Ab initio determination of the ionization potentials of DNA and RNA nucleobases

Daniel Roca-Sanjuán, Mercedes Rubio, Manuela Merchán, and Luis Serrano-Andrés[a]

Instituto de Ciencia Molecular, Universitat de València, Apartado 22085, ES-46071 Valencia, Spain

Quantum chemical high level *ab initio* coupled-cluster and multiconfigurational perturbation methods have been used to compute vertical and adiabatic ionization potentials of the five canonical DNA and RNA nucleobases: uracil, thymine, cytosine, adenine, and guanine. Several states of their cations have been also calculated. The present results represent a systematic compendium of these magnitudes, establishing theoretical reference values at a level not reported before, calibrating computational strategies, and guiding the assignment of the features in the experimental photoelectron spectra. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2336217]

I. INTRODUCTION

Despite the crucial information that intrinsic properties of DNA and RNA nucleobases as ionization potentials (IPs) and electron affinities (EAs) can provide to elucidate the role of the nucleobases electronic structure and interactions in the genetic action, there is a too scattered set of experimental values available in the literature, whereas the theoretical determination has been so far carried out at modest computational levels.[1–22] Both properties are linked to the electrostatic interactions that nucleobases share within the DNA/RNA chain and with other reactive molecules. The interest to get accurate values grows with the increased knowledge of phenomena related with the electron acceptor and donor abilities of the target molecules, such as those involving charge transfer and transport along the DNA strand,[23] radiation damage and repair of the genetic material,[24–26] DNA-protein interaction,[27,28] DNA-based phototherapy,[29] and the new molecular-level biotechnologies.[30] The present paper reports calculations on the IPs of the five DNA and RNA nucleobases (see Fig. 1): cytosine (C), thymine (T), uracil (U), adenine (A), and guanine (G), at the highest level of theory available, in order to establish a benchmark reference for the analysis of these properties, allowing the selection of the best available experimental data, and also to analyze the adequacy of different computational strategies employed nowadays in the theoretical study of these and similar systems. A parallel study on the EAs of nucleobases is being undertaken.[31] Also, additional calculations have been performed on corresponding nucleosides and nucleotides, showing that the most favorable oxidation site in nucleotides is placed at the nucleobase π structure,[32] and not at the phosphate group, as recently suggested.

The ionization potential of a neutral molecule is the energy required to remove one electron from the molecule, and it can be seen as the energy change involved in the reaction $B \rightarrow B^+ + e^-$, where here B denotes a nucleic base, B^+ its cation, and e^- the electron. The scheme on Fig. 2 displays the different theoretical magnitudes related to an ionization potential. The vertical electronic energy difference (VIP_e) corresponds to the electronic transition from the initial (ground) state of the neutral system at its equilibrium geometry to the lowest-energy state of the cation, whereas the so-called adiabatic energy difference (AIP_e) is the energy gap between the minima of the states, and reflects the geometry relaxation taking place in the excited state, setting the lowest energetic bound or electronic band origin of the transition. Adding the zero-point vibrational energy corrections (ZPE) to AIP_e, the AIP_0 value is obtained. Strictly, only AIP_0 have experimental counterparts to be compared with, the 0-0 (also T_0) band origins, whereas the vertical magnitudes are typically related to the band maxima when assuming the Franck-Condon approximation, which is only reliable if the geometry of both initial and final electronic states is similar enough.[33,34]

As stated above, reported experimental values for vertical and adiabatic IPs of nucleobases show a large dispersion of data, reaching up to 0.75 eV,[1–15] mainly because of the experimental difficulties caused by the small photoionization yields[35,36] and the high reactivity of the resulting radical cation.[23] A good set of computed values shall help to discard less reliable data and to reduce the uncertainty in the values of the magnitude, even for measurements in the gas phase. Taking into account that the energy required to remove an electron from a system can be directly related to their reductive power, to know the sequence for the electron detachment energies for the nucleobases is clearly essential. From the theoretical viewpoint, both medium-level *ab initio* and density functional theory (DFT) methods have been widely employed. Taking into account the size of the molecules under analysis, a restricted number of *ab initio* methods have been used from Hartree-Fock (HF) descriptions in which Koopmann's theorem was applied[16,17,20] to second-order perturbative treatments (MP2),[17,20] and electron propagator methods such as the outer-valence Green's function (OVGF) and the partial third-order (P3) approximations,[37–41] employing basis sets up to 6-31++G(d,p). Regarding the most recent DFT
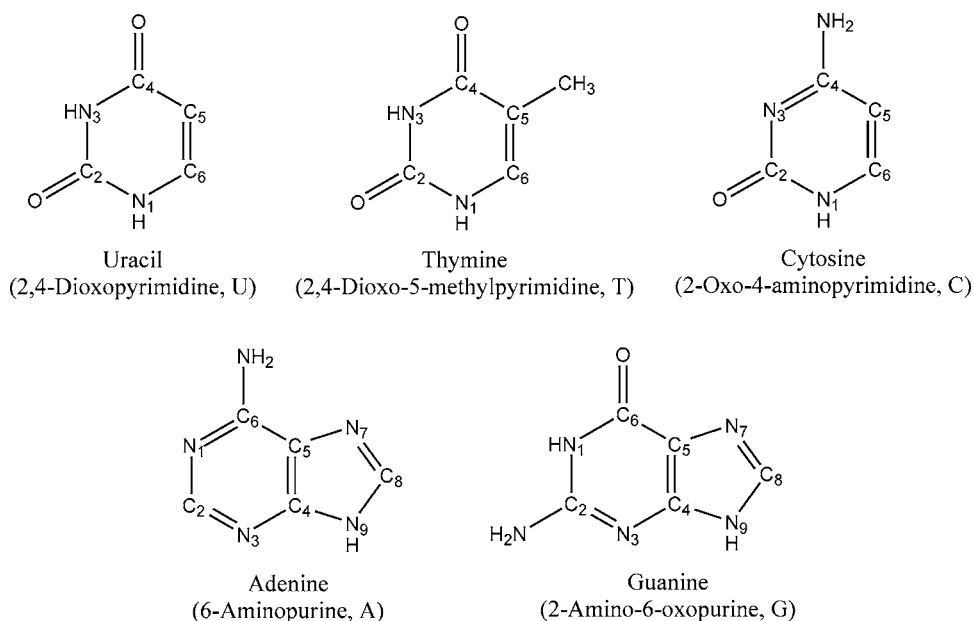
FIG. 1. DNA and RNA nucleobases structure and labeling with their conventional name and, within parentheses, the IUPAC name and the abbreviation.

calculations, many functionals and basis sets have been used, in particular, B3LYP.[18,19,21] Combining experimental and theoretical studies, the basic sequence for IP values has been established as $U > T > C > A > G$, concluding, therefore, that purine bases, adenine, and guanine, mainly the latter, are more easily oxidized than pyrimidine bases, being uracil and thymine those with the largest IPs. There are, however, many suspicious experimental values and, with regards to the theoretical treatment, the presence of spin contamination in the open-shell situations or the uncertain role of the basis set makes it necessary to carry out calculations able to provide quantitative account of these magnitudes.

The present paper reports computations on the lowest IPs of the mentioned nucleobases performed with different *ab initio* methods: Möller-Plesset second-order singly (MP2) and multiconfigurational (CASPT2) perturbation theory[42] and coupled cluster singles and doubles (CCSD), or singles, doubles, and triples [CCSD(T)] approaches, both for vertical and adiabatic values, and employing different types of one-electron basis sets. Also, the ultraviolet photoelectron spectra (UPS) of the nucleobases have been analyzed at the CASPT2 level by computing several states of the corresponding nucleobases cations. As mentioned above, the obtained results will allow us to discard odd experimental values to set up accurate data at the highest theoretical levels for future reference, and also to check the accuracy on the IP values yielded by different computational strategies, such as the use of atomic natural orbital (ANO)-type basis sets within the CASPT2//complete active space self-consistent field (CASSCF) approach, the inclusion of the recent IPEA correction to the CASPT2 methodology,[43] and the effect of the modification of the polarization exponent in a small basis such as 6-31G(d) in order to be used in larger systems.[28] The calibration of a flexible method such as CASPT2, especially suited to compute excited states in a general fashion,[32–34,44–49] with computationally expensive methods such as CCSD(T), known to perform accurately for well-defined ground states, will be extremely useful to establish the expected precision of the method when approaching the photochemistry of nucleobases and related systems.

II. COMPUTATIONAL DETAILS

In order to compare different computational strategies, geometry optimizations were carried out at different levels of theory. Neutral nucleobases were optimized at the MP2/6-31G(d), CASSCF/ANO-L 431/21, and CCSD/aug-cc-pVDZ levels of theory. The lowest state of each nucleobase cation was also optimized at the CASSCF level with a one-electron basis set of the type ANO-L contracted to C,N,O [$4s3p1d$]/H[$2s1p$] (hereafter ANO-L 431/21) to be used in the calculation of adiabatic IPs. No symmetry restriction was imposed, whereas all the minima were characterized by calculation of second-order derivatives at the same level of theory. At the optimized structures, additional CASPT2 and CCSD(T) calculations were performed. ZPE corrections have been included at different levels using the harmonic approach.[33] The multiconfigurational CASSCF calculations for the geometry optimizations used an active space comprising all π electrons and orbitals, which was extended to include the lone pair electrons and orbitals of the heteroatoms for the final CASSCF and CASPT2 calculations of VIPs and AIPs. In order to check a recent modification of the zeroth-order Hamiltonian within the CASPT2 approach,[42] the IPEA correction,[43] especially suited to correct systematic errors in open-shell cases, calculations were performed both with the original and the modified version, with a selected value for the level shift parameter of 0.25 a.u., as recommended.[43] Finally, in order to test the ability to compute IPs of a recently proposed modification of the d-polarization function in the 6-31G(d) one-electron basis sets,[28] MP2 calculations have been performed for VIPs in which the standard Gaussian-d exponent 0.8 has been replaced by the more diffuse exponent 0.2. Both GAUSSIAN-03 (Ref. 50) and MOLCAS-6.0 (Ref. 51) quantum chemistry programs have been employed. Opti-
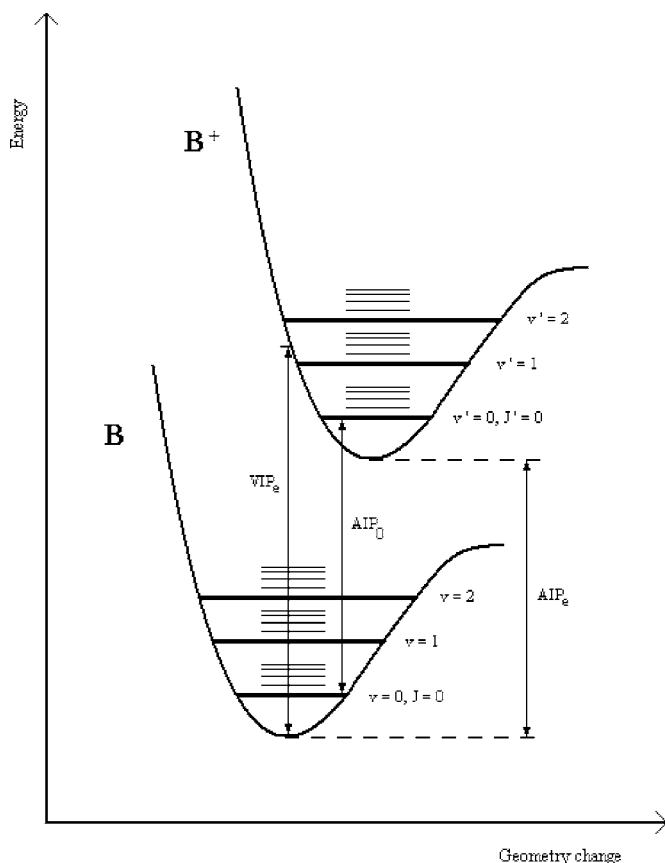
FIG. 2. IP diagram. Definitions of the theoretical magnitudes related to IP are graphically shown through the electronic, vibrational, and rotational potential energy levels. Those magnitudes are VIP_e (vertical electronic ionization potential, from the ground state minimum), AIP_e (adiabatic electronic ionization potential, from minimum to minimum), and AIP_0 (adiabatic ionization potential including the zero-point vibrational corrections of the minima).

mized geometries can be obtained from the authors upon request.

III. RESULTS AND DISCUSSION

A. Lowest-energy vertical ionization potential of the nucleobases

Table I compiles the available experimental data for the lowest-energy VIPs of the five nucleobases together with earlier and present theoretical results. If the old photoionization mass-spectrometry data by Lifschitz et al.[1] were added, the dispersion observed in the experimental estimation would reach 0.5 eV in some cases. However, in view of the obtained results, these old measurements can be clearly ruled out reducing the overall dispersion to 0.2 eV. Independently of the experimental or theoretical techniques employed to determine the lowest-energy VIPs, the observed trend in the nucleobases indicates, as mentioned above, that purines are more easily oxidized than pyrimidines. Regarding the theoretical results listed in Table I, they include HF, MP2, DFT, and P3 studies collected from literature together with presently computed CCSD, CCSD(T), and CASPT2 VIPs. A qualitative description of the trends of the VIPs in nucleobases can be already obtained by using single-reference methods such as HF and MP2 [employing here a

6-31++G(d,p) one-electron basis sets], which are, however, of less quantitative value. The results obtained by using Koopman's theorem at the HF level underestimate VIPs more than 1 eV with respect to the more accurate CCSD(T) and CASPT2(IPEA) results. On the other hand correlated MP2/6-31++G(d,p) calculations overestimate the values near 0.4–1.0 eV. This indicates that the correlation energy introduced at the MP2 level stabilizes more the ground state of the neutral molecule than that of the cation. Unrestricted HF wave functions typically have spin contamination which can be partially removed by using spin annihilation or projected techniques, such in the projected HF (PHF) and projected MP2 (PMP2) approaches. The large differences obtained from nonprojected to projected results, in particular, in the perturbative calculations ranging from 0.55 eV in guanine to 0.92 eV in uracil, reflect important spin-contamination problems and point out that low-level single-reference methods will be of minor quantitative value to obtain accurate predictions. PMP2 VIPs for pyrimidine nucleobases are, however, favorably compared with the benchmark values we will consider as reference, CCSD(T) and CASPT2(IPEA), whereas in purine DNA bases PMP2 results are up to 0.24 eV too high. Calculations performed using a DFT method including the B3LYP functional give similar results to PMP2 employing the 6-31++G(p,d) basis set for the pyrimidine bases, whereas the VIPs decrease by near 0.4 eV for adenine and guanine, representing a slight underestimation of 0.1 eV with respect to the reference values. As regards the P3/6-311G(d,p) propagator approach,[37–41] specifically designed to determine ionization potentials, whereas it systematically increases the VIP values for pyrimidine bases with respect to the PHF, PMP2, and B3LYP results, it provides intermediate VIP values for the purine bases, generally closer to the reference values, although it depends strongly on the type of basis set.[37] Other propagator approaches, such as the OVGF method, give much less accurate values.[37]

CCSD and CCSD(T) calculations were performed here at the neutral ground-states geometries optimized at the CCSD/aug-cc-pVDZ level of calculation. The triples correction is shown here to be of less importance, considering that it changes the values of VIPs not more than 0.04 eV. The T_1 diagnostic gave always values lower than 0.02, whereas the analysis of the CASSCF wave functions proved their single-reference character. Therefore, the computationally demanding CCSD(T) approach can be considered extremely accurate and the computed value can be used as a reference, taking also into account that the aug-cc-pVDZ basis set gives also good results in most of cases. In order to compare the frequently used CASPT2//CASSCF strategy, that is, CASPT2 energies at the CASSCF optimized geometries, we computed VIPs at these level with the accurate ANO-L 431/21 basis set. Two types of CASPT2 calculations were performed, with the standard zeroth-order Hamiltonian[42] and with the new IPEA correction.[43] The application of the CASPT2/IPEA approach leads to a general increase of almost 0.2 eV in the VIPs values. This is expected taking into account that for the lowest state of the cation, an open-shell case, the IPEA correction avoids the usual overestimation of the correlation en-

TABLE I. Low-lying vertical ionization potentials (eV) of DNA and RNA nucleobases obtained by different experimental and theoretical methods.

Method	Uracil	Thymine	Cytosine	Adenine	Guanine
Experimental range[a]	9.4–9.6	9.0–9.2	8.8–9.0	8.3–8.5	8.0–8.3
HF/6-31++G(d,p)[b]	8.77	8.16	7.55	7.45	6.90
PHF//HF/6-31++G(d,p)[b]	8.48	8.21	7.69	7.36	6.97
MP2/6-31++G(d,p)[b]	10.35	9.50	9.44	9.38	8.88
PMP2//MP2/6-31++G(d,p)[b]	9.43	9.07	8.69	8.62	8.33
B3LYP/6-31G++G(d,p)[c]	9.47	9.01	8.69	8.26	7.98
P3//MBPT(2)/6-311G(d,p)[d]	9.54	9.14	8.79	8.49	8.13
CCSD//CCSD/aug-cc-pVDZ[e]	9.40	9.01	8.72	8.39	8.07
CCSD(T)//CCSD/aug-cc-pVDZ[e]	9.43	9.04	8.76	8.40	8.09
CASPT2//CASSCF/ANO-L 431/21[e]	9.22	8.87	8.54	8.18	7.91
CASPT2(IPEA=0.25)//CASSCF/ANO-L 431/21[e]	9.42	9.07	8.73	8.37	8.09

[a]References 2–14.
[b]Reference 20.
[c]Reference 21.
[d]References 37–41.
[e]Present work.

ergy observed in CASPT2. The final CASPT2(IPEA =0.25)//CASSCF/ANO-L 431/21 results are in almost perfect correspondence with the CCSD(T) results, within 0.03 eV, confirming the suitability of the former computational strategy to deal with this type of systems, both for excitation energies and geometry optimizations.

Table II includes results at the PMP2 level where the 6-31G(d) basis set has been employed to compute VIPs both with the standard 0.8 Gaussian-d exponent and with the modified 0.2 value proposed recently, which was found, after a careful analysis, particularly suited to compute cation-π energies.[28] VIPs obtained with the modified 6-31G(d=0.2) basis set systematically increase in all nucleobases with respect to those of the PMP2/6-31G(d=0.8) basis set, undergoing changes ranging from 0.13 eV in adenine to 0.22 eV in guanine. Although the final values are accurate only within 0.2 eV with respect to the reference values, the performance of the modified small-size basis set is promising if one thinks in using it for larger compounds.

The general conclusions one can obtain by all the previous results is that, if CCSD(T) values are taken as reference, an accurate determination of VIPs, within 0.1 eV, requires the use of high-level methods such as specifically designed P3 techniques or the more general CASPT2 method, where the use of the recent IPEA approach seems necessary, whereas DFT-based techniques do not reach the required accuracy. Also, CASSCF determined geometries have shown to adequately describe the ground states of the nucleobases. Differences between the CASSCF and CCSD structures amount to less than 0.03 Å, 2°, and 7° for the bond lengths,

bond angles, and dihedral angles, respectively. The addition of diffuse functions to the one-electron basis set seems to be of secondary importance, as proved in previous calculations.[37] ANO-type basis sets, on the other hand, already contain a large diffuse character because of their general contracted nature and are known to provide very good results in all cases.[33,34,45,48] Regarding the comparison with the experimental values, obtained from band maxima in different gas-phase photoelectron spectra, it can be concluded that the vertical excitation energies correspond in all cases to the lower-energy side of the range of data. In any case, the energy order for the VIPs in the nucleobases is clearly confirmed as U>T>C>A>G, ranging from 9.43 to 8.09 eV, with successive decreases of 0.3–0.4 eV between the systems.

B. Low-lying excited states for the cations of DNA nucleobases

Once established the accuracy of the CASPT2//CASSCF approach to compute ionization potentials in the nucleobases it is possible to study the lowest IPs of these systems and help to interpret the recorded photoelectron spectra. Table III compiles the six low-lying ionization potentials of the nucleobases obtained by computing vertically, at the neutral ground state geometry, the six lowest states of their cations at the CASPT2(IPEA=0.25)//CASSCF/ANO-L 431/21 level of theory. The character of the related state of the cation is also described by the type of the molecular orbital where the unpaired electron is basically placed. All computed states can

TABLE II. Low-lying vertical ionization potential (eV) of DNA and RNA nucleobases obtained at the PMP2 level by using the standard 6-31G(d) basis set (Gaussian-d exponent 0.8) and its modification (0.2) (see also Refs. 27 and 28).

Method	Uracil	Thymine	Cytosine	Adenine	Guanine
PMP2/6-31G(d=0.8)//MP2/6-31G(d)	9.08	8.72	8.38	8.25	7.97
PMP2/6-31G(d=0.2)//MP2/6-31G(d)	9.22	8.86	8.53	8.38	8.19

TABLE III. Experimental (Expt. photoelectron spectroscopy) and computed (PT2, CASPT2(IPEA)//CASSCF/ANO-L 431/21 level) low-lying excited states (eV) for the cations of DNA and RNA nucleobases. Proposed assignments are included.

Uracil				Thymine				Cytosine		Adenine		Guanine	
PT2	Expt.[a]	Expt.[b]	Expt.[c]	PT2	Expt.[a]	Expt.[b]	Expt.[c]	PT2	Expt.[d]	PT2	Expt.[b,d]	PT2	Expt.[d]
9.42	9.60	9.59	9.45	9.07	9.20	9.18	9.02	8.73	8.82	8.37	8.48	8.09	8.30
(π)	(π)	(π)	(π)	(π)	(π)	(π)	(π)	(π)	(π)	(π)	(π)	(π)	(π)
9.83	10.13	10.11	10.02	9.81	10.05	10.03	9.95	9.42	9.45	9.05		9.56	9.90
($n_O C_4$)	(n)	(n)	(n)	($n_O C_4$)	(n)	(n)	(n)	(n_O)	(n)	($n_{N1,N3,N7}$)		(n_O)	(n)
10.41	10.55	10.56	10.51	10.27	10.44	10.39	10.40	9.49		9.54	9.6	9.61	9.90
(π)	(π)	(π)	(π)	(π)	(π)	(π)	(π)	(π)		(π)	(n,π)	($n_{N3,N7}$)	(n)
10.86	11.00	11.16	10.90	10.49	10.88	10.82	10.80	9.88	9.90	9.96		10.05	
($n_O C_2$)	(n)	(n)	(n)	($n_O C_2$)	(n)	(n)	(n)	(n_N)	(π)	($n_{N1,N7}$)		(π)	
12.59	12.7	12.63	12.50	12.37	12.30	12.27	12.10	11.84	11.8	10.38	10.5	10.24	10.45
(π)	(π)	(π)	(π)	(π)	(π)	(π)	(π)/(π)	(π)		(π)	(n,π)	(π)	(n)
13.57				13.40				12.71		11.06	11.39	10.90	11.15
(π)				(π)				(π_{NH2})		($n_{N1,N3,N7}$)	(n)	(π)	(π)

[a]Reference 12.
[b]Reference 8.
[c]Reference 3.
[d]Reference 5.
[e]Reference 7.

be essentially described as hole states with a singly excited configuration involving the occupied-like molecular orbitals (MOs), whereas higher-lying shake-up states, not included here, represent rearrangements of the remaining electrons. In all five molecules the low-lying first and second VIPs are related to a π and a lone pair MOs, respectively. For adenine, all three ring nitrogen atoms containing in-plane electron lone pairs contribute to the description of the corresponding n orbital, whereas in the other molecules the excitation takes place from the oxygen lone-pair MO of the keto group (on C_4 position in uracil and thymine). Except for guanine, in most of the nucleobases there is a sequence of well spaced higher VIPs, differing in nature only in the sixth VIP, which originates from a lone-pair MO in adenine and from π MOs in the pyrimidine bases. The fourth VIP in uracil and thymine can be related to the oxygen lone-pair orbital of the other keto group (on C_2), whereas in cytosine and adenine the excitations corresponds to nitrogen lone-pair orbitals. It is worth noticing the low energy computed for the second VIP in adenine, corresponding to nitrogen lone-pair orbitals, and the close degeneration found in guanine between the second and third VIPs, both related with lone-pair orbitals of the oxygen and nitrogen atoms, respectively, whereas the remaining computed VIPs are of π character. This analysis of the electronic structure in the nucleobases can be of great help to understand important photoinduced biological processes.

If the experimental UPS of the literature are compared,[3,5–8,10–12] it is clear that the agreement with the theoretical CASPT2 results and assignments is very good for uracil, thymine, and cytosine, whereas the discrepancies are larger for the purine bases: adenine and guanine. Similar conclusions are obtained when compared to previous theoretical P3 results.[37–40] The experimental features are deduced from the maxima of the bands of the UPS spectra which are more clear and resolved in the low-energy range of the spec-

tra of the pyrimidine bases. The spectra of purine bases are dense and the bands are supposed to overlap. In adenine, whereas previous theoretical P3 results[37] provide two pairs of degenerated n,π transitions near 9.5 and 10.5 eV corresponding here to the second to the fifth VIPs, the present calculations split the transitions by near 0.4–0.5 eV. Although the excitations computed at 9.05 and 9.96 eV and related to nitrogen lone-pair orbitals cannot be characterized in the spectrum, they are most probably hidden as shoulders of more intense π-type bands. In guanine the UPS spectra is very dense above 9.5 eV. A broad band with three unresolved peaks extends up to 12 eV. The CASPT2 results confirm the expectations of two degenerate transitions related with the oxygen and nitrogen lone pairs placed in the low-energy tail of the main band, whereas three π-type transitions follow. These assignments are also obtained in previous theoretical P3 calculations.[38]

C. Adiabatic ionization potentials of the nucleobases

Table IV compiles the experimental, gas phase, and present computed lowest-energy adiabatic ionization potentials ($AIPs_0$) of DNA nucleobases. The theoretical results have been obtained with the CASPT2 method using the standard and the modified IPEA, 0.25 a.u., zeroth-order Hamiltonian, the latter being considered to yield the most accurate results. All geometries of the ground states of the nucleobases neutral and cation molecules have been optimized at the CASSCF level, whereas the zero-point vibrational energy corrections at the respective minima have been introduced at the same level within the harmonic approach.[33]

Changes in geometry from neutral to cationic structures are moderate. Regarding pyrimidine bases the ionization leads to a slight increase, smaller than 0.08 Å, of the ethylenic C_5-C_6 bond length, while the C_6-N_1 bond enlarges and the remaining bond lengths do not basically change. Bond

TABLE IV. Experimental and present theoretical lowest-lying adiabatic ionization potentials (AIPs$_0$,eV) of DNA and RNA nucleobases.

Method	Uracil	Thymine	Cytosine	Adenine	Guanine
Expt.[a]	9.32	8.87	8.68	8.26	7.77
Expt.[b]	9.20	8.80	8.45	7.80	7.85
Expt.			...	8.55[d]	7.85[c]
CASPT2//CASSCF/ANO-L 431/21/ZPE[e]	8.99	8.69	8.40	7.94	7.51
CASPT2(IPEA=0.25)//CASSCF/ANO-L 431/21/ZPE[e]	9.12	8.84	8.56	8.11	7.65

[a]Photoionization mass spectroscopy in the gas phase. Band origin (Ref. 4).
[b]Photoelectron spectra. Band onset (Ref. 13).
[c]Ultraviolet photoelectron spectra (Ref. 5).
[d]Ion resonance ionization spectroscopy (Ref. 15).
[e]Zero-point energy (ZPE) correction included with the harmonic correction using CASSCF frequencies.

angles are just slightly modified, up to 5°, in particular, those related to the keto group C_2=O. All three nucleobases become essentially planar in the ground state of their cations (except for the methyl group in thymine). That means that the pyramidalization obtained for the NH$_2$ group in neutral cytosine because of the presence of the lone π pair on the nitrogen atom, 28°, vanishes when the cation reorganizes its geometry, although the charge withdrawal from that nitrogen atom is just minor. Also in purine nucleobases the largest geometry change upon ionization relates to the amino group, which becomes coplanar with the ring in the cation, decreasing the corresponding dihedral angle by 15° in adenine and by 43° in guanine. The latter undergoes then the largest geometrical change with respect to the neutral system, what it will be reflected in the energy difference between the vertical and adiabatic IPs.

The sequence of adiabatic ionization potentials in Table IV keeps the energy order obtained for the VIPs: U>T>C>A>G. The IPEA correction to CASPT2 leads to increases in AIPs from 0.13 to 0.17 eV. These results will be considered the most accurate theoretical data obtained. If we focus in the difference between the vertical and adiabatic IPs, they are (in eV) U(−0.30), T(−0.23), C(−0.17), A(−0.26), and G(−0.44). As expected from the small geometrical changes computed, the relaxation energy upon excitation is not large, whereas it is similar in all compounds. Comparing to the experimental results, the reported AIPs in uracil, from 9.20 to 9.32 eV, are slightly higher than the computed value of 9.12 eV, which supports the most modern determination, 9.20 eV.[13] Other theoretical values, essentially using different DFT functionals, do not help to get quantitivity because of the large dispersion of results obtained.[18] This comment is generally applicable to all the AIPs of the nucleobases. In thymine the lowest AIP, 8.84 eV at the CASPT2 level, is safely placed within the experimental range, between 8.80 and 8.87 eV. In cytosine, the computed value, 8.56 eV, lies intermediate among the experimental range of values, 8.45–8.68 eV. Regarding the purine nucleobases, the determination of the lowest AIP in adenine is more problematic. The experimental values are reported from 7.80 to 8.55 eV, whereas the theoretical determination places the band origin at and intermediate position, 8.11 eV. In this case, also other calculations at different levels, MP2 and DFT,[18,19] place the band in the low-energy side of the experimental range, 7.9–8.1 eV, therefore the high and discrepant value of

Hwang *et al.*,[15] 8.55 eV, obtained from analysis of adiabatic recombination energy in adenine radical cations, can be safely ruled out. Finally, the computed value for the AIP in guanine, 7.65 eV, is just slightly lower than the experimental range, 7.77–7.85 eV.

IV. CONCLUSIONS

Quantum chemical *ab initio* coupled-cluster and multiconfigurational perturbation methods have been used to compute the ionization potentials of the five canonical DNA and RNA nucleobases: uracil, thymine, cytosine, adenine, and guanine. By comparing the results of different computational strategies it is concluded that CCSD(T) and CASPT2 excitation energies provide the most accurate set of values for these magnitudes. The recently introduced IPEA definition of the zeroth-order Hamiltonian in the CASPT2 method has been proved crucial to obtain an accuracy of 0.03 eV as compared to the CCSD(T) results, which, on the other hand, have been shown to be reliable by analysis of the corresponding T_1 parameter and the single reference character of the wave functions. The CASPT2 method has also been employed to compute six states of the respective cations. For all molecules except guanine, the character of the six lowest states of the corresponding cation has been determined as related to π, n, π, n, π, and π orbitals, respectively, being the lowest lone-pair ionized those of the oxygen (nitrogen in adenine). In guanine the almost degenerate second and third VIPs correspond to n orbitals, whereas the others belong to the π type. The obtained assignments lead to proper understanding of the nature of the peaks measured in the gas-phase photoelectron spectra. Finally, adiabatic ionization potentials have been computed and compared with recorded band origins. The overall set of results obtained allows establishing reference values for the VIPs and AIPs of the nucleobases, and helps to rule out less reliable theoretical and experimental data.

ACKNOWLEDGMENTS

[1] C. Lifschitz, E. D. Bergmann, and B. Pullman, Tetrahedron Lett. **46**, 4583 (1967).

[2] N. S. Hush and A. S. Cheung, Chem. Phys. Lett. **34**, 11 (1975).

[3] G. Lauer, W. Schäfer, and A. Schweig, Tetrahedron Lett. **45**, 3939 (1975).

[4] V. M. Orlov, A. N. Smirnov, and Y. M. Varshavsky, Tetrahedron Lett. **48**, 4377 (1976).

[5] D. Dougherty, E. S. Younathan, R. Voll, S. Abdulnur, and S. P. McGlynn, J. Electron Spectrosc. Relat. Phenom. **13**, 379 (1978).

[6] A. Padva, T. J. O'Donnell, and P. R. LeBreton, Chem. Phys. Lett. **41**, 278 (1976).

[7] J. Lin, C. Yu, S. Peng, I. Akiyama, K. Li, L. K. Lee, and P. R. LeBreton, J. Am. Chem. Soc. **102**, 4627 (1980).

[8] S. Urano, X. Yang, and P. R. LeBreton, J. Mol. Struct. **214**, 315 (1989).

[9] D. Dougherty and S. P. McGlynn, J. Chem. Phys. **67**, 1289 (1977).

[10] J. Lin, C. Yu, S. Peng, I. Akiyama, K. Li, L. K. Lee, and P. R. LeBreton, J. Phys. Chem. **84**, 1006 (1980).

[11] C. Yu, S. Peng, I. Akiyama, J. Lin, and P. R. LeBreton, J. Am. Chem. Soc. **100**, 2303 (1978).

[12] D. Dougherty, K. Wittel, J. Meeks, and S. P. McGlynn, J. Am. Chem. Soc. **98**, 3815 (1976).

[13] S. G. Lias, J. E. Bartmess, J. F. Liebman, J. L. Holmes, R. D. Levin, and W. G. Mallard, J. Phys. Chem. Ref. Data **17**, 1 (1988).

[14] B. I. Verkin, L. F. Sukodub, and I. K. Yanson, Dokl. Akad. Nauk SSSR **228**, 1452 (1976).

[15] C. T. Hwang, C. L. Stumpf, Y.-Q. Yu, and H. I. Kenttämaa, Int. J. Mass. Spectrom. **182**, 253 (1999).

[16] A. O. Colson, B. Besler, and M. D. Sevilla, J. Phys. Chem. **96**, 9787 (1992).

[17] M. D. Sevilla, B. Besler, and A. O. Colson, J. Phys. Chem. **99**, 1060 (1995).

[18] N. Russo, M. Toscano, and A. Grand, J. Comput. Chem. **21**, 1243 (2000).

[19] S. D. Wetmore, R. J. Boyd, and L. A. Eriksson, Chem. Phys. Lett. **322**, 129 (2000).

[20] C. E. Crespo-Hernández, R. Arce, Y. Ishikawa, L. Gorb, J. Leszczynski, and D. M. Close, J. Phys. Chem. A **108**, 6373 (2004).

[21] D. M. Close, J. Phys. Chem. A **108**, 10376 (2004).

[22] P. Bernhardt and H. G. Paretzke, Int. J. Mass. Spectrom. **223**, 599 (2003).

[23] S. Steenken, Chem. Rev. (Washington, D.C.) **89**, 503 (1989).

[24] A. Gräslund, A. Ehrenberg, A. Rupprecht, G. Ström, and H. Crespi, Int. J. Radiat. Biol. **28**, 313 (1975).

[25] M. Yan, D. Becker, S. Summerfield, P. Renke, and M. D. Sevilla, J. Phys. Chem. **96**, 1938 (1992).

[26] A. O. Colson and M. D. Sevilla, J. Phys. Chem. **99**, 3867 (1995).

[27] R. Wintjens, J. Liévin, M. Rooman, and E. Buisine, J. Mol. Biol. **302**, 395 (2000).

[28] R. Wintjens, C. Biot, M. Rooman, and J. Liévin, J. Phys. Chem. A **197**, 6249 (2003).

[29] R. Bonnett, *Chemical Aspects of Photodynamic Therapy* (Gordon and Breach, Amsterdam, 2000).

[30] P. N. Prasad, *Introduction to Biophotonics* (Wiley, Hoboken, NJ, 2003).

[31] D. Roca-Sanjuán, M. Merchán, and L. Serrano-Andrés (unpublished).

[32] M. Rubio, D. Roca-Sanjuán, M. Merchán, and L. Serrano-Andrés, J. Phys. Chem. B **110**, 10234 (2006).

[33] L. Serrano-Andrés and M. Merchán, in *Encyclopedia of Computational Chemistry*, edited by P. v. R. Schlegel, P. R. Schreiner, H. F. Schaefer III, W. L. Jorgensen, W. Thiel, and R. C. Glen (Wiley, Chichester, 2004).

[34] M. Merchán and L. Serrano-Andrés, in *Computational Photochemistry*, edited by M. Olivucci (Elsevier, Amsterdam, 2005), Vol. 16, Chap. 2, p. 35.

[35] H. Görner, J. Photochem. Photobiol., B **26**, 117 (1994).

[36] C. E. Crespo-Hernández and R. Arce, Photochem. Photobiol. **76**, 259 (2002).

[37] O. Dolgounitcheva, V. G. Zakrzewski, and J. V. Ortiz, Int. J. Quantum Chem. **80**, 831 (2000).

[38] O. Dolgounitcheva, V. G. Zakrzewski, and J. V. Ortiz, J. Am. Chem. Soc. **122**, 12304 (2000).

[39] O. Dolgounitcheva, V. G. Zakrzewski, and J. V. Ortiz, J. Phys. Chem. A **106**, 8411 (2002).

[40] O. Dolgounitcheva, V. G. Zakrzewski, and J. V. Ortiz, J. Phys. Chem. A **107**, 822 (2003).

[41] D. M. Close, J. Phys. Chem. B **107**, 864 (2003).

[42] K. Andersson, P.-A. Malmqvist, and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).

[43] G. Ghigo, B. O. Roos, and P.-Å. Malmqvist, Chem. Phys. Lett. **396**, 142 (2004).

[44] L. Serrano-Andrés, M. Merchán, I. Nebot-Gil, R. Lindh, and B. O. Roos, J. Chem. Phys. **98**, 3151 (1993).

[45] B. O. Roos, M. P. Fülscher, P.-Å. Malmqvist, L. Serrano-Andrés, K. Pierloot, and M. Merchán, Adv. Chem. Phys. **93**, 219 (1996).

[46] L. Serrano-Andrés, N. Forsberg, and P.-Å. Malmqvist, J. Chem. Phys. **108**, 7202 (1998).

[47] L. Serrano-Andrés and M. P. Fülscher, J. Am. Chem. Soc. **120**, 10912 (1998).

[48] L. Serrano-Andrés and M. Merchán, J. Mol. Struct.: THEOCHEM **729**, 99 (2005).

[49] M. Rubio, B. O. Roos, L. Serrano-Andrés, and M. Merchán, J. Chem. Phys. **110**, 7202 (1999).

[50] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Inc., Wallingford, CT, 2004.

[51] K. Andersson, M. Barysz, A. Bernhardsson *et al.*, MOLCAS, Version 6.0, Department of Theoretical Chemistry, Chemical Centre, University of Lund, Sweden, 2004.