman Spectroscopy. Raman spectra were recorded on a multichannel Raman spectrometer based on a Monospec 600 monochromator (Hilger & Watts), equipped with a 1200 grooves/mm grating and a liquid nitrogen-cooled CCD detection system (Princeton Instruments) having 1024 pixels along the dispersion axis. Spectra of solutions and polycrystalline samples were excited with the 488 nm line of an argon laser, using approximately 200 mW of radiant power at the sample and collected in the ca. 350−1800 and 2200−3500 cm^{-1} regions, with an accumulation time of 200 s and a 6 or 3 cm^{-1} spectral slit width, respectively. Rayleigh scattering was suppressed by means of a holographic edge filter (POC) placed in front of the entrance slit. The spectral frequencies were calibrated with argon-ion laser plasma lines and are believed to be accurate to ± 1 cm^{-1}. For collection of the parallel and perpendicular components of the polarized Raman spectra of solution samples, the plane of polarization of the incident beam was rotated and the orientation of a polarizing analyzer, located between the sample and the entrance slit of the spectrograph, was kept fixed. The solvent spectrum was carefully subtracted, and a baseline correction was made. Peak positions and intensities were determined by a curvefit procedure using SpectraCalc software (Galactic Industries).

Infrared Spectroscopy. Infrared spectra of anhydrous cytosine and cytosine hydrochloride in KBr pellets were recorded with a Nicolet Impact 400 FTIR spectrometer. The spectra in the region from 400 to 4000 cm^{-1} were collected with a standard Globar source, a KBr beam splitter, and a DTGS detector.

IR spectra of cytosine monohydrate dispersed in nujol mull were measured in the 400−4000 cm^{-1} region using a JASCO 5300 FTIR spectrometer and 2 cm^{-1} spectral resolution. The nujol bands were subtracted.

Spectra of a 0.05 M solution of cytosine in H_2O at different pH values were measured in a demountable cell (Graseby Specac) consisting of a pair of CaF_2 windows separated by a 12 mm Mylar spacer. Generally, 512 scans were collected using a spectral resolution of 4 cm^{-1} and a Happ−Genzel apodization function. The solvent spectrum was carefully subtracted; however, due to the high solvent absorbance at 1640 cm^{-1} ($\sim$2 absorbance units) and low sample concentration, the signal to noise (S/N) ratio in the carbonyl region is low. No baseline

corrections were made. Absorption peak positions were determined by a curvefit procedure using SpectraCalc software.

Computational Methods

Quantum chemical calculations were carried out using the Gaussian 92/DFT and Gaussian 94 codes.[41,42] The HF and post-HF *ab initio* methods as well as density functional theory (DFT) were employed. For DFT calculations, the hybrid gradient-corrected exchange functional proposed by Becke[43] was combined with the gradient-corrected correlation functional of Lee, Yang, and Parr.[44] Henceforth we will denote this functional as B3-LYP.

The geometries of the two major tautomers of protonated cytosine were fully optimized at the HF/3-21G, HF/6-31G(d), B3-LYP/6-31G(d), and MP2/6-31G(2d,2p) (frozen core) levels. Harmonic vibrational frequencies and force fields were evaluated at the HF/6-31G(d) and B3-LYP/6-31G(d) levels. The electron correlation and basis set extension effects were studied by the single-point MP4, MP3, MP2, and HF calculations using 6-311G(d,p) and 6-311G(2d,2p) basis sets.

The relative tautomeric stabilities (free energies) in aqueous solution were evaluated using two polarizable continuum models. First, self-consistent reaction field (SCRF) calculations were carried out by the Onsager reaction field model.[45,46] In the Onsager model, the solute occupies a spherical cavity in a continuous dielectric medium of uniform dielectric constant ϵ_r. Basically, the permanent electric dipole of the solute induces a dipole in the continuum. The induced dipole (reaction field) then interacts with the permanent molecular dipole, influencing the electronic structure of the solute. The cavity radii for the SCRF calculations were determined from the estimated molecular volumes determined by the 0.001 e·bohr^{-3} isodensity surface.[46] Such predicted radii were then increased by 0.5 Å, to account for the nearest possible approach of solvent molecules. The molecular volumes were determined at the HF/6-31G(d)//MP2/6-31G(d,p) level. The SCRF calculations were carried out at the HF/6-31G(d) level using SCRF HF/6-31G(d)-optimized geometries. In addition, we carried out polarized continuum model (PCM) calculations using the apparent surface charge method of Tomasi.[47,48] The cavity surface was defined in terms of atom-centered interlocking spheres, with the radii determined from the 0.001 e·bohr^{-3} isodensity envelope, as implemented in the Gaussian 94 program. PCM calculations were carried out at the HF/6-31G* and MP2/6-31G** levels with the gas-phase MP2/6-31G* geometries. The relative dielectric constant of water ($\epsilon = 80$) was used for both the SCRF and PCM calculations.

For the sake of comparison, solvation free energies were evaluated by using the Langevin dipole (LD) model. In this microscopic model, water molecules are represented by grid-centered point dipoles that can be reoriented (polarized) by the electrostatic field of the solvent. Mutual electrostatic interactions between individual Langevin dipoles are treated by an iterative manner. The solute is approximated by a set of effective charges centered at positions of individual atoms. The solvation free energy is evaluated within this model as a sum of solute−dipole and dipole−dipole electrostatic energies, plus empirically parametrized terms that account for van der Waals solute−solvent interactions and the hydrophobic contribution from solvation entropy. The solvation entropy has been parametrized in terms of the relation between the magnitude of the electric field on the surface of the solute and the hydrophobic energy. The LD method was the first microscopic method that was able to reproduce quantitatively the solvation free energies.[49] Since then, it has been used extensively in studies of the solvation problem.[50] It was implemented in several versions, with and without self-consistent treatment of the solvent field in the solute Hamiltonian.[51] The present calculations were carried out with the Polaris program package,[52,53] using an approach that involves frozen solute charges. These charges were determined by fitting to the MP2/6-31G** and HF/6-31G* electrostatic potential. The grid points were chosen by using the method of Singh and Kollman[54,55] implemented in the Gaussian 92 program.[41]

The normal modes, vibrational frequencies, IR intensities, and Raman scattering activities were calculated from the SQM HF/6-31G(d) and SQM B3-LYP/6-31G(d) force-constant matrices and HF atomic polar tensors and Cartesian polarizability derivatives using the double harmonic approximation.[56−60] For spectral interpretation, the Cartesian force-constant matrix was transformed into standard internal coordinates observing local symmetry[19] and differentially scaled by the scale factors, which were adjusted to reproduce experimental spectra from crystalline neutral and protonated cytosine. Off-diagonal elements of the force-constant matrix were multiplied (scaled) by a scale factor equal to the geometric mean of the corresponding diagonal scale factors. The program MOLVIB[61] and our program SQMVIB[62] were used for this purpose. The consistent use of only one set of scale factors for all molecules ensures that spectral variations caused by protonation are correctly reproduced; i.e., that they do not originate in side effects of the scaling procedure. For ring-stretching scale factors, which largely depend on the bond order, consistency was achieved by the use of the linear scale factor−force constant relationship. A similar relationship was previously suggested for the STO-3G force fields of nucleic acid bases.[22]

The standard double-harmonic algorithm and nonresonant approximation were utilized for calculation of IR intensities[58] and Raman scattering activities.[59] In the right-angle scattering geometry with the exciting light polarized perpendicularly to the scattering plane and without an analyzer, Raman scattering activity S_j of the jth normal mode was calculated as

$$S_j = 45\alpha_j^2 + 7\beta_j^2 \qquad (1)$$

where α and β are the isotropic and anisotropic invariants of the tensor of polarizability derivatives with respect to the jth normal coordinate.[59,60] Differential Raman scattering cross sections that correspond more closely to the measured Raman spectra were determined by multiplying scattering activities by the standard temperature- and exciting frequency-dependent factor.[60,63] Raman differential cross sections calculated for a temperature of 300 K and an excitation wavelength of 488 nm are presented in this paper.

The intensities in the neutron inelastic scattering spectra were simulated by using the first-order low-temperature approximation.[64,65] The calculated vibrational bands were interpreted according to the normal-mode potential energy distribution (PED)[21] and graphically simulated by Gaussian curves having a half-width of 10 (IR and Raman) and 20 cm^{-1} (NIS).

Results and Discussion

Relative Energies. The relative gas-phase energies and dipole moments calculated for the molecules depicted in Figure 1 are given in Table 1. Obviously, the Cyt3+ and Cyt2t+ tautomers are much more stable than the other species. Given the accuracy obtainable at the HF/6-31G*//HF/3-21G level (HF/6-31G* energies evaluated at the HF/3-21G geometries), this energy separation is large enough to allow us to neglect other tautomeric forms in more accurare (and computationally de-

TABLE 1: Energies and Dipole Moments of Different Tautomers of Protonated Cytosine

molecule	total energy (au) 3-21G[a]	relative energy (kcal/mol) 3-21G[a]	relative energy (kcal/mol) 6-31G*[b]	dipole moment (D) 3-21G[a]	dipole moment (D) MP2[c]
Cyt3$^+$	−390.818397	0	0	5.9	5.9
Cyt2t$^+$	−390.810427	5.0	1.0	2.2	2.3
Cyt2c$^+$	−390.791751	16.7	12.6	4.6	
Cyt23t$^+$	−390.764519	33.8	31.7	2.6	
Cyt23c$^+$	−390.761746	35.5	33.1	5.8	
Cyt4$^+$	−390.750147	42.8	38.3	10.3	

[a] HF/3-21G. [b] HF/6-31G*//HF/3-21G. [c] MP2/6-31G(d,p).

TABLE 2: Relative Stability[a] of Two Major Tautomers (Cyt3$^+$, Cyt2t$^+$) of Protonated Cytosine

method	ΔE_0	$\Delta\Delta G_{hydr}$
HF/3-21G	5.0	4.7[b]
HF/6-31G(d)	0.4	4.3[b]
HF/6-31G(d)[d]	0.0	4.2,[b] 4.7,[c] 4.9[e]
HF/6-31G(d,p)[d]	−1.7	4.8[c]
HF/6-311G(d,p)[d]	−1.7	
HF/6-311G(2d,2p)[d]	−2.8	
MP2/6-31G(d,p)[d]	−0.5	3.2,[c] 3.5[e]
MP2/6-311G(d,p)[d]	−1.0	
MP2/6-311G(2d,2p)[d]	−1.4	
MP4/6-311G(d,p)[d]	−0.1	

[a] ΔE_0 denotes the energy difference (kcal/mol) between the Cyt3$^+$ and Cyt2t$^+$ tautomers (Figure 1) at 0 K: $\Delta E_0 = E(\text{Cyt2t}^+) - E(\text{Cyt3}^+)$. Difference in zero point energies, $\Delta\text{ZPE} = \text{ZPE}(\text{Cyt2t}^+) - \text{ZPE}(\text{Cyt3}^+)$, evaluated at the HF/6-31G* level amounts to −0.001 kcal/mol. $\Delta\Delta G_{hydr}$ denotes the difference in solvation free energies of Cyt3$^+$ and Cyt2t$^+$ at 298 K. The total free energy difference can be evaluated as $\Delta G_{298} = \Delta E_0 + \Delta\Delta G_{hydr}$. [b] Onsager reaction field model. [c] PCM model of Tomasi. [d] MP2/6-31G(d,p) geometry. [e] LD model. The following potential-derived atomic charges (atom (Cyt3$^+$, Cyt2t$^+$) [au]) were used in this calculation: MP2/6-31G**: N1 (−0.486, −0.519), C2 (0.709, 0.848), N3 (−0.584, −0.777), C4 (0.850, 1.024), C5 (−0.540, −0.559), C6 (0.224, 0.191), O7 (−0.415, −0.556), N8 (−0.957, −0.979), H9 (0.487, 0.491), H10 (0.485, 0.468), H11 (0.252, 0.255), H12 (0.182, 0.184), H13 (0.396, 0.435), H14 (0.397), H15(0.494). HF/6-31G*//MP2/6-31G**: N1 (−0.611, −0.640), C2 (0.866, 1.017), N3 (−0.713, −0.911), C4 (1.085, 1.223), C5 (−0.759, −0.746), C6 (0.393, 0.311), O7 (−0.513, −0.625), N8 (−1.100, −1.115), H9 (0.522, 0.522), H10 (0.515, 0.500), H11 (0.288, 0.290), H12 (0.187, 0.197), H13 (0.418, 0.458), H14 (0.423), H15(0.518). For atom numbering see Figure 1.

manding) studies. Thus, we investigated by more sophisticated methods only the relative stability of the Cyt3+ and Cyt2t+ tautomers (Table 2). The use of the larger basis set results in the increased stabilization of the enol tautomer. This trend is partly compensated by the inclusion of electron correlation in the calculation. Taking into account the negligible difference in zero-point vibrational energy between these two tautomers, the energy difference between the keto and enol tautomers falls in the 0.0−1.0 kcal/mol range. In other words, the enol form should slightly predominate in the gas phase at a temperature of 0 K.

While the temperature dependence of this result is small, environmental effects influence the relative stabilities of these two tautomers significantly. Indeed, the tautomeric stability is reversed in a polar continuum. SCRF and PCM models provide similar results, which indicates that the most significant part of the solvent stabilization originates from the difference in dipole moments of the Cyt3$^+$ and Cyt2t$^+$ tautomers. Because this difference is decreased when electron correlation is accounted for by MP2 calculations, a PCM calculation based on the MP2 wave function results in a somewhat smaller (2.7 kcal/mol) stabilization of the N3-protonated (keto) cytosine. This value consists of the gas-phase free energy difference at 298 K and

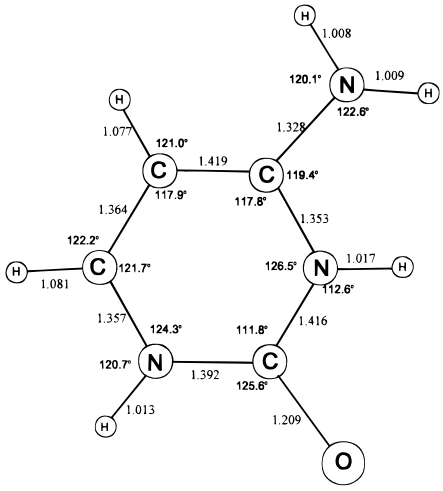

Figure 2. MP2/6-31G(d,p) geometry (Å, deg) of N3-protonated cytosine (Cyt3$^+$).

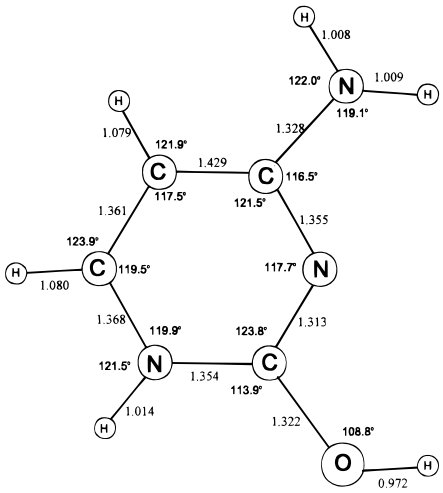

Figure 3. MP2/6-31G(d,p) geometry (Å, deg) of O2-protonated cytosine (Cyt2t$^+$).

the electrostatic part of the solvation free energy difference, $\Delta\Delta G_{el}$.[48] However, the approximate relationship

$$\Delta\Delta G_{hydr} = \Delta\Delta G_{el} \quad (2)$$

can be assumed between the total hydration free energy difference ($\Delta\Delta G_{hydr}$) and its electrostatic part ($\Delta\Delta G_{el}$), calculated by the PCM method.[48] According to this relationship, non-electrostatic contributions to the difference in solvation free energies of Cyt3$^+$ and Cyt2t$^+$ efficiently cancel each other. Considering eq 2, a 20% error bar in the solvation free energy difference, and a 0.5 kcal/mol uncertainty of the gas-phase energy difference, our estimate of ΔG amounts to 2.7 ± 1.1 kcal/mol, which in turn implies a 0.2−8% ratio of cytosine molecules protonated at the O2 position to those protonated at the N3 position, in acidic aqueous solution. This result is supported by the calculations employing a discrete (Langevine dipole) model of the solvent that result in $\Delta\Delta G_{hydr} = 5.5$ kcal/mol and $\Delta G = 3.0$ kcal/mol (Table 2).

Geometry. The calculated geometries of the Cyt3+ and Cyt2t+ tautomers are presented in Figures 2 and 3. The geometry of neutral cytosine obtained at the same computational level (MP2/6-31G(d,p) has been published previously by Sponer and Hobza.[66] The largest structural differences between Cyt3+ and Cyt2t+ tautomers involve the lengths of the CO and the C2N3 bonds. The mean experimental CO and C2N3 bond lengths obtained by statistical averaging over 14 crystal

TABLE 3: Comparison of Diagonal Force Constants (FC) of Neutral and Differently Protonated Cytosine

FC[a]	Cyt	Cyt3$^+$	Cyt2t$^+$	FC[a]	Cyt	Cyt3$^+$	Cyt2t$^+$
νN1C2	6.207	7.256	8.660	δNHS	0.590	0.604	0.585
νC2N3	7.494	6.522	10.477	δNHR	0.685	0.686	0.674
νN3C4	10.362	8.629	8.600	δC5H	0.585	0.610	0.601
νC4C5	6.286	6.709	6.478	δC6H	0.651	0.667	0.655
νC5C6	10.100	9.503	10.063	δN1H	0.612	0.648	0.649
νC6N1	8.291	8.245	7.754	δN3H		0.661	
νC2O	15.205	16.939	9.593	δOH			0.887
νC4N	8.411	10.105	10.082	γC2O	0.991	0.988	0.897
νC5H	6.385	6.469	6.431	γC4N	0.829	0.875	0.819
νC6H	6.353	6.437	6.474	γNH2	0.032	0.086	0.087
νN1H	8.334	8.103	8.089	γN1H	0.241	0.279	0.279
νNH9	8.406	8.272	8.258	γC5H	0.425	0.442	0.491
νNH10	8.508	8.230	8.354	γC6H	0.605	0.637	0.606
νN3H		8.006		γN3H		0.272	
νOH			9.032	τ6Pu	0.276	0.290	0.307
δ6Tr	1.662	1.720	1.677	τ6E1	0.195	0.234	0.230
δ6E1	1.444	1.518	1.484	τ6E2	0.303	0.304	0.309
δ6E2	1.768	1.689	1.783	τNH2	0.278	0.287	0.370
δC2O	1.390	1.254	1.383	τOH			0.152
δC4N	1.296	1.237	1.249				

[a] Unscaled HF/6-31G* force constants in internal coordinates. Stretching FCs are given in mdyn/Å, bending and torsional FC are given in mdyn·Å. For definition of internal coordinates, see Table 4.

structures containing protonated cytosine amount to 1.211 and 1.387 Å, respectively, with the maximum deviation <0.02 Å.[2] Comparison with the calculated bond lengths confirms the assumption that only N3 protonation occurs in these crystals. Also, the magnitude of the C2−N3−C4 angle, a parameter very sensitive to protonation at the N3 position,[2] confirms that N3 protonation is dominant in crystal structures. The value of 126.5°, calculated for N3-protonated cytosine, agrees well with its experimental counterpart of 125.1 ± 0.7°. In agreement with experimental findings, N3 protonation widens this angle by 6.6°. On the other hand, protonation at the cytosine oxygen results in a 2.2° decrease of the C2−N3−C4 angle with respect to its value in neutral cytosine. The geometry of the amino group is similar in both tautomers. Compared to neutral cytosine, the protonated forms possess a shorter C−NH2 bond. Such an increase in the bond order is sufficient to maintain planarity of the amino group. For neutral nucleobases, nonplanar, pyramid-like arrangements of amino groups have been reported.[66,67]

Force Constants. The unscaled diagonal HF/6-31G(d) force constants of neutral and protonated cytosine are given in Table 3. These force constants are defined as second derivatives of the total energy with respect to internal coordinate displacements. The definition of internal coordinates can be found in Table 4. Obviously, differences in stretching force constants among neutral and protonated molecules mimic variations in MP2/6-31G(d,p) bond lengths. Indeed, the largest force constant differences involve the C2O and C2N3 bonds. Also of importance is a notable increase in the C4N and γNH2 (wagging vibration of the amino group) force constants upon protonation.

Scale Factors. The scale factors determined by us for the HF/6-31G* and B3-LYP/6-31G* force fields are given in Table 5. Unlike in our previous studies, these scale factors were not fitted by the least-squares method to the set of experimental frequencies of the parent molecule and its isotopic derivatives. Instead, we adjusted them manually, utilizing not only frequency but also IR and Raman intensity information. We have chosen this procedure because the quality of the unscaled HF/6-31G* and B3-LYP spectra was good enough to reproduce the observed intensity pattern and to keep small the number of adjustable parameters (scale factors).

Individual scale factors were introduced for the wagging vibration of the amino group (γNH2) and, in the HF/6-31G* force field, for the stretching vibration of the carbonyl group. The need for individual scaling was substantiated by a high sensitivity of these vibrations to intermolecular interactions. In addition, the γNH2 mode is known to be strongly anharmonic in partly conjugated systems.[68,69] For example, the experimental (argon matrix) frequency of the γNH2 vibration of formamide amounts to 303 cm^{-1},[70] whereas the corresponding harmonic value is only 186 cm^{-1} (MP2/6-311(3df,2p))[71] or 91 cm^{-1} (MP2/6-31G(d,p)),[72] depending on the basis set extension. When involved in hydrogen bonding, the amino group becomes stiffer. As a result, the γNH2 vibrational frequency is increased. Again, for the formamide molecule, which represents a suitable model system for quantitative estimates, the γNH2 diagonal force constant increases from 0.01 to 0.1 mdyn/Å, and the γNH2 harmonic frequency increases from 91 to 480 cm^{-1} (MP2/6-31G(d,p)) when the cyclic formamide dimer is formed.[72,73] Obviously, both these effects, which increase the frequency of the γNH2 mode from its value calculated for an isolated molecule in the harmonic approximation, necessitate the use of rather high γNH2 scale factors.

Another important feature of our scaling methodology is the use of a quadratic scaling scheme for bonds formed by C, N, and O atoms:

$$\mathrm{FC}'(i,i) = a(\mathrm{FC}(i,i))^2 + b\mathrm{FC}(i,i), \quad a < 0,\ b > 0, \quad i \ni \Omega \tag{3}$$

FC(i,i) and FC′(i,i) denote unscaled and scaled diagonal force constants, respectively, and a and b are empirical parameters common for a subset Ω of internal coordinates. This relationship falls within a standard scaling scheme proposed by Fogarasi and Pulay:

$$\mathrm{FC}'(k,k) = S_k\mathrm{FC}(k,k), \quad k = 1, 3N-6 \tag{4}$$

$$\mathrm{FC}'(k,l) = (S_kS_l)^{1/2}\mathrm{FC}(k,l), \quad k,l = 1, 3N-6 \tag{5}$$

which is modified in such a way that the scale factors S_i depend linearly on the corresponding force constants FC(i,i),

$$S_i = a\mathrm{FC}(i,i) + b \tag{6}$$

For the HF/6-31G(d) force field, only CC and CN bonds were involved in Ω, whereas CO bonds were scaled separately.

The described technique can be rationalized by noting the tendency of HF force fields to overestimate the frequencies of valence vibrations of double bonds to a larger extent than they do those corresponding to single bonds. To put it in a more general way, spectra of heterocyclic molecules calculated at the HF level seem to be stretched in frequency scale, in comparison with spectra measured from aqueous solutions or crystalline samples. If a uniform scaling is employed (one scale factor identical for all FCs), then the frequencies of low-lying bands become underestimated, while vibrational bands observed in the 1500−1700 cm^{-1} region are predicted to be about 100 cm^{-1} too high.

Similar trends but smaller frequency deviations are also obtained when the universal set of scale factors for B3-LYP/6-31G(d) force fields developed recently by Rauhut and Pulay[74] is used. These scale factors were fitted to reproduce experimental frequencies of the "training" set of 20 organic molecules, involving a large range of functional groups containing C, H, N, and O atoms, with the accuracy of 12.8 cm^{-1}. When their scale factors were transferred to a set of 11 medium-sized molecules, including uracil (argon matrix data), a mean deviation

TABLE 4: Definition of Internal Coordinates

coord[a]	definition[b]	coord[a]	definition[b]
νN1C2	$\Delta r_{1,2}$	δNHS	$2\Delta\alpha_{9,8,10} - \Delta\alpha_{4,8,9} - \Delta\alpha_{4,8,10}$
νC2N3	$\Delta r_{2,3}$	δNHR	$\Delta\alpha_{4,8,9} - \Delta\alpha_{4,8,10}$
νN3C4	$\Delta r_{3,4}$	δC5H	$\Delta\alpha_{4,5,11} - \Delta\alpha_{6,5,11}$
νC4C5	$\Delta r_{4,5}$	δC6H	$\Delta\alpha_{5,6,12} - \Delta\alpha_{1,6,12}$
νC5C6	$\Delta r_{5,6}$	δN1H	$\Delta\alpha_{6,1,13} - \Delta\alpha_{2,1,13}$
νC6N1	$\Delta r_{6,1}$	δN3H	$\Delta\alpha_{2,3,14} - \Delta\alpha_{4,3,14}$
νC2O	$\Delta r_{2,7}$	δOH	$\Delta\alpha_{2,7,15}$
νC4N	$\Delta r_{4,8}$	γC2O	$\Delta\beta_{7,2,3,1}$
νNH9	$\Delta r_{8,9}$	γC4N	$\Delta\beta_{8,4,5,3}$
νNH10	$\Delta r_{8,10}$	γNH2	$\Delta\beta_{4,8,9,10}$
νC5H	$\Delta r_{5,11}$	γC5H	$\Delta\beta_{11,5,6,4}$
νC6H	$\Delta r_{6,12}$	γC6H	$\Delta\beta_{12,6,1,5}$
νN1H	$\Delta r_{1,13}$	γN1H	$\Delta\beta_{13,1,2,6}$
νN3H	$\Delta r_{3,14}$	γN3H	$\Delta\beta_{14,3,4,2}$
νOH	$\Delta r_{7,15}$	τOH	$(\Delta\tau_{15,7,2,1} + \Delta\tau_{15,7,2,3})/2$
δ6Tr	$\Delta\alpha_{6,1,2} - \Delta\alpha_{1,2,3} + \Delta\alpha_{2,3,4} - \Delta\alpha_{3,4,5} + \Delta\alpha_{4,5,6} - \Delta\alpha_{5,6,1}$	τ6Pu	$\Delta\tau_{6,1,2,3} - \Delta\tau_{1,2,3,4} + \Delta\tau_{2,3,4,5} + \Delta\tau_{3,4,5,6} + \Delta\tau_{4,5,6,1} + \Delta\tau_{5,6,1,2}$
δ6E1	$2(\Delta\alpha_{2,3,4} + \Delta\alpha_{5,6,1}) - \Delta\alpha_{6,1,2} - \Delta\alpha_{1,2,3} - \Delta\alpha_{3,4,5} - \Delta\alpha_{4,5,6}$	τ6E1	$2(\Delta\tau_{3,4,5,6} + \Delta\tau_{6,1,2,3}) - \tau_{1,2,3,4} - \Delta\tau_{2,3,4,5} - \Delta\tau_{4,5,6,1} - \Delta\tau_{5,6,1,2}$
δ6E2	$\Delta\alpha_{6,1,2} - \Delta\alpha_{1,2,3} + \Delta\alpha_{3,4,5} - \Delta\alpha_{4,5,6}$	τ6E2	$\Delta\tau_{5,6,1,2} - \Delta\tau_{1,2,3,4} + \Delta\tau_{2,3,4,5} - \Delta\tau_{4,5,6,1}$
δC2O	$\Delta\alpha_{1,2,7} - \Delta\alpha_{3,2,7}$	τNH2	$(\Delta\tau_{3,4,8,10} + \Delta\tau_{3,4,8,9} + \Delta\tau_{5,4,8,10} + \Delta\tau_{5,4,8,9})/4$
δC4N	$\Delta\alpha_{3,4,8} - \Delta\alpha_{5,4,8}$		

[a] The stretching, in-plane bending, out-of-plane bending, and torsional coordinates are denoted by ν, δ, γ, and τ, respectively. [b] For numbering, see Figure 1. $\Delta r_{A,B}$ denotes variation of the distance between atoms A and B. $\Delta\alpha_{A,B,C}$ denotes the change in magnitude of the A−B−C angle. $\Delta\beta_{A,B,C,D}$ denotes the change of the angle formed by the normal vector to the B−C−D plane and the A−B bond. $\Delta\tau_{A,B,C,D}$ denotes the change in the A−B−C−D torsional angle.

TABLE 5: Scale Factors for the HF/6-31G* and B3-LYP/6-31G* Force Fields of Cytosine Derivatives in Solid State and in Aqueous Solution

	scale factor	
coordinate type	HF/6-31G*	B3-LYP/6-31G*
CN, CC, CO stretch		1.30−0.041FC[c]
CN, CC stretch	1.265−0.05FC[c]	
C═O stretch	0.64	
C−O stretch	0.75	
NH, OH stretch	0.69	0.79
CH stretch	0.81	0.90
in-plane bending,[a] γCH	0.88	
δ-ring, δCN, δCO, δCH		1.06
δNH		0.98
γCH		1.10
out-of-plane bending[b]	1.00	1.30
γNH2	1.40	1.90

[a] Includes all in-plane bending coordinates (δ). [b] Includes all out-of-plane bending (γ) and torsional (τ) coordinates except for γNH2 and γCH. [c] FC denotes the unscaled force constant (mdyn/Å) (see eq 6).

between calculated and experimental frequencies of only 13.4 cm^{-1} was obtained. These results have clearly demonstrated the power of the scaling approach in combination with the B3-LYP/6-31G(d) force field. For scaling of CC, CN, and CO stretching force constants, a common scale factor of 0.922 was suggested by Rauhut and Pulay.[74] The mean value of the scale factors determined by us for cytosine and N3-protonated cytosine using eq 6 amounts to 0.98. The large difference between these values reflects systematic differences between the spectra of molecules in the gaseous state (prevailing in Rauhut and Pulay's training set) and and those in aqueous solutions or crystals. Also, a small number of heterocycles (only 2 of 21) were included in Rauhut and Pulay's training set.

Because independent scale factors had to be introduced for the HF/6-31G(d) CO stretching FCs, the number of HF/6-31G-(d) scale factors was kept small by correcting the rest of the HF/6-31G(d) vibrational spectra using only four scale factors. Two of these determine the linear relationship between ring stretching scale factors and their corresponding diagonal force constants. The other two were used to separately scale out-of-plane and in-plane bending vibrations, with the exception of the γCH force constants, which were scaled by the same scale factor as the δCH (in-plane) modes. In the B3-LYP force field, δCH, δNH, and γCH FCs were scaled by different scale factors.

Finally, CH and NH stretching scale factors were easily determined from observed frequencies in the spectral region above 3000 cm^{-1}. These modes are little coupled with vibrations of the cytosine ring. Moreover, CH and NH stretching vibrations are easily distinguishable, since they are known to be predominant in Raman and IR spectra, respectively.

Given the good transferability of B3-LYP scale factors, demonstrated by Rauhut and Pulay,[74] and the fact that off-diagonal force constants determined by nonlocal density functional methods can be consistently improved[73] by Pulay's type of scaling (eq 2), the scale factors presented in Table 5 are directly transferable to other nucleic acid bases and their derivatives. Because the B3-LYP/6-31G(d) and HF/6-31G(d) scale factors of the phosphodiester linkage are also known,[75] calculations of vibrational spectra of nucleotides using the SQM method might be feasible in the near future.

Vibrational Spectra of Neutral Cytosine. IR and Raman spectra of neutral cytosine in aqueous solution are compared with the spectra measured from polycrystalline sample of cytosine hydrate and with the calculated spectra in Figures 4 and 5. These results are augmented in Table 6 by the data measured for anhydrous cytosine (polycrystalline) and by SQM B3-LYP/6-31G(d) band assignments. For the sake of brevity, the SQM HF/6-31G* results are included in supporting information (Table 1S). While spectra of anhydrous cytosine have been reported before,[23,34,76] IR and Raman spectra of a cytosine aqueous solution at neutral pH and of the crystal of cytosine monohydrate are presented for the first time in this study. Obviously, by comparison among spectra of neutral cytosine in different environments, valuable information about spectral variations induced by varying intermolecular interactions can be obtained. The observed spectral variations also impose a limitation on the accuracy achievable within the model of the isolated molecule, in which only mean intermolecular effects are taken into account by the scaling procedure.

In the crystal of *cytosine monohydrate*,[39] cytosine molecules form parallel hydrogen-bonded ribbons that are linked into a three-dimensional network by water molecules. In addition,

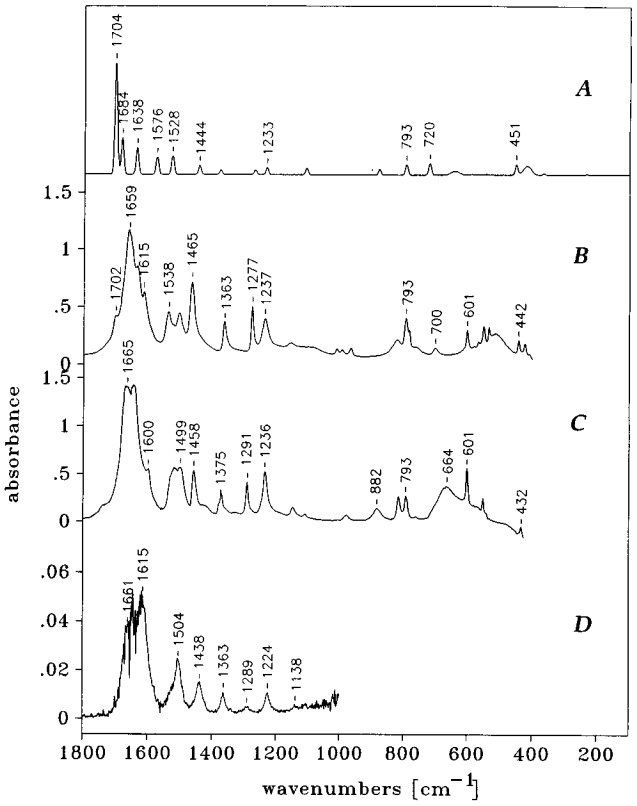

Figure 4. IR spectra of neutral cytosine. (A) Calculated spectrum (SQM B3-LYP/6-31G*). (B,C) Spectra of polycrystalline cytosine−anhydrous cytosine in KBr pellet (B) and cytosine monohydrate dispersed in nujol (C). (D) FTIR spectrum of 0.05 M aqueous solution (pH = 7.3).

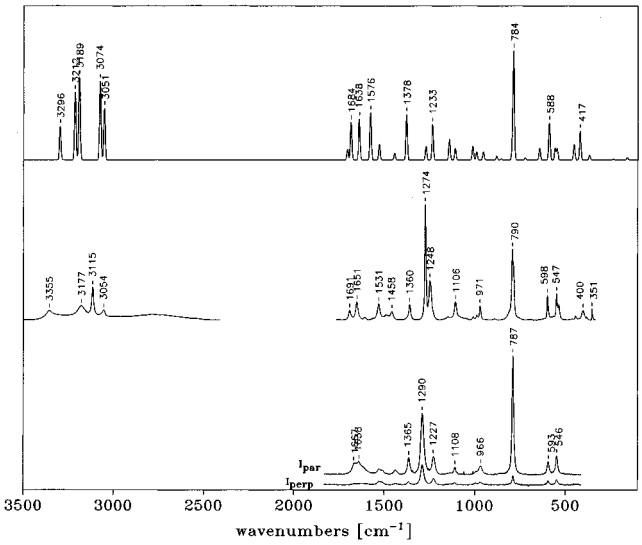

Figure 5. Raman spectra of neutral cytosine. (Top) Calculated spectrum (determined from SQM B3-LYP/6-31G* force field and HF/6-31++G** polarizability derivatives). (Middle) Polycrystalline anhydrous cytosine. (Bottom) Polarized spectra of aqueous solution (0.05 M, pH = 7.3).

cytosine molecules are slightly distorted from planarity by repulsive interactions between water molecules and methylene carbon atoms. The lengths of N1H···N3, N7H···O8, and N7H···water hydrogen bonds amount to 2.95, 2.99, and 2.97 Å, respectively. In the crystal structure of *anhydrous cytosine*,[77,78] amino groups form hydrogen bonds 2.98 and 3.03 Å long with carbonyl oxygens of neighboring molecules. In addition, N1(H)···N3 hydrogen bonds (2.84 Å) are present in this crystal structure.

The largest crystal-structure-dependent spectral differences can be found in the *low-frequency (*100−700 cm^{-1}) spectral region. IR spectra in this region are dominated by a broad band with maxima at 513 and 664 cm^{-1} for anhydrous cytosine and cytosine monohydrate, respectively. This band is shifted to 388 cm^{-1} upon N deuteration.[23,76] The high intensity, sensitivity to crystal structure, and large isotopic shift support our assignment of this band as mixed wagging and torsion of the amino group (γNH2, τNH2). The same interpretation was previously given for this mode by Susi et al.[23] and Putnam and Van Zandt.[24] The other mode, corresponding to coupled γNH2 and τNH2 vibrations, was predicted by our calculations to lie near 417 cm^{-1}. Because only weak IR bands can be found near 400 cm^{-1}, proper assignment of the out-of-plane vibrational modes of the cytosine amino group remains uncertain, mainly because of their strongly anharmonic character.

In Figure 6 we present the simulated neutron inelastic scattering (NIS) spectrum of cytosine. In our opinion, the anharmonic spectral position of τNH2 vibrations could be resolved by NIS spectroscopy, which provides valuable information about normal modes that involve large displacements of hydrogen atoms. Unfortunately, no NIS experimental data on cytosine have been reported to date.

On the other hand, the assignment of the most intense low-frequency Raman band at 787 cm^{-1} (solution) to a ring-breathing mode is done with a high degree of certainty. In IR spectra this band possesses low intensity, and it is overlapped by the γC5H (wagging) vibration.

A somewhat less successful prediction of the Raman intensity pattern can be noted for the 1100−1800 cm^{-1} frequency region. Namely, the intensity of the very strong 1290 cm^{-1} band is largely underestimated, while the calculated intensities of the bands lying above 1300 cm^{-1} are systematically too high. We attribute this phenomenon to insufficient basis set extension. Similar basis set-related Raman intensity overestimation was found for CH bending modes in ethane[79] and sodium dimethyl phosphate.[75] Because CH bending vibrations significantly contribute to the PED of most bands above 1200 cm^{-1} and to the 1138 cm^{-1} band, the intensity of which is overestimated as well, it is probable that CH bending modes are the primary cause of the observed discrepancy between calculated and experimental Raman intensities. According to our benchmark calculations for ethane, improved Raman intensities can be obtained only after inclusion of f-polarization functions or field-induced polarization functions into the basis set[79] used for calculation of the polarizability tensor and its derivatives. Unfortunately, this is highly demanding computational procedure.

The bands lying in the 1600−1800 cm^{-1} region are of considerable interest in biophysically oriented applications of vibrational spectroscopy, mainly for their sensitivity to intermolecular interactions. The theoretical interpretation of this spectral region is complicated by the strong coupling among νCO, νC5C6, and NH2-scissoring vibrations. In addition, strong Fermi resonance occurs between the νCO fundamental and second harmonics of the γCO mode in other pyrimidine nucleobases.[64] Therefore, the agreement of the calculated and experimental IR frequencies is quite encouraging. Because the frequency of the γCO fundamental is larger in cytosine than in uracil and thymine, only weak 1702 and 1740 cm^{-1} bands (IR, solid), attributable to 2 × γCO overtone, are observed on the high-frequency tail of the very strong ∼1660 cm^{-1} band.

To further assist in the interpretation of experimental data measured from DNA or RNA, with the N1−H group of cytosine replaced by N1−sugar, we included in supporting information (Tables 2S and 3S) SQM/B3-LYP results for the cytosine

TABLE 6: IR and Raman[a] Spectral Data of Cytosine. Comparison of Experimental and SQM B3-LYP/6-31G* Results

solid				aqueous solution[d]			calculation[g]				
IR[b]	R[b]	IR[c]	R[c]	IR	R	ρ^e	freq	I_{IR}	I_R	ρ^e	assignment, PED (%)
	162m[f]						157	1.5	6.2	0.75	τ6E1(48) − γN1H(26) + τ6Pu(21)
							232	7.7	2.5	0.45	τ6Pu(47) + τ6E2(36)
	350w		369w				365	10	13	0.64	δC4N(53) − δC2O(12) − δ6E1(11) + δNHR(8)
422w	400m	415w	411m				417	183	71	0.48	γNH2(47) + τNH2(27) − νC4N(9)
442w	442w	432w	445m				451	69	38	0.72	τ6E2(41) + γN1H(23) + γC5H(15) − τ6Pu(11)
513bt											coupled NH2 torsion and wagging vibration[h]
533m	535m	sh	541m				544	4.5	28	0.69	δC2O(49) + δC4N(21)
549m	547m	551w	553m		546m	0.30	556	1.8	29	0.36	δ6E2(56) − δ6E1(14)
566w											not assigned
578vw											not assigned
601m	598m	601m	604m		594m	0.33	588	5.2	90	0.41	δ6E1(55) + δ6E2(19)
		664m,br	vw				641	78	29	0.72	τNH2(35) − γNH2(25) + γN1H(15)
700m	696w						720	78	6.5	0.42	γN1H(64) − γC5H(17)
758w	755w	760vw	757vw								not assigned
782sh	784vs		793vs		787vs	0.07	784	2.1	262	0.05	νN1C2(24) + νC4C5(13) − δ6E1(11) + δ6Tr(10)
793m	791vs	793m					793	69	17	0.10	γC5H(63) + γC6H(19)
820m	803w	815m					854	2.6	3.2	0.75	γC4N(47) + τ6Pu(14) − τ6E1(12) − γC5H(11) − γC2O(9)
	890w	882m	889w		862vw		878	39	9.3	0.73	τ6Pu(41) + γC2O(40)
966w	971m	977w	979w		966w	0.28	952	3.0	21	0.39	νN1C2(26) − νC4C5(17) − δNHR(14) − δ6Tr(13)
944w	989w		1003w		994vw	0.77	990	1.4	21	0.74	γC6H(80) − γC5H(10)
1010w	1009w		1045vw				1011	1.0	34	0.23	δ6Tr(50) − νC4C5(24)
	1045vw										overtone (2 × 535 cm^{-1})
	1080vw										overtone (2 × 547 cm^{-1})
	1106m	1109w	1117m	1106	1109w	0.29	1108	44	30	0.13	δNHR(38) − δC2O(17) + νN1C2(16)
1155w	1146w	1148w	1155m	1138			1140	1.9	507	0.29	δC5H(28) − νC6N1(27) − νC5C6(14) − δN1H(12) +δNHR(8)
1237m	1245s	1236m	1255m	1224	1226m	0.36	1233	48	87	0.37	δN1H(16) − δC6H(14) + δC5H(13) + νC6N1(13) − νN1C2(11)
1277m	1273s	1291m	1295s	1289	1290s	0.33	1270	31	34	0.16	νC2N3(35) + δC6H(14) − νN1C2(10) − νN3C4(7) − νC4N(7)
1363m	1359m	1375m	1379m	1363	1365m	0.18	1378	31	110	0.10	νC4N(17) − δC6H(17) − δC5H(15) − νC5C6(14) − δN1H(12)
	1420w	1421w									overtone or combination mode
1465s	1458w	1458s	1451w	1438	1437w	0.31	1444	61	17	0.27	δN1H(25) − νC6N1(20) + νC2O(9) − νC4C5(9) + νC4N(7)
1504m	1490w	1499m	1495w	1504	1509w	0.51	1528	125	38	0.10	νN3C4(31) − δC6H(18) − νC4N(15) − δC5H(10) + δNHS(8)
1538m	1531m	1518m	1537w	1528sh	1528w	0.72	1576	117	115	0.68	νC4C5(30) − νN3C4(14) + νC2O(11) − δC5H(10) − νC6N1(7)
1615sh	1606w	1600sh	1601w	1616	1615w	0.24	1638	181	100	0.12	δNHS(65) − νC5C6(9) − νC2O(7)
1636m	1651m	1645vs	1651m	1646	1641w	0.09	1684	250	92	0.11	δC2O(27) + δNHS(16) + νC5C6(12) + δC4N(11) − δN1H(8)
1659vs	1691m	1665vs	1685w	1661	1665w	0.07	1704	761	26	0.48	νC5C6(21) − νC2O(16) − νC6N1(13) + νC2N3(12)
1702m		1740vw	1727vw								overtone (2 × γC2O)
	3053m						3051	5.4	126	0.55	νC6H(84) − νC5H(15)
	3115s				3108		3074	4.9	190	0.21	νC5H(84) + νC6H(15)
3170s	3178m						3189	69	200	0.12	νNH9(60) + νNH10(40)
3356sh	3355m						3212	60	164	0.21	νN1H(99)
3382s							3296	34	82	0.71	νNH10(60) − νNH9(40)

[a] Raman spectra (R) were excited with 488 nm radiation. Intensities are abreviated as follows: vs (very strong), s (strong), m (medium), w (weak), vw (very weak), sh (shoulder), d (doublet). [b] Polycrystalline cytosine (anhydrous). [c] Cytosine monohydrate, polycrystalline, Raman data taken from ref 81. [d] 0.05 M cytosine in aqueous solution at pH = 7.3. [e] Depolarization ratio ($I_{perp}/I_{parallel}$). [f] From Susi & Ard, 1973. [g] Frequency (cm^{-1}), IR intensity (km/mol) and Raman differential cross section (10^{-36} m^2/sr), and potential energy distribution (%), respectively. Only PED contributions larger than 7% are given. [h] See the text.

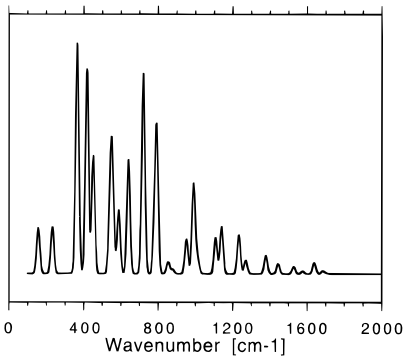

Figure 6. Calculated NIS spectrum of neutral cytosine (SQM B3-LYP/6-31G*; relative intensities).

residue and its N-deuterated isotopomer. In these model systems, the mass of the H-atom attached to N1 was increased to simulate the kinetic effects of ribose upon vibrational spectra. A similar replacement was also made for N3-protonated cytosine (Tables 4S and 5S). No other geometry or force field adjustments were done. The simple simulation of an effective mass increase is helpful in removing the spectral effects caused by coupling among the δN1H vibration and other cytosine stretching and bending modes present in the 1200−1700 cm^{-1} spectral region. However, the results of these calculations should be taken with caution because interactions between cytosine and ribose were completely neglected. For a more detailed discussion on this methodology we refer the reader to our analysis of the vibrational spectra of the guanine residue.[64]

Vibrational Spectra of Protonated Cytosine. Results of our SQM B3-LYP/6-31G* calculations of the IR and Raman spectra of N3-protonated cytosine are compared with experimental spectra measured for polycrystalline samples of cytosine hydrochloride in Table 7 and in Figures 7 and 8. The crystal structure of cytosine hydrochloride was determined by Mandel.[80] In this crystal, cytosine is protonated at the N3 position. The crystal structure consists of centrosymmetric dimers of protonated cytosine molecules linked by a pair of N1···H−O2 hydrogen bonds (2.79 Å). These dimers are interconnected by Cl···H−C and Cl···H−N multicenter hydrogen bonds involving the N3, N4, C5, and C6 atoms. Because of the ionic character of this crystal, its hydrogen bonds are stronger than those present in the crystals of cytosine monohydrate and anhydrous cytosine.

Obviously, agreement between the calculated and experimental spectra of N3-protonated cytosine is better than for neutral cytosine, especially in the low-frequency region. We

TABLE 7: IR and Raman[a] Spectral Data of N3-Protonated Cytosine. Comparison of Experimental and SQM B3-LYP/6-31G(d) Results

solid[b]		solution[c]			calculation[f]				
infrared	raman	infrared	raman	ρ^e	freq	I_{IR}	I_R	ρ^e	assignment, PED (%)
	208m[d]				182	0.6	4.8	0.75	τ6Pu(39) + τ6E2(17) − γN3H(17) + τ6E1(16) − γN1H(11)
	234m[d]				212	6.6	2.8	0.75	τ6E1(48) − τ6E2(17) − τ6Pu(9) + γN3H(9) + γC5H(8)
	368s[d]				363	5.1	46	0.73	δC4N(52) − δC2O(20) − δ6E1(10)
422m	425m				471	38	27	0.75	τ6E2(44) + γN1H(17) + γC5H(14) − τ6Pu(13)
533m	522w				527	6.9	1.7	0.67	δC2O(38) − δ6E2(28) + δC4N(19)
	544s		543m	0.35	547	3.9	100	0.42	δ6E2(41) + δC2O(14) + δC4N(9)
551m					535	5.5	0.8	0.75	τNH2(66) + γNH2(13) + γN3H(13)
588m	590m,d		581w	0.65	579	7.9	74	0.34	δ6E1(57) + δ6E2(13)
641s	639vw				631	228	1.8	0.75	γNH2(86)
	690vw								overtone (2 × δC4N)
699m	718w				747	52	4.1	0.75	γN1H(46) − γC4N(20) + γC2O(12)
757m	760w				767	27	11	0.75	γN3H(60) − γN1H(22) + τ6Pu(6)
785w,d	790vs		788vs	0.08	793	0.1	279	0.06	νN1C2(22) − δ6E1(16) + νC4C5(15) + δ6Tr(13) + νC4N(8)
826s					810	96	0.8	0.75	γC5H(33) + γC6H(23) + τ6Pu(18) + γN1H(12) + γC4N(10)
849w	844w,d		843vw		856	13	5.8	0.75	γC4N(35) − γC5H(33) + τ6Pu(17) − τ6E1(9)
887s	891w				863	117	5.7	0.75	γC2O(40) + τ6Pu(30) − γN3H(12) − γN1H(10)
	947w								combination tone
976m	977m		981w	0.42	978	8.7	6.0	0.57	νN1C2(20) + δNHR(18) − δ6Tr(15) − νC4C5(14) + νC2N3(14)
1005w	1000w		1009w	0.17	1032	4.4	28	0.09	δ6Tr(53) − νC4C5(22)
					1034	0.0	16	0.75	γC6H(79) − γC5H(10)
1076vw	1080w				1079	2.4	1.5	0.64	δNHR(50) + νN3C4(20) − δN3H(9)
1108vw	1109m		1132m	0.40	1149	5.2	76	0.23	δC5H(25) − νC6N1(25) − νC5C6(19) − δN1H(12)
	1123sh								overtone (2 × δ6E1)
1232s	1240vw	1236s			1274	103	77	0.45	δC6H(23) − δN1H(17) − δC5H(17) − νC6N1(16) + νC5C6(8)
	1277vs		1267vs	0.32	1238	4.1	152	0.28	νC2N3(32) − νN1C2(23) − νN3C4(11) + δC2O(8)
	1304vw								overtone (2 × γNH2)
1376m*	1388sh		1384m	0.12	1370	13	30	0.40	δN3H(43) + νC4N(27) − νC2O(14)
	1408sh		1418sh						overtone or combination tone
1415m	1417m	1438w	1436m	0.63	1433	11	31	0.16	δC5H(25) + δN1H (22) + δC6H(19) + δN3H(16)
1491w	1493m		1497w	0.44	1491	18	25	0.70	δN1H(30) − νC6N1(17) − δC5H(14) + νC2O(12)
1543m	1546w	1554m	1553w	0.78	1565	37	66	0.74	νN3C4(33) − νC2O(12) − νC4C5(9) − νC4N(8)
1573vw	1567w								overtone (2 × ring breathing mode)
∼1610sh	1612w				1625	33	53	0.74	νC4C5(20) − δC6H(18) − νC6N1(15) + δNHS(9) + δN3H(8)
1672vs	1657s	∼1655	1656s	0.13	1673	36	127	0.09	δNHS(47) − νC5C6(22) + δC6H(9)
1708s	1697m	1691vs	1692s	0.15	1707	653	85	0.22	δNHS(26) + νC4N(24) + νC5C6(13) − νC4C5(8)
1725s		1733vs	1734m	0.11	1724	785	25	0.23	νC2O(39) + δN3H(10) − νN1C2(10) − δN1H(8)
1762sh									overtone, (2 × γC2O)
2660br					3083	2.6	111	0.49	νC6H(90) − νC5H(8)
3090m	3094m		3096		3101	12	152	0.19	νC5H(91) + νC6H(8)
			3122		3163	81	50	0.45	νN3H(95)
					3178	379	112	0.10	νNH10(48) + νNH9(37) − νN1H(11)
					3185	64	244	0.16	νN1H(89)
3360m	3358vw				3272	101	70	0.74	νNH9(56) − νNH10(44)

[a] 488 nm excitation. [b] Polycrystalline cytosine hydrochloride, for intensity abbreviations see Table 6. [c] 0.05 M cytosine in aqueous solution at pH = 1.7. [d] Reference 81. [e] Depolarization ratio ($I_{perp}/I_{parallel}$). [f] I_{IR} denotes IR intensities (km/mol), I_R denotes Raman differential cross sections (10^{-36} m^2/sr). Only such contributions to the PED that were larger than 7% are presented. For definition of internal coordinates see Table 4.

believe that this improvement is not fortuitous, but it is connected with increased rigidity of the amino group in protonated cytosine. In this way, protonation improves the validity of the harmonic approximation for out-of-plane deformation vibrations of the amino group. The τNH2 (631 cm^{-1}) and γNH2 (535 cm^{-1}) vibrations are also notably decoupled by N3 protonation.

The good agreement among the spectra of crystalline samples, the spectra calculated for N3-protonated cytosine, and the spectra measured from acidic aqueous solution (pH = 1.7) demonstrates that protonation occurs in solution predominantly at the N3 position. However, to exclude the presence of a small amount (∼10%) of minor tautomeric forms of protonated cytosine, more detailed spectral examination (including the calculation of the spectra of minor tautomers) is needed.

The *Raman spectrum* of N3-protonated cytosine is dominated by strong bands at 788 and 1267 cm^{-1} (solution spectra). We interpret the 788 cm^{-1} band as a ring-breathing vibration. (In this mode all the bonds of the six-membered ring are stretching with the same phase, but the magnitudes of individual contributions largely differ). In contrast, the 1267 cm^{-1} band belongs to an out-of-phase stretching vibration of the bonds in the N1−C2−N3−C4 moiety. The Raman spectrum of N3-protonated cytosine is reminiscent of the spectrum of neutral cytosine, wherein these bands lie at 787 and 1290 cm^{-1}, respectively. Obviously, protonation at the N3 position has only a negligible effect on the breathing vibration. Accordingly, our calculation predicts zero protonation-induced frequency shift for this mode. The observed −23 cm^{-1} shift of the 1290 cm^{-1} band of neutral cytosine upon N3 protonation is also well reproduced by our SQM force field. The predicted shift amounts to −32 cm^{-1}, supposing that the character of this band is similar (out-of-phase ring stretching vibration with the dominant νC2N3 contribution) in both molecules. This frequency shift can be explained by weakening the C2N3 bond upon N3 protonation. The agreement of calculated and measured Raman intensities above 1300 cm^{-1} is also somewhat better for protonated than for neutral cytosine. In the context of our previous discussion, such an improvement might originate from better performance of the given basis set for cationic molecules. Indeed, the electron densities of positively charged systems generally exhibit less diffuse char-

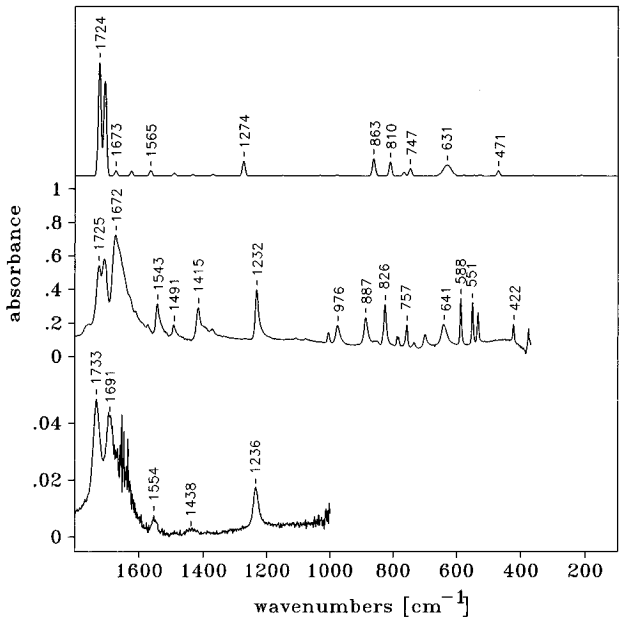

Figure 7. IR spectra of protonated cytosine. (Top) Calculated spectrum of N3-protonated cytosine (SQM B3-LYP/6-31G*). (Middle) Polycrystalline cytosine hydrochloride in nujol mull. (Bottom) FTIR spectrum of 0.05 M aqueous solution (pH = 1.7).

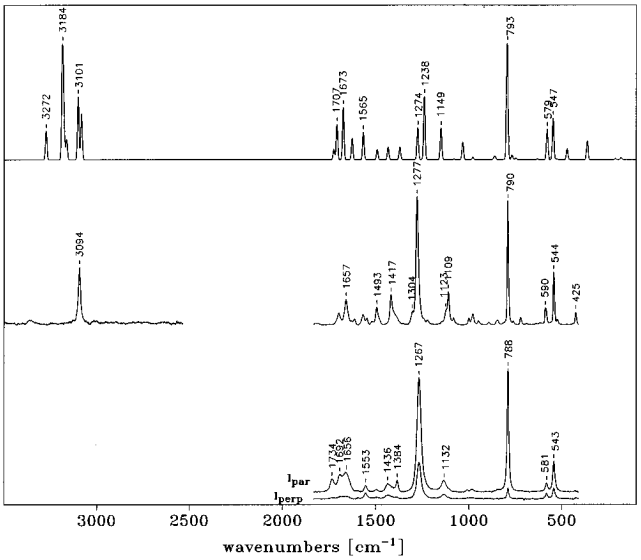

Figure 8. Raman spectra of protonated cytosine. (Top) Calculated spectrum of N3-protonated cytosine (determined from SQM B3-LYP/6-31G* force field and HF/6-31++G** polarizability derivatives). (Middle) Polycrystalline cytosine hydrochloride. (Bottom) Polarized spectra of aqueous solution (0.05 M, pH = 1.3).

acter. This property probably results in smaller basis set truncation errors obtained for the Raman spectrum of protonated cytosine.

In the 1600−1750 cm^{-1} frequency region, three medium Raman bands and three very strong IR bands were observed in the spectra of cytosine hydrochloride. When comparison with the IR spectrum of neutral cytosine is made, one notices that upon protonation vibrational bands are shifted to higher wavenumbers. The calculated force constant and frequency variations qualitatively agree with the observed frequency increase. In particular, the νCO and νC4N stretching and δNH2-scissoring force constants are increased by protonation. The largest protonation-induced shift occurs for the C═O stretching vibration, which is predicted to move to 1724 cm^{-1} in N3-protonated cytosine. This vibration contributes predominantly to the very strong IR band lying near 1730 cm^{-1}, seen also in the solution Raman spectrum of cytosine at pH = 1.7 (Figure 8).

TABLE 8: Prediction of the Vibrational Spectra of O2-Protonated Cytosine[a]

freq	I_{IR}	I_R	ρ	assignment[b]
194	3.1	3.7	0.75	τ6E1(75)
254	8.1	7.0	0.75	τ6Pu(40) + τ6E2(36)
357	0.5	28	0.74	δC4N(43) − δC2O(28)
494	16	17	0.75	τ6E2(46) + γC5H(16)
509	23	13	0.28	δC2O(46) + δC4N(33)
567	0.1	70	0.37	δ6E2(67)
595	1.7	77	0.26	δ6E1(60)
620	135	14	0.75	τOH(93)
634	117	4.8	0.75	τNH2(46) + γNH2(43)
712	178	2.2	0.75	τNH2(30) − γNH2(48)
759	67	0.2	0.75	γN1H(84)
801	7.6	249	0.06	breathing vibration
811	22	1.5	0.75	γC2O(29) − γC4N(21)
863	4.5	4.4	0.75	γC2O(36) + τ6Pu(28)
893	44	6.7	0.75	τ6Pu(36) + γC4N(29) − γC5H(23)
990	1.4	9.6	0.49	δNHR(30) + νC4C5(23) − νN1C2(18)
1014	3.2	55	0.10	δ6Tr(58)
1015	0.1	13	0.75	γC6H(71) − γC5H(20)
1083	49	27	0.07	δNHR(23) + δOH(19)
1149	0.6	32	0.40	δC5H(24)-νC6N1(22)
1227	223	15	0.13	δOH(45) − νC2O(18)
1269	16.3	29	0.40	δC6H(24) − δC5H(23)
1392	4.6	1.2	0.48	δC6H(25) + δC5H(16)
1408	31	100	0.40	νN3C4(19) − νC2N3(16)
1481	9.6	78	0.36	νC2O(16) − νC5C6(16)
1563	78	42	0.25	νC4N(20) − δNHS(16)
1570	99	13	0.05	νC2N3(19)
1653	446	66	0.66	δNHS(29) + δN1H(16)
1684	150	24	0.50	δNHS(32)
1711	641	88	0.23	νC5C6(28)
3084	3.2	95	0.49	νC5H(87)
3100	5.4	174	0.21	νC6H(86)
3177	404	5.8	0.26	νN1H(56)
3182	4.3	321	0.16	νN1H(43) + νNH9(35)
3278	194	95	0.22	νOH(100)
3284	104	73	0.73	νNH10(61) − νNH9(38)

[a] SQM B3LYP/6-31G(d) force field. Frequencies and IR and Raman intensities are given in cm^{-1}, km/mol, and 10^{-36} m^2/sr, respectively. ρ denotes the calculated depolarization ratio. [b] Only contributions to the PED larger than 15% are given.

The band at 1730 cm^{-1} was also observed in the resonance Raman spectra of cytidine monophosphate dissolved in acidic H_2O solution.[16] On the basis of the measured excitation profiles, its sensitivity to deuteration, and an increase in its relative intensity with increasing temperature, Purrello et al.[16] interpreted this band as originating from the νC2N3 stretching vibration of O2-protonated cytosine.

Our SQM B3-LYP calculation of the *vibrational spectrum of O2-protonated cytosine* (tautomer Cyt2t$^+$ in Figure 1) shows that this spectrum may indeed feature a vibrational band in the 1700−1730 cm^{-1} range (Table 8). However, our interpretation of this band (calculated at 1711 cm^{-1}) differs from that proposed by Purrello et al. The calculated D_2O frequency shift of the 1711 cm^{-1} band of Cyt2t$^+$ amounts to −10 cm^{-1}, which is close to the value of −15 cm^{-1} observed by Purrello et al. In contrast, the 1724 cm^{-1} band of N3-protonated cytosine is predicted to shift to 1691 cm^{-1} upon N deuteration. Among other vibrations, only the 1408 cm^{-1} band seems to be intense enough to be a likely candidate for being the Cyt2t$^+$ marker band in the Raman spectrum of an aqueous solution of protonated cytosine, wherein N3-protonated forms strongly dominate. Indeed, the weak shoulder, which could correspond to this marker, was measured at 1418 cm^{-1}. There is no fundamental vibration of Cyt3$^+$ that could be assigned to this shoulder. However, a similar 1408

cm^{-1} shoulder can be found in the spectrum of the crystal, in which the presence of $Cyt2^+$ tautomer is highly improbable.

Conclusions

The most probable protonation sites of cytosine involve the O2 and N3 atoms. The predicted gas-phase relative stabilities of these protonated forms slightly favor O2-protonated (enol) form. For an acidic aqueous solution modeled by polarized continuum, the equilibrium shifts significantly toward the N3-protonated form, but our calculations indicate that the enol tautomer of protonated cytosine may be present in observable quantities (0.2−8%). While this amount falls below the detection limit of the IR and Raman techniques we used, a recent Resonance Raman study[16] used the anomalous behavior of the marker band near 1730 cm^{-1} to prove the presence of the O2-protonated cytidine monophosphate in acidic aqueous solution. Indeed, a band lying at 1711 cm^{-1} was predicted by us for O2-protonated cytosine. However, vibrational analysis of N3-protonated cytosine reveals that vibrational bands of both tautomers should overlap in the 1700−1740 cm^{-1} range, which makes plausible interpretation of vibrational spectra in this region very difficult. Thus, further independent experimental examination of tautomeric equilibria of protonated cytosine in aqueous solution is needed.

We measured IR and Raman spectra of neutral and protonated cytosine in aqueous solution and the solid state. Also, we interpreted these spectra using the scaled HF/6-31G* and B3-LYP/6-31G* force fields, including intensity predictions. We have shown that scale factors can be determined in such a way that they can correct calculated spectra for the mean frequency shifts caused by hydrogen bonds present in the condensed phase. Using this methodology, we obtained frequency deviations between experimental and calculated frequencies of the same magnitudes as those observed between different crystals and aqueous solution. Thus, the accuracy limits of the scaled quantum mechanical methodology, which uses force fields determined for isolated molecules along with the internal-coordinate-dependent scale factors, were reached. The scale factors of the HF/6-31G* and B3-LYP/6-31G* force fields determined in this study can be used for scaling force fields of other cytosine derivatives and other nucleobases.

Acknowledgment. This work was supported in part by NIH Grant #332090, ONR Grant #N00014-95-1-0049, and by a contract (DAAH 04-95-C-0008) between the Army Research Office and the University of Minnesota for the Army High Performance Computing Research Center. The Mississippi Center for Supercomputing Research is acknowledged for a generous allotment of computer time. Kind help of Michael J. Stewart, Dr. Jiří Šponer, Dr. A. Tinti, Prof. A. Bertoluza, Prof. I. Barvík, P. Gajdůšek, Dr. J. Štěpánek, and Prof. A. Warshel is greatly appreciated. We also thank Dr. J. Zachová for growing the single crystals used in this study.

Supporting Information Available: SQM HF/6-31G* harmonic vibrational frequencies, IR and Raman intensities, Raman depolarization ratios, and band assignments for neutral cytosine. SQM B3-LYP/6-31G* harmonic vibrational frequencies, IR and Raman intensities, Raman depolarization ratios, and band assignments for neutral and N3-protonated cytosine residue and their N-deuterated isotopomers (5 pages). Ordering information is given on any current masthead page.

References and Notes

(1) Saenger, W. *Principles of Nucleic Acid Structure*; Springer Verlag: New York, 1984.
(2) Taylor, R.; Kennard, O. *J. Mol. Struct.* **1982**, *78*, 1.
(3) DelBene, J. E. *J. Phys. Chem.* **1983**, *87*, 367.
(4) Landridge, L.; Rich, A. *Nature* **1963**, *198*, 725.
(5) Hartman, K. A.; Rich, A. *J. Am. Chem. Soc.* **1965**, *87*, 2033.
(6) Inman, R. B. *J. Mol. Biol.* **1964**, *9*, 624.
(7) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059.
(8) Rajagopal, P.; Feigon, J. *Nature* **1989**, *339*, 637.
(9) Gehring, K.; Leroy, J.-L.; Guéron, M. *Nature* **1993**, *363*, 561.
(10) Chen, L.; Cai, L.; Zhang, X.; Rich, A. *Biochemistry* **1994**, *33*, 13540.
(11) Sowers, L. C.; Shaw, B. R.; Veigl, M. L.; Sendvick, W. D. *Mutat. Res.* **1987**, *177*, 201.
(12) Sowers, L. C.; Fazakerley, G. V.; Eritja, R.; Kaplan, B. E.; Goodman, M. F. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 5434.
(13) Hunter, W. N.; Brown, T.; Anand, N. N.; Kennard, O. *Nature* **1986**, *320*, 553.
(14) Wang, C.; Gao, H.; Gaffney, B. L.; Jones, R. A. *J. Am. Chem. Soc.* **1991**, *113*, 5486.
(15) Sponer, J.; Leszczynski, J.; Vetterl, V.; Hobza, P. *J. Biomol. Struct. Dyn.*, in press.
(16) Purrello, R.; Molina, M.; Wang, Y.; Smulevich, G.; Fossella, J.; Fresco, J. R.; Spiro, T. G. *J. Am. Chem. Soc.* **1993**, *115*, 760.
(17) Pulay, P.; Fogarasi, G.; Pongor, G.; Boggs, J. E.; Vargha, A. *J. Am. Chem. Soc.* **1983**, *105*, 7037.
(18) Fogarasi, G.; Pulay, P. In *Vibrational Spectra and Structure*; Durig, J. R., Ed.; Elsevier: Amsterdam, 1985; Vol. 14; p 125.
(19) Pongor, G.; Pulay, P.; Fogarasi, G.; Boggs, J. E. *J. Am. Chem. Soc.* **1984**, *106*, 2765.
(20) Pongor, G.; Fogarasi, G.; Boggs, J. E. *J. Am. Chem. Soc.* **1985**, *107*, 6487.
(21) Florián, J. *J. Mol. Struct. (THEOCHEM)* **1992**, *253*, 83.
(22) Florián, J.; Baumruk, V. *J. Phys. Chem.* **1992**, *96*, 9283.
(23) Susi, H.; Ard, J. S.; Purcell, J. M. *Spectrochim. Acta* **1973**, *29A*, 725.
(24) Putnam, B. F.; VanZandt, L. L. *J. Comput. Chem.* **1982**, *3*, 305.
(25) Letellier, R.; Ghomi, M.; Taillandier, E. *Eur. Biophys. J.* **1987**, *14*, 227.
(26) Lagant, P.; Derreumaux, P.; Vergoten, G.; Peticolas, W. L. *J. Comput. Chem.* **1991**, *12*, 731.
(27) Nishimura, Y.; Tsuboi, M. *Chem. Phys.* **1985**, *98*, 71.
(28) Szczesniak, M.; Kwiatkowski, J. S.; KuBulat, K.; Szczepaniak, K.; Person, W. B. *J. Am. Chem. Soc.* **1988**, *110*, 8319.
(29) Gould, I. A.; Vincent, M. A.; Hillier, I. H.; Lapinski, L.; Nowak, M. *Spectrochim. Acta* **1992**, *48A*, 811.
(30) Person, W. B.; Szczepaniak, K. In *Vibrational Spectra and Structure*; Durig, J. R., Ed.; Elsevier: Amsterdam, 1993; Vol. 20, p 239.
(31) Kwiatkowski, J. S.; Leszczynski, J. *J. Phys. Chem.* **1996**, *100*, 941.
(32) Nowak, M. J.; Lapinski, L.; Fulara, J. *Spectrochim. Acta* **1989**, *45A*, 229.
(33) Szczesniak, M.; Leszczynski, J.; Person, W. B. *J. Am. Chem. Soc.* **1992**, *114*, 2731.
(34) Lord, R. C.; Thomas, G. J., Jr. *Spectrochim. Acta* **1967**, *23A*, 2551.
(35) Hrouda, V.; Florián, J.; Polásek, M.; Hobza, P. *J. Phys. Chem.* **1994**, *98*, 4742.
(36) Zhang, Q.; Bell, R.; Truong, T. N. *J. Phys. Chem.* **1995**, *99*, 592.
(37) Estrin, D. A.; Paglieri, L.; Corongiu, G. *J. Phys. Chem.* **1994**, *98*, 5653.
(38) Hall, R. J.; Burton, N. A.; Hillier, I. H.; Young, P. E. *Chem. Phys. Lett.* **1994**, *220*, 129.
(39) Jeffrey, G. A.; Kinoshita, Y. *Acta Crystallogr.* **1963**, *16*, 20.
(40) Zachová, J.; Moravcová, M. *Czech. J. Phys.* **1977**, *B27*, 1402.
(41) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Wong, M. W.; Foresman, J. B.; Robb, M. A.; Head-Gordon, M.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 92/DFT, Revision G*; Gaussian, Inc.: Pittsburgh, PA, 1992.
(42) 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. J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94, Revision A.1*; Gaussian, Inc.: Pittsburgh, PA, 1995.
(43) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(44) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(45) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486.
(46) Wong, M. W.; Frisch, M. J.; Wiberg, K. B. *J. Am. Chem. Soc.* **1991**, *113*, 4776.
(47) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.
(48) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.
(49) Warshel, A. *J. Phys. Chem.* **1979**, *83*, 1640.

(50) Warshel, A.; Chu, Z. T. In *ACS Symposium Series: Structure and Reactivity in Aqueous Solution. Characterization of Chemical and Biological Systems*; Cramer, C. J., Truhlar, D. G., Eds.; American Chemical Society: Washington, DC, 1994; p 71.
(51) Luzhkov, V.; Warshel, A. *J. Comput. Chem.* **1992**, *13*, 199.
(52) Lee, F. S.; Chu, Z. T.; Warshel, A. *J. Comput. Chem.* **1993**, *14*, 161.
(53) Warshel, A. *Polaris Operation Manual, Version 6.2*; University of Southern California: Los Angeles, 1995.
(54) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
(55) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.
(56) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; McGraw-Hill: New York, 1955.
(57) Gwinn, W. D. *J. Chem. Phys.* **1971**, *55*, 477.
(58) Yamaguchi, Y.; Frisch, M.; Gaw, J.; Schaefer, H. F., III; Binkley, J. S. *J. Chem. Phys.* **1986**, *84*, 2262.
(59) Amos, R. D. *Chem. Phys. Lett.* **1986**, *124*, 376.
(60) Polavarapu, P. L. *J. Phys. Chem.* **1990**, *94*, 8106.
(61) Sundius, T. *J. Mol. Struct.* **1990**, *218*, 321.
(62) Florián, J. *SQMVIB: A computer program for calculation of scaled quantum mechanical vibrational spectra of molecules*; Charles University: Prague, 1994.
(63) Florián, J.; Leszczynski, J. *Int. J. Quantum. Chem.: Quantum. Biol. Symp.* **1995**, *22*, 207.
(64) Florián, J. *J. Phys. Chem.* **1993**, *97*, 10649.
(65) Kearley, G. J. *J. Chem. Soc., Faraday Trans. 2* **1986**, *82*, 41.
(66) Sponer, J.; Hobza, P. *J. Phys. Chem.* **1994**, *98*, 3161.
(67) Leszczynski, J. *Int. J. Quantum. Chem.: Quantum Biol. Symp.* **1992**, *19*, 43.
(68) McCarthy, W. J.; Lapinski, L.; Nowak, M. J.; Adamovicz, L. *J. Chem. Phys.* **1995**, *103*, 656.
(69) Bour, P.; Bednárová, L. *J. Phys. Chem.* **1995**, *99*, 5961.
(70) Rasanen, M. *J. Mol. Struct.* **1983**, *101*, 275.
(71) Kwiatkowski, J. S.; Leszczynski, J. S. *J. Mol. Struct.* **1993**, *297*, 277.
(72) Florián, J.; Johnson, B. G. *J. Phys. Chem.* **1995**, *99*, 5899.
(73) Florián, J.; Johnson, B. G. *J. Phys. Chem.* **1994**, *98*, 3681.
(74) Rauhut, G.; Pulay, P. *J. Phys. Chem.* **1995**, *99*, 3093.
(75) Florián, J.; Baumruk, V.; Strajbl, M.; Bednárová, L.; Stepánek, J. *J. Phys. Chem.* **1996**, *100*, 1559.
(76) Beetz, C. P., Jr.; Ascarelli, G. *Spectrochim. Acta* **1980**, *36A*, 1.
(77) Barker, D. L.; Marsh, R. E. *Acta Crystallogr.* **1964**, *17*, 1581.
(78) McClure, R. J.; Craven, B. M. *Acta Crystallogr.* **1973**, *B29*, 1234.
(79) Johnson, B. J.; Florián, J. *Chem. Phys. Lett.* **1995**, *247*, 120.
(80) Mandel, N. S. *Acta Crystallogr.* **1977**, *B33*, 1079.
(81) Gajdusek, P. MS-Thesis, Charles University, Prague, 1993.

JP953284W