Ionization energies of aqueous nucleic acids: Photoelectron spectroscopy of pyrimidine nucleosides and ab initio calculations

Petr Slavíček,[1] Bernd Winter,[2*] Manfred Faubel,[3] Stephen E. Bradforth,*[4] and Pavel Jungwirth[5*]

[1]*Department of Physical Chemistry, Institute of Chemical Technology, Technická 5, Prague 6, Czech Republic and J. Heyrovský Institute of Physical Chemitsry, Academy of Sciences of the Czech Republic, Dolejškova 3, 18223 Prague 8, Czech Republic*

[2]*Helmholtz-Zentrum Berlin für Materialien und Energie, Albert-Einstein-Strasse 15, D-12489 Berlin, Germany, and Max-Born-Institut für Nichtlineare Optik und Kurzzeitspektroskopie, Max-Born-Strasse 2A, D-12489 Berlin, Germany*

[3]*Max-Planck-Institut für Dynamik und Selbstorganisation, Bunsenstrasse 10, D-37077 Göttingen, Germany.*

[4]*Department of Chemistry, University of Southern California, Los Angeles, California 90089, USA.*

[5]*Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, and Center for Biomolecules and Complex Molecular Systems, Flemingovo nám. 2, 16610 Prague 6, Czech Republic.*

[*]*Corresponding authors: bernd.winter@bessy.de (B.W.), stephen.bradforth@usc.edu, and Pavel.Jungwirth@uochb.cas.cz (P.J.)*

Abstract

Vertical ionization energies of the nucleosides cytidine and deoxythimidine in water, the lowest ones amounting in both cases to 8.3 eV, are obtained from photoelectron spectroscopy measurements in aqueous microjets. Ab initio calculations employing a non-equilibrium polarizable continuum model quantitatively reproduce the experimental spectra and provide molecular interpretation of the individual peaks of the photoelectron spectrum showing also that lowest ionization originates from the base. Comparison of calculated vertical ionization potentials of pyrimidine bases, nucleosides, and nucleotides in water and in the gas phase underlines the dramatic effect of bulk hydration on the electronic structure. In the gas phase, the presence of sugar and, in particular, of phosphate has a strong effect on the energetics of ionization of the base. Upon bulk hydration, the ionization potential of the base in contrast becomes rather insensitive to the presence of the sugar and phosphate, which indicates a remarkable screening ability of the aqueous solvent. Accurate aqueous phase vertical ionization potentials provide a significant improvement to the corrected gas phase values used in the literature and represent important information in assessing the threshold energies for photooxidation and oxidation free energies of solvent-exposed DNA components. Likewise, such energetic data should allow improved assessment of delocalization and charge hopping mechanisms in DNA ionized by radiation.

Introduction

One of the principal effects on cells of high-energy radiation such as that encountered in radiation therapy or from cosmic rays is the ionization of DNA (1). Similarly, high energy photons can directly ionize a base within DNA; in both cases the impact of the radiation is to lead to an oxidized center or "hole". While any DNA component can be initially ionized, typically, the hole ends up localized on a guanine base as it bears the lowest *in situ* oxidation potential (2). The early processes leading to such radiation induced damage are not well understood (3-6). Surprisingly this includes primary experimental information on the ionization energetics of each of the four constituent bases of aqueous DNA (7,8). Moreover, in double stranded DNA, the sugar phosphate backbone is the most exposed to the solvent. Direct ionization, as well as oxidative damage by hydroxyl radicals also produced by ionizing radiation (1) can occur at this locus, too and eventually leads to strand breaks (9). Therefore, the ionization energies from not only the molecular orbitals on the bases, but also on the sugar and phosphate backbone components of nucleotides are relevant for unraveling mechanisms of DNA damage.

In order to understand charge mobility along the double helix, determinations of both the vertical and adiabatic (or threshold) ionization potentials of each of the bases within realistic environments are required. Attention has focused on establishing the ionization threshold in solution and whether direct one-photon oxidation can occur at the edge of the solar UV-B region. Deoxyguanosine, which contains the most readily ionized base guanine, is the only system with a clear-cut measurement of the ionization threshold. Using low intensity illumination and electron scavenging, LeBreton and coworkers established a threshold at ~4.8 eV (10). In laser excitation experiments, because the electron and ionization product yields are very low, high intensity excitation is needed, thus requiring difficult separation of one versus two-photon

processes (12,13). Most thresholds currently cited actually come from semi-empirical estimates (8) based on gas-phase vertical ioniziation potentials (IP) and the application of large thermodynamic corrections. Alternatively, calculations starting with energy changes in vacuum and then employing various reaction field and empirical corrections for the hydration of products have been published (7). Direct measurement of the solution vertical IP as presented here can be expected to substantially improve threshold energies estimates for DNA bases.

Early gas-phase photoelectron spectra of the bases by Hush and Cheung (14) as well as of a range of base (15-17) and nucleoside (18) analogs have been reported by LeBreton and coworkers and Kubota and Kobayashi (19). The most relevant to the current study are those reported by Yu *et al.* for thymine and cytosine bases and *O*-methyl derivatives of deoxythymidine and cytidine nucleosides (18). There have been several more recent measurements exploring microhydration on the ionization energy of bases in gaseous molecular beams. Rather than the photoelectron spectrum which reports on electron removal from each of the highest-lying molecular orbitals, these experiments only provide the appearance threshold for observing cations in a mass spectrometer (20-22). Recently, Wang and coworkers have succeeded in getting all four deoxynucleotides (*d*NMPs) into the gas phase and recording their photoelectron spectra (23). Their supporting hybrid density functional calculations indicate that for *d*NMPs in vacuum, the lowest ionization potential arises for removing an electron from the negatively charged phosphate, except for $dGMP^-$ where the base orbital is the lowest (23). Using a partial third order self-energy approximation method, Ortiz *et al.* concluded, that states originating from ionization of phosphate and the base are almost degenerate for $dAMP^-$ and $dTMP^-$, while the lowest ionization of $dCMP^-$ is phosphate centered (24). Serrano-Andres and coworkers then demonstrated that at the CASPT2 level the lowest ionization of $dTMP^-$ comes

from the base, with the state originating from phosphate ionization lying only 0.2 eV higher (25). Additionally, these computational studies show for gas phase pyrimidine systems that the base orbital energies change on substitution of the aromatic ring, *e.g.,* on addition of the sugar moiety, on addition of the unscreened charge of the phosphate group, or on microhydration around the base.

The effects of bulk hydration and of counter-ions are expected to lead to large changes, possibly re-ordering the orbital energies. This might be expected to be most dramatic for the vertical IP of orbitals on the negatively charged phosphate. However, there should also be changes in the orbital energies of the base (and sugar) components connected with the longer-range polarization of the water solvent or due to π-stacking of bases in double stranded DNA. Here we explore the former effect, i.e., that of the full aqueous environment, on ionization energetics. This is done via a combination of *liquid phase* photoelectron spectroscopy of two nucleosides and electronic structure calculations of DNA components that explore a non-equilibrium polarizable continuum model (NEPCM) (26-29) for the water environment. These are the first reports of experimental solution phase ionization potentials of the pyrimidine DNA building blocks, and we can use the experiment to appraise the accuracy of the various theoretical solvation models on this important biochemical system.

Ab initio electronic structure calculations have been extensively applied for ionization of DNA components in the gas phase (24,25,30-32) and in microhydrated environments containing a small number of water molecules (33,34). Extensions toward bulk solvation employ a standard (equilibrium) polarizable continuum model (PCM) of the solvent which, however, does not correspond to vertical but to adiabatic ionization of the solute (35-36). Alternatively, a thermodynamic cycle, which is also pertinent to an adiabatic ionization process, is used (8). Very

recently, QM/MM methods have been applied to obtain vertical ionization energies of cytosine with water and counter-ions included from molecular dynamics simulations (37). Here, we show that ab initio electronic structure calculations employing a non-equilibrium version of a polarizable continuum model can efficiently and with sufficient accuracy account for bulk solvent effects on ionization of DNA building blocks. Such calculations not only quantitatively reproduce the experimental vertical ionization potentials but also allow for a detailed interpretation of the observed photoelectron spectra in terms of the underlying electronic states.

Methods and Materials

Experimental details

Photoemission (PE) measurements from aqueous solutions of cytidine ($C_9H_{13}N_3O_5$) and of deoxythymidine ($C_{10}H_{14}N_2O_5$) were performed at the U41 PGM undulator beamline of the synchrotron radiation facility BESSY, Berlin. PE spectra were collected from a 15 μm liquid microjet traveling at a velocity of 120 m/s with a temperature of 4 ^{o}C. In the present valence photoelectron study we have used a photon energy of 200 eV. Details of the experimental setup have been described previously (38,39). Briefly, excitation is carried out with the synchrotron light polarization vector parallel to the flow of the liquid microjet, while the mean electron detection angle is normal to the polarization vector. Photoelectrons are collected through an orifice 200 μm in diameter, which acts as a conductance limiting aperture to differentially pump the main chamber housing the liquid micro-jet (operating at 10^{-5} mbar) from the hemispherical energy analyzer (operating at 10^{-9} mbar). Energy resolution of the beamline was better than 200 meV at the incident photon energy used here. The resolution of the hemispherical electron-energy analyzer is constant with kinetic energy (200 meV at a constant analyzer energy of 20

eV). Typical signal count rates were 10^3-10^4 s^{-1}. The small focal size (23 x 12 μm^2) of the incident photon beam at the U41 PGM beamline allows for matching spatial overlap with the liquid microjet, which limits the contribution of gas-phase components in the collected photoemission spectra to less than 5% for liquid water. Electron kinetic energies are calibrated based on the $1b_1$ binding energy of liquid water (40).

Highly demineralized water was used for preparing the 0.2m cytidine and deoxythymidine aqueous solutions. The purity of both chemicals was >99% (Sigma-Aldrich). The pH value of the as-prepared solutions was 7.56 and 7.05, respectively.

Computational details

Ground state geometry of all investigated systems (i.e., CMP^-, $dTMP^-$, CMP^{2-}, $dTMP^{2-}$, cytidine, deoxythimidine, cytosine, thymine, ribose, deoxyribose, and dihydrogenphospate anion) were optimized in the gas phase at the MP2/aug-cc-pVDZ level. We have considered only canonical forms of both cytosine and thymine molecules. Canonical tautomers dominate in water environment for both molecules and it is also the most stable structure for thymine in the gas phase (41, 42). (For cytosine, one of the enol forms is more stable in the gas phase, however already two solvating water molecules change the energy ordering in favor of the canonical form (41)). The chemical structures of the pyrimidine bases, nucleosides, and nucleotides are presented in Figure 1. At optimized geometries, vertical ionization potentials (VIP) were evaluated. To calculate the lowest ionization potential, we used the unrestricted version of the MP2 method. Higher spin components were anihilated via the Schlegel's projection method (PMP2). The PMP2 approach was demonstrated to provide an excellent values of ionization energies for isolated gas phase nucleic acid bases (7).

To evaluate higher ionization potentials, i.e., to simulate the photoemission from orbitals energetically bellow HOMO, we utilized the time dependent version of the density functional theory (TDDFT). The n-th ionization potential has been then calculated as the sum of the first ionization potential IP_1 (calculated at PMP2/aug-cc-pVDZ level) and the (n-1)-th excitation energy of the ionized species (calculated at TDDFT/6-31g* level). This level of theory allows for a combination with non-equilibrium PCM to account for solvent effects (*vide infra*). Photoionization probabilities are considered to be the same for all these ionized states, which is a reasonable approximation for high energy photons used in the experiment.

Note that density functional methods suffer less from the spin contamination than the MP2 method, however, they can predict spurious charge transfer states due to the self-interaction error (43). This is an issue particularly relevant for larger species such as CMP^- or $dTMP^-$. Functionals with a low fraction of exact exchange provide low ionization potentials for states with a charge transfer between base, sugar and phosphate moieties. This can be rectified by employing hybrid functionals with a large portion of exact exchange, such as the new and accurate BMK functional (44). Table 1 shows benchmark calculations for the lowest four IPs of gas phase cytosine, thymine. We see that the TDDFT method with the BMK functional performs very well compared to experiment, being of comparable accuracy to the computationally much more demanding CASPT2 method. Note that other hybrid functionals, B3LYP and BHandH, proved to be somewhat less accurate.

The effect of the aqueous solvent on ionization was accounted for using the polarizable continuum model (45-47). The standard PCM procedure applied both before and after ionization corresponds to adiabatic IPs. Since our principal aim was to calculate vertical IPs, we optimized the reaction field only for the parent species before ionization. In a subsequent step we have used

the non-equilibrium PCM protocol as implemented in the Gaussian 03 code (26, 27, 48). In this way, only the fast electronic component of the solvent response to ionization is accounted for and the slow component of this response (corresponding to nuclear relaxation) is left out. The non-equilibrium PCM approach combined with the PMP2 method was thoroughly tested for aqueous anions, neutral molecules, and cations in our previous studies (28, 29).

Results

Valence photoelectron spectra of aqueous cytidine and deoxythymidine, measured at 200 eV photon energy, are presented in Figures 2a,b. For comparison, in the inset the PE spectrum of pure water is also depicted, the effect of the added solutes being highlighted by the differential spectra shown at the bottom of each inset and (in more detail) in the main panel of the graphs. Before subtraction, the spectra are normalized at the background intensity around -25 eV BE. The most interesting signature of cytidine or deoxythymidine is the relatively broad spectral feature between roughly -6.5 and -10.5 eV, i.e., before the rise of the first water peak ($1b_1$). Photoelectron intensities from these solutes obtained in the experiment are relatively high, even though the concentrations were only 0.2m (which corresponds to almost maximum solubility). Note that for alkali halide aqueous solutions at the same low concentration solute signal is barely detectable. Since for the photon energy used here PE spectroscopy is surface selective (38), this behavior reflects a considerable surface activity of cytidine and deoxythymidine (being larger for the latter case). For both species, the spectral signature consists of a lower intensity region at smaller binding energies and a more intense region at higher energies. Both these features are broad and they overlap to some extent. They can be fitted by two Gaussians, centered at -8.3 and -9.6 eV for cytidine, and -8.3 and -9.5 eV for deoxythymidine (Figures 2a,b). However, since

the features are broad and overlapping, they are difficult to assign without the help of electronic structure calculations which provide further analysis and interpretation of the underlying states.

The calculated spectra for aqueous cytidine and deoxythimidine in the low electron binding energy region, corresponding to the part of the experimental signal not obscured by water features, are shown in Figures 2c,d. Vertical bars depict the individual vertical ionization transitions of aqueous cytidine and deoxythimidine, evaluated within the non-equilibrium polarizable continuum model (the lowest ionized state being evaluated at the PMP2 level, while the TDDFT method is employed for the higher states). The letter assignment above each bar indicates whether the electron is ejected from an orbital dominantly situated at the base (B) or sugar (S). The spectral envelope in Figures 2c,d is obtained by broadening the calculated discrete spectra using Gaussians of full-width-at-half-maximum (FWHM) of 0.9 eV, which corresponds to a typical vibrational and inhomogeneous broadening in polyatomic aqueous solute systems (48). We have assigned equal intensities to all lines, a reasonable assumption born out by the agreement with experiment showing that, with incident photon energies well above threshold, the photoionization cross sections do not vary dramatically between states.

We see from Figure 2 that the low energy feature is indeed composed from a single peak, as suggested by the fit to the experimental spectra. This peak reflects the formation of the ground electronic state of the cation of cytidine or deoxythimidine with the solute and solvent geometry corresponding to the neutral species as pertinent to vertical ionization. The calculated values of the lowest electron binding energy are -7.78 eV for cytidine and -7.74 eV for deoxythimidine. For both aqueous cytidine and deoxythimidine this lowest ionization takes place at the pyrimidine base. This is demonstrated in Figures 3a,b which show the highest occupied

molecular orbital (HOMO) of the two species in water prior to ionization, localized at the respective base. Qualitatively the same picture is also seen in the unpaired spin density after ionization, which is depicted for both species in Figures 3c,d. We note here that upon adding the phosphate group, the lowest ionized state remains localized on the base, a point relevant to later discussion.

The agreement between the theoretical and experimental spectra over the range reported is excellent (Figure 2). The calculations are correctly reproducing the two overlapping spectral features, including their relative intensities. Qualitatively, there is little difference between the photoelectron spectra of cytidine and deoxythymidine, quantitatively the low energy peak is slightly more pronounced and the higher energy peak broader and lower for the latter species. A great asset of the calculations is that they allow for interpretation of the observed spectral features in terms of electronic states localized on different parts of the solute. To this end, we have modeled (in addition to aqueous cytidine and deoxythymidine) also the photoelectron spectra of aqueous cytosine, thymine, ribose, deoxyribose, dihydrogenphosphate anion, together with CMP^- and $dTMP^-$ (Figure 4). The picture emerging from these calculations is as follows. To a good approximation, the aqueous spectrum of CMP^- or $dTMP^-$ can be decomposed into additive contributions from their building blocks, which indicates a relatively weak electronic interaction between them in water. The weakest binding energy feature originates solely from vertical ionization of the base (cytidine or deoxythymidine). In contrast, the higher binding energy feature is composed from states localized on the base, sugar, and phosphate. This region is, therefore, characterized by a mixture of states from all constituent parts of the solutes.

The lowest IPs of all aqueous species considered so far and additional data for 1-methylthymine, 1-methylcytosine, and the nucleotide dianions are summarized in Table 2. Note both mono- and di-anionic forms of nucleotides (albeit with transient counter-ion binding) are present at neutral pH, therefore, both were considered. The table presents vertical values obtained using NEPCM, together with equilibrium PCM. These are two very different quantities, the former best representing vertical ionization with only fast electronic relaxation of the solvent and the latter adiabatic ionization which additionally models the slow solvent nuclear relaxation. We see that solvent nuclear relaxation is connected with a sizable decrease of about 1.1 eV of the IPs of all species containing the pyrimidine bases. For the sugars and dihydrogenphosphate this reduction is even larger, 0.9 – 1.4 eV in the former and 1.3 eV in the latter case. Further relaxation of the solute intramolecular framework in systems containing pyrimidine base and sugars is less important leading to an additional decrease of IPs of less than 0.3 eV. However, to relax intramolecular modes for $H_2PO_4^-$ is more important, leading to an additional decrease of IP of more than 0.9 eV. We note that in water the vertical IPs of ribose, deoxyribose, and dihydrogenphosphate are more than 1.2 eV higher than those of the bases. Even in the adiabatic framework this gap persists, albeit being slightly reduced to ~1 eV.

Table 2 underlines a rather unexpected result – there is virtually no change (i.e., < 0.15 eV) in the vertical or adiabatic IPs of aqueous pyrimidine bases upon substitution at the 1-position, either by adding a methyl or sugar group, or by adding a sugar-phosphate group. Moreover, in water the difference in IPs of nucleotides with a monobasic versus dibasic phosphate group is also below 0.15 eV. This stability of IPs is in stark contrast to the situation in the gas phase, as shown for the lowest IPs in Table 3. The difference between aqueous and gas phase environments can be also easily seen upon comparing the calculated spectra in Figures

4a,b with Figures 4c,d. Note that in order to directly compare aqueous with vacuum systems we applied the same 0.9 eV broadening although actual experimental gas phase peak widths are substantially narrower (8). Upon dehydration, electron binding energies of CMP^- and $dTMP^-$ decrease by 2 eV primarily due to destabilization of the parent anion. More strikingly, the spectrum of CMP^- or $dTMP^-$ in vacuum can no longer be viewed as a sum of its constituent parts, of which the smallest IP in the gas phase is that of $H_2PO_4^-$, while all the other species exhibit electron binding energies larger by several eV (Table 3). Analysis of the molecular orbitals of the parent species in the gas phase (figures not shown here) reveals that for CMP^- and $dTMP^-$ the lowest ionization still comes from the base, while for gaseous CMP^{2-} and $dTMP^{2-}$ it comes from the phosphate with a dramatically reduced IP due to the presence of the additional electron (Table 3). Returning to gas phase CMP^- or $dTMP^-$, the base-centered HOMO is strongly destabilized by the presence of the phosphate. The next IP, which is almost degenerate with the lowest one, then originates from the phosphate group.

The monobasic phosphate group is responsible for lowering the gas phase IP of the pyrimidine by 2 eV, and the dibasic group by a further 4.5 eV. Compared to this, the influence of the sugar is more modest. Nevertheless, it still leads to a 0.3-0.5 eV shifts toward smaller electron binding energies in vacuum (Table 3). Unlike in water, the gas-phase IPs of the pyrimidine bases thus change appreciably upon substitution at the 1-position on the pyrimidine ring. This effect is stronger for thymine than for cytosine. It is important to point out that calculated lowest IPs of the constituent parts are in excellent agreement with gas phase experiments. This is also true for CMP^- and $dTMP^-$ where the present calculations fall within the large uncertainty of the experimental values (Table 3).

Discussion

The present experimental and computed photoelectron spectra for the aqueous pyrimidine nucleosides are in a very good agreement with each other. In particular, calculations are capable of faithfully modeling the shape of the experimental spectra. Taken together with the excellent performance of the theoretical models employed in reproducing gas phase vertical IPs (31) including the lowest states of the base and phosphate moieties we have established a reliable and computationally inexpensive approach for evaluating and interpreting photoionization experiments on aqueous DNA components. Note, however, that the values of vertical IPs of aqueous cytidine and deoxythimidine based on NEPCM calculations are about 0.5 eV smaller than those obtained from the two-Gaussian fit to the experimental photoelectron spectra (Table 2 and Figure 2). The calculations somewhat underestimate the ionization potentials due to the neglect of the detailed molecular structure of the first solvent shell around the solutes within the polarizable continuum solvent model. Indeed, our previous calculations of aqueous imidazole (28) showed that explicit inclusion of up to five water molecules in the immediate solvation shell into the NEPCM calculation leads to an increase of the vertical IP by 0.2 eV with the effect of more distant explicit solvent molecules estimated to account for additional 0.2 eV. Extrapolating analogous corrections also for aqueous cytidine and deoxythimidine brings the calculation within 0.1 eV of the fit to the experiment.

On the experimental side, as commented earlier, the recorded nucleosides signals are stronger than those for inorganic ions at the same concentration presumably due to enhanced surface activity. The photoelectron spectroscopy experiment preferentially probes the interfacial region (38), particularly for the photon energies employed here. This is due to the fact that the

electron inelastic mean free path in water is short amounting to only several water layers for the present kinetic energies of the photoelectrons (38). Some of the discrepancy between theory and experiment may thus also come from differences between interfacial and bulk solvation, the former involving an incomplete solvation sphere. However, our previous study of aqueous iodide shows that the difference between IPs of bulk and interfacial ions is only of the order of 0.1 eV (50). An additional subtlety results from the way the experimental spectra are fitted to a reduced set of two Gaussians. Although justified by the level of detail resolved in the experiment, if a set of Gaussians corresponding to the actual number of the calculated states were used, the lowest fitted Gaussian would be narrower and shifted toward the calculated value. The inherent vibrational and inhomogeneous broadening of the experimental spectra highlights the point that a detailed analysis of the electronic states of the solutes can only be achieved with an underlying theoretical model. Calculations clearly interpret the higher binding energy part of the spectrum as originating from a multitude of electronically excited states of the nascent cation, while the low binding energy part is due to a single (ground) state localized on the base.

For comparison, the vertical ionization energies of uracil as well as the RNA polynucleotides (poly rU and poly rC) in cryogenic thin films on a pyrolytic graphite substrate have been measured recently. These reveal a binding energy of 8.05 eV for uracil, 8.1 eV for cytidine in poly rC, and 8.5 eV for uridine in poly rU (11). These results are in qualitative agreement with those presented here, although the immediate environment of the base clearly differs. This good agreement can also partly be attributed to the relatively small change of the lowest IP of cytidine and deoxythymidine upon moving from the gas phase into water (compare Tables 2 and 3). Note that the solvent effect on the bare base is larger and comparable to that observed in our previous study on imidazole (28). However, the sugar and phosphate moieties

act as "presolvating" agents reducing thus the effect of the aqueous solvent on the ionization potential (*vide infra*).

Arguably the most dramatic result of the present NEPCM calculations is the prediction of a negligible (< 0.15 eV) effect on the IPs of the aqueous bases of substitution on the base by adding a methyl, sugar, or sugar-phosphate group. This situation is in strong contrast with that in the gas phase, where the ionization potential of the base is highly sensitive to both substitution and the local electrostatic environment. In terms of electronic structure, 1-methyl cytosine and 1-methyl thymine are better analogs to the nucleosides than the free bases as they include the electronic effect of substitution at 1-position on the pyrimidine ring. There is a marked decrease (~0.4 eV) in the gas phase vertical IP particularly for thymine on methyl substitution at the 1-position nitrogen (8), which is reproduced well in our calculations. This shift is largely removed by the screening effect of the solvent, as is also the case on substituting with a sugar group at the 1-position. This screening is even more striking in the case of adding a charged phosphate group which leads to lowering of the IP in the gas phase by 2 – 2.5 eV (51). Even deprotonating the phosphate group, so that it bears a double negative charge, has only a minor effect on the IP after dielectric screening by the aqueous solvent. This latter computational result is confirmed by our preliminary photoelectron spectra for aqueous CMP^{2-} which can be fit with two Gaussian peaks at 8.5 and 9.3 eV. These very modest shifts compared to aqueous cytidine are consistent with the spectral simulation shown in Figure 4 for CMP^{-} and our computed results for CMP^{2-}. The present calculations therefore highlight the limitations of gas phase data to provide a predictive capability for ionization of DNA components in bulk water. Further, these observations suggest a remarkable ability of bulk water to screen the electronic effects of components attached to the DNA bases.

It is worth stressing that standard (equilibrium) PCM calculations, which effectively relaxes the solvent after ionization, underestimate the present experimental IPs by more than 1.5 eV (see Table 2). Such PCM calculations (33-35) are adiabatic with respect to the solvent coordinates and do not represent true vertical ionization. Therefore, standard PCM calculations should not be directly compared to experimental vertical IPs but rather to the ionization onsets. A similar argument also applies to estimates of IPs combining gas phase calculations with hydration free energies (7). Experimental measurements for the ionization onsets of the free bases have been made for microhydrated clusters and compared with ab initio cluster calculations (33, 36). Herschbach and coworkers estimate the adiabatic IP of thymine changing from 9.15 to 8.5 eV upon moving from a bare base to a cluster with three water molecules (20). Belau et al. have observed somewhat smaller (~0.3 eV) shifts in ionization onsets for hydrated thymine and cytosine clusters of similar size (22). From the current experiments, we can establish only an upper bound for the ionization onset in bulk water (with respect to the vacuum level) at ~ 6.8 eV for both deoxythymidine and cytidine (Figure 2), which is at least 1.5 eV lower than that for the base$(H_2O)_3$ clusters. Essentially, increasing hydration on the pyrimidine nucleosides leads to a relatively small shift in the vertical ionization potential but shifts the adiabatic onset to much lower energies.

The vertical IPs presented also provide valuable information on the limiting energetics of the Rydberg manifold of electronically excited states of the closed shell aqueous bases. The upper Rydberg-like states may dominate the excited state absorption observed in femtosecond studies following the electronic relaxation of nucleobases (52). To evaluate charge hopping mechanisms in oligomeric DNA, it will be interesting to see if the ionization energetics of bases change markedly when sequestered in the hydrophobic environment of the double helix (37).

Our present study serves as a possible prelude to direct measurements of ionization energetics within oligomeric segments of DNA of increasing length and within the native aqueous environment.

Conclusions

Photoelectron spectroscopy from an aqueous microjet has been used together with ab initio calculations employing a non-equilibrium polarizable continuum model for characterizing ionization of pyrimidine components of DNA in bulk water. The lowest vertical ionization energies of aqueous cytidine and deoxythimidine were determined experimentally as 8.3 eV for both species. Calculations, which are in quantitative agreement with the experiment, allow for a detailed interpretation of the photoelectron spectrum in terms of the underlying electronic states. A dramatic effect of the aqueous environment is also revealed from the ab initio computations. Namely, bulk water not only modestly lowers the ionization potential of the DNA bases but it also makes it insensitive to the presence of sugar or phosphate. This is a very different situation from the gas phase where the other DNA components (phosphate in particular) strongly influence the ionization process at the base. Further, the present results suggest that the lowest energy ionization process originates by electron ejection from the base even in the aqueous pyrimidine nucleotides.

The present vertical ionization potentials and photoelectron band shapes for the aqueous pyrimidine DNA components provide important energetics data for modeling photoionization and charge-transfer processes in DNA. We therefore expect these results, along with future data employing the techniques reported here to purine and oligomeric DNA, to enable a better understanding of radiation induced processes that lead to DNA damage in cells.

Acknowledgment

Support from the Czech Ministry of Education (grant LC512 and Research Project No. 6046137307) and from the Czech Science Foundation (grants 203/08/0114 and 203/07/P449) is gratefully acknowledged. B.W. gratefully acknowledges support by the Deutsche Forschungsgemeinschaft (Project WI 1327/3-1) and thanks the BESSY staff for assistance. S.E.B. gratefully acknowledges the support of the U.S. National Science Foundation through grant CHE-0617060 and the donors of the American Chemical Society Petroleum Research Fund. Part of the work in Prague was supported via Project Z40550506.

References

1. Friedberg, E. C., Walker, G. C., W., S., Wood, R. D., Schultz, R. A., & Ellenberger, T. (2005) DNA Repair and Mutagenesis (ASM Press, Washington, D.C.).
2. Ward, J. F. (1998) in DNA Damage and Repair, eds. Nickoloff, J. A. & Hoekstra, M. F. (Humana Press, Totowa, N.J.), pp. 65-84.
3. von Sonntag, C. & John, R. S. a. E. B. (2007) Free-Radical-Induced DNA Damage as Approached by Quantum-Mechanical and Monte Carlo Calculations: An Overview from the Standpoint of an Experimentalist. *in Advances in Quantum Chemistry* (Academic Press), pp. 5-20.
4. Turecek, F. & John, R. S. a. E. B. (2007) Computational Studies of Radicals Relevant to Nucleic Acid Damage. *in Advances in Quantum Chemistry* (Academic Press), pp. 89-120.
5. Cauet, E., Lievin, J., & John, R. S. a. E. B. (2007) Radical Cations of the Nucleic Bases and Radiation Damage to DNA: Ab Initio Study, in Advances in Quantum Chemistry. *in Advances in Quantum Chemistry* (Academic Press), pp. 121-147.
6. Candeias, L. P., Stenken, S. (1992) Ionization of Purine Nucleosides and Nucleotides and Their Components by 193-Nm Laser Photolysis in Aqueous-Solution - Model Studies for Oxidative Damage of DNA. *J. Am. Chem. Soc.*, *114*, 699-704
7. Crespo-Hernandez, C. E., Arce, R., Ishikawa, Y., Gorb, L., Leszczynski, J. & Close, D. M. (2004) Ab Initio Ionization Energy Thresholds of DNA and RNA Bases in Gas Phase and in Aqueous Solution. *J. Phys. Chem. A* **108,** 6373-6377.

8. Fernando, H., Papadantonakis, G. A., Kim, N. S., & LeBreton, P. R. (1998) Conduction-band-edge ionization thresholds of DNA components in aqueous solution. *Proc. Natl. Acad. Sci. USA* **95,** 5550–5555.

9. Naegeli, H. (1997) Mechanisms of DNA Damage Recognition in mammalian cells. *Mechanisms of DNA Damage Recognition in mammalian cells* (Landes Bioscience, Austin, TX).

10. Papadantonakis, G.A., Tranter, R., Brezinsky, K., Yang, Y. N., van Breemen, R. B. & LeBreton, P. R. (2002) Low-energy, low-yield photoionization, and production of 8-oxo-2'-deoxyguanosine and guanine from 2 '-deoxyguanosine. *J. Phys. Chem. B* **106,** 7704-7712.

11. Magulick, J., Beerbom, M. M., Schlaf, R. (2008) Investigation of adenine, uracil, and ribose phosphate thin films prepared by electrospray in vacuum deposition using photoemission spectroscopy. *Thin Solid Films*, **516**, 2396-2400.

12. Crespo-Hernandez, C. E. & Arce, R. (2002) Photoionization of DNA and RNA bases, nucleosides and nucleotides through a combination of one- and two-photon pathways upon 266 nm nanosecond laser excitation. *Photochem. Photobiol.* **76,** 259-267.

13. Marguet, S., Markovitsi, D., & Talbot, F. (2006) One- and two-photon ionization of DNA single and double helices studied by laser flash photolysis at 266 nm. *J. Phys. Chem. B* **110,** 11037-11039.

14. Hush, N. S. & Cheung, A. S. (1975) Ionization-Potentials and Donor Properties of Nucleic-Acid Bases and Related Compounds. *Chem. Phys. Lett.* **34,** 11-13.

15. Padva, A., O'Donnell, T. J., & LeBreton, P. R. (1976) UV photoelectron studies of biological pyrimidines: the valence electronic structure of methyl substituted uracils. *Chem. Phys. Lett.* **41,** 278.

16. Peng, S., Padva, A., & LeBreton, P. R. (1976) Ultraviolet photoelectron studies of biological purines: The valence electronic structure of adenine. *Proc. Natl. Acad. Sci. USA* **73,** 2966-2968.

17. Urano, S., Yang, X., & LeBreton, P. R. (1989) UV photoelectron and quantum mechanical characterization of DNA and RNA bases: valence electronic structures of adenine, 1,9-dimethyl-guanine, 1-methylcytosine, thymine and uracil. *J. Molec. Spec.* **214,** 315-328.

18. Yu, C., O'Donnell, T. J., & LeBreton, P. R. (1981) Ultraviolet Photoelectron Studies of Volatile Nucleoside Models. Vertical Ionization Potential Measurements of Methylated Uridine, Thymidine, Cytidine, and Adenosine. *J. Phys. Chem.* **85,** 3851-3855.

19. Kubota, M. & Kobayashi, T. (1996) Electronic structure of uracil and uridine derivatives studied by photoelectron spectroscopy. *J. Electron Spec. Rel. Phenomena* **82,** 61-70.

20. Kim, S. K., Lee, W., & Herschbach, D. R. (1996) Cluster Beam Chemistry: Hydration of Nucleic Acid Bases; Ionization Potentials of Hydrated Adenine and Thymine. *J. Phys. Chem.* **100,** 7933-7937.

21. Jochims, H. W., Schwell, M., Baumgärtel, H., & Leach, S. (2005) Photoion mass spectrometry of adenine, thymine and uracil in the 6-22 eV photon energy range. *Chem. Phys.* **314,** 263-282.

22. Belau, L., Wilson, K. R., Leone, S. R., & Ahmed, M. (2007) Vacuum-ultraviolet photoionization studies of the microhydration of DNA bases (Guanine, cytosine, adenine, and Thymine). *J. Phys. Chem. A* **111,** 7562-7568.

23. Yang, X., Wang, X.-B., Vorpagel, E. R., & Wang, L.-S. (2004) Direct experimental observation of the low ionization potentials of guanine in free oligonucleotides by using photoelectron spectroscopy. *Proc. Natl. Acad. Sci. USA* **101,** 17588–17592.

24. Zakjevskii, V. V., King, S. J., Dolgounitcheva, O., Zakrzewski, V. G., & Ortiz, J. V. (2006) Base and Phosphate Electron Detachment Energies of Deoxyribonucleotide Anions. *J. Am. Chem. Soc.* **128,** 13350-13351.

25. Rubio, M., Roca-Sanjuan, D., Merchan, M., & Serrano-Andres, L. (2006) Determination of the lowest-energy oxidation site in nucleotides: 2 '-Deoxythymidine 5 '-monophosphate anion. *J. Phys. Chem. B* **110,** 10234-10235.

26. Cossi, M., & Barone, V. (2000) Separation between fast and slow polarizations in continuum solvation models. *J. Phys. Chem. A* **104,** 10614-10622.

27. Cossi, M., & Barone, V. (2000) Solvent effect on vertical electronic transitions by the polarizable continuum model. *J. Chem. Phys.* **112,** 2427-2435.

28. Jagoda-Cwiklik, , *et al.* (2008) Ionization of imidazole in the gas phase, microhydrated environments, and in aqueous solution. *J. Phys. Chem. A* **112,** 3499-3505.

29. Jagoda-Cwiklik, B., Slavicek, P., Nolting, D., Winter, B., & Jungwirth, P. (2008) Ionization of aqueous cations: Photoelectron spectroscopy and ab initio calculations of protonated imidazole. *J. Phys. Chem. B* **112,** 7355-7358.

30. Roca-Sanjuan, D., Rubio, M., Merchan, M., & Serrano-Andres, L. (2006) Ab initio determination of the ionization potentials of DNA and RNA nucleobases. *J. Chem. Phys.* **125,** 084302.

31. Cauet, E., Dehareng, D., & Lievin, J. (2006) Ab initio study of the ionization of the DNA bases: Ionization potentials and excited states of the cations. Journal of Physical Chemistry A, 2006. *J. Phys. Chem. A* **110,** 9200-9211.

32. Colson, A.-O., Besler, B., & Sevilla, M.D. (1993) Ab initio molecular-orbital calculations on DNA radical ions. 3. Ionization potentials and ionization sites in components of the DNA sugar-phosphate backbone. *J. Phys. Chem.* **97**, 8092-8097.

33. Close, D. M., Crespo-Hernandez, C. E., Gorb, L., & Leszczynski (2006) Influence of microhydration on the ionization energy thresholds of thymine: Comparisons of theoretical calculations with experimental values. *J. Phys. Chem. A* **110,** 7485-7490.

34. Close, D. M., Crespo-Hernandez, C. E., Gorb, L., & Leszczynski (2008) Theoretical elucidation of conflicting experimental data on vertical ionization potentials of microhydrated thymine. *J. Phys. Chem. A* **112,** 4405-4409.

35. Close, D. M. (2004) Calculation of the ionization potentials of the DNA bases in aqueous medium. *J. Phys. Chem. A* **108,** 10376-10379.

36. Close, D. M., Crespo-Hernandez, C. E., Gorb, L., & Leszczynski (2005) The influence of microhydration on the ionization energy thresholds of uracil and thymine. *J. Phys. Chem. A* **109,** 9279-9283.

37. Cauet, E., Valiev, M. & Weare, J. H. *J. Am. Chem. Soc.* (2008), in press.

38. Winter, B. & Faubel, M. (2006) Photoemission from liquid aqueous solutions. *Chem. Rev.* **106**, 1176-1211.

39. Faubel, M., Steiner, B., Toennies, P. (1997) Photoelectron spectroscopy of liquid water, some alcohols, and pure nonane in free micro jets. *J. Chem. Phys.* **106**, 9013-9031.

40. Winter, B., Weber, R., Widdra, W., Dittmar, M., Faubel, M. & Hertel, I. V. (2004) Full Valence Band Photoemission from Liquid Water Using EUV Synchrotron Radiation. *J. Phys. Chem. A* **108**, 2625-2632.

41. Trygubenko, S. A., Bogdan, T. V., Rueda M., Orozco M., Luque F. J., Sponer J., Slavicek, P. & Hobza, P. (2002) Correlated ab initio study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution - Part 1. Cytosine. *Phys. Chem. Chem. Phys.* **4**, 4192-4203.

42. Rejnek, J., Hanus, M., Kabelac, M., Ryjacek, F., & Hobza, P. (2005) Correlated ab initio study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution. Part 4. Uracil and thymine. *Phys. Chem. Chem. Phys.* **7**, 2006-2017.

43. Magyar, R. J. (2007) Dependence of spurious charge-transfer excited states on orbital exchange in TDDFT: Large molecules and clusters. *J. Chem. Theory Comput.* **3**, 976-987.

44. Boese, A. D., & Martin, J. M. L. (2004) Development of density functionals for thermochemical kinetics. *J. Chem. Phys.* 121, 3405-3414.

45. Tomasi, J., Mennucci, B., & Cammi, R. (2005) Quantum mechanical continuum solvation models. *Chem. Rev.* **105**, 2999-3093.

46. Pavanello, M., Mennucci, B., & Tomasi, J. (2006) DFT calculation of deuterium quadrupolar tensor in crystal anthracene. *Theor. Chem. Acc.* **116**, 711-717.

47. Mennucci, B., Cammi, R., & Tomasi, J. (1998) Excited states and solvatochromic shifts within a nonequilibrium solvation approach: A new formulation of the integral equation formalism method at the self-consistent field, configuration interaction, and multiconfiguration self-consistent field level. *J. Chem. Phys.* **109**, 2798-2807.

48. Frisch, M. J. et al., *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

49. Poterya, V., Farnik, M., Oncak, M., & Slavicek, P. (2008) Water photodissociation in free ice nanoparticles at 243 nm and 193 nm. *Phys. Chem. Chem. Phys*., **10,** 4835-4842.

50. Bradforth, S. E., Jungwirth, P. (2002) Excited states of iodide anions in water: A comparison of the electronic structure in clusters and in bulk solution. *J. Phys. Chem. A* **106**, 1286-1298.

51. Wang, X.-B., Vorpagel, E. R., Yang, X., Wang, L. S. (2001) Experimental and theoretical investigations of the stability, energetics, and structures, of $H_2PO_4^-$, $H_2P_2O_7^{2-}$, and $H_3P_3O_{10}^{2-}$ in the gas phase. *J. Phys. Chem. A*, **105**, 10468-10474.

52. Crespo-Hernandez, C. E., Cohen, B., Hare, P. M., & Kohler, B. (2004) Ultrafast Excited-State Dynamics in Nucleic Acids. *Chem. Rev.* **104**, 1977-2020.

53. Dougherty, D.; Younathan, E. S.; Voll, R.; Abdulnur, S.; McGlynn, S. P. (1978) Photoelectron-spectroscopy of some biological molecules. *J. Electron Spectrosc. Relat. Phenom.*, **13**, 379.

54. Dougherty, D.; Wittel, K; Meeks, J.; McGlynn P. (1975) Photoelectron Spectroscopy of Carbonyls. Ureas, Uracils, and Thymine. *J. Am. Chem. Soc*., **98**, 3815.

Table 1 Lowest four ionization energies (in eV) of gas phase pyrimidine bases. The table provides benchmarking of the applied method against CASPT2 calculations and experiment.

	MP2 + BMK	CASPT2 (ref. 30)	experiment (refs. 56, 57)
cytosine	8.84	8.73	8.82
	9.38	9.42	9.45
	9.66	9.49	---
	10.00	9.88	9.90
thymine	9.15	9.07	9.20
	10.09	9.81	10.05
	10.45	10.27	10.44
	10.92	10.49	10.88

Table 2 Calculated vertical (NEPCM) and adiabatic (PCM and PCM with relaxed solute) ionization energies of aqueous pyrimidine DNA components together with current experimental values for aqueous pyrimidine nucleosides (in eV).

		PMP2 NEPCM	PMP2 PCM	PMP2 PCM with relaxed solute
cytosine		7.86	6.59	6.34
1-methylcytosine		7.75	6.60	6.36
cytidine	*exp.* 8.3	7.77	6.64	6.32
CMP^-		7.83	6.72	6.12
CMP^{2-}		7.70	6.63	6.04
thymine		7.87	6.65	6.39
1-methylthymine		7.75	6.62	6.41
deoxythymidine	*exp.* 8.3	7.76	6.66	6.32
$dTMP^-$		7.70	6.67	6.14
$dTMP^{2-}$		7.68	6.62	6.08
ribose		8.72	7.85	7.32
deoxyribose		9.10	7.75	7.43
$H_2PO_4^-$		8.94	7.61	6.68

Table 3 Calculated and experimental gas phase vertical ionization energies of pyrimidine DNA components (in eV).

	PMP2	experiment
Cytosine	8.84	8.82[u]
1-methylcytosine	8.79	8.65[v]
Cytidine	8.56	8.46[w]
CMP^-	6.37	5.80±0.5[x]
CMP^{2-}	1.62	-
Thymine	9.15	9.20[y]
1-methylthymine	8.74	8.79[v]
deoxythymidine	8.58	-
$dTMP^-$	6.35	5.85±0.5[x]
$dTMP^{2-}$	1.54	-
Ribose	10.21	-
Deoxyribose	10.07	-
$\mathbf{H_2PO_4^-}$	5.83	5.06[u]

u ref. 53

v ref. 8

w ref. 18. Experimental data are for cytidine analog (2',3',5'-tri-*O*-methylcytidine) that can be volatilized. Substitution on the ribose ring is not expected to shift lowest IP.

x ref. 23

y refs. 17 and 54.

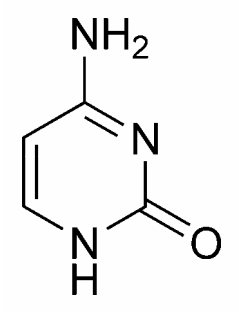
Figure Captions

Figure 1 Standard representation of the chemical structures of the pyrimidine bases, nucleosides, and nucleotides.

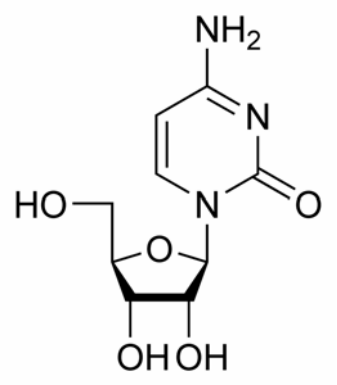
Figure 2 Experimental and computed photoelectron spectra of cytidine and deoxythymidine. Insets: Experimental spectra of (a) 0.2m cytidine and (b) 0.2m deoxythymidine aqueous solution (black curves) and of neat liquid water (blue curve), both measured at photon energy of 200 eV and collected over 10 individual runs. The difference spectrum is shown at the bottom; the positive signal (in red) arises from solute contributions. Photoemission from the three water valence orbitals $1b_2$, $3a_1$, and $1b_1$, is labeled. The main panels in (a) and (b) show increased detail of the difference spectra including 2-Gaussian fits with lowest ionization peak at 8.3 eV for both species. (c, d) Computed photoelectron spectra of aqueous cytidine and deoxythimidine respectively (red bars), broadened by Gaussians with FWHM of 0.9 eV (thick blue line). Origin of each ionization transition is labeled as B = base or S = sugar.

Figure 3 The highest occupied molecular orbitals (HOMO) of (a) cytidine and (b) deoxythimidine in water prior to ionization and the spin densities of (c) cytidine and (d) deoxythimidine cations in water after ionization.

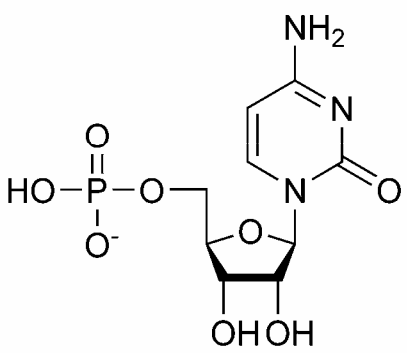
Figure 4 Comparison of simulated photoelectron spectra in liquid water and gas phase. Computed spectra of (a) & (c) CMP^-, cytidine, cytosine, ribose, and the dihydrogenphosphate anion and (b) & (d) $dTMP^-$, deoxythimidine, thymine, deoxyribose, and the dihydrogenphosphate anion. Computed spectra in water are in (a) & (b), while those in the gas phase in (c) & (d). All stick spectra were broadened by gaussians with FWHM of 0.95 eV.



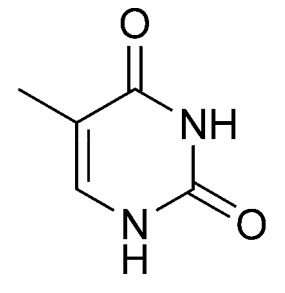
cytosine



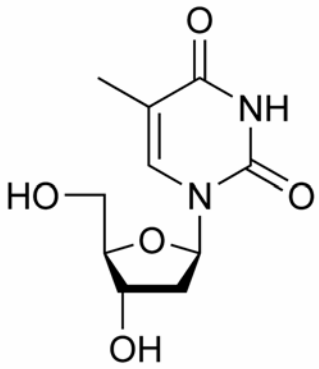
cytidine



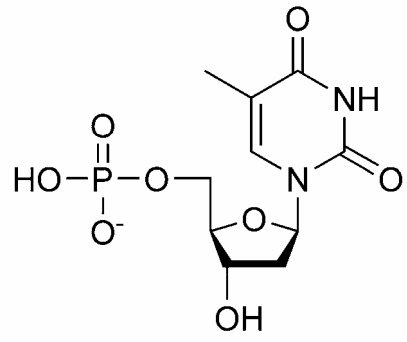
CMP^-



thymine



deoxythymidine



$dTMP^-$

Figure 1

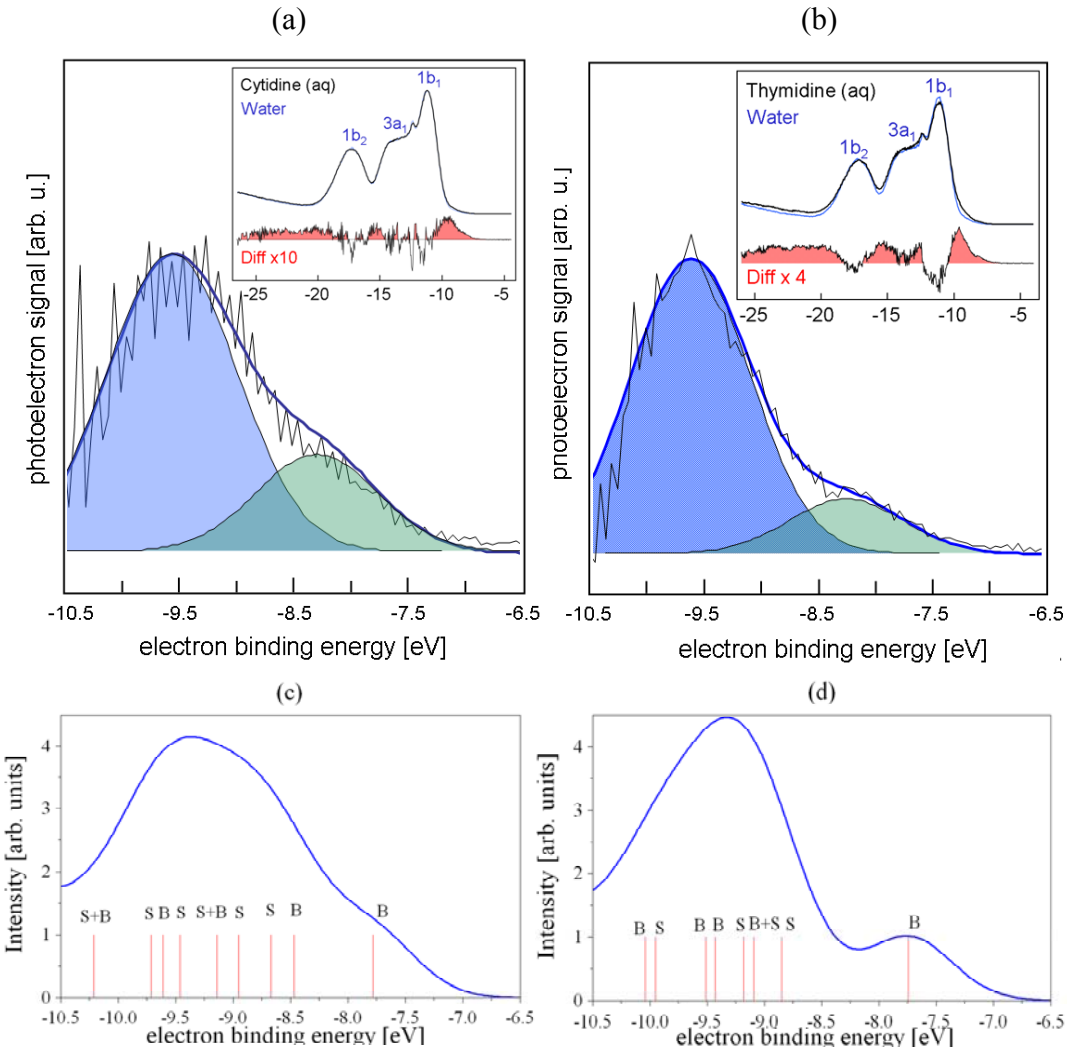

(a)
(b)
Cytidine (aq)
Water
1b1
3a1
1b2
Diff x10
Thymidine (aq)
Water
Diff x 4
photoelectron signal [arb. u.]
electron binding energy [eV]
(c)
(d)
Intensity [arb. units]
S+B
S B S S+B S S B
B
B S
B B S B+S S
electron binding energy [eV]

Figure 2

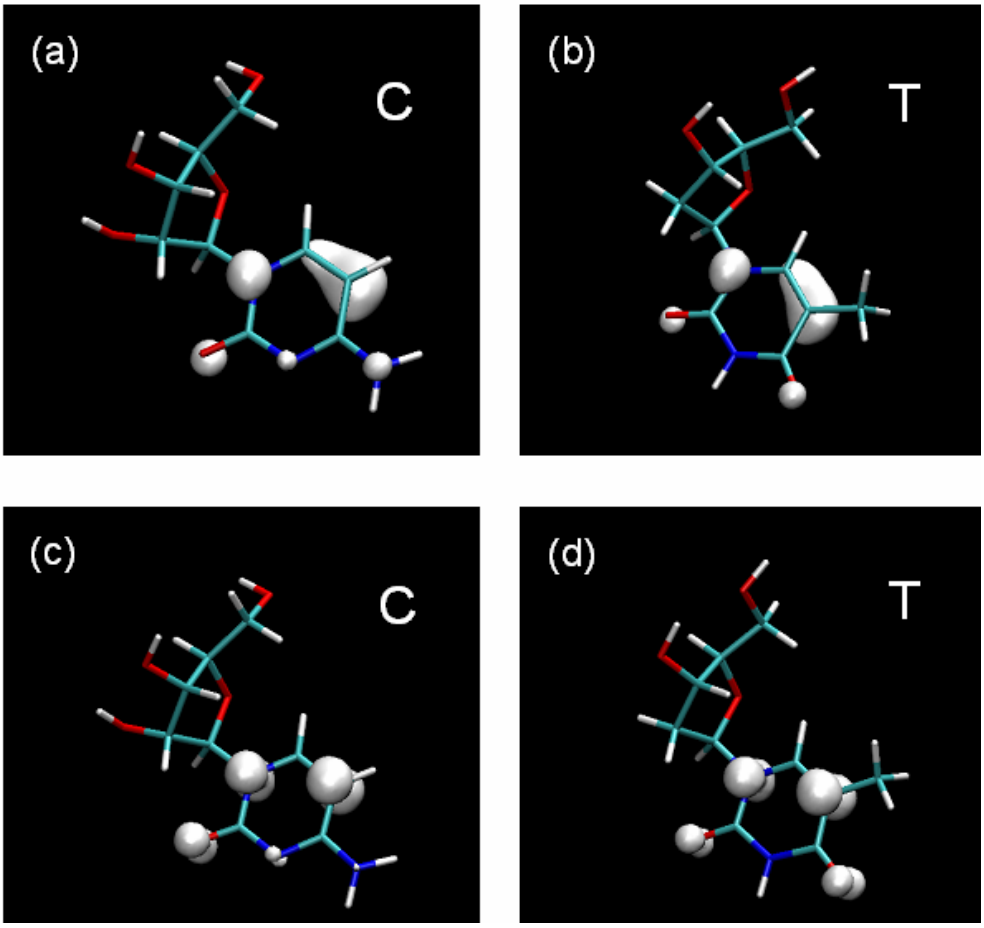


Figure 3

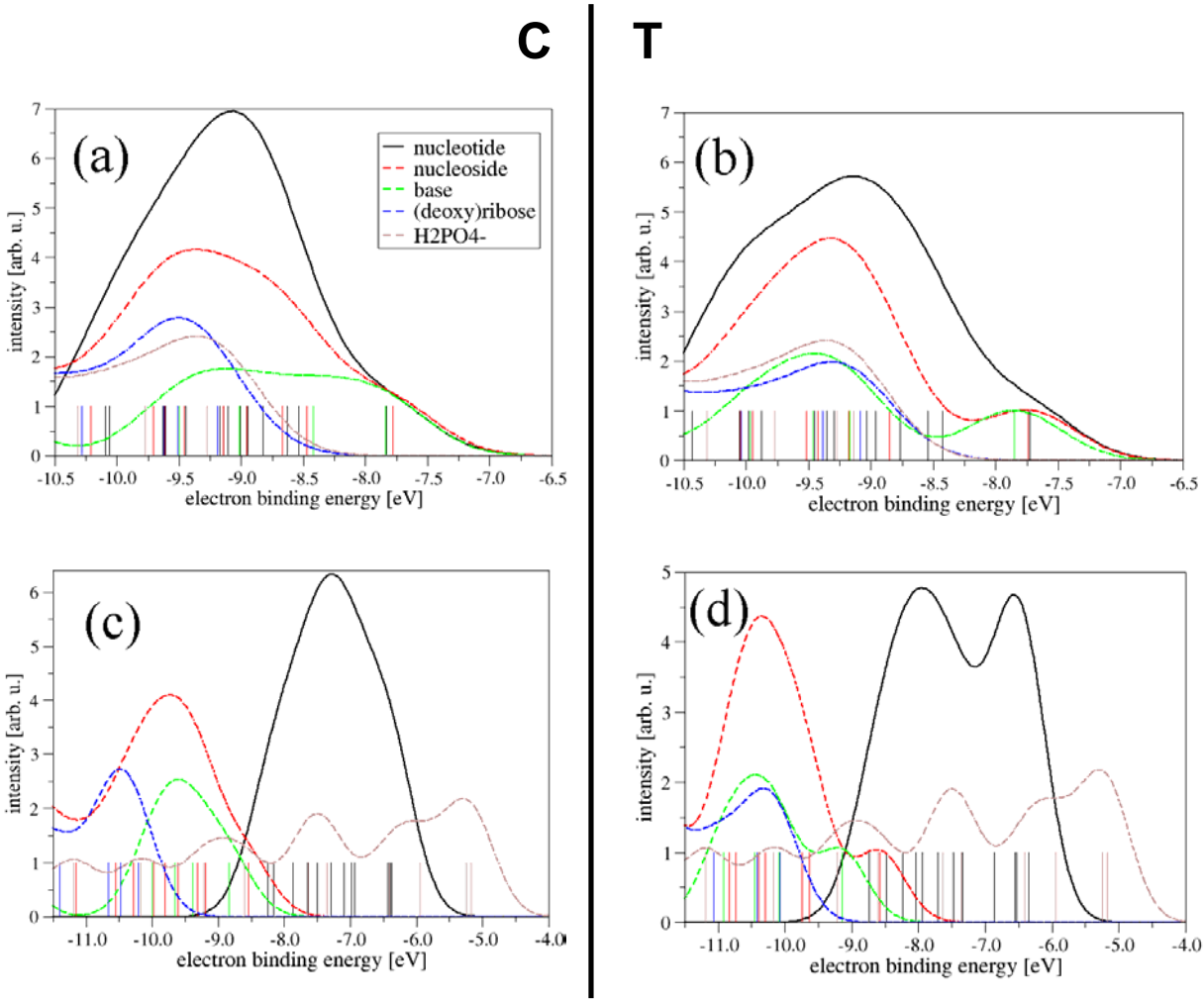

C
T
(a)
(b)
(c)
(d)
nucleotide
nucleoside
base
(deoxy)ribose
H2PO4-
intensity [arb. u.]
electron binding energy [eV]

Figure 4

TOC

X-ray photoelectron spectroscopy measurements from nucleosides C and T in water are complemented with ab initio calculations employing a non-equilibrium polarizable continuum model, which reveal that their ionization potentials are much less sensitive to the presence of the sugar and phosphate as compared to their gas-phase counter parts. This indicates a remarkable screening ability of the aqueous solvent.