

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/47642271>

Electronic Structure and Spectroscopy of Nucleic Acid Bases: Ionization Energies, Ionization-Induced Structural Changes, and Photoelectron Spectra

ARTICLE in THE JOURNAL OF PHYSICAL CHEMISTRY A · NOVEMBER 2010

Impact Factor: 2.69 · DOI: 10.1021/jp1063726 · Source: PubMed

CITATIONS

47

READS

22

6 AUTHORS, INCLUDING:



[Oleg Kostko](#)

Lawrence Berkeley National Laboratory

59 PUBLICATIONS 1,146 CITATIONS

SEE PROFILE



[Musahid Ahmed](#)

Lawrence Berkeley National Laboratory

127 PUBLICATIONS 2,323 CITATIONS

SEE PROFILE



[Anna I Krylov](#)

University of Southern California

178 PUBLICATIONS 6,980 CITATIONS

SEE PROFILE

Electronic Structure and Spectroscopy of Nucleic Acid Bases: Ionization Energies, Ionization-Induced Structural Changes, and Photoelectron Spectra

Ksenia B. Bravaya,[†] Oleg Kostko,[‡] Stanislav Dolgikh,[†] Arie Landau,[†] Musahid Ahmed,[‡] and Anna I. Krylov^{*,†}

Department of Chemistry, University of Southern California, Los Angeles, California 90089-0482, United States, and Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

Received: July 9, 2010; Revised Manuscript Received: October 7, 2010

We report high-level *ab initio* calculations and single-photon ionization mass spectrometry study of ionization of adenine (A), thymine (T), cytosine (C), and guanine (G). For thymine and adenine, only the lowest-energy tautomers were considered, whereas for cytosine and guanine we characterized the five lowest-energy tautomeric forms. The first adiabatic and several vertical ionization energies were computed using the equation-of-motion coupled-cluster method for ionization potentials with single and double substitutions. Equilibrium structures of the cationic ground states were characterized by DFT with the ω B97X-D functional. The ionization-induced geometry changes of the bases are consistent with the shapes of the corresponding molecular orbitals. For the lowest-energy tautomers, the magnitude of the structural relaxation decreases in the following series, $G > C > A > T$, the respective relaxation energies being 0.41, 0.32, 0.25, and 0.20 eV. The computed adiabatic ionization energies (8.13, 8.89, 8.51–8.67, and 7.75–7.87 eV for A, T, C, and G, respectively) agree well with the onsets of the photoionization efficiency (PIE) curves (8.20 ± 0.05 , 8.95 ± 0.05 , 8.60 ± 0.05 , and 7.75 ± 0.05 eV). Vibrational progressions for the S_0 – D_0 vibronic bands computed within double-harmonic approximation with Duschinsky rotations are compared with previously reported experimental photoelectron spectra and differentiated PIE curves.

I. Introduction

Ionization of DNA bases is one of the processes involved in DNA radiation and photodamage resulting in dangerous mutations with potential risk for cancer and neurodegenerative diseases.

In a realistic environment, ionization of DNA bases triggers several processes. Based on the gas-phase IEs, G is the easiest to oxidize; however, the positive charge does not necessarily remain on G and can migrate along the DNA strand as far as 200 \AA^{1-3} to the other low IE sites (i.e., extended G runs). Charge migration is coupled with ionization-induced proton transfer,^{4–10} which separates the hole and the radical center. These processes ultimately lead to the formation of oxo-guanine (8-oxo-7,8-dihydroguanine) and fused thymine dimers.¹¹

The interactions between the bases (and with solvent and backbone) affect the hole migration and proton transfer, as well as the character of the initial hole. Recently, we characterized the effect of hydrogen bonding and π -stacking on the ionized states in a combined experimental and theoretical study of the gas-phase dimers of AA, TT, AT, and CC.^{12,13} These studies compared and contrasted the electronic structure of the dimers with that of the monomers; however, the monomers' data have not been analyzed in detail. This work reports the electronic structure calculations and the photoionization efficiency (PIE) curves of the monomers.

Being of primary biological importance, ionized states of the DNA bases have been investigated both theoretically^{14–24} and experimentally;^{14–16,25–32} however, some questions are still unresolved. For example, the assignment of the bands in

photoelectron spectra (PES) as corresponding to π - or σ -like states of the cations is still controversial. A recent angle-resolved photoelectron study¹⁴ has shown that the five lowest ionized states of adenine, thymine, and cytosine are of π , σ , π , σ , π character. However, a study based on time-resolved PES has reported a different assignment for adenine.¹⁵ Although the relative order of the π - and σ -states was concluded to be the same, the 9.5 eV band in PES was assigned to an isolated D_1 (σ) ionized state, in contrast to the earlier photoelectron study of Trofimov et al.¹⁴ in which the 9.5 eV band was assigned to the closely lying σ - and π -ionized states.

The analysis of the PES of G and C is complicated by the presence of different tautomeric forms, which have different IEs and vibrational progressions. Thus, understanding the effect of tautomerization on the ionized states of nucleic acid bases (NABs) is important for interpretation of the experimental results.

Although the first PES of isolated DNA bases were obtained in the 1970s,^{25,26} the vibrationally resolved spectrum has been reported only for thymine.³³ The theoretical studies focused on assigning different bands in the PES based on vertical ionization energies and pole strengths calculations, while the Franck–Condon factors (FCFs) have not received much attention. The vibrational progressions contain information about ionization-induced geometric relaxation. Thus, they are related to the intrinsic response of the bases to ionization, which is a part of ionization-induced structural changes of DNA.

We present combined theoretical and experimental study of ionization of the NABs including all relevant low-energy tautomers of C and G (Figure 1). We report raw and differentiated PIE curves recorded in the range of 7.5–11.5 eV, as well as vertical (VIE) and adiabatic (AIE) ionization energies

[†] University of Southern California.

[‡] Lawrence Berkeley National Laboratory.

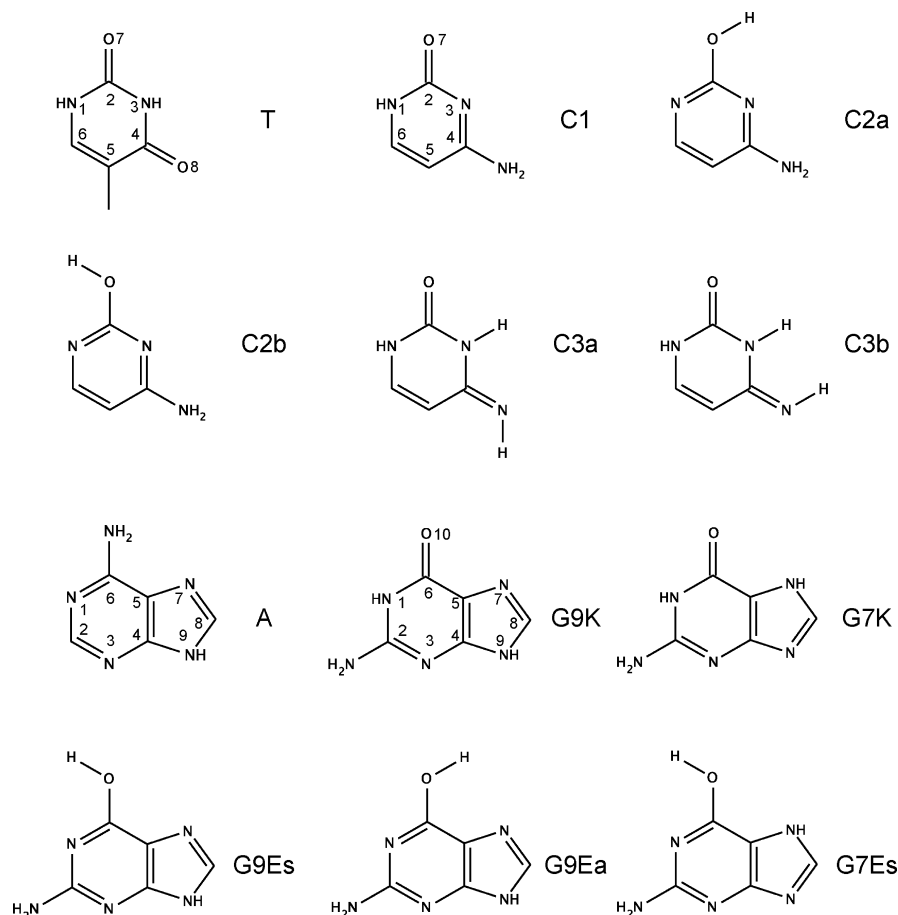


Figure 1. Structures of the thymine, adenine, cytosine, and guanine bases considered in this study.

computed using the equation-of-motion coupled-cluster method with single and double substitutions (EOM-IP-CCSD). We also report calculations of FCFs for the first band in photoelectron spectra.

The focus of the paper is on the ionization-induced geometry relaxation of the bases. We analyze patterns in structural relaxation in terms of molecular orbitals (MOs) and quantify its consequences in terms of FCFs, vibrational structure of the PES, and relaxation energies. The computed FCFs help to interpret the shape of PIEs and low-resolution PES. We hope that these computed spectra will be useful for future spectroscopic studies of the gas-phase NABs.

The structure of the paper is as follows. Section II describes the details of quantum chemistry methods and experimental techniques used in this work. Section III presents the results and discussion. We first discuss the geometries of the neutral bases and relative populations of tautomers based on our calculations and previously published experimental results. We then present VIEs and analyze the wave functions of the cationic states in terms of MOs. Next, we discuss structural relaxation of the cations and AIEs and present the FCFs calculations. The performance of double-harmonic approximation and electronic structure methods used for the PES calculations is validated by comparing the computed FCFs of thymine with an experimental mass-analyzed threshold ionization (MATI) spectrum.³³ Finally, we compare the computed spectra with the differentiated PIEs and previously published spectra.

II. Computational and Experimental Methods

A. Computational Methods. The geometries of the neutral nucleobases were optimized using MP2 with resolution of

identity approximation (RI-MP2).^{34–36} No symmetry constraints were imposed in geometry optimizations. For A, T, and C, Dunning's cc-pVTZ basis set³⁷ was used, whereas for guanine we employed an extended cc-pVTZ+diff basis³⁸ augmented by additional s- and p-diffuse functions for the second-row elements, and s-functions for hydrogen. The exponents were taken from the 6-311(2+,+)G basis set.^{39,40} RI-MP2 optimized geometries were used for IEs calculations. ω B97X-D/6-31+G(d,p) optimized geometries of the neutral bases were used for the analysis of the ionization-induced geometry changes and frequency calculations. For detailed comparison of the two sets of the optimized geometries, see the Supporting Information.

Accurate description of ionized states with open-shell character presents a challenge for quantum chemistry.^{41–43} Standard single-reference methods based on unrestricted Hartree–Fock (UHF) references, e.g., UMP2, suffer from spin contamination and tend to overestimate charge localization.⁴² The performance of DFT for this type of system is affected by notorious self-interaction error, which results in charge overdelocalization.

An elegant solution for the problematic open-shell states in ionized systems is provided by the EOM-IP-CCSD^{41,42,44–47} and IP-CISD⁴⁸ methods. In these approaches, the open-shell wave functions of the ionized states are derived from the well-behaved closed-shell wave function of the neutral. As in other EOM and CC methods, the so-called “orbital relaxation” is described by excitation operators (T_1 and R_2) rather than by tedious state-specific orbital optimization (the EOM and CC methods are invariant with respect to unitary transformations among occupied and/or among virtual orbital manifolds). This results in a balanced description of several target states and convenient formalism for calculation of interstate properties. Both IP-CCSD

and IP-CISD are size-intensive and spin-pure. IP-CISD employs uncorrelated Hartree–Fock determinant as a reference, in contrast to EOM-IP-CCSD which uses correlated CCSD reference. EOM-IP-CCSD is of course more accurate than IP-CISD, but the computational cost of the latter is N^5 .

Equilibrium structures of the cations in the ground electronic states were computed by IP-CISD, DFT with the long-range corrected functional ω B97, ω B97X, and ω B97X-D (the latter includes empirical correction for dispersion interactions⁴⁹), and EOM-IP-CCSD. The IP-CISD/6-31+G(d) geometries of the cations were used in calculations of the first AIEs for A, T, and C, whereas ω B97/cc-pVTZ(+dff) geometries were employed for G.

To compare the structural parameters for the neutral and ionized bases, we first analyzed the performance of the ω B97X-D functional and the IP-CISD method for geometry optimization of ionized species. The first approach may be affected by spin contamination, whereas the latter accounts only for the small portion of electron dynamical correlation. The benchmark calculations for the five ionized states of uracil⁴⁸ have shown that IP-CISD/6-31+G(d) mean absolute errors in bond lengths are about 0.015 Å. The performance of different functionals from the ω B97 family was investigated for G tautomers by comparing the results of ω B97X-D/6-31+G(d,p) with ω B97/6-31+G(d,p) and ω B97/cc-pVTZ+dff. The accuracy of IP-CISD and ω B97X-D structural parameters was verified by comparison with EOM-IP-CCSD/6-31+G(d) for the most problematic purine bases (A and G). We found that the ω B97X-D/6-31+G(d,p) geometry is closer to the EOM-IP-CCSD/6-31+G(d) one (see the Supporting Information) than the IP-CISD one. Thus, we chose to use the ω B97X-D/6-31+G(d,p) geometries for the analysis of the ionization-induced geometry changes. The ω B97X-D/6-31+G(d,p) structures of the cations were also used in frequency computations and vibrational analysis. A fine EML(75, 302) grid consisting of 75 points in the Lebedev⁵⁰ radial grid and 302 points in the Euler–Maclaurin⁵¹ angular grid was used for all DFT calculations.

VIEs and AIEs of A, T, and C were computed by EOM-IP-CCSD with the cc-pVTZ basis set.³⁷ For guanine, an augmented basis, cc-pVTZ+dff, and frozen natural orbitals (FNO) approximation were employed.³⁸ The IEs were computed by using four-point extrapolation (99.00%, 99.15%, 99.30%, and 99.45% thresholds for FNO population criteria) to the full virtual space values. The extrapolated results are expected to be within 1 kcal/mol relative to the full valence space calculations based on benchmark calculations for the guanine's G9K tautomer and other model systems.³⁸

Harmonic frequencies were computed at the ω B97X-D/6-31+G(d,p) level of theory for both the neutral and the first ionized states of the bases (at the equilibrium geometries optimized at the same level). The unscaled frequencies were used in calculations of zero-point energy (ZPE) correction for AIEs. The frequencies were scaled by 0.97 to obtain the positions of the vibronic lines in FCF calculations. The FCFs were obtained within double-harmonic approximation with Duschinsky rotations⁵² at the C_s geometry using normal modes and frequencies for the nonplanar structures. This can be justified by low pyramidalization barriers for the bases suggesting delocalized vibrational wave functions and effectively planar vibrationally averaged structures. The barriers of 0.050, 0.832, and 0.060 kcal/mol have been reported using the MP2/aug-cc-pVQZ method⁵³ for canonical forms of the A, G, and C, bases, respectively. These values are less than or at least comparable with the NH_2 out-of-plane vibrational frequencies, which lie

above 300–350 cm^{-1} . Thus, the maximum of the anharmonic ground-state vibrational wave function is located above the transition-state planar structure, and the vibrationally averaged geometries are planar. The Duschinsky rotations are important because the normal-mode overlap matrix for the neutral and the ionized states is significantly nondiagonal; e.g., we observed multiple nondiagonal elements of the normal mode overlap matrix up to 0.4–0.7 for all bases. The ezSpectrum program⁵⁴ was used to compute the spectra. All quantum chemistry calculations were performed with the Q-Chem package of quantum chemistry programs.⁵⁵

B. Experimental Setup. The experiments were performed on a molecular beam apparatus coupled to a 3 m VUV monochromator on the Chemical Dynamics Beamline at the Advanced Light Source. The gas-phase NABs are prepared by thermal desorption ($T = 495$ K for A and T, $T = 582$ K for C, and $T = 698$ K for G) and cooled in a supersonic beam. The thermal vaporization source has been described recently in a publication detailing the microsolvation of DNA bases.⁵⁶ In the current experiments, the backing pressure of Ar expanded through a 100 μm diameter nozzle was 35 kPa. The bases are ionized by tunable synchrotron radiation in the 7.4–11.5 eV region, and the ions are detected by a time-of-flight mass spectrometer. For each mass, the yield of the ions is measured as a function of photon energy, which produces PIE spectra. The typical step size for the PIE scans is 50 meV, and a dwell time of 10 s at a repetition rate of 10 kHz. The differentiation of the PIE curves following the method used by Berkowitz in interpreting photoionization of methanol⁵⁷ produces a spectrum similar to a photoelectron spectrum from which information about vibrational progressions and other electronic states can be extracted. The differentiation is performed numerically after taking a two-points nearest-neighbor average to reduce the effects of noise in the PIE. The accuracy of reported onset energies in the PIE spectra is 0.05 eV.

III. Results and Discussion

A. Neutral Bases: Relative Energies of Different Tautomers. Gas-phase NABs exist in a number of tautomeric forms. Thus, the interpretation of the experimental data involves the analysis of the composition of the gas-phase mixture. Although molecular beam experiments involve several nonequilibrium steps, the relative energies and the respective Boltzmann populations provide a useful estimate of the population of tautomers.

A thorough study of the ground-state structures of NABs has been reported by Wand and Schaefer.⁵³ Analysis of the geometries optimized at a different level of theory used in this work and comparison to the results of the previous studies is presented in the Supporting Information. A systematic study of the relative energies of adenine tautomers has been performed by Guerra et al.⁵⁸ and by Hanus et al.⁵⁹ The two tautomers closest to 9H-adenine (the lowest-energy one) are 3H-adenine and 7H-adenine; both forms lie higher by 7.0–7.5 kcal/mol.^{58,59} The energy difference between the lowest-energy canonical form of T and other tautomers is even higher. Piacenza and Grimme⁶⁰ have reported 9.4 and 9.9 kcal/mol energy gap between the lowest energy and the second stable tautomer [computed by QCISD(T)/TZV(2df,2pd) and MP2/TZV(2df,2dp), respectively].

The relative energies of the lowest tautomers of G and C, which are considered in this study, are given in Table 1. Table 1 also lists dipole moments of the tautomers. Interestingly, biologically relevant forms have the largest dipole moments and become the most stable in solution.^{59,61,62} Moreover, large dipole

TABLE 1: Relative Energies of Different Tautomers of NABs and Absolute Value of the Dipole Moment Computed at the RI-MP2 Optimized Equilibrium Structures^a

NAB	ΔE , kcal/mol		D , debye
	RI-MP2	CCSD//RI-MP2	RI-MP2
T	—	—	4.106
A	—	—	2.606
C1	2.01	1.35 (0.29)	6.213
C2a	0.71	0.69 (0.20)	4.472
C2b	0.00	0.00 (0.38)	3.186
C3a	3.21	1.15 (0.10)	4.543
C3b	4.90	2.75 (0.03)	2.317
G9K	0.63	0.76 (0.25)	6.296
G7K	0.00	0.00 (0.42)	1.860
G9Es	0.50	1.28 (0.16)	3.022
G9Ea	0.68	1.43 (0.15)	3.811
G7Es	3.39	4.41 (0.02)	4.218

^a Boltzmann populations for different tautomers computed for the free energy differences obtained within rigid-rotor harmonic-oscillator approximation (ω B97X-D/6-31+G(d,p) frequencies) using CCSD relative energies are given in parentheses. cc-pVTZ (A, T, C) and cc-pVTZ-dff (G) basis sets are used for geometry optimization.

moments contribute toward stability of H-bonded and stacked dimers of the biologically relevant forms. Even in the gas phase, the hydrogen-bonded dimers of the higher energy tautomers can be lower in energy than the dimers of the most stable tautomers; see, for example, a recent cytosine study.¹³ Both solvent stabilization and efficient dimer formation might have important implications for the biological functioning of these tautomers.

Boltzmann populations of tautomeric forms for the experimental vaporization temperature ($T = 582$ K for C and $T = 698$ K for G) based on free energies obtained within the rigid-rotor harmonic-oscillator approximation are given in Table 1. Assuming thermal equilibrium, only the canonical forms of A and T can be populated. In agreement with this energetics, a single tautomer of T (canonical form) was observed in molecular beam Fourier transform microwave spectroscopy⁶³ and infrared laser spectroscopy studies of thymine in helium nanodroplets.⁶⁴ Nir et al.⁶⁵ reported that two tautomers of adenine are present in the gas phase based on IR-UV spectroscopy. In this experiment, laser desorption was used to generate gas-phase species, which may result³² in higher-energy isomers being populated relative to thermal desorption employed in our work.

Relative populations of C and G tautomers suggest that several tautomeric forms might be present in the molecular beam (see Table 1). Moreover, tautomeric composition of the gas phase depends on the vaporization method. In particular, it has been shown using VUV-SPI mass spectrometry analysis that different tautomeric forms of G are populated depending on whether thermal vaporization or laser desorption is used to evaporate the sample.³² Nir et al. reported that four and two tautomers of G and C, respectively, are populated based on R2PI and IR-UV spectra combined with the analysis of the spectra of methylated derivatives. The four tautomers for G were reported to be two enol and two oxo forms.⁶⁵ One oxo and one enol form of C were registered.⁶⁵ It should be noted that results of R2PI studies depend strongly on the corresponding excited state lifetime and relative absorption cross sections of different tautomeric forms. Marian⁶⁶ suggested that the bands corresponding to the oxo forms in the IR-UV experiments⁶⁵ originate from the imino-oxo isomers, whereas the amino-oxo forms do not appear in the spectra due to poor FCFs and a conical intersection located in the vicinity of the Franck-Condon region of vertical $\pi-\pi^*$ excitation, which results in the short excited-

state lifetime. Assignment of the IR spectra for guanine in He nanodroplets pointed to the presence of G7K, G9K, G9Es, and G9Ea tautomers.⁶⁷ Recent Fourier transform microwave spectroscopy studies using laser ablation molecular beam (LA MB FTMW) supported this assignment.⁶⁸ Moreover, the estimated relative populations of the four tautomers based on the integrated intensities of IR bands 1.00:0.76:0.35:0.28 for G7K:G9K:G9Es:G9Ea⁶⁷ are very close to those estimated assuming thermal equilibrium (see Table 1). The existence of three tautomers of C was confirmed by Choi et al.⁶⁹ based on IR laser spectroscopy in He nanodroplets and analysis of vibrational transition moment angles. The spectrum was assigned to C2b, C2a, and C1 cytosine tautomers,⁶⁹ which is also consistent with the estimated Boltzmann populations (Table 1). Our previous study of cytosine¹³ also suggested that at least three tautomers are present in the beam. Recent core level photoionization studies of cytosine reported a 0.26:0.11:0.63 ratio for the C1:(C3):(C2a+C2b) tautomers based on the fitting of the experimental O(1s) ionization spectrum by three Gaussians.⁷⁰ Although C3b was not considered in this study in view of its higher energy, one can expect the ionization energy from the O(1s) level to be close to that of the C3a tautomer. Therefore, C3a and C3b will contribute to the same peak. The gas-phase mixture in these experiments was obtained using thermal vaporization ($T = 450$ K).⁷⁰ Our C1:(C3a+C3b):(C2a+C2b) ratio computed with Boltzmann populations is 0.29:0.13:0.58 and 0.29:0.10:0.61 for $T = 582$ and 463 K, respectively. Both are in perfect agreement with the experimental ratio and estimated Boltzmann populations for $T = 450$ K reported in ref 70.

To summarize, the two experimental estimates of the gas-phase G and C tautomers populations are based on two very different techniques, i.e., IR spectroscopy in the helium nanodroplets⁶⁷ and core level photoionization spectroscopy.⁷⁰ In both cases, the estimated populations comply with the Boltzmann distribution. Our thermal vaporization conditions are comparable to the He droplet and effusive sources mentioned above. Contrary to typical sources used to generate biomolecules in the gas phase, in our source, the biomolecules are entrained in carrier gas prior to expansion in the nozzle.

Therefore, we do not anticipate significant deviations from the Boltzmann distribution even under supersonic expansion conditions. Indeed, assuming thermal equilibrium for the gas-phase monomers prior to supersonic expansion, the changes in the relative population can occur only via tautomerization reactions, which can proceed only under collisional conditions or upon reaction in the clusters and following dissociation. Both are unlikely processes under the supersonic cooling conditions of our molecular beam.

B. Vertical Ionization Energies and the Character of Ionized States. 1. Thymine. VIEs of thymine computed at different levels of theory are listed in Table 2. The corresponding MOs are shown in Figure 2. The EOM-IP wave functions of the six lowest ionized states are dominated by a single determinant derived by removing an electron from a canonical HF orbital. The lowest R_1 amplitude, 0.804, was observed in the highest ionized state derived by ionization from the MO comprised of symmetry-adapted combination of the bonding C-H orbitals. The wave functions of lower ionized states corresponding to ionization from π - or lone-pair orbitals are even more Koopmans-like, i.e., the R_1 amplitudes are ≥ 0.934 . This means that correlated Dyson orbitals⁷¹ are very similar to the canonical Hartree-Fock MOs and that the perturbative methods like OVGF and P3 should be able to describe these states accurately.

TABLE 2: Vertical Ionization Energies of the Six Lowest Ionized States of Thymine (eV)

	D_0	D_1	D_2	D_3	D_4	D_5	ref
P3//MBPT(2)/6-311G(d,p)	9.14 (π)	9.95 (σ)	10.43 (π)	10.99 (σ)	12.52 (π)	—	17
OVGF/6-311++G(d,p)// B3LYP/6-311G(d,p)	8.85 (π)	10.46 (π)	10.46 (σ)	11.36 (σ)	12.52 (π)	13.92 (σ)	14
ADC(3)/6-31G// B3LYP/6-311G(d,p)	8.59 (π)	10.00 (π)	10.12 (σ)	10.95 (σ)	12.31 (π)	13.68 (σ)	14
CASPT2(IPEA=0.25)// CASSCF/ANO-L 431/21	9.07 (π)	9.81 (σ)	10.27 (π)	10.49 (σ)	12.37 (π)	13.82 (π)	21
EOM-IP-CCSD/ cc-pVTZ// RI-MP2/cc-pVTZ	9.13 (π)	10.13 (σ)	10.52 (π)	11.04 (σ)	12.67 (π)	13.82 (σ)	this work
experiment	9.19 (π)	10.14 (σ)	10.45 (π)	10.89 (σ)	12.27 (π)	13.31	14
	9.20 (π)	10.05 (σ)	10.44 (π)	10.88 (σ)	12.30 (π)	—	74
	9.18 (π)	10.03 (σ)	10.39 (π)	10.82 (σ)	12.27 (π)	—	28
	9.02 (π)	9.95 (σ)	10.40 (π)	10.80 (σ)	12.10 (π/π)	—	29

The order of the six lowest ionized states is in agreement with the results of Trofimov et al.¹⁴ For π ionized states, EOM-IP-CCSD VIEs for thymine agree well with the P3 propagator calculations¹⁷ and with the CASPT2 values.²¹ OVGF and ADC(3) yield a qualitatively different picture by changing the order of the ionized states¹⁴ (see Table 2). The quantitative deviations from EOM-IP-CCSD are also large. For example, OVGF VIE for the first ionized σ -like state is 0.33 eV higher than the EOM-IP-CCSD value, whereas the VIE for the first π -ionized state is 0.28 eV lower. ADC(3) exhibits a similar trend overestimating IEs of σ -like states relative to the π -like ones. Small basis sets used in ADC calculations result in strong underestimation of VIEs. The CASPT2 results show an opposite trend²¹ underestimating the IEs of the σ states, likely due to a small number of active σ -orbitals (i.e., only the heteroatoms lone pairs were included in the CASSCF active space). Note that the differences between CASPT2 and EOM-IP-CCSD (and the experimental values) increase for higher ionized states, possibly due to limited active space, which favors the states derived by ionization from higher orbitals. Note that all the methods (except CASPT2) consistently overestimate IEs corresponding to the D_4 and D_5 states of the cation by 0.3–0.5 eV relative to the experimental values.

2. Adenine. Relevant MOs and VIEs of A are given in Figure 3 and Table 3. As in the thymine case, the propagator methods

yield similar errors for ionization from the σ - and π -orbitals. CASPT2 again underestimates VIEs for the σ -ionized states. The order of the states (which are all Koopmans-like) is π , σ , π , σ , π , σ . The lowest R_1 amplitude for the EOM-IP-CCSD wave functions is 0.938 for the D_2 ionized state. The relative state ordering for five lowest ionized states is consistent for all methods except OVGF and ADC(3). Moreover, VIEs and the assignments agree with that reported by Trofimov and co-workers,¹⁴ contrary to the results of Stolow and co-workers.¹⁵ The discrepancy with the TR-PES data could be explained by a low cross section for the $S_2(\pi\pi^*) \rightarrow D_2$ ionization; however, a detailed explanation of this missing D_2 band requires further investigation. The EOM-IP-CCSD VIEs are in excellent agreement with the experimental¹⁴ VIE values.

3. Cytosine. Similar trends in performance of different ab initio methods as described for T and, to a lesser extent, for A were also observed for the G and C bases. The comparison of the EOM-IP-CCSD ionization energies of the G and C canonical tautomers with those computed with other methods^{14,16,19–21,72} is given in the Supporting Information. The effects of the tautomerization on the order and the character of the ionized states of the noncanonical tautomers of G and C were previously discussed by Dolgounitcheva et al.^{19,20} on the basis of P3 calculations. As was discussed above, the EOM-IP-CCSD VIEs are close to the previously reported P3 results as all the states are of Koopmans-like character.

Figure 4 and Table 4 present the MOs and VIEs of the lowest tautomers of cytosine. Tautomerization has a prominent effect on the ionized states.¹³ As follows from Figure 4, the state

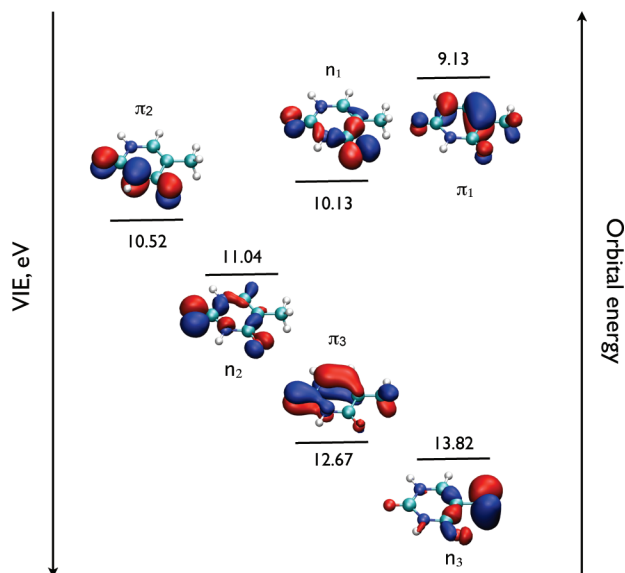


Figure 2. Vertical ionization energies (EOM-IP-CCSD/cc-pVTZ) and the corresponding MOs for thymine.

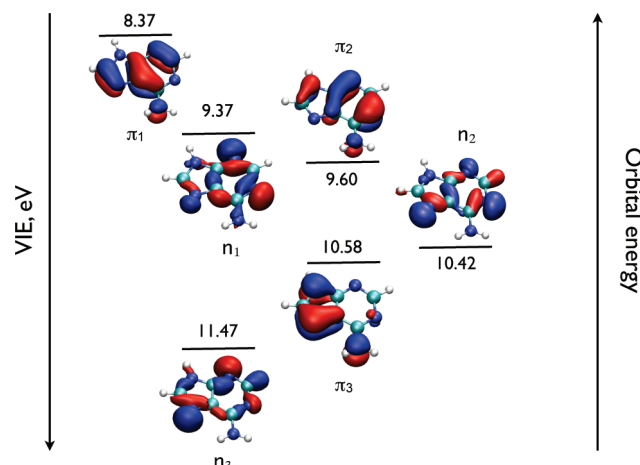


Figure 3. Vertical ionization energies (EOM-IP-CCSD/cc-pVTZ) and the corresponding MOs for adenine.

TABLE 3: Vertical Ionization Energies of the Six Lowest Ionized States of Adenine (eV)

state	D_0	D_1	D_2	D_3	D_4	D_5	ref
P3//MBPT(2)/6-311G(d,p)	8.49 (π)	9.45 (σ)	9.56 (π)	10.48 (σ)	10.50 (π)	11.54 (σ)	18
OVGF/6-311++G(d,p)// B3LYP/6-311G(d,p)	8.32 (π)	9.40 (π)	9.45 (σ)	10.50 (π)	10.53 (σ)	11.61 (σ)	14
ADC(3)/6-31G// B3LYP/6-311G(d,p)	7.93 (π)	9.20 (σ)	9.36 (π)	10.23 (σ)	10.26 (π)	11.20 (σ)	14
ADC(3)/6-311++G(d,p)// MP2/6-311G(d,p)	8.33 (π)	9.63 (π)	9.78 (σ)	10.65 (π)	10.81 (σ)	11.84 (σ)	23
CASPT2(IPEA=0.25)// CASSCF/ANO-L 431/21	8.37 (π)	9.05 (σ)	9.54 (π)	9.96 (σ)	10.38 (π)	11.06 (π)	21
EOM-IP-CCSD/ cc-pVTZ// RI-MP2/cc-pVTZ	8.37 (π)	9.37 (σ)	9.60 (π)	10.42 (σ)	10.58 (π)	11.47 (σ)	this work
experiment	8.47 (π)	9.45 (σ)	9.54 (π)	10.45 (σ)	10.51 (π)	11.35 (σ)	14
	8.5 (π)	9.6 (σ)	—	10.5 (π, σ)	—	—	15
	8.48 (π)	9.6 (σ, π)	—	10.5 (σ, π)	—	11.39 (σ)	28, 75

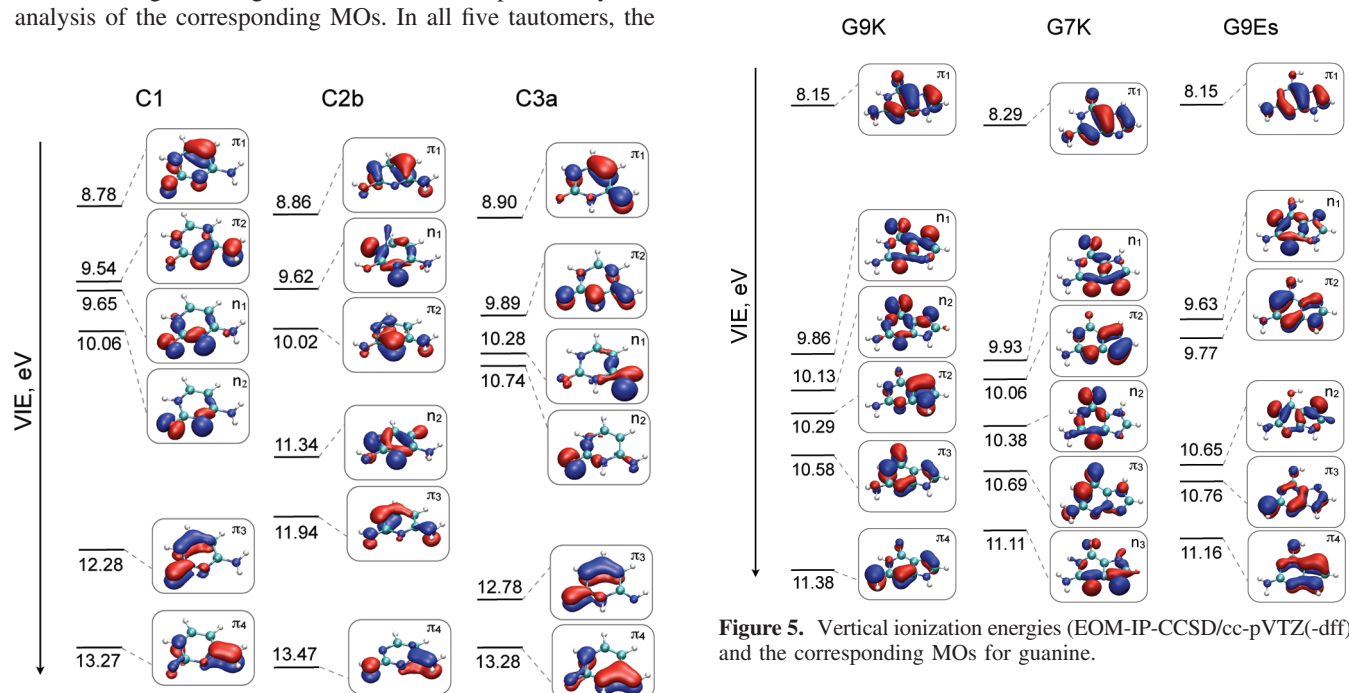
ordering in the most stable oxy tautomer (C2b) is different from that in the C1 oxo form. The predicted order of the six lowest ionized states for C2b is π , σ , π , σ , π , σ , in agreement with the experimental assignment.¹⁴

4. Guanine. Both VIEs and the character of the ionized states of G are affected by tautomerization, as follows from Table 5 and Figure 5. Although the character of the lowest ionized state is similar in all five tautomeric forms (see Figure 5 and the Supporting Information), the first VIE for G7K is noticeably shifted to higher energies, which can be explained by the analysis of the corresponding MOs. In all five tautomers, the

TABLE 5: EOM-IP-CCSD/cc-pVTZ-dff//RI-MP2/cc-pVTZ-dff Vertical Ionization Energies for the Six Lowest Ionized States of Guanine Tautomers^a

state	D_0	D_1	D_2	D_3	D_4	D_5
G9K ³⁸	8.15 (π)	9.86 (σ)	10.13 (σ)	10.29 (π)	10.58 (π)	11.38 (π)
G7K	8.29 (π)	9.93 (σ)	10.06 (π)	10.38 (σ)	10.69 (π)	11.11 (σ)
G9Es	8.15 (π)	9.63 (σ)	9.77 (π)	10.65 (σ)	10.76 (π)	11.16 (π)
G9Ea	8.17 (π)	9.58 (σ)	9.79 (π)	10.72 (π)	10.82 (σ)	11.28 (π)
G7Es	8.17 (π)	9.52 (σ)	9.65 (π)	10.52 (σ)	10.74 (π)	11.27 (π)
experiment ¹⁶		8.26 (A)	9.81 (B)	10.36 (C)	11.14 (D)	

^a IEs for lowest energy G7K tautomer are marked in bold.

**Figure 5.** Vertical ionization energies (EOM-IP-CCSD/cc-pVTZ(-dff)) and the corresponding MOs for guanine.**Figure 4.** Vertical ionization energies (EOM-IP-CCSD/cc-pVTZ) and the corresponding MOs for cytosine.**TABLE 4: EOM-IP-CCSD/cc-pVTZ//RI-MP2/cc-pVTZ Vertical Ionization Energies (eV) for the Six Lowest Ionized States of Cytosine Tautomers^a**

state	D_0	D_1	D_2	D_3	D_4	D_5
C1	8.78 (π)	9.55 (π)	9.65 (σ)	10.06 (σ)	12.28 (π)	13.27 (π)
C2a	8.86 (π)	9.58 (σ)	10.12 (π)	11.38 (σ)	11.91 (π)	13.52 (π)
C2b	8.86 (π)	9.62 (σ)	10.02 (π)	11.34 (π)	11.94 (π)	13.48 (π)
C3a	8.90 (π)	9.89 (π)	10.28 (σ)	10.74 (σ)	12.78 (π)	13.28 (π)
C3b	8.88 (π)	10.01 (π)	10.19 (σ)	11.06 (σ)	12.75 (π)	13.31 (π)
experiment ¹⁴	8.89 (π)	9.55 (σ)	9.89 (π)	11.20 (σ)	11.64 (π)	12.93 (π, σ)

^a IEs for lowest energy C2b tautomer are marked in bold.

HOMO involves the bonding N7–C8 orbital. Neutral G9K has a double bond between these atoms, whereas in G7K this bond is single. The N7–C8 π bond appears only in the resonance structures of G7K that involve positively charged imine fragment at N7. Thus, ionization from this orbital requires higher energy. Additionally, N7 in G9K is not protonated and its lone pair, which lies in the plane of the molecule, does not participate in the delocalized π -system, whereas in the G7K isomer N7 donates its pair to the conjugated system thus acquiring a partial positive charge. This also contributes to the increase of IE in the G7K tautomer. The opposite effect is observed for the π_2 -orbital involving N9. In G7K, the lone pair on N9, which lies in the molecular plane, pushes this orbital to higher energies

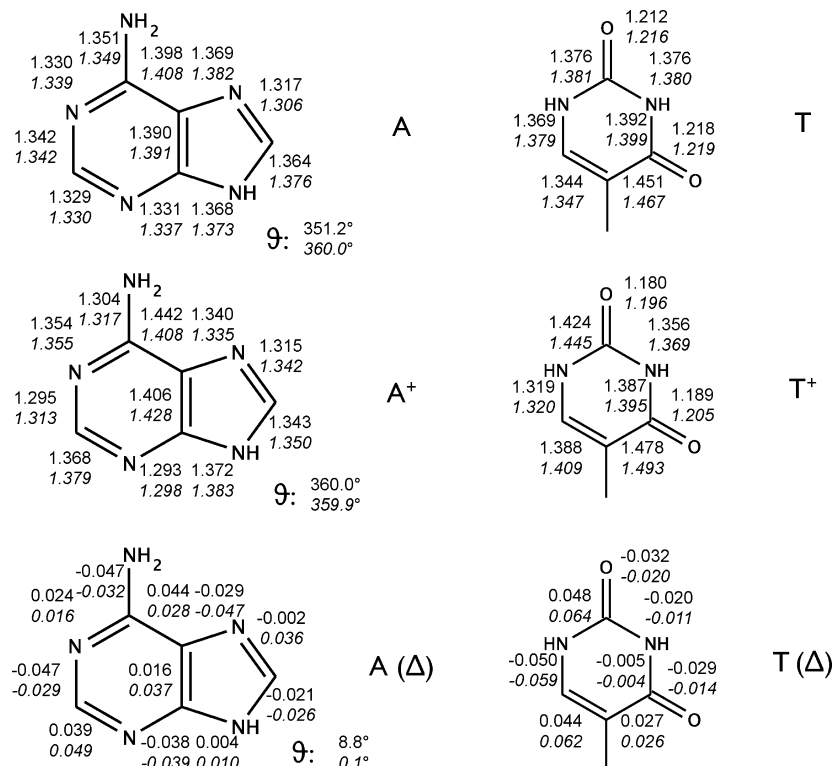


Figure 6. Relevant bond lengths for A and T and the ionization-induced geometry changes. Neutrals: RI-MP2/cc-pVTZ and ω B97X-D/6-31+G(d,p) (italics) optimized geometries. Cations: IP-CISD/6-31+G(d) and ω B97X-D/6-31+G(d,p) (italics) optimized geometries. Ionization-induced bond length changes (Δ): RI-MP2/cc-pVTZ-IP-CISD/6-31+G(d) and ω B97X-D/6-31+G(d,p) (italics).

thus decreasing the IE relative to the G9K form. As in the case of the C tautomers, tautomerization of G also changes the order of the ionized states. Note that the analysis of the tautomerization effect on VIEs is more complicated for higher ionized states as the shapes of the MOs become different for different tautomers.

C. Ionization-Induced Structural Relaxation and Adiabatic Ionization Energies. The ionization-induced geometry changes of NABs are summarized in Figures 6, 7, and 8. The corresponding Cartesian geometries are given in the Supporting Information. Because the character of ionized states is different for different tautomeric forms, the structural relaxation patterns also vary. Computed AIEs and the reference experimental data are given in Table 6.

As a quantitative measure of the overall geometry change upon ionization, we consider the difference between the geometries of the neutral and the cation positioned such that their centers of mass coincide and their principal axes are optimally aligned. The magnitude of the geometry change is defined as

$$dQ = \|\mathbf{Q}_f - \mathbf{Q}_i\| \quad (1)$$

where \mathbf{Q}_f and \mathbf{Q}_i are 3*N*-dimensional vectors of atomic coordinates of the optimized cation and the neutral structures, respectively. The values of *dQ* correlate with the relaxation energies, i.e., difference between vertical and electronic (i.e., non-ZPE-corrected) AIEs computed with EOM-IP-CCSD/cc-pVTZ. The smallest geometry change (*dQ* = 0.17 Å) was observed in T, and the corresponding relaxation energy for the first ionized state equals 0.20 eV.

The ionization-induced geometry changes can be explained by the shapes of canonical HF MOs. The distance between the atoms for which the bonding or antibonding overlap is present

in the HOMO should increase or decrease, respectively. The HOMO of T has the strongest contribution from the bonding C5–C6 and N1–C6 antibonding π -orbitals. Indeed, ionization leads to the increase of C5–C6 and the decrease of the N1–C6 distances; i.e., C5–C6 and N1–C6 distances are 1.344 and 1.379 Å for the neutral, whereas in the cation these distances are 1.409 and 1.320 Å. The *dQ* value for A is 0.21 Å resulting in larger relaxation energy (0.25 eV) relative to T. The removal of an electron from the HOMO of A should lead to the increase of the interatomic C2–N3, C5–C6, and N7–C8 distances and shortening of the N3–C4, C5–N7, and C6–N(H₂) bonds. As follows from Figure 6, the bond lengths from the first group increase by more than 0.028 Å, whereas the interatomic distances from the second group decrease by more than 0.032 Å upon ionization. Such qualitative considerations are valid for all the species considered in this study.

Another important change induced by ionization is planarization of the NH₂ group due to electronic density transfer from the nitrogen lone pair to the main π -system compensating for the loss of a π -electron. This change was suggested as a possible explanation for the lower FCFs at the PIE onset of amino-guanine tautomers relative to the imino ones.³² Planar configuration of the NH₂ group in the ionized state is common for all bases with the NH₂ moiety (Figures 6–8). However, the degree of nonplanarity in the neutral bases varies. The θ value, which is the sum of three valence angles at nitrogen atoms, is 351°, 350°, and 338° for the optimized RI-MP2 geometries of neutral A, C2b, and G7K, respectively. Note that degree of nonplanarity is the highest for guanine.

Cytosine shows moderate structural changes upon ionization; i.e., the *dQ* values for C1, C2a, C2b, C3a, and C3b tautomers are 0.28, 0.27, 0.23, 0.10, and 0.15 Å. The relaxation energy for the lowest energy C2b tautomer is higher than that for A and T, in agreement with a larger *dQ* value. The HOMO of

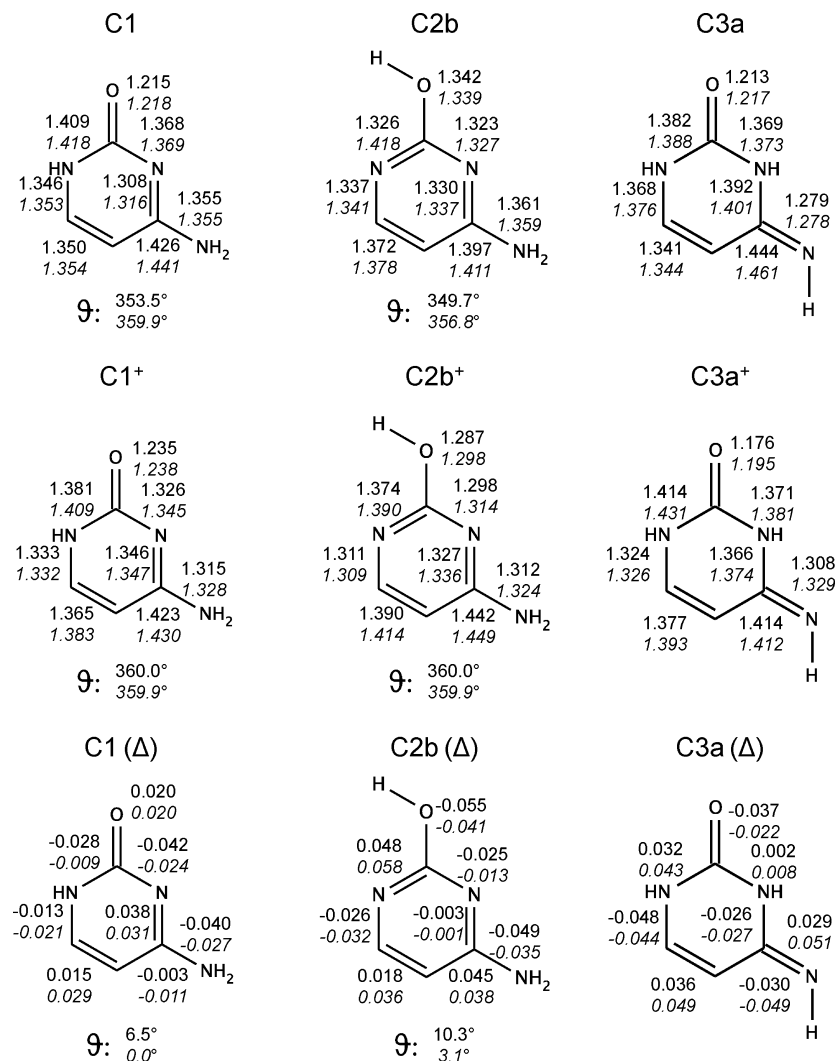


Figure 7. Relevant bond lengths for cytosine's C1, C2b, and C3a tautomers and the ionization induced-geometry changes. Neutrals: RI-MP2/cc-pVTZ and ω B97X-D/6-31+G(d,p) (*italics*) optimized geometries. Cations: IP-CISD/6-31+G(d) and ω B97X-D/6-31+G(d,p) optimized geometries. Ionization-induced bond length changes (Δ): RI-MP2/cc-pVTZ-IP-CISD/6-31+G(d) and ω B97X-D/6-31+G(d,p) (*italics*).

C2b is of bonding character for C4–C5, C5–C6, and N1–C2 bonds. Thus, the distances between the corresponding atoms increase upon ionization.

The guanine tautomers have the highest relaxation energy among the four NABs. The HOMO is very similar in the G9K and G7K tautomers and has N7–C8, C4–C5, C5–C6, and C2–N3 bonding and C5–N7, N3–C4, and C6–O antibonding character. Thus, one may expect elongation for the first group and contraction for the second group of bonds. The observed ionization-induced changes in the interatomic distances for G9K are 0.033, 0.056, 0.025, and 0.057 Å for the N7–C8, C4–C5, C5–C6, and C2–N3 bonds, respectively. The changes in the interatomic distances for antibonding C5–N7, N3–C4, and C6–O orbitals are -0.051 , -0.057 , and -0.016 Å, respectively. Similar trends are observed for the G7K tautomer. The shape of the HOMO for the enol tautomers is slightly different; see, for example, the G9Es HOMO shown in Figure 5. It has the N1–C2, C2–N3, C4–C5, C5–C6, and N7–C8 bonding and NH₂–C2, N1–C6, N3–C4, and C5–N7 antibonding character. The changes in the corresponding bond lengths for the bonding orbitals are 0.010, 0.043, 0.009, 0.047, and 0.038 Å, respectively. For the antibonding orbitals, the changes are -0.049 , -0.017 , -0.038 , and -0.057 Å.

In sum, for all NABs considered in this study, we observe an excellent correlation between the geometry relaxation and

the shape of the MO from which an electron is removed at the geometry of the neutral. The largest structural changes leading to the highest relaxation energy were observed for guanine. Thus, not only is the VIE of guanine the lowest among the four bases, but the relaxation further increases the gap in the respective AIEs.

D. Franck–Condon Factors for the S_0 – D_0 Transition in Thymine. In addition to relaxation energy, the character and the magnitude of the ionization-induced geometry changes determine the shape of vibrational progressions in photoelectron spectra. A vibrationally resolved photoionization spectrum is only available for thymine;³³ thus, we use this example to validate the accuracy of the computed FCFs.

The ionization-induced changes in bond lengths (described above) along with the changes in valence angles, which are also perturbed, result in the prominent displacements along eight normal modes shown in Figure 9. In addition, normal mode ν_{39} is active due to a large frequency change from 200 cm⁻¹ in the neutral to 58 cm⁻¹ in the cation and appears as an overtone in S_0 – D_0 spectra (Figure 10a). A similar spectrum has been reported previously by Zhou et al.⁷³

The modes are labeled according to symmetry and energy order for the cation. The resolution of the experimental PES and differentiated PIE spectra is not sufficient for resolving vibrational structure. To compare the computed spectra with

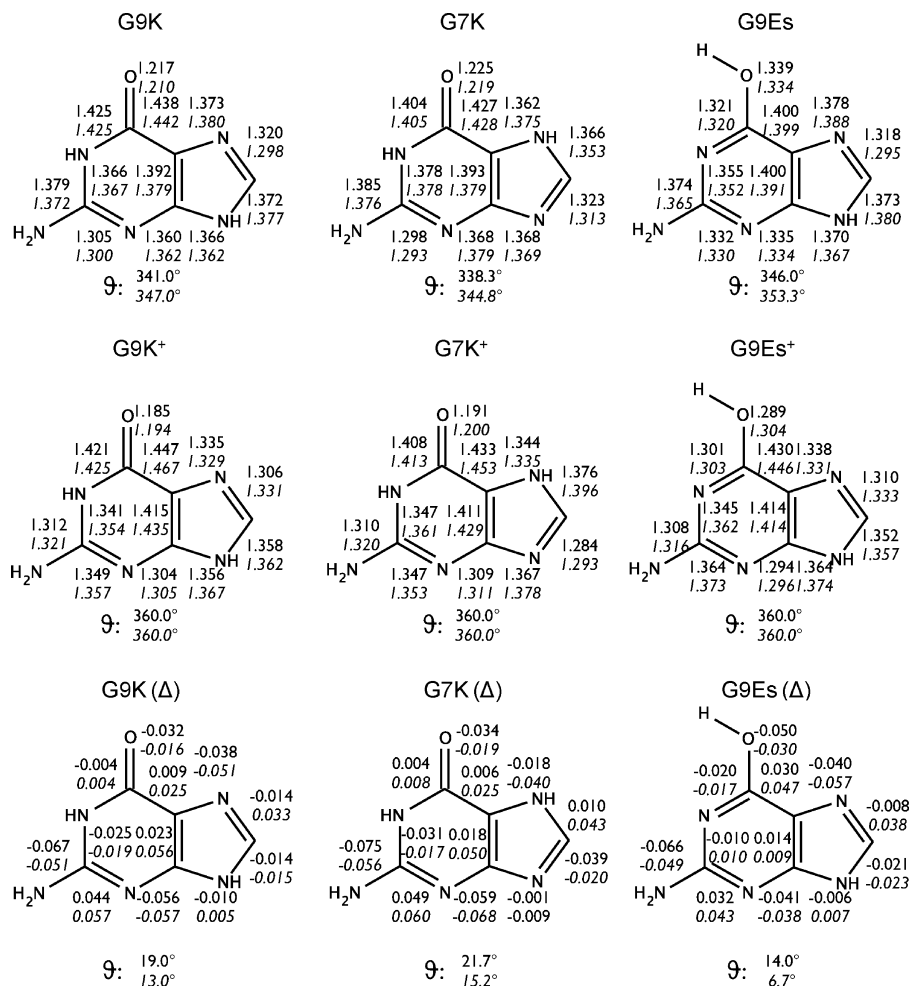


Figure 8. Relevant bond lengths for guanine's G9K, G7K, and G9Es tautomers and the ionization-induced geometry changes. Neutrals: RI-MP2/cc-pVTZ and ω B97X-D/6-31+G(d,p) (italics) optimized geometries. Cations: IP-CISD/6-31+G(d) and ω B97X-D/6-31+G(d,p) optimized geometries. Ionization-induced bond length changes (Δ): RI-MP2/cc-pVTZ-IP-CISD/6-31+G(d) and ω B97X-D/6-31+G(d,p) (italics).

TABLE 6: Computed Adiabatic IEs and Appearance Energies

method	T	A	C1	C2a	C2b	C3a	C3b	G9K	G7K	G9Es	G9Ea	G7Es
theory												
B3LYP ^{a, 22}	8.81	8.11	8.63	—	—	—	—	7.69	—	—	—	—
B3LYP ^{b, 76}	—	—	8.58	8.47	8.48	8.44	8.50	—	—	—	—	—
B3LYP ^{c, 32}	—	—	—	—	—	—	—	7.68	7.80	7.70	7.75	7.76
CASPT2 ^{d, 21}	8.69	8.11	8.56	—	—	—	—	7.65	—	—	—	—
CCSD(T) ^{e, 76}	—	—	8.71	8.62	8.64	8.58	8.64	—	—	—	—	—
EOM-IP-CCSD ^f	8.89	8.13	8.64	8.51	8.53	8.67	8.64	7.75	7.87	7.78	7.87	7.79
experiment												
	T	A	C	G								
VUV-SPI, this work	8.95 ± 0.05	8.20 ± 0.05	8.60 ± 0.05	7.75 ± 0.05								
MATI ³³	8.9178 ± 0.0010	—	—	—								
VUV-SPI ²⁵	8.87 ± 0.05	8.26 ± 0.05	8.68 ± 0.05	7.77 ± 0.05								

^a B3LYP/6-311++G(d,p)//B3LYP/6-31G(2d(0.8,0.1),p). ^b B3LYP/6-311++G(3df,2p)/B3LYP/6-31+G(d,p). ^c B3LYP/TZVP(+diff)//B3LYP/TZVP(+diff). ^d CASPT2(IPEA=0.25)//CASSCF/ANO-L 431/21/ZPE. ^e Extrapolated to CCSD(T)/6-311++(3df,2p)//B3LYP/6-31+G(d,p) using energy additivity scheme. ^f A, T, C: EOM-IP-CCSD/cc-pVTZ//IP-CISD/6-31+G(d), ZPE-corrected [ω B97X-D/6-31+G(d,p) frequencies]. G: EOM-IP-CCSD/cc-pVTZ(-diff)// ω B97/cc-pVTZ(-diff), ZPE-corrected [ω B97X-D/6-31+G(d,p) frequencies].

the experimental PES, each vibronic line was represented by a Lorentzian with the maximum corresponding to the FCF for the particular transition and the HWHM of 0.05 eV, which corresponds to the energy width of the light. The spectrum is shown in Figure 10a. We also attempted a Gaussian-based fit; however, the agreement was found to be considerably worse (see the Supporting Information).

The maximum of the spectrum is ≈ 0.2 eV higher than the 0–0 transition, and the band has two shoulders. The lowest-energy shoulder is due to the 0–0 transition. The maximum corresponds to strong bands 14¹ and 8¹, which can be described as coupled stretching and bending motions that are strongly affected by the radical relaxation. The shoulder lying 0.4 eV higher than the 0–0 transition is due to the combination bands

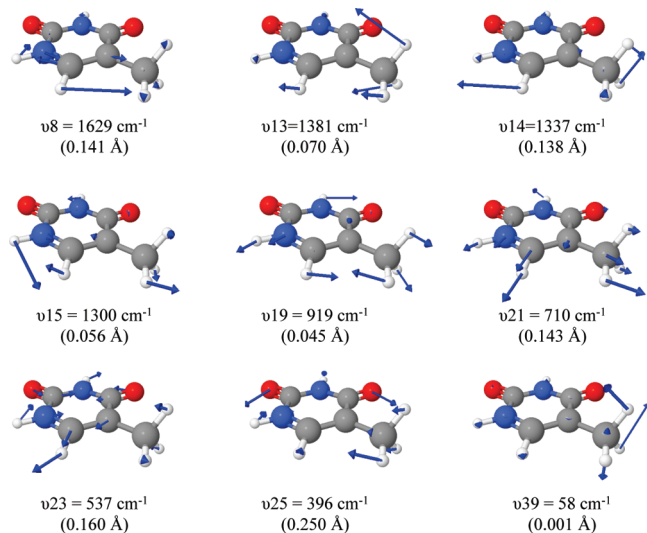


Figure 9. Franck-Condon active normal modes of thymine in the target (ionized) state representation.

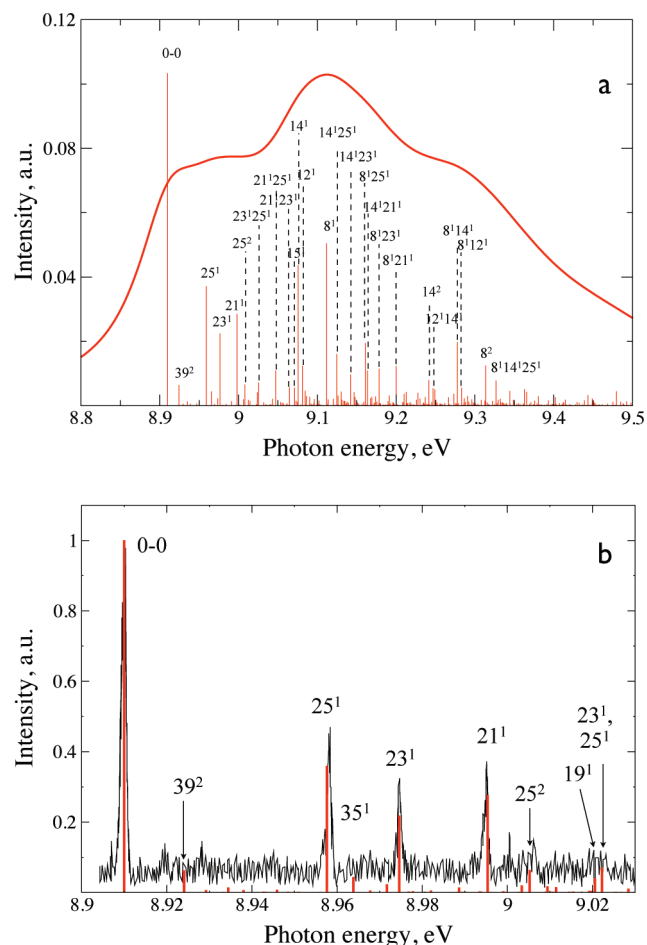


Figure 10. S_0 - D_0 vibronic spectra of thymine. (a) Complete S_0 - D_0 band computed within double-harmonic approximation with Duschinsky rotation and convoluted by Lorentzians. (b) Computed low-energy vibronic bands compared with a high-resolution MATI spectrum.³³

and overtones of these two normal modes. We found that this pattern in the band shape with two nearly symmetrically spaced shoulders is common for all NABs.

To validate the computed spectrum, we compare it with a high-resolution MATI spectrum for thymine. We observe an excellent agreement for the absolute value of the computed AIE

8.89 eV and experimental 0-0 transition $8.9178 \pm 0.0010 \text{ eV}$.³³ To check whether the intensities pattern is reproduced within the double-harmonic/Duschinsky rotation scheme, we compared the FC intensities obtained using ω B97X-D frequencies (scaled by 0.97) to the MATI spectra³³ (Figure 10b). The 0-0 transition energy of the computed spectrum was shifted to coincide with the experimental one. One can see that the modeled spectrum is in good agreement with the experiment, and, therefore, the double-harmonic/Duschinsky rotation scheme can be used for interpretation and prediction of the experimental PES spectra.

E. Photoelectron Spectra and Differentiated Photoionization Efficiency Curves: Theory versus Experiment.

1. S_0 - D_0 Vibrationally Resolved Bands. Figure 11 compares the computed S_0 - D_0 bands and the experimental photoelectron spectra of T, A, C,¹⁴ and G.²⁷ For T and A, there is a reasonable agreement with the experimental spectra, and the deviations at higher energies might be due to the overlap of the S_0 - D_0 and S_0 - D_1 bands.

C and G can be present in several tautomeric forms, each having its unique FC progression. To estimate the overall shape of the S_0 - D_0 band, we computed a sum of the Lorentzian-fitted profiles weighted by the Boltzmann populations at the PES experimental evaporation temperature ($T = 463 \text{ K}$ for C¹⁴ and $T = 613 \text{ K}$ for G²⁷): 0.29:0.19:0.42:0.08:0.02 and 0.25:0.45:0.15:0.14:0.01 for C1:C2a:C2b:C3a:C3b and G9K:G7K:G9Es:G9Ea:G7Es, respectively. The resulting band for cytosine is in reasonable agreement with the experimental spectrum. However, the agreement for G is worse. The computed spectrum has much lower onset for each of the five low-energy tautomers.

Another interesting feature of G concerns the difference between the differentiated PIE curves and the PES. The experimental PES of A, T, and C have much lower onsets than the differentiated PIE curves, which was previously ascribed to a wider bandwidth of light sources.¹³ However, the S_0 - D_0 bands in the differentiated PIE and the PES of guanine are nearly identical.

The differences may be due to different light sources used in two sets of photoelectron spectroscopy experiments.^{14,27} Different populations of tautomers can also alter the spectrum. However, the vaporization conditions are similar in the PES and VUV-SPI (this work) experiments, and, as discussed above, we expect only the five lowest tautomers of guanine and cytosine to be noticeably populated. An unambiguous determination of IEs of various guanine and cytosine tautomers and their relative populations will require MATI-like measurements.

A possible explanation for the deviation of the computed spectra from the experimental one is stronger anharmonic effects, as might be inferred from higher relaxation energy for G relative to other bases. To test this assumption, we performed a series of single-point calculations using ω B97X-D/6-31+G(d,p) along two FC active normal modes of the G9K tautomer (modes 32 and 35 in the cation representation). We found only minor negative anharmonicity in the cation's PES. Taking into account negative anharmonicity of active stretching/bending modes in the target state can shift the computed S_0 - D_0 harmonic spectra to lower energies, but not to higher energies (as needed for better agreement with the experiment). Another vibration that may affect the shape of the reported photoelectron band is the NH_2 wagging mode. As discussed above, the NH_2 inversion barriers are small resulting in planar vibrationally averaged structures. Importantly, the frequencies used for calculations of the FCFs were obtained with ω B97X-D/6-31+G(d,p), which for G yields (incorrect) planar geometry. Therefore, ω B97X-D/6-31+G(d,p) does not reproduce the

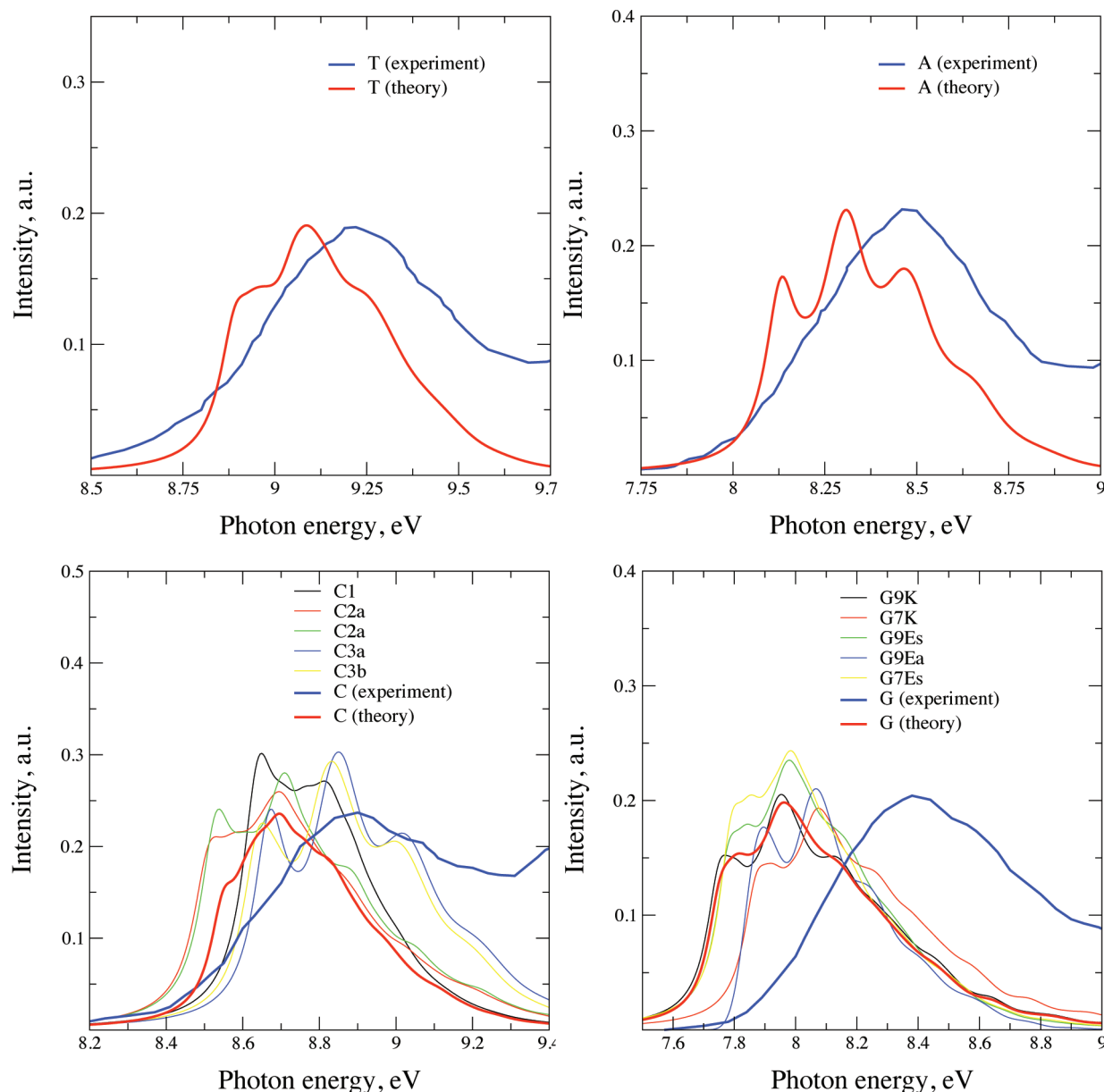


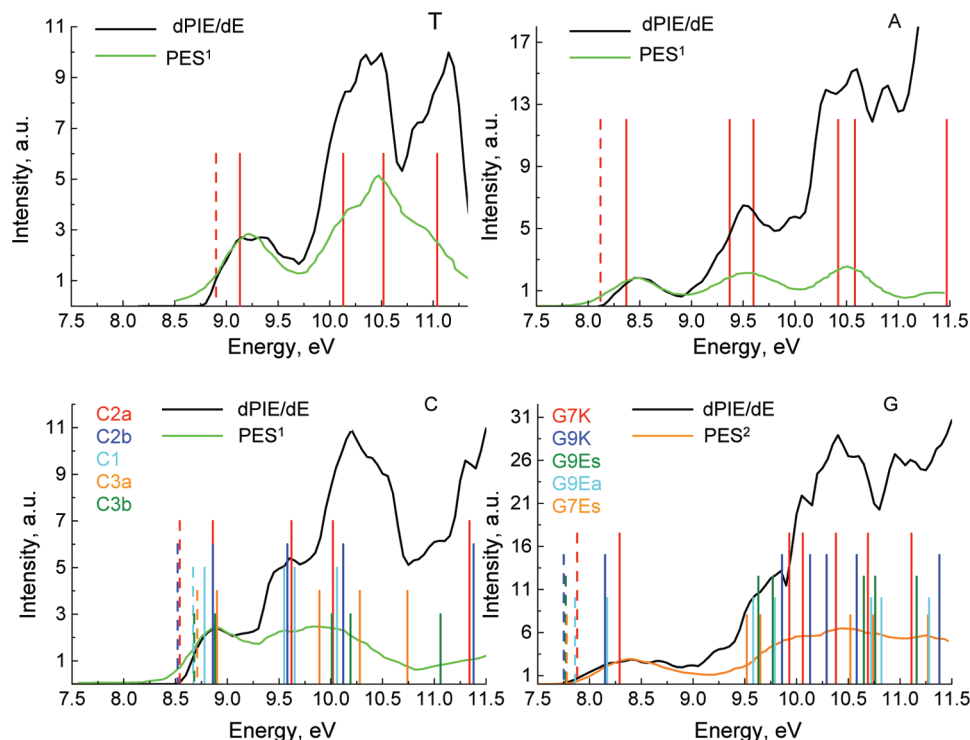
Figure 11. Comparison of the computed photoelectron spectra of NABs and PES (A, T, C ref 14; G ref 27). The computed PES for C and G were obtained as a sum of tautomers spectra, weighted according to Boltzmann populations.

correct curvature of the potential energy surface along the NH_2 out-of-plane vibrations and overestimates the localization of the nuclear wave function. The true vibrational wave function corresponding to the zero level in the double-well potential along NH_2 inversion should be more delocalized, and this may result in higher FC intensities for transitions to higher vibrational levels along this mode and in the respective combination bands. In sum, more accurate description of the NH_2 out-of-plane motion may lead to the blue shift of the computed $S_0\text{--}D_0$ band and, consequently, better agreement with the experiment. To summarize, the vibronic structure of guanine's PES remains an open question requiring more accurate theoretical treatment including anharmonic effects in both states, normal mode mixing, evaluation of the electronic part of the photoionization cross section, possibly more sophisticated ab initio methods, and, even more important, isomer-specific experimental data.

2. Photoelectron Spectra for the $S_0\text{--}D_n$ Transitions. Figure 12 shows the PES, differentiated PIE curves, and computed VIEs for all four NABs. We note a very good agreement between the computed AIEs and VIEs and the onsets and band maxima

of the experimental spectra. The thymine and adenine spectra are clearly dominated by ionization of a single tautomeric form. The first bands for A and T correspond to ionization from the π -type HOMO. The second band is derived from ionization from π - and σ -like orbitals. The third band of A also covers σ - and π -states, whereas for thymine there is only one state below 11.5 eV (the maximum energy accessed in the experiment). G and C present a more difficult case due to a complex composition of the gas-phase mixture. As discussed elsewhere,¹⁴ the PES of cytosine to a high degree can be explained by ionization of the lowest-energy C2b tautomer. For guanine, the experimental spectra are broad. The first band covers the first ionized state of all five tautomers. The second broad band (9.0–11.5 eV) is nearly structureless and does not allow any assignment.

The most prominent difference between the differentiated PIE curves and PES is the rise in signal with photon energy in the former. The computed adiabatic IEs agree well with the onsets of the differentiated PIE curves. However, the PES onsets are shifted to the lower energies, which is possible due to wider source energy gap. Thus the differentiated PIE curves provide



¹ PES from Ref. [14]; ² PES from Ref. [27].

Figure 12. Differentiated PIE curves and the PES. Solid lines denote VIEs computed at the EOM-IP-CCSD//RI-MP2 level of theory (A, T, C cc-pVTZ; G: cc-pVTZ(-dff)). Dashed lines mark AIEs computed with EOM-IP-CCSD/cc-pVTZ//IP-CISD/6-31+G(d) (A, T, C) and EOM-IP-CCSD/cc-pVTZ(-dff)// ω B97/cc-pVTZ(-dff) (G).

better values of adiabatic IEs as compared to the PES. However, the VIEs are better represented by the PES band maxima. The differentiated PIE curves may have contributions from ionization-induced fragmentation of larger clusters, which could lead to the enhancement of the signals in the higher-energy region.

IV. Conclusions

We present a combined theoretical and experimental study of the ionized states of NABs. The differentiated PIE curves provide a good approximation of the true photoelectron spectra, in particular, for low photon energies, where the spectra are not strongly perturbed by cluster fragmentation. We have shown that PIE onsets better describe adiabatic IEs than the onsets of the true photoelectron spectra, in contrast to the VIEs, which are better described by the latter.

The measured onsets (AIEs) are 8.20 ± 0.05 , 8.95 ± 0.05 , 8.60 ± 0.05 , and 7.75 ± 0.05 eV for A, T, C, and G, respectively. The computed adiabatic and VIEs are in good agreement with the experimental data. We demonstrated that tautomerization affects IEs, relative order, and the characters of ionized states. Analysis of ionization-induced geometry changes reveals correlation between the magnitude of the geometry changes and relaxation energy for most of the bases.

The structural changes between the neutrals and the cations are substantial and involve more than one bond, in agreement with the delocalized character of the HOMO. The most prominent change involves N1–C2, C2–N3, N3–C4, C5–N7, and N7–C8 bonds in purines and N1–C2, N1–C6, C4–C5, and C5–C6 bonds in pyrimidines. C2–O7 and C4–N(H₂) are also notably affected in cytosine.

The structural differences between the neutrals and the cations give rise to Franck–Condon progressions. Overall, there are 8–10

modes active in each of the four bases. Most of the modes are skeletal deformations, but some involve rocking and other out-of-plane motions of substituent groups, e.g., mode 39 for thymine, and appear as overtones due to strong frequency changes.

The shape of the computed S_0 – D_0 band of the PES follows a similar pattern in all four NABs; i.e., the maxima of the spectra are ≈ 0.2 eV higher than the 0–0 transition, and the band has two nearly symmetrical shoulders. The lowest-energy shoulder is due to the 0–0 transition. The maximum corresponds to the coupled stretching and bending motions that are strongly affected by the radical's relaxation. The shoulder lying 0.4 eV higher than the 0–0 transition is due to the combination bands and overtones of these two normal modes. For T, the computed FCFs are in excellent agreement with a high-resolution MATI spectra validating the theoretical methods employed. Overall, the computed FCFs agree well with the experimental PES for A, T, and C. For G, the agreement is worse and more accurate treatment of the anharmonicities is required for better description of the PES, which will be the subject of future work.

Acknowledgment. This work was conducted in the framework of the iOpenShell Center for Computational Studies of Electronic Structure and Spectroscopy of Open-Shell and Electronically Excited Species (iopenshell.usc.edu) supported by the National Science Foundation through the CRIF:CRF CHE-0625419 + 0624602 + 0625237 and CHE-0951634 (A.I.K.) grants. O.K. and M.A. acknowledge support by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

Supporting Information Available: (i) Structural parameters for optimized geometries of A and G9K (cations and

neutrals); comparison of VIEs of G9K and C1 computed with different methods; relevant MOs for C2a, C3b, G9Ea, and G7Es; modeled S_0 – D_0 PES assuming Gaussian line shape for individual vibronic transition; PIE curves and mass spectra; (ii) Cartesian coordinates; (iii) harmonic frequencies; (iv) FC intensities. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- Núñez, M. E.; Hall, D. B.; Barton, J. K. *Chem. Biol.* **1999**, *6*, 85.
- Henderson, P. T.; Jones, D.; Hampikian, G.; Kan, Y.; Schuster, G. B. *Proc. Nat. Acad. Sci.* **1999**, *96*, 8353.
- Lewis, F. D.; Letsinger, R. L.; Wasielewski, M. R. *Acc. Chem. Res.* **2001**, *34*, 159.
- Candeias, L. P.; Steenken, S. J. *Am. Chem. Soc.* **1989**, *111*, 1094.
- Steenken, S. *Chem. Rev.* **1989**, *89*, 503.
- Colson, A.-O.; Besler, B.; Sevilla, M. D. *J. Phys. Chem.* **1992**, *96*, 9787.
- Hutter, M.; Clark, T. *J. Am. Chem. Soc.* **1996**, *118*, 7574.
- Ghosh, A. K.; Schuster, G. B. *J. Am. Chem. Soc.* **2006**, *128*, 4172.
- Kumar, A.; Sevilla, M. D. *J. Phys. Chem. B* **2009**, *113*, 11359.
- Bertran, J.; Oliva, A.; Rodríguez-Santiago, L.; Sodupe, M. *J. Am. Chem. Soc.* **1998**, *120*, 8159.
- Kanvah, S.; Joseph, J.; Barnett, R. N.; Schuster, G. B.; Cleveland, C. L.; Landman, U. *Acc. Chem. Res.* **2010**, *43*, 280.
- Bravaya, K. B.; Kostko, O.; Ahmed, M.; Krylov, A. I. *Phys. Chem. Chem. Phys.* **2010**, *12*, 2292.
- Kostko, O.; Bravaya, K. B.; Krylov, A. I.; Ahmed, M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 2860.
- Trofimov, A. B.; Schirmer, J.; Kobaychev, V. B.; Potts, A. W.; Holland, D. M. P.; Karlsson, L. J. *Phys. B* **2006**, *39*, 305.
- Satzger, H.; Townsend, D.; Stolow, A. *Chem. Phys. Lett.* **2006**, *430*, 144.
- 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.
- Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. *J. Phys. Chem. A* **2002**, *106*, 8411.
- Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. *Int. J. Quantum Chem.* **2000**, *80*, 831.
- Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. *J. Am. Chem. Soc.* **2000**, *122*, 12304.
- Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. *J. Phys. Chem. A* **2003**, *107*, 822.
- Roca-Sanjuán, D.; Rubio, M.; Merchán, M.; Serrano-Andrés, L. *J. Chem. Phys.* **2006**, *125*, 084302.
- Cauët, E.; Dehareng, D.; Lievin, J. *J. Phys. Chem. A* **2006**, *110*, 9200.
- Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. *J. Phys. Chem. A* **2009**, *113*, 14630.
- Zhang, L.; Pan, Y.; Qi, F. *J. Theor. Comput. Chem.* **2009**, *8*, 1103.
- Orlov, V. M.; Smirnov, A. N.; Varshavsky, Ya. M. *Tetrahedron Lett.* **1976**, *17*, 4377.
- Peng, S.; Padvá, A.; LeBreton, P. R. *Proc. Natl. Acad. Sci.* **1976**, *73*, 2966.
- Lin, J.; Yu, C.; Peng, S.; Akiyama, I.; Li, K.; Li, K.; LeBreton, P. R. *J. Phys. Chem.* **1980**, *84*, 1006.
- Urano, S.; Yang, X.; LeBreton, P. R. *J. Mol. Struct.* **1989**, *214*, 315.
- Dougherty, D.; Wittel, K.; Meeks, J.; McGlynn, S. P. *J. Am. Chem. Soc.* **1976**, *98*, 3815.
- Hush, N. S.; Cheung, A. S. *Chem. Phys. Lett.* **1975**, *34*, 11.
- Zavilopulo, A. N.; Sheprik, O. B.; Agafonova, A. S. *J. Phys. B* **2009**, *42*, 025101.
- Zhou, J.; Kostko, O.; Nicolas, C.; Tang, X.; Belau, L.; de Vries, M. S.; Ahmed, M. *J. Phys. Chem. A* **2009**, *113*, 4829.
- Choi, K.-W.; Lee, J.-H.; Kim, S. K. *J. Am. Chem. Soc.* **2005**, *127*, 15674.
- Weigend, F.; Häser, M. *Theor. Chim. Acta* **1997**, *208*, 359.
- Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143.
- Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175.
- Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
- Landau, A.; Khistyayev, K.; Dolgikh, S.; Krylov, A. I. *J. Chem. Phys.* **2010**, *132*, 014109.
- Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
- Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
- Krylov, A. I. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433.
- Pieniazek, P. A.; Arnstein, S. A.; Bradforth, S. E.; Krylov, A. I.; Sherrill, C. D. *J. Chem. Phys.* **2007**, *127*, 164110.
- Yau, A. D.; Perera, S. A.; Bartlett, R. J. *Mol. Phys.* **2002**, *100*, 835.
- Pal, S.; Rittby, M.; Bartlett, R. J.; Sinha, D.; Mukherjee, D. *Chem. Phys. Lett.* **1987**, *137*, 273.
- Stanton, J. F.; Gauss, J. *J. Chem. Phys.* **1994**, *101*, 8938.
- Kamiya, M.; Hirata, S. *J. Chem. Phys.* **2006**, *125*, 074111.
- Pieniazek, P. A.; Bradforth, S. E.; Krylov, A. I. *J. Chem. Phys.* **2008**, *129*, 074104.
- Golubeva, A. A.; Pieniazek, P. A.; Krylov, A. I. *J. Chem. Phys.* **2009**, *130*, 124113.
- Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.
- Lebedev, V. I. *Zh. Vychisl. Mat. Mat. Fiz.* **1975**, *15*, 48.
- Murray, W. C.; Handy, N. C.; Laming, G. J. *Mol. Phys.* **1993**, *78*, 997.
- Berger, R.; Fischer, C.; Klessinger, M. *J. Phys. Chem. A* **1998**, *102*, 7157.
- Wang, S.; Schaefer, H. F., III. *J. Chem. Phys.* **2006**, *124*, 044303.
- Mozhayskiy, V. A.; Krylov, A. I. *ezSpectrum*, <http://iopencell.usc.edu/downloads/>.
- Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neil, D. P.; Distasio, R. A., Jr.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Bird, E. F. C.; Daschel, H.; Doerksen, R. J.; Drew, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G. S.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Liang, W. Z.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L., III; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Herberich, W. J.; Schaefer, H. F., III; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172.
- Belau, L.; Wilson, K. R.; Leone, S. R.; Ahmed, M. *J. Phys. Chem. A* **2007**, *111*, 7562.
- Berkowitz, J. *J. Chem. Phys.* **1978**, *69*, 3044.
- Guerra, C. F.; Bickelhaupt, F. M.; Saha, S.; Wang, F. *J. Phys. Chem. A* **2006**, *110*, 4012.
- Hanus, M.; Kabeláč, M.; Rejnek, J.; Ryjék, F.; Hobza, P. *J. Phys. Chem. B* **2004**, *208*, 2087.
- Piacenza, M.; Grimme, S. *J. Comput. Chem.* **2004**, *25*, 83.
- Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811.
- Broo, A.; Holmén, A. *Chem. Phys.* **1996**, *211*, 147.
- López, J. C.; Peña, M. I.; Sanz, M. E.; Alonso, J. L. *J. Chem. Phys.* **2007**, *126*, 191103.
- Choi, M. Y.; Miller, R. E. *J. Phys. Chem. A* **2007**, *111*, 2475.
- Nir, E.; Plützer, Ch.; Kleinerhanns, K.; de Vries, M. *Eur. Phys. J. D* **2002**, *20*, 317.
- Marian, C. M. *J. Phys. Chem. A* **2007**, *111*, 1545.
- Choi, M. Y.; Miller, R. E. *J. Am. Chem. Soc.* **2006**, *128*, 7320.
- Alonso, J. L.; Peña, I.; Lopéz, J. C.; Vaquero, V. *Angew. Chem., Int. Ed. Engl.* **2009**, *48*, 6141.
- Choi, M. Y.; Dong, F.; Miller, R. E. *Phil. Trans. R. Soc. London, Ser. A* **2005**, *363*, 393.
- Feyer, V.; Plekan, O.; Richter, R.; Coreno, M.; Vall-Ilosera, G.; Prince, K. C.; Trofimov, A. B.; Zaytseva, I. I.; Moskovskaya, T. E.; Gromov, E. V.; Schirmer, J. *J. Phys. Chem. A* **2009**, *113*, 5736.
- Oana, C. M.; Krylov, A. I. *J. Chem. Phys.* **2007**, *127*, 234106.
- Jones, D. B.; Wang, F.; Winkler, D. A.; Brunger, M. J. *Biophys. Chem.* **2006**, *121*, 105.
- Zhou, J.; Takahashi, L. K.; Wilson, K. R.; Leone, S. R.; Ahmed, M. *Anal. Chem.* **2010**, *82*, 3905.
- Lauer, G.; Schäfer, W.; Schweig, A. *Tetrahedron Lett.* **1975**, *16*, 3939.
- Dougherty, D.; Younathan, E. S.; Voll, R.; Abdunur, S.; McGlynn, S. P. *J. Electron Spectrosc. Relat. Phenom.* **1978**, *13*, 379.
- Wolken, J. K.; Yao, C. Y.; Tureček, F.; Polce, M. J.; Wesdemiotis, C. *Int. J. Mass. Spectrom.* **2007**, *267*, 30.