Tautomerism in Cytosine and Uracil: A Theoretical and Experimental X-ray Absorption and Resonant Auger Study

Vitaliy Feyer,[†] Oksana Plekan,[†,‡] Robert Richter,[†] Marcello Coreno,[§] Monica de Simone,[||] Kevin C. Prince,*[,†,||] Alexander B. Trofimov,[⊥,#] Irina L. Zaytseva,[⊥] and Jochen Schirmer[▽]

Sincrotrone Trieste, in Area Science Park, I-34012 Basovizza (Trieste), Italy, CNR-IMIP, Montelibretti (Rome), I-00016 Italy, Laboratorio Nazionale TASC, INFM-CNR, 34012 Trieste, Italy, Laboratory of Quantum Chemistry, Irkutsk State University, 664003 Irkutsk, Russian Federation, Favorsky Institute of Chemistry, SB RAS, 664033 Irkutsk, Russia, and Theoretische Chemie, Physikalisch-Chemisches Institut, Universität Heidelberg, Im Neuenheimer Feld 229, D-69120 Heidelberg, Germany

Received: June 2, 2010; Revised Manuscript Received: August 3, 2010

The core level photoabsorption spectra of the nucleobases cytosine and uracil in the gas phase have been measured and the results interpreted with theoretical calculations using an ab initio Green's function approach. A single tautomer of uracil is populated, in agreement with previous work, while three tautomers of cytosine are clearly identified, whose identity and relative populations at the temperature of the experiment were reported previously. The second-order ADC approach to polarization propagator was employed in calculations of X-ray photoabsorption energies and intensities. The theoretical spectra have been constructed as Boltzmann-factor-weighted sums of individual tautomer spectra. These theoretical spectra are in good agreement with the experimental photoabsorption results at the oxygen, nitrogen, and carbon edges. In addition we report resonant Auger spectra of the valence band of cytosine, which support previous assignments of the character of the valence band states.

I. Introduction

The properties of the five nucleobases of DNA and RNA—adenine, thymine, guanine, cytosine, and uracil—are of fundamental chemical, physical, and biological importance, and we have recently published their core level photoelectron spectra[1−4] and some X-ray absorption spectra. In this paper we complete the data set with the core level X-ray absorption spectra of cytosine and uracil, which have not yet been published for the free molecules. A key issue addressed in our previous papers has been tautomerism: a single tautomer of uracil and three tautomers of cytosine are populated at the temperature of our experiment. In this work we characterize the electronic structure of cytosine and uracil via X-ray absorption and investigate the influence of tautomerism on the spectra.

In addition we report resonant Auger spectra of cytosine at the oxygen K edge. In resonant Auger spectroscopy, a core electron is excited to an empty orbital and the electron spectrum due to decay of the excited state is measured, for example, ref 5. The lowest energy states, which we concentrate on here, are those corresponding to a single hole in the valence band, that is, the states observed in direct photoemission from the valence band. Because the excitation is resonant, the relative intensities and Franck−Condon envelopes of the states differ from those observed in direct photoemission. In particular, interference occurs and is strongest for ionization of orbitals localized on the same site as the core hole excitation. Thus the technique serves as a method for identifying which molecular orbitals are strongly localized on the site of the core excitation.

II. Experimental and Theoretical Methods

Experimental Section. The experimental methods have been described in detail elsewhere.[1,2] As in that work, the evaporation temperatures were 450 K for cytosine and 405 K for uracil. The spectra were taken at the Gas phase Photoemission Beamline, Elettra, Trieste,[6] using the same apparatus. For the photoabsorption spectra, the resolution was 70, 60, and 100 meV at the C, N, and O edges, respectively. The energy scales of the spectra were calibrated as in[7] with the following calibrant gases and energies: 290.77 eV (C 1s → π, CO_2);[8] 400.87 eV (N 1s → π, N_2);[9] 535.4 eV (O 1s → π, CO_2).[10] The accuracy is estimated to be 100 meV at the carbon and nitrogen edges and 150 meV at the oxygen edge. Although the resonances of the calibrant gases are broad (about 0.7−1.2 eV) the centers of the structures can be located to high accuracy.

Theoretical Approach. Essentially the same theoretical approach as in our previous study of thymine and adenine[1] was employed, and we summarize only the main details. The energies (Ω) and the optical oscillator strengths (f) of the vertical K-shell excitations, corresponding to the near edge X-ray absorption fine structure (NEXAFS) experimental spectra, were evaluated using the so-called extended second-order algebraic-diagrammatic construction (ADC(2)) approximation scheme for the polarization propagator[11−13] also including the CVS approximation.[14] The 6-31+G basis sets[15,16] including diffuse basis functions were used in these calculations. The calculations were performed using the ADC polarization propagator code[17] interfaced to the GAMESS (US) program.[18]

The ground-state molecular structures of cytosine (Figure 1a) and uracil (Figure 1b) were obtained as previously.[2] The

* Corresponding author, Prince@Elettra.Trieste.It.
[†] Sincrotrone Trieste.
[‡] Present address: Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, 8000, Aarhus C, Denmark.
[§] CNR-IMIP.
[||] Laboratorio Nazionale TASC, INFM-CNR.
[⊥] Laboratory of Quantum Chemistry, Irkutsk State University.
[#] Favorsky Institute of Chemistry.
[▽] Theoretische Chemie, Physikalisch-Chemisches Institut, Universität Heidelberg.

10.1021/jp105062c
Published on Web 08/26/2010

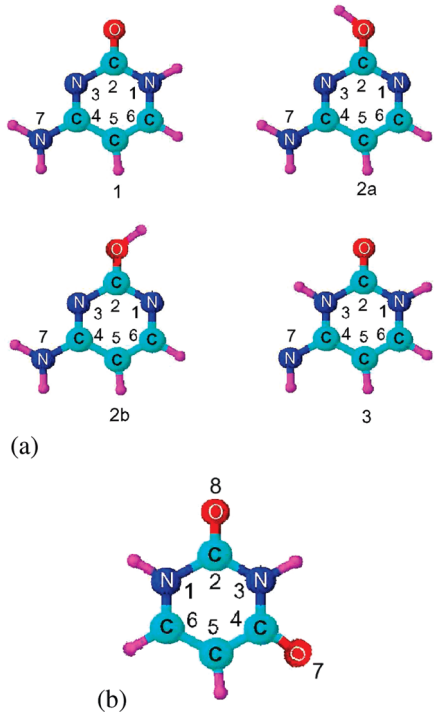

Figure 1. (a) Low-energy tautomers of cytosine and (b) canonical form of uracil.

interpretation of the ADC(2) polarization propagator calculations for the X-ray absorption spectra is greatly simplified when cytosine is considered as a planar molecule. For this reason the molecular structures of cytosine tautomers were constrained to C_s point group symmetry and constrained geometry optimization was performed. This approximation is considered to be an insignificant perturbation to the final K-shell excitation spectra. The B3LYP/6-311G** calculations were carried out using the GAUSSIAN program.[19]

The theoretical spectra of cytosine were constructed as Boltzmann population ratios (BPR) weighted sums of the spectra of all four tautomers. The BPRs were derived from the previous theoretical data[2,20,21] and extrapolated to the temperature of the experiment (see section II).

The theoretical intensities, $I = fBPR$, for photoabsorption were normalized to unit intensity for the strongest peak. The spectral envelopes were generated by convoluting the discrete transition lines (histograms) with Lorentzian functions. The values of the Lorentzian full widths at half-maximum (fwhm) were selected to best match the experimental intensities, and in an approximate manner they account for experimental resolution, unresolved vibrational structure, and natural lifetimes of the respective core-hole states.

The ADC(2) computations reported here are expected to afford an at least qualitatively good description of the excitation energies and spectral intensities in the low-energy regions of the 1s-electron excitation spectra in the cytosine and uracil nucleobases. An obvious limitation of the present theoretical model is the use of the relatively small 6-31+G basis set, enforced by limitations of the available ADC(2) code. The basis set error could (at least partly) be assessed by comparing the results obtained for the uracil N 1s excitations at the simpler ADC(2)-s (strict) level (refs 11−13) using both the 6-31+G and the larger 6-311+G*[22] basis sets (see Table 6 in the Supporting Information). The larger basis resulted in an essentially uniform lowering of the 1s−π* excitation energies by 1.5 eV and a similar, although somewhat smaller, uniform lowering of the 1s Rydberg excitations in the range of 1.2−1.3 eV.

The spectral intensities, by contrast, proved to be only modestly susceptible to basis set changes, at least in the low-energy part of the excitation spectrum. As is well understood, a sufficiently large and flexible basis set is indispensable for an adequate description of Rydberg-type excitations at higher energy. Clearly, here the use of a basis set larger than 6-31+G would be desirable.

III. Results: X-ray Photoabsorption of Cytosine and Uracil

Preliminary Considerations. For further discussion it is useful to consider the structure of the core−hole electron excitation manifold of cytosine and uracil. The approximation of planar molecular structure allows one to assign molecular orbitals and excited states using the irreducible representations of the C_s point group (the molecules are assumed to be in the xy plane) which is a great advantage with respect to the "no-symmetry" case.

The strongest spectral maxima can typically be explained by one or more core-to-valence transitions, many of which are vacant π* orbitals. The number of such orbitals (and corresponding valence states) can be calculated from the following considerations. In tautomers (1−3) of cytosine, eight atomic p_z orbitals lead to eight a″ MOs. The occupied MOs here comprise three π-orbitals due to double bonds and two MOs due to lone pairs. The three vacant π*-type MOs constitute three A″ states with electronic configurations labeled V_1−V_3 upon excitation of core electrons. Similarly one finds for uracil three π-type occupied A″ orbitals due to double bonds, two A″ MOs due to lone pairs, and three vacant π*-type MOs, which contribute three A″ core-excited states with configurations also labeled V_1−V_3.

The rest of the excitations can be considered as various members of Rydberg series, with varying admixtures of antibonding character. This admixture is usually most pronounced for low-lying σ*-type Rydberg states, while here all states of A′ symmetry are labeled as ns and $np_{x,y}$ Rydberg states, connected to specific core holes and converging to their respective ionization thresholds. The nd and higher angular momentum series are not considered in the present work due to the lack of appropriate diffuse basis functions in our calculations. Accordingly, only np_z-type Rydberg states can be expected among the A″ excited states. At this point it has to be recalled that the division of excitations into Rydberg and valence is not rigorous theoretically, and the assignments proposed in the present work are only approximate. Moreover, the analysis of A″ states is complicated by possible interaction of Rydberg and π*-valence electronic configurations. The cases where such interaction cannot be excluded are indicated by question marks in the Tables in the Supporting Information.

Near Edge X-ray Absorption Spectra of Cytosine: O 1s Spectrum. The O 1s excitation spectrum of cytosine, Figure 2, is rather similar to that of guanine.[3] The presence of the two maxima A and B in the experimental spectrum at about 532.0 and 535.4 eV indicate excitations of oxygen atoms in different chemical states. The two bands are assigned to the oxo and hydroxy tautomers, as confirmed by the present calculations which reproduce the experimental spectrum quite well, and are consistent with our O 1s photoemission data.[2]

As seen from Figure 2 (Table 1 of the Supporting Information), the oxo tautomers 1 and 3 are responsible for the low energy band A, while excitations of hydroxy tautomers 2a and 2b contribute to band B. In each case the main contributions

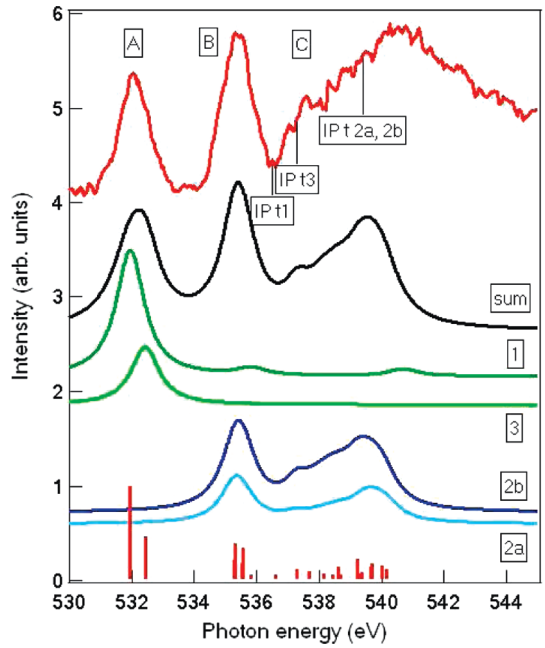

Figure 2. Experimental (top curve) and theoretical O 1s photoabsorption spectra of the tautomers of cytosine. Histogram: theoretical energy positions and relative oscillator strengths of all tautomers. Theoretical curves are shifted by −2.05 eV. Lorentzian fwhm = 1.1 eV.

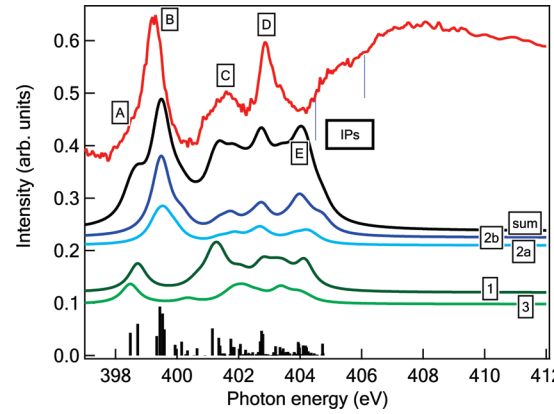

Figure 3. Experimental (upper curve) and theoretical N 1s photoabsorption spectra of cytosine. The range of the experimental ionization potentials (IPs) is indicated by vertical bars. Histogram: theoretical energy positions and relative oscillator strengths of all tautomers. Theoretical curves are shifted by −2.3 eV. Lorentzian fwhm = 0.7 eV.

are due to transitions O 1s → V_1, though a part of the spectral intensity of band B comes from the strong Rydberg-like excitations O 1s → 3s of tautomers 2a and 2b. In the tail of the experimental spectra, there is a weak feature C and this is reproduced by the calculations. This feature is reminiscent of one seen in the spectrum of formic acid (which contains oxo and hydroxy groups) and probably has the same origin.[23] The O 1s ionization thresholds are located in the NEXAFS spectrum between 536 and 540 eV. This part of the spectrum is not very well reproduced by our calculations because our method does not model continuum states. However qualitatively we may characterize this part of the O 1s spectrum as consisting of various Rydberg transitions with the most intense lines related to the hydroxy tautomers, 2a and 2b.

McNaughton et al.[24] calculated the spectra of the oxo form of cytosine and compared the spectra with solid state data. They predicted a single strong peak, as the present calculations do for tautomer 1. Since they did not consider the other tautomers, and were concerned with condensed cytosine, they did not predict the extra features due to tautomers 2a and 2b.

Near Edge X-ray Absorption Spectra of Cytosine: N 1s Spectrum. The N 1s NEXAFS spectrum of cytosine is shown in Figure 3, together with the calculated spectrum. The agreement between the experimental and theoretical spectral profiles is less satisfactory than in the O 1s and C 1s (see below) spectra, but still acceptable. The features A and B are well reproduced and the appearance of two distinct features is characteristic of the presence of both oxo and hydroxy tautomers. The broad feature C, which appears to consist of two peaks, is also present, but with different relative intensity in the theoretical spectrum. Similarly, there is a feature corresponding to D, but with reduced intensity and slightly shifted energy. The theoretical feature E is not observed, but it represents the features just below the first ionization potential, and as we do not calculate the continuum states, it is probably not a peak but simply the density of states just below the IP.

The N 1s spectrum contains distinctly more transitions than the O 1s excitation spectrum. The absolute values of oscillator strengths for the individual transitions are within the usual range but many valence type transitions appear to be equally strong, Table 2, Supporting Information. This increases the number of transitions that are important for the overall spectral envelope. In such a situation, already small shifts in the calculated excitation energies, e.g., due to basis set errors, will lead to significant changes in the shape of the spectral envelope, which explains the worse agreement of the N 1s experimental and theoretical spectra (for more details see the discussion of the N 1s excitations in uracil below). As seen from Figure 3, the spectra of all tautomers contribute to all parts of the theoretical spectrum. For a few bands however the contributions can be disentangled and an approximate assignment of certain features observed in the experiment becomes possible.

The lowest transitions N_7 1s → V_1 and N_3 1s → V_1 of tautomers 3 and 1, respectively, are responsible for maximum A of the theoretical spectrum which corresponds to the low energy shoulder A of the experimental band B. Peak B at about 399.20 eV of the experimental spectrum originates essentially from excitations N_3 1s → V_1 and N_1 1s → V_1 of both tautomers 2a and 2b. Interestingly, the calculations predict that the small conformational difference between 2a and 2b is sufficient to alter the relative oscillator strengths of these two main transitions, from 0.532/0.859 for 2a to 1.0/0.397 for 2b. Unfortunately the overlap of spectral features and vibrational envelopes do not permit us to verify this prediction experimentally. Nevertheless, the observation of the two features A and B is a clear signature of the presence of the two tautomeric forms, oxo and hydroxy.

Bands C−E of the experimental spectrum are very complex. Our calculations show these bands originate from strong overlap of all four tautomer spectra. For this reason only the low-energy flank of the experimental peak C at 401.64 eV can be assigned confidently. Our calculations indicate that this part of the spectrum is shaped mainly by excitations of tautomer 1, where the most important transitions involve the atoms N_7 and N_1 and the final states V_1 and 3s. The higher energy part of feature C is dominated by contributions from tautomer 3. This feature is the only one which provides some evidence for separate spectral signatures of the tautomers 1 and 3.

The higher energy regions of the N 1s X-ray absorption spectrum can be assigned only cautiously. Band D of the experimental spectrum at about 402.90 eV can be tentatively assigned to N_7 1s → 3p transitions of tautomers 1, 2a, and 2b,

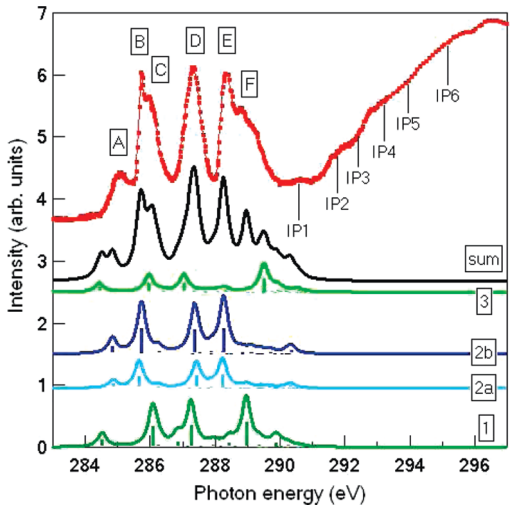

Figure 4. Experimental (upper curve) and theoretical (b) C 1s photoabsorption spectra of cytosine. Theoretical curves are shifted by −2.65 eV. Experimental ionization potentials are indicated. Lorentzian fwhm = 0.32 eV.

while its high energy shoulder E at about 403.35 eV can be attributed to transitions N_3 1s → V_2 and N_1 1s → V_2 of tautomers 1 and 3, respectively.

As for the oxygen edge, MacNaughton et al.[24] calculated the N K edge spectrum of one oxo tautomer of cytosine and compared it with solid state data. The solid state spectrum shows a much stronger peak A compared with the gas phase but much less detail at higher energy. Our calculations correctly predict the strong peak B due to the oxo tautomers, but it is difficult to make comparisons as at higher energies, solid state effects appear to broaden the structures.

Concluding our discussion of the N 1s spectrum, we note that although the spectrum is rather complex the tautomerism of cytosine is moderately well represented here. The explicit signatures of different tautomers here are shoulder A (tautomers 1 and 3), peak B (tautomers 2a and 2b), and also the shape of band C (tautomers 1 and 3 as separate contributions).

Near Edge X-ray Absorption Spectra of Cytosine: C 1s Spectrum. The agreement of the C 1s NEXAFS spectrum of cytosine, Figure 4, with theory is much better than that of the N 1s spectrum, and in particular, the six salient features A−F are well reproduced. The structure of the spectrum can be qualitatively understood as follows. The four nonequivalent carbon atoms C_5, C_6, C_4, and C_2 give rise to the four main bands A, B−C, D, and E−F, respectively, where the ordering of the excited states is subject to chemical shift arguments. Such a view of the C 1s spectrum is of course oversimplified and the actual situation is more intricate as can be seen from the results of our calculations in Table 3 in Supporting Information and Figure 4. Also, in general each of the cytosine tautomers contributes to all C 1s excitation bands and there are no bands due to individual tautomers. This is because the valence configurations of carbon do not change from one tautomer to another and the C 1s excitation spectra of different tautomers are similar, although some assignments to individual tautomers seem to be possible as demonstrated below on the basis of our theoretical results.

The experimental band A at 285.04 eV comprises excitations C_5 1s → V_1 of all four tautomers. The low-energy flank of band A should be more influenced by excitations of tautomers 1 and 3, while the intensity of its high-energy flank should be more dependent on tautomers 2a and 2b, but the experimental splitting is not as great as that predicted by theory. The relatively low overall spectral intensity of band A is due to the low values of the optical oscillator strengths of the C_5 1s → V_1 valence-type transitions. The latter can be explained by the specific properties of the V_1 state in which the excited electron populates π^* orbitals with significant localization at the C_6 site.

Band B with maximum at about 285.70 eV and its shoulder C at about 285.98 eV are also related to all four tautomers. Here the main contribution to spectral intensity is due to the transition C_6 1s → V_1. The high-energy flank of shoulder C also contains contributions of C_5 1s → 3s Rydberg excitations. Theory shows that peak B can be attributed to tautomers 2a and 2b, whereas shoulder C is attributed to tautomers 1 and 3.

The next band E−F has its maximum E at about 288.34 eV and a shoulder F at 288.78. The theoretical results reproduce the main features of band E−F though feature F appears in the theoretical envelope as two distinct peaks rather than as a shoulder. Peak E is essentially due to excitations C_2 1s → V_1 of tautomers 2a and 2b, whereas shoulder F is due to the same excitation of tautomers 1 and 3. A large number of less intense Rydberg and valence transitions can be identified within these spectral regions on the background of the C_2 1s → V_1 excitations but all these transitions change the overall spectral profile only slightly, making it more diffuse.

Both the experimental and theoretical spectra are in good overall agreement with the results of MacNaughton et al.,[24] although peak F is missing in their work. Thus it appears that the C K edge is less sensitive to the effects of tautomerism.

The above spectra can be compared with those of cytosine in the solid state,[25] where cytosine exists as tautomer 1.[26] The carbon edge spectrum is quite similar with the peaks A, B, D, and F clearly preserved, while C becomes a shoulder. Energies agree to within a few tenths of an electronvolt. Our peak E is suppressed, consistent with the prediction that it is substantially due to tautomers 2a and 2b. At the nitrogen edge, two peaks were observed in the solid state at 399.1 and 400.4 eV, compared with our theoretical prediction for tautomer 1 of 398.7 and 400.65 eV (see Supporting Information). Zubavichus et al. did not comment on a broad structure in their spectrum at about 403 eV, but this is also predicted by our calculations. Finally at the O edge, a single peak at 532.2 eV was observed, in good agreement with our prediction of 531.95 eV for tautomer 1. Overall, the features observed in condensed cytosine agree well with the predicted spectra for tautomer 1, and the differences in energy of a few tenths of an electronvolt can be attributed to effects such as hydrogen bonding in the solid state.

Core Excitations of Uracil. For uracil only one tautomer is populated, so the interpretation of its near edge X-ray absorption spectra is easier. The electronic structure of uracil is rather similar to that of thymine, studied previously.[1] Here we only very briefly describe the assignment of the O 1s, N 1s, and C 1s NEXAFS spectra of uracil.

In the O 1s spectrum, the two intense peaks A and B at 531.37 and 532.35 eV of the experimental O 1s excitation spectrum, Figure 5 (Table 4 of Supporting Information), are assigned to the valence transitions O_7 1s → V_1 and O_8 1s → V_1, respectively. The next broad band C with maximum at about 535.38 eV comprises various excitations from the 1s orbitals of the O_7 and O_8 atoms to Rydberg states. Like thymine, the O 1s photoemission spectrum of uracil shows a single peak for the two chemically distinct oxygen atoms, and like thymine, the NEXAFS spectrum shows well-resolved resonances.

The experimental N 1s NEXAFS spectrum of uracil, Figure 6, is in only fair agreement with theory at the ADC(2)/6-31+G

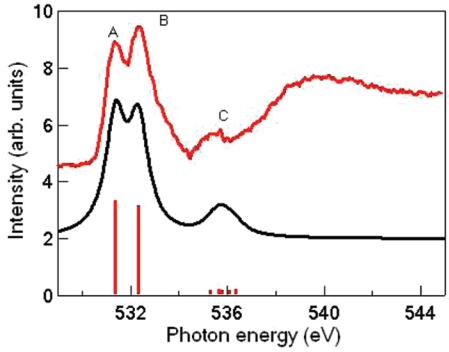

Figure 5. Experimental (upper curve) and theoretical (lower curve) O 1s photoabsorption spectrum of uracil. Histogram: theoretical energy positions and relative oscillator strengths. Theoretical curve is shifted by −2.15 eV, Lorentzian fwhm = 0.5 eV.

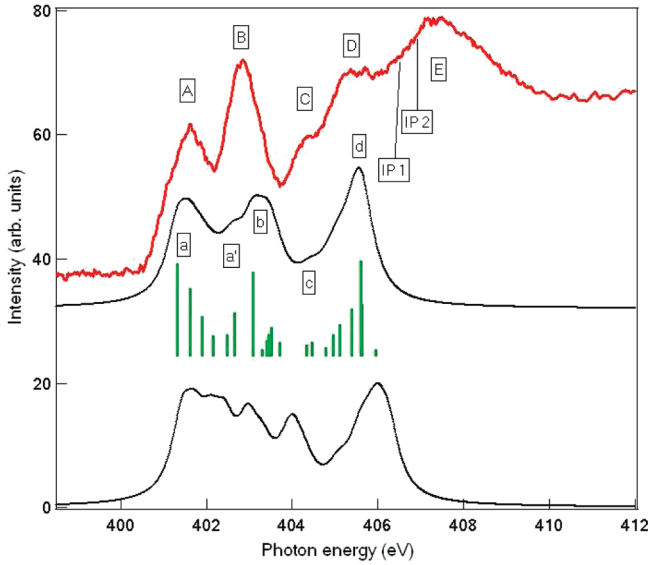

Figure 6. Experimental (upper curve) and theoretical ADC(2)s/6-311+G* (middle curve) and ADC(2)/6-31+G (lower curve) N 1s photoabsorption spectra of uracil. Histogram: theoretical resonance energies and relative oscillator strengths. Theoretical curves are shifted by −3.80 eV (ADC(2)s/6-311+G*) and by −2.5 eV (ADC(2)/6-31+G). Lorentzian fwhm = 0.6 eV.

level. As in the case for cytosine, there are groups of closely lying transitions having comparable oscillator strengths. This means that the shape of the spectral profile is very susceptible to small errors in the individual excitation energies, which may be caused, for example, by basis set limitations. To demonstrate the basis set effects, we have performed calculations at the methodologically simpler ADC(2)s (strict) level, where the larger 6-311+G* basis set could be used, being of triple-ζ quality augmented by polarization and diffuse functions.[22,16] The corresponding spectral profile is seen to be in much better agreement than the small basis set result. It should be noted, though, that some energy and intensity modulations are due to the different ADC(2) variants (for details see Tables 5 and 6 in the Supporting Information).

The ADC(2)s/6-311+G* results can readily be used to assign the experimental data (see Table 5 in the Supporting Information). With regard to the results obtained at the ADC(2)/6-31+G level (Figure 6 and Table 6 in the Supporting Information), we note that the basic composition of the spectrum is consistent with that of the large basis results. However, there are small energy shifts and also some intensity modulations, leading to distinct modifications in the spectral envelope.

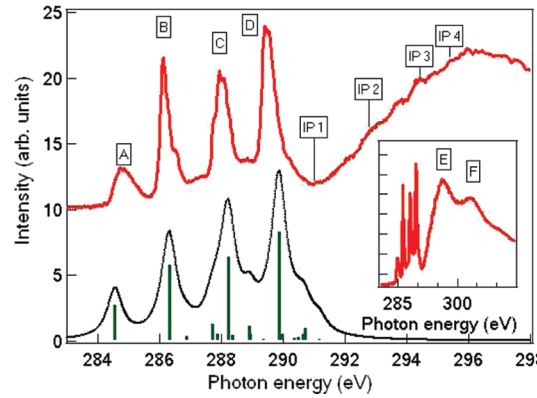

Figure 7. Experimental (upper curve) and theoretical (lower curve) C 1s photoabsorption spectra of uracil. Histogram: theoretical energy positions and relative oscillator strengths. Inset: broad scan, showing resonances above threshold. Theoretical curves are shifted by −2.46 eV. Lorentzian fwhm = 0.55 eV.

According to the ADC(2)s/6-311+G* results (see Table 5 in the Supporting Information), the maximum A observed in the experiment at about 404.4 eV arises from a group of transitions comprising the valence-type excitations V_1 and 3s Rydberg excitation from the 1s orbitals of the N_1 and N_3 atoms.

The next maximum B at about 402.62 eV in the experimental spectrum is due to excitations from the 1s orbitals of N_1 and N_3 atoms to the valence V_2 and various 3p Rydberg levels. Also the N_3 1s → 4s transition contributes to peak B. The N_1 1s → 4s transition together with the two N_1 1s → 4p transitions can be attributed to shoulder C, seen in the experimental spectrum at about 404.1 eV. Finally, experimental maximum D at about 405.8 eV is due to valence V_3 and various 5s, 5p Rydberg transitions.

The C 1s NEXAFS spectrum of uracil, Figure 7, is broadly similar to that of cytosine, and can be understood very well already at a quite basic theoretical level. The four bands A−D observed in the experimental spectrum are associated with the four nonequivalent carbon atoms of uracil, C_5, C_6, C_4, C_2, respectively. The ordering of the C 1s excitations is defined by the respective chemical shifts, which in turn, depend in a rather clear manner on the local environment of the carbon atoms in the molecule.[1] More specifically, the spectral intensity of bands A−D results from the valence transitions C 1s → V_1 responsible for the experimental peaks at 284.79, 286.11, 288.00, 289.50 eV, respectively (Figure 7 and Table 7 of Supporting Information). The less intense transitions to other final states give rise to minor features observed in the experiment. Shoulders on the low energy side of C at 287.78 and the high energy side of D at 290.24 eV originate essentially from the valence transitions C_5 1s → V_2 and C_2 1s → V_2, respectively. Also, our results indicate than many additional weak contributions mainly of Rydberg nature can be identified within the areas of bands C and D. The region above threshold is shown in the inset of Figure 6, and two shape resonances, E and F are clearly identified.

As for cytosine, our spectra can be compared with those of condensed uracil.[25] All C K edge peaks are reproduced with an average difference in energy of about 200 meV. The N K edge features A and B are reproduced although in the solid state there appears to be a rigid shift of half an electronvolt to lower energy. Although not commented on by Zubavichus et al., the feature D at 405 eV is also present in their spectra and predicted by our calculations. For oxygen we observed a double peak structure whose average energy was 531.81 eV, whereas in the

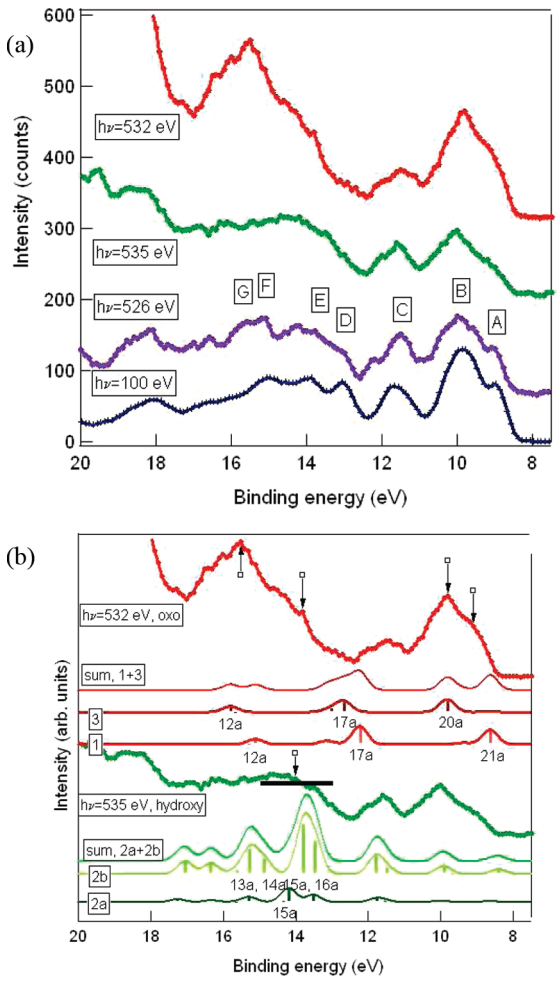

Figure 8. (a) Resonant Auger spectra of the valence band of cytosine at the photon energies indicated: 100 eV, far below resonance; 526 eV, just below resonance; 535 eV, at the O 1s → V_1, 3s (OH) resonance; 532 eV, at the O 1s → V_1 (C═O) resonance. (b) Resonant Auger spectra at 532 and 535 eV. Below each spectrum are histograms and Gaussian broadened curves for each tautomer (1, 2a, 2, 3), equal to the product of the Mulliken populations on the oxygen atoms, multiplied by the population of the respective tautomer. For the 532 eV spectrum only A″ states are included. Vertical arrows indicate positions of resonances. The horizontal line represents the approximate spread of the resonant enhancement for the hydroxy tautomers.

solid state a single peak at 532.0 eV was observed. It appears that in the solid state, broadening due to hydrogen bonding and different screening blurs the spectrum and the two peaks are not resolved. However the agreement within 200 meV is completely acceptable and the overall the agreement is very satisfactory.

IV. Resonant Auger Experiment

As discussed above, peak A in the O 1s photoabsorption spectrum of cytosine (Figure 2) is due to O 1s → V_1 transitions of the oxo tautomers, and peak B is due to O 1s → V_1 and O 1s → 3s transitions of tautomers 2a and 2b. The resonant Auger spectra in Figure 8a were measured at the energy of these two resonances, while the other two spectra were measured off-resonance at 100 eV (far below resonance) and at 526 eV (just below resonance). The photoelectron energy scale was calibrated by adding a small quantity of neon and measuring the Ne 2p emission. The valence band spectrum has previously been measured and calculated theoretically by Trofimov et al.,[27] and the present spectrum at 100 eV is in agreement with that work. The peaks are labeled following that work.

In Figure 8b the measured spectra are compared with curves derived from the present theoretical calculations. A full calculation of the resonant Auger spectra for all four tautomers is beyond the scope of this work. Instead we have taken a simplified approach similar to that of Piancastelli et al.,[28,29] aimed at locating the energies at which we expect resonances to occur. The Mulliken populations were calculated with the 6-31G basis set and the true (nonplanar) B3LYP/6-311G** geometry; see Tables 8−11 in Supporting Information. The orbital energies calculated at this level do not agree well with experimental values, so the energies calculated at the OVGF level were used; these calculations are accurate for the outer valence region with which we are concerned. The Mulliken populations projected on the oxygen atoms were multiplied by the tautomer populations and are shown in Figure 8b as histograms and curves (Gaussian broadened histograms.) This is a rough method of estimating the energy position of resonances, but it should be remembered that symmetry plays a role. Selection rules restrict the number of orbitals which resonate; for example in the case of C_s symmetry, only orbitals of the same symmetry as the resonance resonate. To take account of this, for the 532 eV spectrum we have included only orbitals of a″ (π like) symmetry. The resonance at 535 eV contains overlapping contributions of A′ and A″ symmetry states, and so we have included all possibilities. In practice, vibronic coupling may break the selection rules. The deviation of the molecular structure of cytosine from C_s symmetry in the initial state, which makes the selection rules less strict, should also be remembered.

Feature C at about 11 eV appears not to resonate, and so we have approximately normalized the intensity in all spectra to this feature. At the resonance energy of the oxo form, 532 eV, the features A and B (9 and 9.8 eV) clearly increase in intensity, Figure 8a. The orbitals which contribute to the resonance in this region are 21a (the HOMO, also denoted π_5) of tautomer 1, and orbital 20a (HOMO-1) of tautomer 3, Figure 8b.

The next strong resonance occurs at about the energy of feature *E*, 14−15 eV, but there are no orbitals with both the required symmetry and sufficient Mulliken population in this range. The resonance may be due to orbital 15a of tautomer 1, which is of a′ symmetry. An even stronger resonance occurs at the energy of peak G and is assigned to directly excited two-hole one-particle states, or so-called spectator states, which have an onset at 15 eV. The very strong intensity at binding energy higher than 18 eV is certainly due to spectator states.

On excitation of the O 1s of the hydroxy tautomers at 535 eV, a weak resonant enhancement of intensity is observed in the energy range of 13−15 eV (peaks D, E, and F), and a stronger enhancement at greater than 18 eV binding energy, Figure 8a. We interpret the spectra in terms of the orbital character of tautomers 2a and 2b.[27] Features A and B do not resonate because they consist of orbitals with low Mulliken populations projected on oxygen. Feature C contains orbitals 17a and 18a of tautomer 2b, which have significant oxygen character, but despite this, the feature does not appear to resonate measurably. The resonance in the 13−15 eV range is associated with the orbitals 15a and 16a of tautomer 2b, and possibly also 13a and 14a. Although they have quite strong oxygen character, the resonant increase in intensity is modest. The enhancement at binding energy higher than 18 eV is again assigned to two-hole, one-particle states.

V. Conclusions

The experimental and theoretical X-ray photoabsorption spectra of cytosine and uracil have been reported, and the main resonances assigned. The spectra are consistent with a single tautomer of uracil and four tautomers of cytosine being populated at the experimental temperatures. However the NEXAFS spectra mostly provide spectral fingerprints of oxo and hydroxy populations, and it is not possible to distinguish the two oxo forms 1 and 3, or the two hydroxy conformers, 2a and 2b. Resonant Auger spectra of cytosine at the O 1s edge have been interpreted in terms of the Mulliken populations of the valence orbitals projected on the oxygen atoms, and the main features have been assigned. The spectra are consistent with calculated valence molecular orbitals and the oxo tautomers give a clear fingerprint of their presence in the spectrum.

Acknowledgment. The theoretical part of this study was supported by a grant of the Russian Foundation for Basic Research (RFBR). We gratefully acknowledge the assistance of our colleagues at Elettra for providing good quality synchrotron light. O. Plekan acknowledges financial support from the Area di Ricerca di Trieste under the Incoming Mobility scheme. V. Feyer thanks the Abdus Salaam International Center for Theoretical Physics for a TRIL (Training and Research in Italian Laboratories) fellowship.

Supporting Information Available: A complete ref 19; a list of calculated energies, pole strengths, and intensities of the vertical O 1s, N 1s, and C 1s excitations of cytosine tautomers and uracil; and Mullikan populations of the outer valence orbitals of tautomers of cytosine. This material is available free of charge via the Internet at http://pubs.acs.org/.

References and Notes

(1) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K. C.; Trofimov, A. B.; Gromov, E. V.; Zaytseva, I. L.; Schirmer, *J. Chem. Phys.* **2008**, *347*, 360–375.

(2) Feyer, V.; Plekan, O.; Richter, R.; Coreno, M.; Vall-llosera, G.; Prince, K. C.; Trofimov, A. B.; Zaytseva, I. L.; Moskovskaya, T. E.; Gromov, E. V.; Schirmer, J. *J. Phys. Chem. A* **2009**, *113*, 5736–5742.

(3) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; Vall-llosera, G.; Prince, K. C.; Trofimov, A. B.; Zaytseva, I. L.; Moskovskaya, T. E.; Gromov, E. V.; Schirmer, J. *J. Phys. Chem. A* **2009**, *113*, 9376–9385.

(4) Zaytseva, I. L.; Trofimov, A. B.; Schirmer, J.; Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; Prince, K. C. *J. Phys. Chem. A* **2009**, *113*, 15142.

(5) Fink, R. H.; Sorensen, S. L.; Naves de Brito, A. *J. Chem. Phys.* **2000**, *112*, 6666.

(6) Prince, K. C.; Blyth, R. R.; Delaunay, R.; Zitnik, M.; Krempasky, J.; Slezak, J.; Camilloni, R.; Avaldi, L.; Coreno, M.; Stefani, G.; Furlani, C.; de Simone, M.; Stranges, S. *J. Synchrotron Radiat.* **1998**, *5*, 565–568.

(7) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K. C.; Carravetta, V. *J. Electron Spectrosc. Relat. Phenom.* **2007**, *155*, 47–53.

(8) Tronc, M.; King, G. C.; Read, F. H. *J. Phys. B: At., Mol. Phys.* **1979**, *12*, 137. ibid. **1980**, *13*, 999.

(9) Sodhi, R. N. S.; Brion, C. E. *J. Electron Spectrosc. Relat. Phenom.* **1984**, *34*, 363.

(10) Wight, G. R.; Brion, C. E. *J. Electron Spectrosc. Relat. Phenom.* **1974**, *3*, 191.

(11) Schirmer, J. *Phys. Rev. A* **1982**, *26*, 2395.

(12) Barth, A.; Schirmer, J. *J. Phys. B: At., Mol. Phys.* **1985**, *18*, 867.

(13) Trofimov, A. B.; Moskovskaya, T. E.; Gromov, E. V.; Vitkovskaya, N. M.; Schirmer, J. *J. Struct. Chem.* **2000**, *41*, 483.

(14) Cederbaum, L. S.; Domcke, W.; Schirmer, J. *Phys. Rev. A* **1980**, *22*, 206–222.

(15) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261.

(16) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; v. R. Schleyer, P. *J. Comput. Chem.* **1983**, *4*, 294–301.

(17) Polarization propagator ADC code written by A. B. Trofimov, G. Stelter, and J. Schirmer.

(18) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.

(19) Frisch, M. J., et al. *Gaussian 98, Revision A.7*; Gaussian, Inc.: Pittsburgh, PA, 1998.

(20) Trygubenko, S. A.; Bogdan, T. V.; Rueda, M.; Orozco, M.; Luque, F. J.; Sponer, J.; Slavıček, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4192.

(21) Fogarasi, G. *J. Phys. Chem. A* **2002**, *106*, 1381.

(22) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(23) Prince, K. C.; Richter, R.; de Simone, M.; Coreno, M. *J. Phys. Chem. A* **2003**, *107*, 1955.

(24) MacNaughton, J.; Moewes, A.; Kurmaev, E. Z. *J. Phys. Chem. B* **2005**, *109*, 7749–7757.

(25) Zubavichus, A.; Shaporenko, V.; Korolkov, M.; Grunze; Zharnikov, M. *J. Phys. Chem. B* **2008**, *112*, 13711–13716.

(26) Barker, D. L.; Marsh, R. E. *Acta Crystallogr.* **1964**, *17*, 1581–1587.

(27) Trofimov, A. B.; Schirmer, J.; Kobychev, V. B.; Potts, A. W.; Holland, D. M. P.; Karlsson, L. *J. Phys. B: At., Mol. Opt. Phys.* **2006**, *39*, 305.

(28) Piancastelli, M. N.; Lischke, T.; Prümper, G.; Liu, X. J.; Fukuzawa, H.; Hoshino, M.; Tanaka, T.; Tanaka, H.; Harries, J.; Tamenori, Y.; Bao, Z.; Travnikova, O.; Céolin, D.; Ueda, K. *J. Electron Spectrosc. Relat. Phenom.* **2007**, *156−158*, 259–264.

(29) Piancastelli, M. N.; Céolin, D.; Travnikova, O.; Bao, Z.; Hoshino, M.; Tanaka, T.; Kato, H.; Tanaka, H.; Harries, J. R.; Tamenori, Y.; Prümper, G.; Lischke, T.; Liu, X. J.; Ueda, K. *J. Phys. B: At., Mol. Opt. Phys.* **2007**, *40*, 3357–3365.

JP105062C