Calculation of the Ionization Potentials of the DNA Bases in Aqueous Medium

David M. Close*

Department of Physics, Box 70652, East Tennessee State University, Johnson City, Tennessee 37614

Received: July 27, 2004; In Final Form: September 2, 2004

Projected MP2/6-31++G(d,p) ionization potential calculations have recently been performed on the DNA bases in the gas phase and in aqueous solution (*J. Phys. Chem. A*, **2004**, 6373). The goal of the present work is to explore methods to perform these same calculations with density functional theory. New results for the vertical ionization potentials for the DNA bases at the B3LYP/6-31++G(d,p) level are close to the PMP2 results and also close to the experimental results. Vertical ionization potentials for the DNA bases in aqueous medium at the PCM/B3LYP/6-31++G(d,p) level are thymine 5.41, (5.36), cytosine 5.32, (5.24), adenine 5.05, (5.08), and guanine 4.71, (4.77) eV. The numbers in parentheses are the previous PCM/PMP2 results. Again, the DFT results are comparable to the PMP2 results. Results are also presented for the vertical ionization potential of 5-MeC in the gas phase and also in aqueous solution. This results in a cytosine base that has an IP more like a purine and may therefore have to be considered to be in competition with guanine as a hole trap.

Introduction

A major goal in photochemistry has been to establish the threshold energies necessary to ionize nucleotides. The experimental determination of the ionization potentials of nucleotides presents two challenging problems. First of all, it is difficult to prepare intact gas-phase nucleotides. Also, one expects that the large number of valance orbitals with similar energy in a nucleotide would give rise to poorly resolved ionization spectra. Previous work has therefore focused on the components of a nucleotide, the bases, and the deoxyribose-phosphates.

Since there are difficulties in determining the experimental ionization potentials (IPs) of nucleotides, one would like to use theoretical calculations to estimate IPs. Theoretical gas-phase vertical IPs have been calculated for the DNA bases by Colson et al.[1] at the HF/6-31+G(d) level on structures optimized at the HF/3-21G level. Overall, Koopmans' theorem IPs compared most favorably with the experimental vertical IPs, though the best fit for the pyrimidine IPs was found for the 6-31+G(d) vertical values. A second study conducted higher level MP2/6-31+G(d) calculations on structures optimized at the ROHF/6-31G(d) level and found that the vertical IPs are well predicted at this level (within 0.2 eV) except for thymine (which is 1.2 eV too high).[2] Projected MP2/6-31++G(d,p) IP calculations have recently been performed on the DNA bases. The average deviation between calculated and experimental values is only about 0.1 eV for the gas-phase IPs. This study also examined the IPs of the bases in aqueous solution.[3]

Accurate theoretical results of IPs can also be obtained using ab initio propagator calculations in the partial third-order (P3) approximation with the 6-311G(d,p) basis set.[4] Gas-phase P3 IP calculations on all the DNA bases have recently been reported.[5] However, this technique is not presently available for calculations in an aqueous medium. Therefore, attempts are made in the present study to examine other methods of performing accurate IP calculations on the DNA bases in both the gas phase and in aqueous solution.

* Phone: 423-439-5646; fax: 423-439-6905; e-mail: closed@etsu.edu.

TABLE 1: Vertical Ionization Potentials for the DNA Bases[d]

molec.	OVGF-MP2/ 6-311G(d,p)[a]	MP2/ 6-31+ +G(d,p)[b]	PMP2/ 6-31+ +G(d,p)[c]	B3LYP/ 6-31+ +G(d,p)	expt	expt ref
Ura	9.54	10.02	9.43	9.47	9.50	(9)
Thy	9.13	9.55	9.07	9.01	~9.1	(9)
Cyt	8.79	9.45	8.69	8.69	8.80	(10)
Ade	8.49	9.42	8.62	8.26	8.44	(11)
Gua	8.13	8.91	8.33	7.98	8.24	(12)
indole	7.75	8.04	7.68	7.66	7.76	(13)

[a] The values for Thy, Cyt, Ade, and Gua were taken from ref 5. [b] These values are similar to those presented in ref 3. Here, however, they are not corrected for the ZPE. [c] These are the values presented in ref 3. [d] Energies in eV.

Theoretical Methods

All of the calculations presented here are based on optimizations with either the second-order Møller−Plesset perturbational theory (MP2), or the hybrid Hartree−Fock density functional theory functional B3LYP, in conjunction with a series of different basis sets. Frequency calculations were performed at the same level of theory to ensure that the systems represent true minima on the potential energy surfaces and to provide corrections for the zero-point vibrational effects. To investigate the effect of the surrounding matrix on these systems, the environment was modeled using the polarized continuum model (PCM) of Tomasi et al.[6] All calculations were performed with the Gaussian 98 (Revision A11.3) suite of programs.[7]

Results and Discussions

IP calculations previously obtained using ab initio electron propagator calculations in the partial third-order (P3) approximation with the 6-311G(d,p) basis set are presented in Table 1. The calculations are labeled OVGF-MP2/6-311G(d,p). Here, one sees the rather remarkable agreement between the calculated and experimental vertical ionizations potentials for the canonical DNA bases.

Barton and coworkers have used indole derivatives as artificial nucleic acid bases.[8] These molecules appear to serve as a hole

10.1021/jp046660y CCC: $27.50
Published on Web 10/16/2004

TABLE 2: Vertical Ionization Potentials for the DNA Bases OVGF-MP2/6-311G(d,p)[a]

molec.	$N_{1,9}-H$	expt[b]	$N_{1,9}-CH_3$	expt	refs
Ura	9.54	9.50	9.14	~9.2	(15)
Thy	9.13	~9.1	8.78	8.79	(15)
Cyt	8.79	8.80	8.53	8.65	(16)
5-MeC	8.50	8.78	8.26	8.50	(16)
Ade	8.49	8.44	8.34	8.39	(11)
Gua	8.13	8.24	7.98	8.02	(12)

[a] Energies in eV. [b] These are the experimental values given in Table 1 for the $N_{1,9}-H$ bases.

TABLE 3: Zero-Point Energy Corrections at B3LYP/6-31+G(d,p)[a]

molec.	B3LYP/6-31+G(d,p)	corrected for ZPE
Thy	9.01	8.97
Cyt	8.69	8.64
Ade	8.26	8.26
Gua	7.98	8.03

[a] Energies in eV.

trap in DNA since they have a lower oxidation potential than guanine. Calculations on indole are also included in Table 1. The results of some more familiar calculations using MP2 and density functional theory are also included in Table 1 for comparison and will be discussed herein.

There are likely problems with some of the experimental data. The experimental paper on ionization of cytosine points out that the photoelectric spectra of the cytosine derivatives are rather broad, something not seen in the other bases.[10] This is most likely due to the presence of several isomers in the sample. Also, higher probe temperatures were required for some of the cytosine derivatives (195 °C versus 152 °C for thymine). This could give rise to partial decomposition. Even though no error limits are given in the experimental paper, it may be that the errors in measuring the IPs for cytosine were greater than for the other bases.

The calculations in Table 1 are for the canonical bases with N1−H for the pyrimidines and N9−H for the purines. To provide a better representation of DNA, it was decided to repeat the calculations with $N1-CH_3$ or $N9-CH_3$ to mimic the glycosidic bond in the nucleosides. The calculations in Table 2 show a small downward shift in IPs for the methylated bases.[14] In this same vein, new calculations have been performed on C5 methlylated cytosine.[5] This results in a cytosine base that has an IP more like a purine and may therefore have to be considered to be in competition with guanine as a hole trap.

Regardless of the overall comparisons between theoretical and experimental IPs, all the calculations presented in Table 2 suggest that the trend in IPs is U > T > C > A > G, with the pyrimidines having significantly higher IPs than the purines. Guanine has the lowest IP and therefore would be the easiest to oxidize.

Most IP calculations in the literature are given with the ZPE correction.[17] The ZPE corrections for the DNA bases at the B3LYP/6-31+G(d,p) level are listed in Table 3. It can be seen for the canonical bases at least that the corrections are very small.

The next step is to ask what are these ionization potentials in a biologically relevant (aqueous) environment. Here, one has only limited experimental data. One knows, for example, that reported ionizations with 250-nm light (which corresponds to energies of 4.9 eV) involve a biophotonic process.[18] On the other hand, 193 light (6.4 eV) does mono-photonically ionize purines.[19] We can use this information to bracket the threshold energies required for DNA ionization in aqueous solution to lie in the range of ~4.9−6.4 eV.

TABLE 4: Polarized Continuum Model Calculations in Water (B3LYP/6-31++G(d,p))[a]

molec.	B3LYP	minus 1.30	PMP2[b]	expt	refs
Ura	7.01	5.71	5.55		
Thy	6.71	5.41	5.36	5.4	(22)
Cyt	6.62	5.32	5.24	5.5	(22)
5-MeC	6.43	5.13			
Ade	6.35	5.05	5.05	5.0	(22)
Gua	6.01	4.71	4.77	4.8	(22)
8-OxG	5.71	4.41			
Indole	5.76	4.46	4.46	4.35	(21)

[a] Energies in eV. Pyrimidines with N1−H and purines with N9−H. [b] The results at the PMP2/6-31++G(d,p) level are taken from ref 3.

To study how an aqueous environment would affect IPs, one needs to repeat these calculations in a dielectric continuum. Presently, the electron propagator methods cannot be used to do these calculations. So, one must look back at MP2 or DFT techniques. Here, one has to look to see which technique yields the best results. Table 1 shows vertical IP calculations using various basis sets with MP2 and B3LYP.

The MP2 calculations in Table 1 seem to overestimate the vertical IPs by slightly more that 0.5 eV. Bertran has pointed out that difficulties can be encountered in doing MP2 calculations of radical cations because of the overestimation of spin polarization (which is related to spin contamination).[20]

The problems with spin contamination arise in computing the energy of the cation. For example in adenine, the cation has an S^2 value of 1.0486. Also, the frequency calculation on the adenine cation has a negative frequency ($-747.8\ cm^{-1}$). This is an indication that there is likely a geometry change in creating the cation (one electron removal). The problem with spin contamination can be addressed by using the projected MP2 (PMP2) energies. These calculations, as previously reported,[3] are presented in Table 1 and show rather good agreement with the experimental IPs.

Good results for the gas-phase IPs have been obtained at the B3LYP/6-31++G(d,p) level. For easy comparisons, these results are presented in Table 1. These calculations slightly underestimate the vertical IPs by about 0.14 eV. This is probably related to the fact that the DFT calculations at the B3LYP level do not overestimate spin polarization. The appeal of the DFT procedures is that one can perform geometry optimizations on the DNA bases in a dielectric continuum. Presently, this cannot be done for the MP2 calculations in the Gaussian suite of programs.

Using B3LYP/6-31++G(d,p), the vertical ionization potentials of the DNA bases were calculated in a dielectric continuum with the polarized continuum model of Tomasi et al.[6] The results are shown in Table 4. The results under the heading B3LYP/6-31++G(d,p) represent calculations on the optimized coordinates (optimized at the HF/6-31++G(d,p) level) in water ($\epsilon = 78.3$). This therefore represents the energy of the solvated ionized radical. To calculate the energy for the IP in water, one must subtract the solvation energy of the electron (which is 1.3 eV).[21] To the calculations on the bases, additional IPs are included for 5-MeC, 8-OxoGuanine, and indole in water.

The results in Table 4 are compared with recent results by LeBreton et al.[22] This group has used data from gas-phase photoelectron experiments, combined with results from self-consistent field and post self-consistent field calculations and with theoretical Gibbs free energies of hydration, to describe the aqueous ionizations of the nucleotide anions. The negative

TABLE 5: Ionization Potentials of Planar and Nonplanar Bases

	nonplanar IP		planar IP		
molec.	OVGF	(Koopmans')	OVGF	(Koopmans')	ΔE kcal/mol[a]
Thy	8.78	(9.29)	8.82	(9.29)	~0.0
Cyt	8.53	(9.07)	8.46	(9.04)	0.29
5-MeC	8.26	(8.76)	8.18	(8.71)	0.96
Ade	8.34	(8.28)	8.18	(8.15)	0.62
Gua	7.98	(8.01)	7.77	(7.83)	1.8

[a] ΔE here is the difference in energy between the nonplanar base and the planar base.

charge on the phosphate does not seem to affect the results as the authors show that the aqueous ionization threshold energy of anionic 5′dTMP (5.4 eV) is nearly the same as that of the neutral nucleoside 2-deoxythymidine (5.3 eV).

Comparing Table 2 (vacuum IPs) and Table 4 (IPs in water), one sees that in water the ordering of the IPs is the same as in a vacuum. Furthermore, the IP of 5-MeC is smaller than that of cytosine. This has important implications to the radiation chemistry of DNA. Once 5-MeC is oxidized, the cation may irreversibly deprotonate at C5−CH_3 producing the 3αH radical.[5] Such an irreversible deprotonation could in principle halt hole transfer in DNA. Therefore, one has to consider the consequences of the oxidation of 5-MeC, and not just guanine oxidation, to understand the radiation chemistry of DNA.[23]

Base Stacks and Base Pairs. It would be desirable to extend the results obtained here to compute the IPs of stacked bases. First of all, it is necessary to use extended basis sets to adequately describe π-electron distributions. Also, one should include the effects of electron correlation to accurately compute the IPs. This could be very computational demanding since one is dealing with up to 24 second-row atoms for a purine:purine stack To save CPU time, previous calculations have been done with smaller basis sets on nonoptimized structures.

Saito and coworkers have computed IPs on the DNA base stacks by using geometries based on standard bond lengths and bond angles taken from X-ray crystal data.[24] IPs were estimated from Koopmans' theorem from single-point calculations at the HF/6-31G* level. Similar results were reported by Prat et al.[25] using geometries optimized with the AMBER force field. Vertical IPs were then estimated from the Koopmans' theorem from calculations performed at the RHF/6-31G* level.

The use of Koopmans' theorem to estimate IPs was discussed by Colson et al.[1] They showed that the best results were obtained with small basis sets. The good fit between the Koopmans' values (using the 3-21G basis set) and the experimental values is believed to result from a cancellation of unaccounted errors in the electron correlation and orbital relaxation energies. The use of higher basis sets did not improve the fit of the calculated IPs to the experimental values.

To study the effects of using nonoptimized structures, one has to consider that A- or B-DNA crystal structures are not minima on the gas-phase potential energy surface. The basic difference is that in the geometry optimized structures of cytosine and guanine the −NH_2 groups are rather nonplanar. Table 5 shows the effect this has in calculating IPs.

There are several interesting features in Table 5. First of all, the IPs calculated with Koopmans' theorem seem to agree with the OVGF calculated IPs (and the experimental values) for the purines but not for the pyrimidines. Also, there are differences between the calculated IPs for the planar and nonplanar calculations. The nonplanarities are basically caused by the out-of-plane bending of the NH_2 groups. It is seen that there is only a small energy difference between the planar and nonplanar geometries.

The works mentioned above on stacked bases include calculations on guanine. The calculated IP of guanine by Saito et al. is 7.75 eV[24] and by Prat et al. is 7.72 eV[25] which is very close to the values in Table 5 for planar guanine. This is because these authors are performing calculations on planar molecules. Saito uses atomic coordinates from crystal structures, and Prat used AMBER force fields to optimize the molecular geometries. Both procedures result in planar bases.[26]

When one sees a calculated IP of guanine to be ~7.75 eV, this may cause some concern since the standard reference has the experimental value as 8.24 eV.[9] However, the experimental value was for normal guanine, not N9−CH_3 guanine. Table 2 shows that the experimental IP for N9 methylated guanine is 8.02 eV. Table 5 shows that one can accurately calculate this IP in N9 methylated guanine (7.77 eV), at least for a calculation involving the fully optimized geometry. Also, perhaps fortuitously, the Koopmans' value for the IP of guanine is very close to the experimental value.

There have also been efforts to calculate the influence of base pairing on IPs. Hutter and Clark[27] have calculated that the vertical IP of G:C is 7.16 eV and 7.74 eV for A:T. They suggest that these values are too low and so have performed linear correlations to experimental IPs, which give corrected values of 7.51 eV for G:C and 8.06 eV for A:T. There are also problems of comparing calculations of the IP on a base pair with the IP of an individual base because of the basis set superposition error. More recent calculations by Li et al.[28] have vertical IPs of 7.80 eV for G:C and 7.23 eV for A:T.

Conclusions

Table 1 lists the vertical ionization potentials for the DNA bases previously calculated at the PMP2/6-31++G(d,p) level.[3] The new results for the vertical ionization potentials for the DNA bases at the B3LYP/6-31++G(d,p) level in Table 1 are seen to be close to the PMP2 results and also close to the experimental results. The density functional theory calculations can be run at considerable savings in CPU time.

Table 4 lists the vertical ionization potentials for the DNA bases in aqueous medium at the B3LYP/6-31++G(d,p) level. These values can be compared with the previous results at the PMP2/6-31++G(d,p) level,[3] which are included in Table 4 (under the heading PMP2). In both cases, these are single-point calculations on geometries first optimized at the PCM/HF/6-31++G(d,p) level. One sees very good agreement between the two sets of calculations.

Therefore, it seems as if one can obtain reliable vertical ionization potentials for the DNA bases with the faster density functional theory calculations. Work is in progress to look at DFT calculations of the ionization potentials of a larger system including the DNA bases with explicit waters of hydration in a PCM cavity and the influence of 5-MeC in a guanine base stack.

Acknowledgment. This work is supported by PHS Grant R01 CA36810-18 awarded by the National Cancer Institute, DHHS. Helpful discussions with Leonid Gorb at Jackson State University are gratefully acknowledged.

References and Notes

(1) Colson, A. O.; Besler, B.; Sevilla, M. D. *J. Phys. Chem.* **1992**, *96*, 9787.

(2) Sevilla, M. D.; Besler, B.; Colson, A. O. *J. Phys. Chem.* **1995**, *99*, 1060.

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

(4) Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. *Int. J. Quantum Chem.* **2002**, *90*, 1547.

(5) Close, D. M. *J. Phys. Chem. B* **2003**, *107*, 864−867. There is a typo in Table 2 of this reference where the vertical IP for 5-MeC with a methyl group on N1. The IP is listed as 8.39 eV. The correct value is 8.26 eV.

(6) Barone, V.; Cossi, M.; Tomasi, J. *J. Comput. Chem.* **1998**, *19*, 404.

(7) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.11.3; Gaussian, Inc.: Pittsburgh, PA, 1998.

(8) Pascaly, M.; Yoo, J.; Barton, J. K. *J. Am. Chem. Soc.* **2002**, *124*, 9083.

(9) Hush, N. S.; Cheung, S. A. *Chem. Phys. Lett.* **1975**, *24*, 11.

(10) Urano, S.; Yang, X.; LeBreton, P. *J. Mol. Struct.* **1898**, *214*, 315.

(11) Lin, J.; Yu, C.; Peng, S.; Akiyama, I.; Li, K.; Lee, L. K.; Lebreton, P. R. *J. Am. Chem. Soc.* **1980**, *102*, 4627.

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

(13) Hagar, J.; Ivanco, M.; Smith, M. A.; Wallace, S. C. *Chem. Phys.* **1986**, *105*, 397.

(14) In a paper by Wetmore, S. D.; Boyd, R. J.; Eriksoon, L. A. (*Chem. Phys. Lett*. **2000**, *322*, 129.) one reads, "It should be noted that although model systems were implemented for the nucleotides, where the sugar group was replaced with hydrogens, then use of large models is not expected to alter the results. In fact, replacing the hydrogens at N1 in uracil by a methyl group is found to increase the EA by 0.016 eV. The IP's are expected to change even less with a similar substitution."

(15) Padva, A.; O'Donnell, T. J.; LeBreton, P. R. *Chem. Phys. Lett.* **1976**, *41*, 278.

(16) Yu, C.; Peng, S.; Akiyama, I.; Lin, J.; LeBreton, P. R. *J. Am. Chem. Soc.* **1978**, *100*, 2303.

(17) Li, X.; Cai, Z.; Sevilla, M. D. say "the vertical EA's of the bases cannot be appropriately corrected for ZPE's since they are not in their relaxed states as anion radicals" (*J. Phys. Chem A* **2002** *106*, 1596). Sevilla also says (private correspondence) the same should apply to IPs. While a ZPE correction to a vertical state does seem to be in the right direction, it does not provide a full correction.

(18) Nikogosyan, D. N. *Int. J. Radiat. Biol*. **1990**, *57*, 233.

(19) Candeias, L. P.; Steenken, S. *J. Am. Chem. Soc.* **1992**, *114*, 699.

(20) Bertran, J.; Oliva, A.; Rodriguez-Santiago, L.; Sodupe, M. *J. Am. Chem. Soc.* **1998**, *120*, 8159.

(21) Bernas, A.; Grand, D.; Amouyal, E. *J. Phys. Chem.* **1980**, *84*, 1262.

(22) Fernando, H.; Papadantonakis, G. A.; Kim, N. S.; LeBreton, P. R. *Proc. Natl. Acad. Sci.* **1998**, *95*, 5550.

(23) Kanvah and Schuster (*J. Am. Chem. Soc*. **2004**, *126*, 7341) have recently questioned whether 5-MeC substituted in DNA can affect radical cation hopping through DNA. Their measurements indicate that the answer may be no. The markers in their experiments are strand cleavage at G sites. Oxidation of 5-MeC provides a stable allyl radical that would not lead to a strand break, so it is not clear how these results test the hypothesis that 5-MeC is a locus of oxidative damage in DNA.

(24) Sugiyama, H.; Saito, I. *J. Am. Chem. Soc.* **1996**, *118*, 7063.

(25) Prat, F.; Houk, K. N.; Foote, C. S. *J. Am. Chem. Soc.* **1998**, *120*, 845.

(26) Reviews of the X-ray crystallographic structures of the nucleic acid bases have led to the conclusion that "The similarity between the dimensions of the standard residues and the average dimensions observed in the crystal structures suggest that the assumption of exact base planarity is justified" (Taylor, R.; Kennard, O. *J. Am. Chem. Soc*. **1982**, *104*, 3209). Widely used empirical force fields such as AMBER and CHARMM also assume planar and rather rigid amino groups. New parametrization of the Cornell force field allows for partial sp^3 hybridization of the amino nitrogen (Rajacek, F.; Kubar, T.; Hobza, P. *J. Comput. Chem*. **2003**, *24*, 1891.).

(27) Hutter, M.; Clark, T. *J. Am. Chem. Soc*. **1996**, *118*, 7574.

(28) Li, X.; Cai, Z.; Sevilla, M. D. *J. Phys. Chem. A* **2002**, *106*, 9345.