Lawrence Berkeley National Laboratory
Chemical Sciences

Title
Electronic structure and spectroscopy of nucleic acid bases: Ionization energies, ionization-induced structural changes, and photoelectron spectra

Permalink
https://escholarship.org/uc/item/3189s5d5

Journal
Lawrence Berkeley National Laboratory

Author
Bravaya, Ksenia B.

Publication Date
2010-12-15

eScholarship.org Powered by the California Digital Library
University of California

Electronic structure and spectroscopy of nucleic acid bases: Ionization energies, ionization-induced structural changes, and photoelectron spectra

Ksenia B. Bravaya[a], Oleg Kostko[b], Stanislav Dolgikh[a],
Arie Landau[a], Musahid Ahmed[b], and Anna I. Krylov[a]

[a] Department of Chemistry, University of Southern California,
Los Angeles, CA 90089-0482, USA

[b] Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

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 five lowest-energy tautomeric forms. The first adiabatic and several vertical ionization energies were computed using 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 ωB97X-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.

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 neurodegenera-

tive deceases.

In the 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 Å[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 electronic structure of the dimers with that of the monomers; however, the monomer's data have not been analyzed in details. 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 [14–23] and experimentally [14–16, 24–31]; 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 study [14] 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 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 $D_1(\sigma)$ ionized state, in contrast to the earlier photoelectron study of Trofimov *et al.* [14] in which the 9.5 eV band was assigned to 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 70s [24, 25], the vibra-

tionally resolved spectrum has been reported only for thymine [32]. 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 (Fig. 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 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 the experimental mass-analyzed threshold ionization (MATI) spectrum[32]. 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) [33–35]. No symmetry constrains were imposed in geometry optimizations. For A, T and C, Dunning's cc-pVTZ basis set[36] was used, whereas for guanine we employed an extended cc-pVTZ+dff basis[37] augmented by additional s and p diffuse functions for the second row elements, and s-functions for hydrogen. The exponents were taken from 6-311(2+,+)G basis set. RI-MP2 optimized geometries were used for IEs calculations. ωB97X-D/6-31+G(d) 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 Supplementary Materials.

Accurate description of ionized states with open-shell character presents a challenge for quantum chemistry[38, 39, 45]. Standard single-reference methods based on unrestricted Hartree-Fock (UHF) references, e.g., UMP2, suffer from spin-contamination and tend to overestimate charge localization[39]. The performance of DFT for this type of systems 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[38–43] and IP-CISD [44] 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 orbitals manifolds). This results in a balanced description of several target states and convenient formalism calculation of inter-state 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 were computed by IP-CISD, DFT with the long-range corrected ωB97, ωB97X and ωB97X-D (the latter includes empirical correction for

dispersion interactions[46]), and EOM-IP-CCSD. The IP-CISD/6-31+G(d) geometries of the cations were used in calculations of AIEs for A, T and C, whereas ωB97/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 ωB97X-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 correlation. The benchmark calculations for the five ionized states of uracil[44] 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 ωB97 family was investigated for G tautomers by comparing the results of ωB97X-D/6-31+G(d,p) with ωB97/6-31+G(d,p) and ωB97/cc-pVTZ+dff. The accuracy of IP-CISD and ωB97(X-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 ωB97X-D/6-31+G(d,p) geometry is closer to the EOM-IP-CCSD/6-31+G(d) one (see Supplementary Materials) than IP-CISD one. Thus, we choose to use the ωB97X-D/6-31+G(d,p) geometries for the analysis of the ionization-induced geometry changes. The ωB97X-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 Lebedev [47] radial grid and 302 points in the Euler-Maclaurin [48] 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[36]. For guanine, an augmented basis, cc-pVTZ+dff, and Frozen Natural Orbitals (FNO) approximation were employed [37]. 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 [37].

Harmonic frequencies were computed at the ωB97X-D/6-31+G(d,p) level of theory for both the neutral and the 1st 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 rotations [49] at the C_s geometry using normal modes and frequencies for the non-planar 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 method [50] for canonical forms of the A, G and C bases, respectively. These values are less or at least comparable with the NH_2 out-of-plane vibrational frequencies, which lie above 300-350 cm^{-1}. Thus, 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 non-diagonal, e.g., we observed multiple nondiagonal elements of the normal mode overlap matrix up to 0.4—0.7 for all bases. The *ezSpectrum* program [51] was used to compute the spectra. All quantum chemistry calculations were performed with the *Q-Chem* package of quantum chemistry programs [52].

B. Experimental setup

The experiments were performed on a molecular beam apparatus coupled to a 3 meter VUV monochromator on the Chemical Dynamics Beamline at the Advanced Light Source (ALS). 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[53]. In the current experiments, the backing pressure of Ar expanded through a 100 μm diameter nozzle was 35 kPa. The clusters 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 methanol[54] 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 non-equilibrium 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[50]. 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 Supplementary Materials. A systematic study of the relative energies of adenine tautomers has been performed by Guerra *et al.*[55] and by Hanus et al.[56]. 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 [55, 56]. The energy difference between the lowest-energy diketo form of T and other tautomers is even higher. Piacenza and Grimme [57] 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 I. Table I also lists dipole moments of the tautomers. Interestingly, biologically relevant forms have the largest dipole moments and become the most stable in solution [56, 58, 59]. Moreover, large dipole moments contribute towards 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 canonical forms, 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 rigid-rotor-

harmonic-oscillator approximation are given in Table I. Assuming thermal equilibrium, only the canonical forms of A and T can be populated. In agreement with this energetics, a single tautomer of T (the diketo form) was observed in molecular beam Fourier transform microwave spectroscopy [60] and infrared laser spectroscopy studies of thymine in helium nanodroplets [61]. Nir *et al.* [62] 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 result[31] 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 I). 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 are used to evaporate the sample [31]. 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 keto- forms [62]. One keto- and one enol- form of C were registered [62]. 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. Marian [63] suggested that the bands corresponding to the oxo-forms in the IR-UV experiments [62] 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 Franck-Condon region of vertical π-π^* 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 [64]. Recent Fourier transform microwave spectroscopy studies using laser ablation molecular beam (LA MB FTMW) supported this assignment [65]. 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:G9Ea [64] are very close to those estimated assuming thermal equilibrium (see Table I). The existence of three tautomers of C was confirmed by Choi *et al.* [66] 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 [66], which is also consistent with the estimated Boltzmann populations

(Table I). Our previous study of cytosine[13] also suggested that at least three tautomers are present in the beam. Recent core photoionization studies of cytosine reported 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 [67]. Although C3b was not considered in this study in view of its higher energy, one can expect ionization energy from 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) [67]. Our C1:(C3a+C3b):(C2a+C2b) ratio computed as 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. [67].

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 nanodroplets [64] and core photoionization spectroscopy [67]. In both cases, the estimated populations comply with the Boltzmann distribution. 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 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.

B. Vertical ionization energies and the character of ionized states

VIEs of thymine computed at different levels of theory are listed in Table II. The corresponding MOs are shown in Fig. 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 orbitals[68] 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.

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 calculations[17] and with the CASPT2 values [20]. OVGF and ADC(3) yield a qualitatively different picture by changing the order of the ionized states [14] (see Table II). 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 VIE for the first π ionized state is 0.28 eV lower. ADC(3) exhibits 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 trend [20] 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, state-averaging procedure and IPEA correction, which all favor the states derided 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.

Relevant MOs and VIEs of A are given in Fig. 3 and Table III. As in the thymine case, the propagator methods yield similar errors for ionizations 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 is consistent for all methods. Moreover, VIEs and the assignments agree with that reported by Trofimov and coworkers[14], contrary to the results of Stolow and coworkers [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 experimental[14] VIE values.

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 (see Supplementary Materials for G and C data).

Fig. 4 and Table IV present the MOs and VIEs of the lowest tautomers of cytosine. Tautomerization has prominent effect on the ionized states[13]. As follows from Fig. 4, the

state 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].

Both VIEs and the character of the ionized states of G are affected by tautomerization, as follows from Table V and Fig. 5. Although the character of the lowest ionized state is similar in all five tautomeric forms (see Fig. 5 and Supplementary Materials), 1st VIE for G7K is noticeably shifted to higher energies, which can be explained by the analysis of the corresponding MOs. In all 5 tautomers, the 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 thus decreasing 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 Fig. 6–8. The corresponding Cartesian geometries are given in Supplementary Materials. 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 VI.

As a quantitative measure of the overall geometry change upon ionization, we consider the difference between the Cartesian 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 3N-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. 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 MOs. The distance between the atoms for which the bonding or anti-bonding overlap is present in HOMO should increase or decrease, respectively. The HOMO of T has the strongest contribution from the bonding C5-C6 and N1-C6 anti-bonding 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 Å and 1.379 Å for the neutral, whereas in the cation these distances are 1.409 Å 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(H_2) bonds. As follows from Fig. 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 NH_2 group due to electronic density transfer from the nitrogen lone pair to the main π-system compensating for the lost 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 [31]. Planar configuration of the of the NH_2 group in the ionized state is common for all bases with the NH_2 moiety. However, the degree of non-planarity 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 non-planarity 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 for A and T, in agreement with a larger dQ value. The HOMO of C2b is of bonding character with respect to the C4-C5-C6 and N1-C2 moieties. 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, C6-O anti-bonding 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 anti-bonding C5-N7, N3-C4, 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 Fig. 5. It has the N1-C2, C2-N3, C4-C5, C5-C6 and N7-C8 bonding and NH2-C2, N1-C6, N3-C4 C5-N7 anti-bonding 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 anti-bonding 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 VIE of guanine is the lowest among the four bases, but the relaxation further increases the gap in 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 determines the shape of vibrational progressions in photoelectron spectra. A vibrationally resolved photoionization spectrum is only available for thymine [32], 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 8 normal modes shown in Fig. 9. In addition, normal mode ν_9 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 overtone in S_0-D_0 spectra [Fig. 10(a)]. A similar spectrum has been reported previously by Zhou *et al.* [69].

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 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 Fig. 10(a). We also attempted a Gaussian-based fit, however, the agreement was found to be considerably worse (see Supplementary Materials).

The maximum of the spectrum is $\approx$ 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 14^1 and 8^1, 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 and overtones of these two normal modes. We found that 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 eV [32]. To check whether the intensities pattern is reproduced within the double-harmonic/Duschinsky rotation scheme, we compared the FC intensities obtained using ωB97X-D frequencies (scaled by 0.97) to the MATI spectra [32] (Fig. 10)(b). The 0-0 transition energy of the computed spectrum was shifted to coincide to the experimental one. One can see that the modeled spectrum is in good agreement with the experiment, and, therefore, 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

Fig. 11 compares the computed S_0-D_0 bands and the experimental photoelectron spectra of T, A, C [14] and G [26]. 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 K for C [14], and T=613 K for G [26]): 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 energy gap of light sources [13]. However, the 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, 26]. 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[31].

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 series of single point

calculations using ωB97X-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 of the computed S_0-D_0 harmonic spectra to lower energies, but not to higher energies (as need 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 ωB97X-D/6-31+G(d,p), which for G yields (incorrect) planar geometry. Therefore, ωB97X-D/6-31+G(d,p) is not reproducing the correct curvature of the potential energy surface along the NH_2 out-of-plane vibrations and overestimates the localization of the nuclear wave function. 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-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-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 the single tautomeric from. 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 a rising with the energy signal of 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 better values of adiabatic IEs as compared to the PES. However, the VIEs are better represented by the PES band maxima. This can be due to a constantly rising signal of the latter. 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 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(H_2) 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 similar pattern in all four NABs, i.e., the maxima of the spectra are $\approx$ 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 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 the 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 is the subject of a future work.

Acknowledgments

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.) grant. 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 neutrals and radicals of A and G9K; comparison of VIEs of G9K and C1 computed with different methods; relevant MOs for C2a,

C3b and G9Ea and GEs; 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. Supporting information is available free of charge via the Internet at http://pubs.acs.org.

TABLES

TABLE I: Relative energies of different tautomers of NABs and absolute value of the dipole moment computed at the RI-MP2 optimized equilibrium structures. Boltzmann populations for different tautomers computed for the CCSD relative energies are given in parenthesis (see Section II). cc-pVTZ (A,T,C) and cc-pVTZ-dff (G) basis sets are used for geometry optimization.

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

TABLE II: Vertical ionization energies of the six lowest ionized states of thymine (eV).

	D_0	D_1	D_2	D_3	D_4	D_5	Reference
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)	[20]
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)	—	[70]
	9.18 (π)	10.03 (σ)	10.39 (π)	10.82 (σ)	12.27(π)	—	[27]
	9.02 (π)	9.95 (σ)	10.40 (π)	10.80(σ)	12.10 (π/π)	—	[28]

TABLE III: 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	Reference
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)	[22]
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)	[20]
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	[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)	[27, 71]

TABLE IV: The EOM-IP-CCSD/cc-pVTZ//RI-MP2/cc-pVTZ vertical ionization energies (eV) for the six lowest ionized states of cytosine tautomers. IEs for lowest energy C2b tautomer are marked in bold.

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 (σ)**	**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)

TABLE V: EOM-IP-CCSD/cc-pVTZ-dff//RI-MP2/cc-pVTZ-dff vertical ionization energies for the six lowest ionized states of guanine tautomers. IEs for lowest energy G7K tautomer are marked in bold.

State	D_0	D_1	D_2	D_3	D_4	D_5
G9K [37]	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(π)**
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)					

TABLE VI: The computed adiabatic IEs and the appearance energies.

Method	T	A	C1	C2a	**C2b**	C3a	C3b	G9K	**G7K**	G9Es	G9Ea	G7Es
theory												
B3LYP [a] [21]	8.81	8.11	8.63	—	—	—	—	7.69				
B3LYP [b] [72]	—	—	8.58	8.47	**8 .48**	8.44	8.50			—		
B3LYP [c] [31]								7.68	**7. 80**	7.70	7.75	7.76
CASPT2 [d] [20]	8.69	8.11	8.56	—	—	—	—	7.65				
CCSD(T) [e] [72]	—	—	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												
VUV-SPI, this work	8.95 ± 0.05	8.20 ± 0.05	8.60 ± 0.05					7.75 ± 0.05				
MATI [32]	8.9178 ± 0.0010	—	—					—				
R2PI [73]	—	8.606 ± 0.006	—					—				
VUV-SPI [24]	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; G:

EOM-IP-CCSD/cc-pVTZ(-dff)//ωB97/cc-pVTZ(-dff), ZPE-corrected

FIGURES

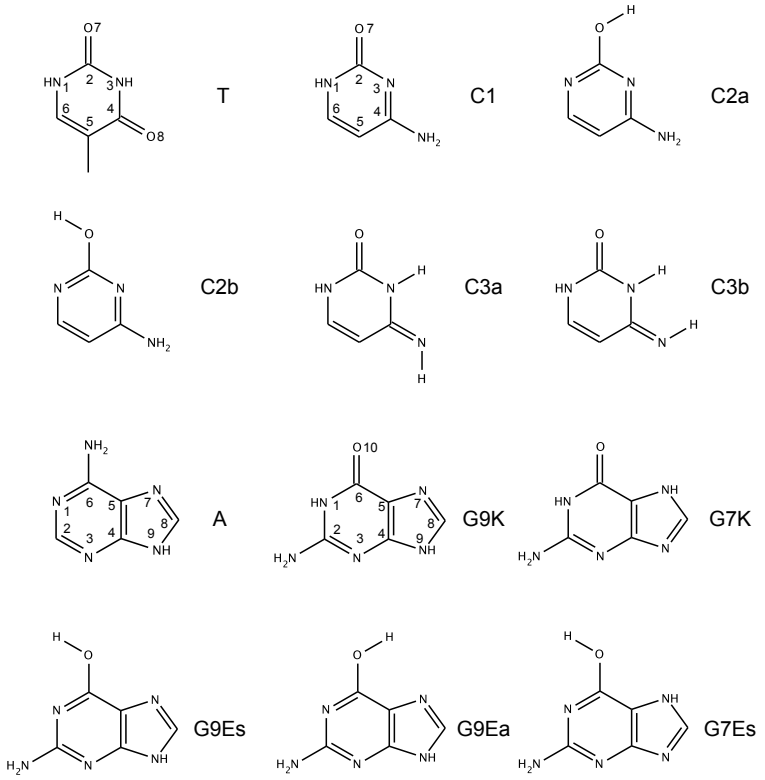


FIG. 1: The structures of thymine, adenine, cytosine and guanine bases considered in this study.

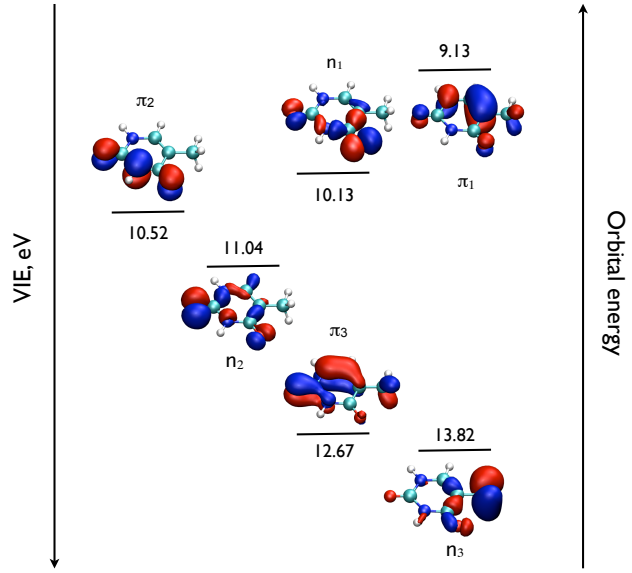


FIG. 2: Vertical ionization energies (EOM-IP-CCSD/cc-pVTZ) and the corresponding MOs for thymine.

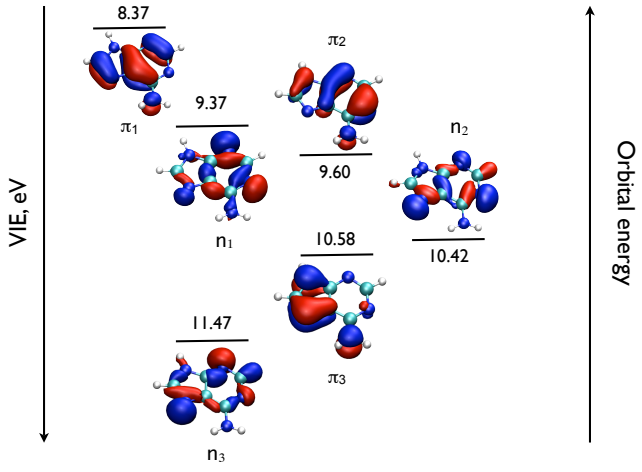


FIG. 3: Vertical ionization energies (EOM-IP-CCSD/cc-pVTZ) and the corresponding MOs for adenine.

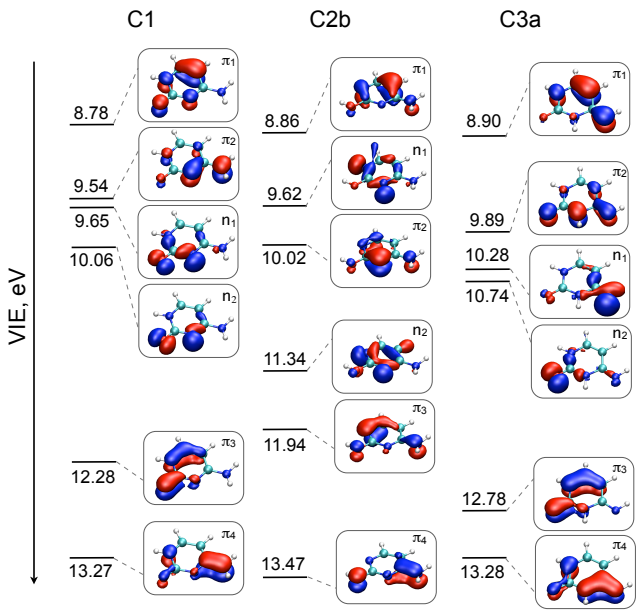


FIG. 4: Vertical ionization energies (EOM-IP-CCSD/cc-pVTZ) and the corresponding MOs for cytosine.

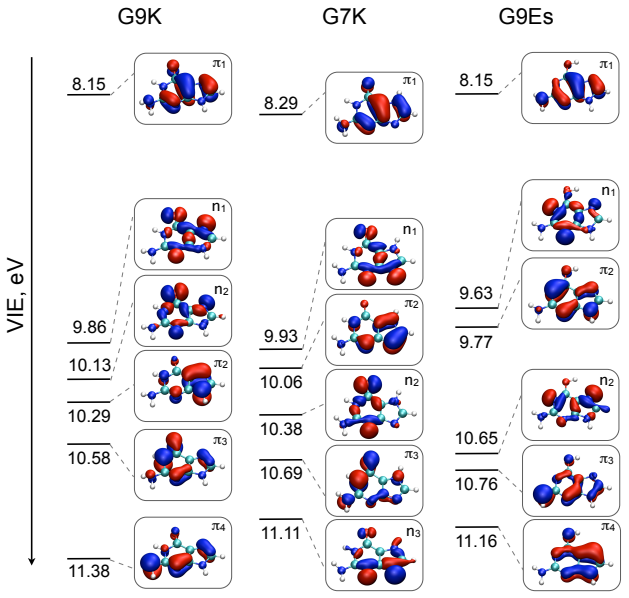


FIG. 5: Vertical ionization energies (EOM-IP-CCSD/cc-pVTZ(-dff)) and the corresponding MOs for guanine.

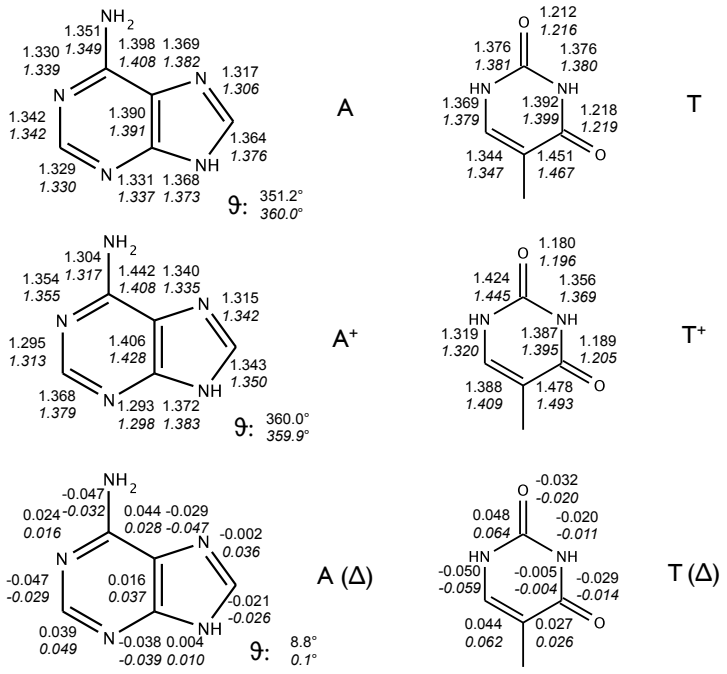


FIG. 6: Relevant bond lengths for A and T and the ionization induced geometry changes. Neutrals: RI-MP2/cc-pVTZ and ωB97X-D/6-31+G(d,p) (italics) optimized geometries; cations: IP-CISD/6-31+G(d) and ωB97X-D/6-31+G(d,p) optimized geometries; ionization induced bond length changes (Δ): RI-MP2/cc-pVTZ - IP-CISD/6-31+G(d) and ωB97X-D/6-31+G(d,p) (italics).

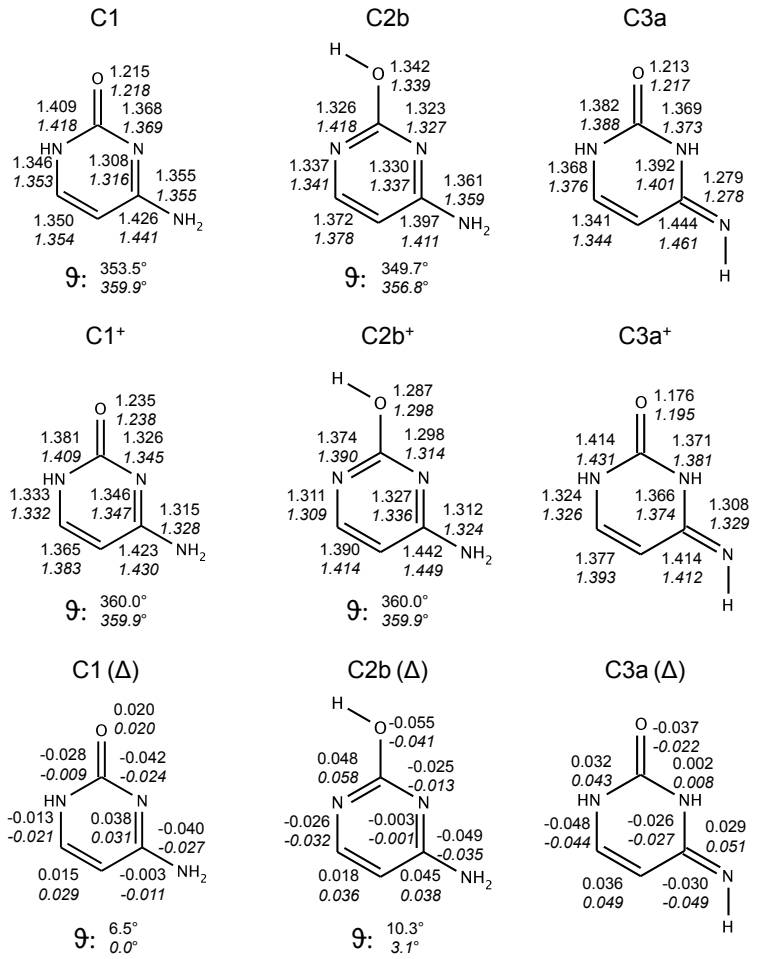


FIG. 7: Relevant bond lengths for cytosine's C1, C2b and C3a tautomers and the ionization induced geometry changes. Neutrals: RI-MP2/cc-pVTZ and ωB97X-D/6-31+G(d,p) (italics) optimized geometries; cations: IP-CISD/6-31+G(d) and ωB97X-D/6-31+G(d,p) optimized geometries; ionization induced bond length changes (Δ): RI-MP2/cc-pVTZ - IP-CISD/6-31+G(d) and ωB97X-D/6-31+G(d,p) (italics).

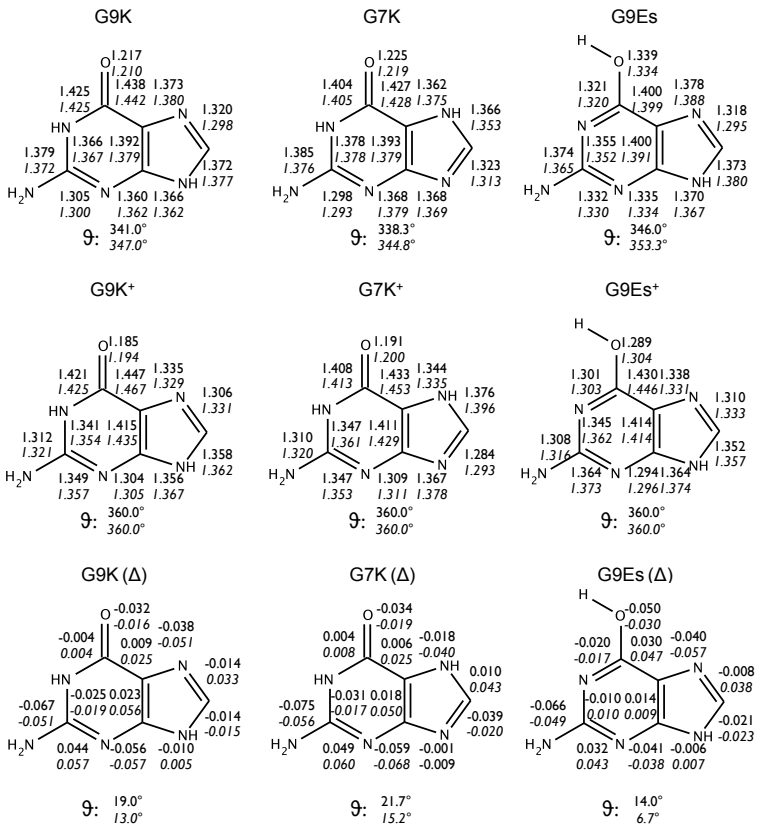


FIG. 8: Relevant bond lengths for guanine's G9K, G7K and G9Es tautomers and the ionization induced geometry changes. Neutrals: RI-MP2/cc-pVTZ and ωB97X-D/6-31+G(d,p) (italics) optimized geometries; cations: IP-CISD/6-31+G(d) and ωB97X-D/6-31+G(d,p) optimized geometries; ionization induced bond length changes (Δ): RI-MP2/cc-pVTZ - IP-CISD/6-31+G(d) and ωB97X-D/6-31+G(d,p) (italics).

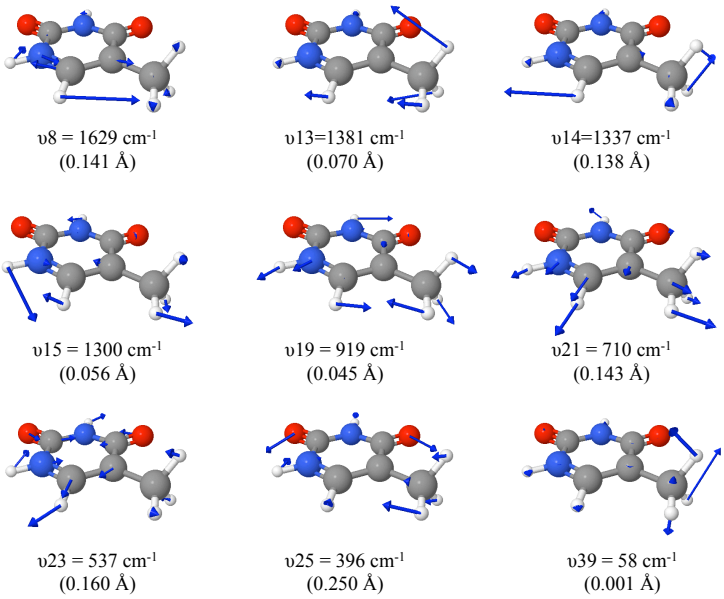


FIG. 9: Franck-Condon active normal modes of thymine in the target (ionized) state representation.

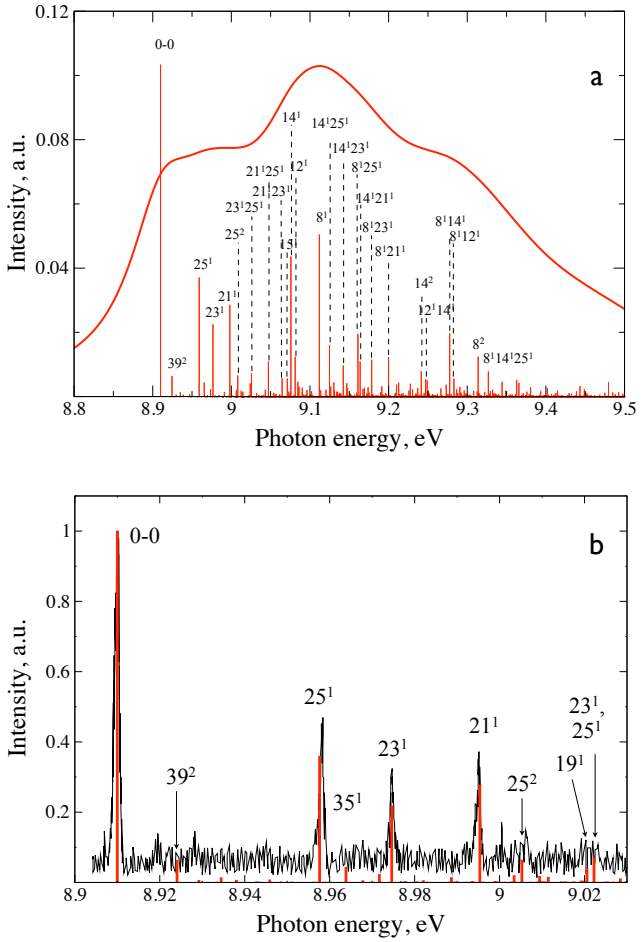


FIG. 10: The S_0-D_0 vibronic spectra of thymine. (a) The complete S_0-D_0 band computed within double-harmonic approximation with Duschinsky rotation and convoluted by Lorentzians. (b) The computed low-energy vibronic bands compared with the high-resolution MATI spectrum [32].

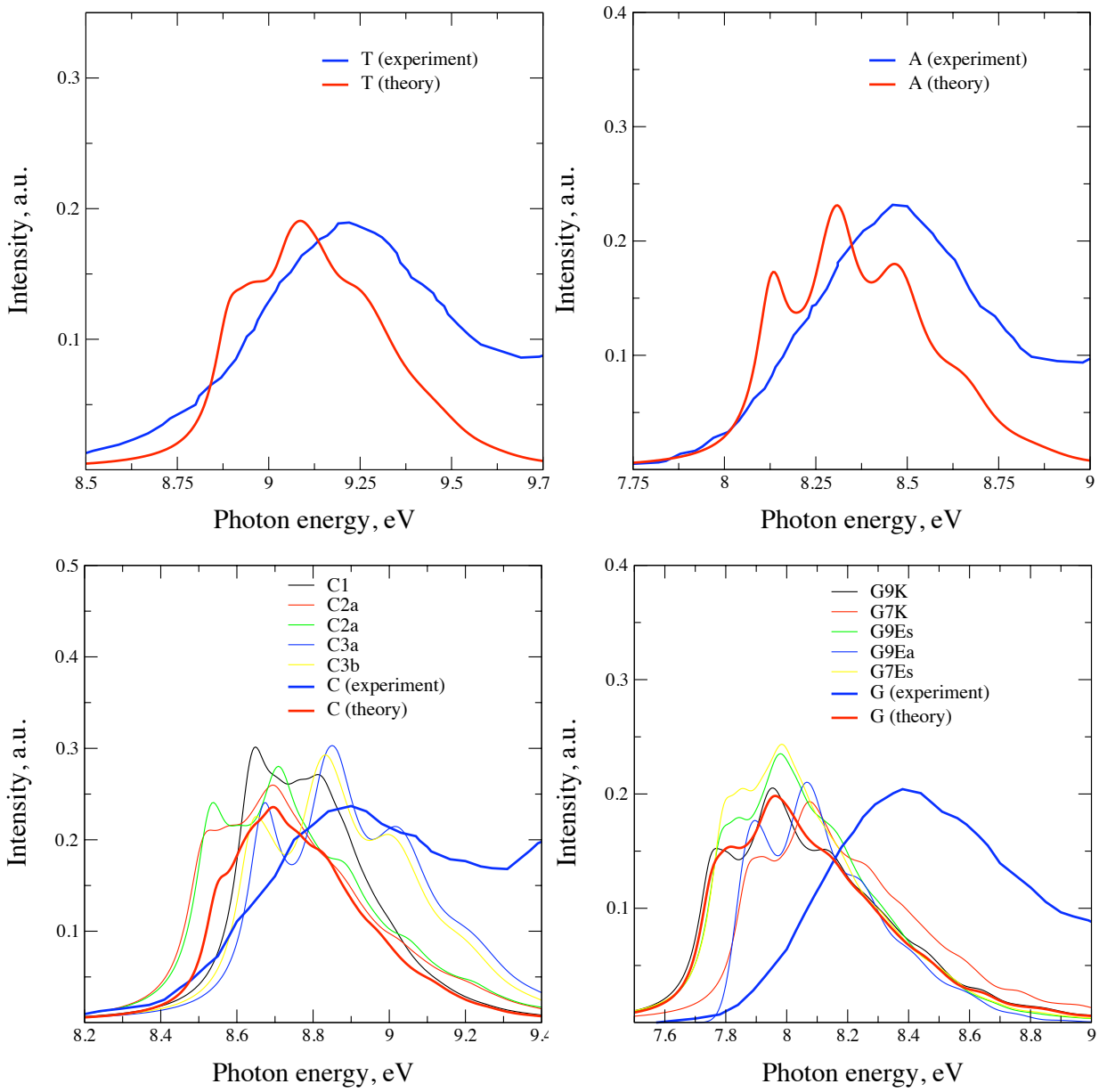
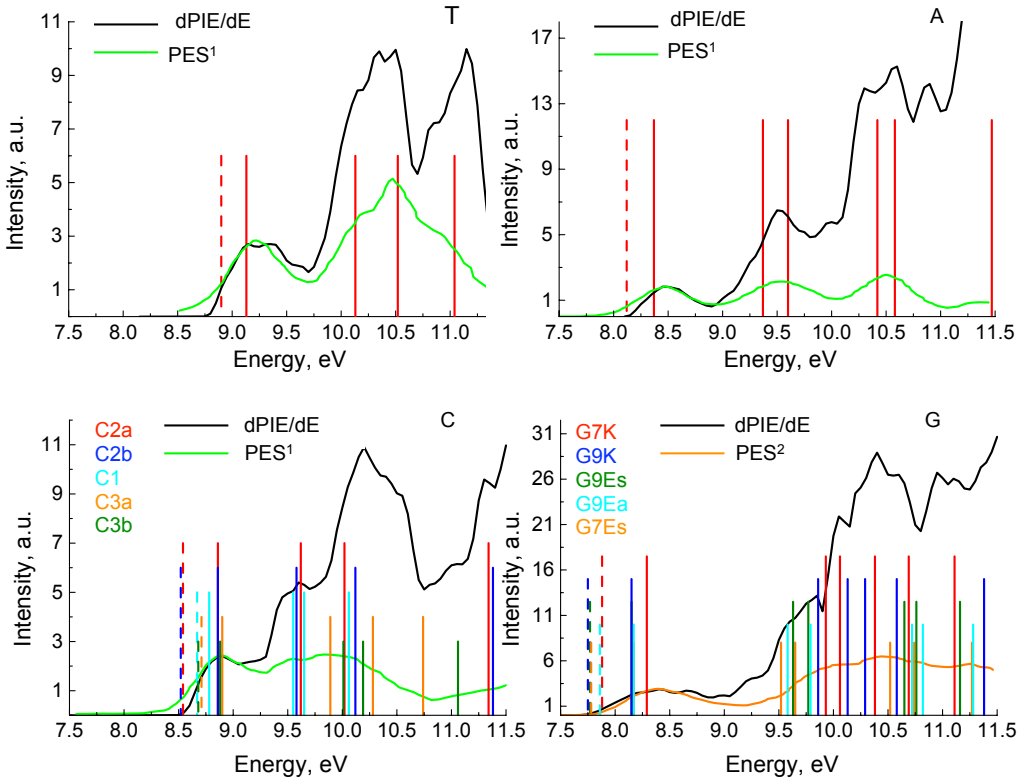


FIG. 11: Comparison of the computed photoelectron spectra of NABs and PES (A,T,C: [14]; G: [26]). The computed PES for C and G were obtained as a sum of tautomers spectra, weighted according to Boltzmann populations.

FIG. 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)//ωB97/cc-pVTZ(-dff) (G).



[1] PES from Ref. [14]; [2] PES from Ref. [26].

[1] Núñez, M.E. ; Hall, D.B. ; Barton, J.K. *Chem. Biol.* **1999**, *6*, 85.

[2] Henderson, P.T. ; Jones, D. ; Hampikian, G. ; Kan, Y. ; Schuster, G.B. *Proc. Nat. Acad. Sci.* **1999**, *96*, 8353.

[3] Lewis, F.D. ; Letsinger, R.L. ; Wasielewski, M.R. *Acc. Chem. Res.* **2001**, *34*, 159.

[4] Candeias, L.P. ; Steenken, S. *J. Am. Chem. Soc.* **1989**, *111*, 1094.

[5] Steenken, S. *Chem. Rev.* **1989**, *89*, 503.

[6] Colson, A.-O. ; Besler, B. ; Sevilla, M.D. *J. Phys. Chem.* **1992**, *96*, 9787.

[7] Hutter, M. ; Clark, T. *J. Am. Chem. Soc.* **1996**, *118*.

[8] Ghosh, A.K. ; Schuster, G.B. *J. Am. Chem. Soc.* **2006**, *128*, 4172.

[9] Kumar, A. ; Sevilla, M.D. *J. Phys. Chem. B* **2009;** asap.

[10] Bertran, J. ; Oliva, A. ; Rodríguez-Santiago, L. ; Sodupe, M. *J. Am. Chem. Soc.* **1998**, *120*, 8159.

[11] Kanvah, S. ; Joseph, J. ; amd R.N. Barnett, G.B. Schuster ; Cleveland, C.L. ; Landman, U. *Acc. Chem. Res.* **2010**, *43*, 280.

[12] Bravaya, K.B. ; Kostko, O. ; Ahmed, M. ; Krylov, A.I. *Phys. Chem. Chem. Phys.* **2010**, *12*, 2292.

[13] Kostko, O. ; Bravaya, K.B. ; Krylov, A.I. ; Ahmed, M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 2860.

[14] Trofimov, A.B. ; Schirmer, J. ; Kobychev, V.B. ; Potts, A.W. ; Holland, D.M.P. ; Karlsson, L. *J. Phys. B* **2006**, *39*, 305.

[15] Satzger, H. ; Townsend, D. ; Stolow, A. *Chem. Phys. Lett.* **2006**, *430*, 144.

[16] 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.

[17] Dolgounitcheva, O. ; Zakrzewski, V.G. ; Ortiz, J.V. *J. Phys. Chem. A* **2002**, *106*, 8411.

[18] Dolgounitcheva, O. ; Zakrzewski, V.G. ; Ortiz, J.V. *Int. J. Quant. Chem.* **2000**, *80*, 831.

[19] Dolgounitcheva, O. ; Zakrzewski, V. G. ; Ortiz, J. V. *J. Phys. Chem. A* **2003**, *107*, 822.

[20] Roca-Sanjuán, D. ; Rubio, M. ; Merchán, M. ; Serrano-Andrés, L. *J. Chem. Phys.* **2006**, *125*, 084302.

[21] Cauët, E. ; Dehareng, D. ; Lievin, J. *J. Phys. Chem. A* **2006**, *110*, 9200.

[22] Dolgounitcheva, O. ; Zakrzewski, V. G. ; Ortiz, J. V. *J. Phys. Chem. A* **2009**, *113*, 14630.

[23] Dolgounitcheva, O. ; Zakrzewski, V. G. ; Ortiz, J. V. *J. Am. Chem. Soc.* **2000**, *122*, 12304.

[24] Orlov, V.M. ; Smirnov, A.N. ; Varshavsky, Ya.M. *Tetrahedron Lett.* **1976**, *17*, 4377.

[25] Peng, S. ; Padva, A. ; LeBreton, P.R. *Proc. Nat. Acad. Sci.* **1976**, *73*, 2966.

[26] Lin, J. ; Yu, C. ; Peng, S. ; Akiyama, I. ; Li, K. ; Li, Li Kao ; LeBreton, P. R. *J. Phys. Chem.* **1980**, *84*, 1006.

[27] Urano, S. ; Yang, X. ; LeBreton, P.R. *J. Mol. Struct.* **1989**, *214*, 315.

[28] Dougherty, D. ; Wittel, K. ; Meeks, J. ; McGlynn, S.P. *J. Am. Chem. Soc.* **1976**, *98*, 3815.

[29] Hush, N.S. ; Cheung, A.S. *Chem. Phys. Lett.* **1975**, *34*, 11.

[30] Zavilopulo, A. N. ; Shepnik, O. B. ; Agafonova, A. S. *J. Phys. B* **2009**, *42*, 025101.

[31] Zhou, J. ; Kostko, O. ; Nicolas, C. ; Tang, X. ; Belau, L. ; de Vries, M.S. ; Ahmed, M. *J. Phys. Chem. A* **2009**, *113*, 4829.

[32] Choi, K.-W. ; Lee, J.-H. ; Kim, S.K. *J. Am. Chem. Soc.* **2005**, *127*, 15674.

[33] Weigend, F. ; Haser, M. *Theor. Chim. Acta* **1997**, *208*, 359.

[34] Weigend, F. ; Häser, M. ; Patzelt, H. ; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143.

[35] Weigend, F. ; Köhn, A. ; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175.

[36] Dunning, T.H. *J. Chem. Phys.* **1989**, *90*, 1007.

[37] Landau, A. ; Khistyaev, K. ; Dolgikh, S. ; Krylov, A.I. *J. Chem. Phys.* **2010**, *132*, 014109.

[38] Krylov, A.I. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433.

[39] Pieniazek, P.A. ; Arnstein, S.A. ; Bradforth, S.E. ; Krylov, A.I. ; Sherrill, C.D. *J. Chem. Phys.* **2007**, *127*, 164110.

[40] Pal, S. ; Rittby, M. ; Bartlett, R.J. ; Sinha, D. ; Mukherjee, D. *Chem. Phys. Lett.* **1987**, *137*, 273.

[41] Stanton, J.F. ; Gauss, J. *J. Chem. Phys.* **1994**, *101*, 8938.

[42] Kamiya, M. ; Hirata, S. *J. Chem. Phys.* **2006**, *125*, 074111.

[43] Pieniazek, P.A. ; Bradforth, S.E. ; Krylov, A.I. *J. Chem. Phys.* **2008**, *129*, 074104.

[44] Golubeva, A.A. ; Pieniazek, P.A. ; Krylov, A.I. *J. Chem. Phys.* **2009**, *130*, 124113.

[45] Yau, A. D. ; Perera, S. A. ; Bartlett, R. J. .

[46] Chai, J.-D. ; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.

[47] Lebedev, V.I. *Zh. Vychisl. Mat. Mat. Fiz.* **1975**, *15*, 48.

[48] Murray, W.C. ; Handy, N.C. ; Laming, G.J. *Mol. Phys.* **1993**, *78*, 997.

[49] Berger, R. ; Fischer, C. ; Klessinger, M. *J. Phys. Chem. A* **1998**, *102*, 7157.

[50] Wang, S. ; III, H. F. Schaefer *J. Chem. Phys.* **2006**, *124*, 044303.

[51] Mozhayskiy, V.A. ; Krylov, A.I. ;
ezSpectrum, http://iopenshell.usc.edu/downloads/.

[52] Y. Shao, L.F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. Brown, A.T.B. Gilbert, L.V. Slipchenko, S.V. Levchenko, D.P. O'Neil, R.A. Distasio Jr, R.C. Lochan, T. Wang, G.J.O. Beran, N.A. Besley, J.M. Herbert, C.Y. Lin, T. Van Voorhis, S.H. Chien, A. Sodt, R.P. Steele, V.A. Rassolov, P. Maslen, P.P. Korambath, R.D. Adamson, B. Austin, J. Baker, E.F.C. Bird, H. Daschel, R.J. Doerksen, A. Drew, B.D. Dunietz, A.D. Dutoi, T.R. Furlani, S.R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G.S. Kedziora, R.Z. Khalliulin, P. Klunziger, A.M. Lee, W.Z. Liang, I. Lotan, N. Nair, B. Peters, E.I. Proynov, P.A. Pieniazek, Y.M. Rhee, J. Ritchie, E. Rosta, C.D. Sherrill, A.C. Simmonett, J.E. Subotnik, H.L. Woodcock III, W. Zhang, A.T. Bell, A.K. Chakraborty, D.M. Chipman, F.J. Keil, A. Warshel, W.J. Herhe, H.F. Schaefer III, J. Kong, A.I. Krylov, P.M.W. Gill, M. Head-Gordon *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172.

[53] Belau, L. ; Wilson, K.R. ; Leone, S.R. ; Ahmed, M. *J. Phys. Chem. A* **2007**, *111*, 7562.

[54] Berkowitz, J. *J. Chem. Phys.* **1978**, *69*, 3044.

[55] Guerra, C.F. ; Bickelhaupt, F.M. ; Saha, S. ; Wang, F. *J. Phys. Chem. A* **2006**, *110*, 4012.

[56] Hanus, M. ; Kabeláç, M. ; Rejnek, J. ; Ryjék, F. ; Hobza, P. *J. Phys. Chem. B* **2004**, *208*, 2087.

[57] Piacenza, M. ; Grimme, S. *J. Comput. Chem.* **2004**, *25*, 83.

[58] Colominas, C. ; Luque, F. J. ; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811.

[59] Broo, A. ; Holmén, A. *Chem. Phys.* **1996**, *211*, 147.

[60] López, J.C. ; Peña, M.I. ; Sanz, M.E. ; Alonso, J.L. *J. Chem. Phys.* **2007**, *126*, 191103.

[61] Choi, M.Y. ; Miller, R.E. *J. Phys. Chem. A* **2007**, *111*, 2475.

[62] Nir, E. ; Plützer, Ch. ; Kleinermanns, K. ; de Vries, M. *Eur. Phys. J. D* **2002**, *20*, 317.

[63] Marian, C. M. *J. Phys. Chem. A* **2007**, *111*, 1545.

[64] Choi, M. Y. ; Miller, R. E. *J. Am. Chem. Soc.* **2006**, *128*, 7320.

[65] Alonso, J. L. ; Peña, I. ; Lopéz, J. C. ; Vaquero, V. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 6141 .

[66] Choi, M.Y. ; Dong, F. ; Miller, R.E. *Phil. Trans. Royal. Soc. A* **2005**, *363*, 393.

[67] Feyer, V. ; Plekan, O. ; Richter, R. ; Coreno, M. ; Vall-llosera, 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.

[68] Oana, C.M. ; Krylov, A.I. *J. Chem. Phys.* **2007**, *127*, 234106.

[69] Zhou, J. ; Takahashi, L. K. ; Wilson, K. R. ; Leone, S. R. ; Ahmed, M. *Anal. Chem.* **2010**, *82*, 3905 .

[70] Lauer, G. ; Schäfer, W. ; Schweig, A. *Tetrahedron Lett.* **1975**, *16*, 3939.

[71] Dougherty, D. ; Younathan, E.S. ; Voll, R. ; Abdulnur, S. ; McGlynn, S.P. *J. Electron. Spectrosc. Relat. Phenom.* **1978**, *13*, 379.

[72] Wolken, J.K. ; Yao, C.Y. ; Turečeka, F. ; Polce, M.J. ; Wesdemiotis, C. *Int. J. Mass. Spectrom.* **2007**, *267*, 30.

[73] Plützer, Chr. ; Kleinermanns, K. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4877.