Ab initio base-pairing energies of uracil and 5-hydroxyuracil with standard DNA bases at the BSSE-free DFT and MP2 theory levels†

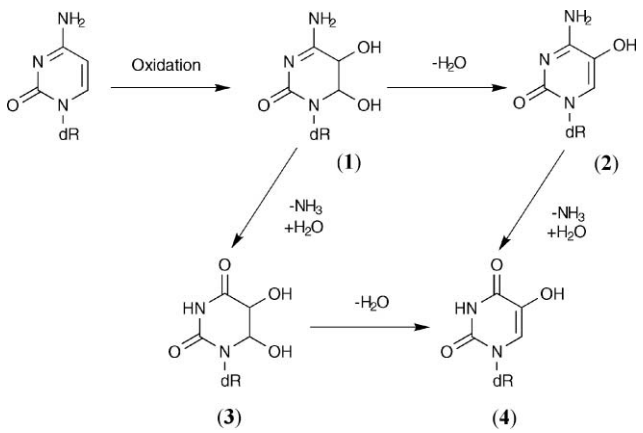
David E. Volk, Varatharasa Thiviyanathan, Anoma Somasunderam and David G. Gorenstein*

Received 15th February 2006, Accepted 10th March 2006
First published as an Advance Article on the web 30th March 2006
DOI: 10.1039/b602263d

Oxidized cytosine product 5-hydroxyuracil has been shown to be the major chemical precursor for the GC to AT transition, the most frequent substitution mutation observed in aerobic organisms. We have calculated the interaction energy of base-pair formation involving uracil or 5-hydroxyuracil, which is formed in cells by oxidative deamination of cytosine, bound to any of the natural DNA bases, A, C, G, and T, and discuss the effects of the hydroxyl group in this respect. The base-pair geometries and energies were calculated using the 6-311G(dp) basis set under four conditions: using density functional theory (DFT) without out basis set super-position error (BSSE) correction, using DFT with BSSE correction of geometries and energies, using Møller–Plesset second order perturbation theory (MP2) without BSSE correction, and using MP2 with BSSE geometry and energy correction. We find that the hydroxyl group of 5-HO-U (relative to U) has little effect on the base-pairs with A, C or one conformation of T, while making a substantial energy difference in base-pairs involving G or a different conformation of T. For most of the complexes studied, the BSSE-corrected energies at the DFT and MP2 levels of theory agreed to within 0.5 kcal.

Introduction

Exposure to reactive oxidation species and subsequent DNA damage has been linked to cancer, aging, rheumatoid arthritis and other diseases.[1,2] Oxidized cytosine products have been shown to be major chemical precursors in DNA for GC to AT transition mutations,[3,4,5] the most abundant base substitution mutation observed in aerobic organisms.[6,7,8] Cytosine oxidation gives rise to a number of products (Scheme 1),[9] including 5,6-dihydroxy-5,6-dihydrocytosine (**1**), which can be dehydrated to form 4-amino-5-hydroxypyrimidine-2-one (5-OH-C) (**2**) or deaminated to form 5,6-dihydroxy-5,6-dihydrouracil (**3**). Dehydration and deamination of (**1**) yields 5-hydroxy-2,4(1*H*,3*H*)-pyrimidinedione (4), (called 5-hydroxyuracil (5-OH-U) when incorporated in nucleotides or DNA) the most mutagenic product of cytosine oxidation. Both (**3**) and (**4**) have been detected in human cells after oxidative damage to DNA at levels comparable to those of 8-oxoguanine, the most frequently observed oxidized purine product in DNA.[10] In humans, at least four enzymes excise 5-OH-U from damaged DNA, suggesting that 5-OH-U is produced in significant amounts in the human genome that requires multiple enzymes for its repair.[11,12] Their presence indicates that 5-hydroxyuracil is produced in significant amounts within cells.



Scheme 1 Oxidative deamination pathways converting cytosine to 5-hydroxyuracil, 5-OH-U (**4**).

The chemical structure of 5-hydroxyuracil is very similar to that of thymidine, suggesting that it would form a normal Watson–Crick (WC) base-pair with adenine (Fig. 1). Although (**4**) can undergo a keto–enol tautomerism *via* a 1,3 hydrogen shift mechanism, the enol form is expected to be more stable because the C5–C6 double bond is conjugated. UV absorption studies showed that 5-OH-U and 5-OH-C retain the enol rather than the keto configuration.[13] Although there is no evidence that C gets mispaired with 5-OH-U *in-vivo*, G, T, and A residues are incorporated, at different frequencies, opposite to 5-OH-U by mammalian DNA polymerases during DNA replication.[14] To understand the base-pairing preferences of 5-OH-U during DNA replication, we are studying the interaction energy of base-pair formation of 5-OH-U with other DNA bases.

Ab initio methods have been widely used to study DNA base stacking and base-pair interactions, and have been the subject of several excellent reviews.[15,16,17] More recent reports concerning AT base-pair patterns,[18] uracil flexibility,[19] tandem GU base-pair patterns (studied by NMR[20] or quantum mechanics[21]) or more general base-pair interaction energies[22] may interest the reader.

Sealy Center for Structural Biology and Molecular Biophysics and the Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, TX, USA 77555-1157. E-mail: dggorens@utmb.edu; Fax: 01 409 747-6850; Tel: 01 409 747-6800

† Electronic supplementary information (ESI) available: PDB coordinates of all complexes at the MP2-BSSE level. See DOI: 10.1039/b602263d

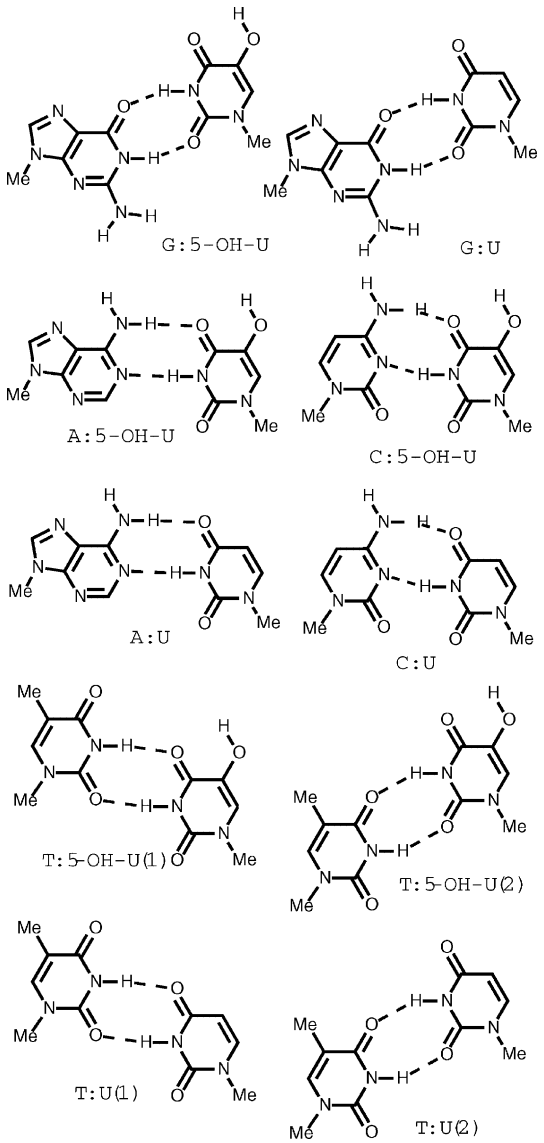


Fig. 1 Geometric arrangements of the base-pairs studied.

Recently, we reported that 5-OH-U forms stable base-pairs with the four standard DNA bases, A, C, G, and T, based on NMR spectroscopy, UV-melting experiments, and low level *ab initio* studies.[23] The reported ordering of base-pair stabilities, namely G : 5-OH-U > A : 5-OH-U > C : 5-OH-U > T : 5-OH-U, is analogous to the ordering of base-pairs involving thymine when in comparable conformations.[17] The NMR data indicated that for all four 5-OH-U mismatches studied, all of the bases remain in the normal *anti* configuration about the glycosidic bond, the imino protons were protected from solvent exchange by hydrogen bonding, and that the base-pair between C and 5-OH-U was less stable, or more dynamic, than the base-pairs formed between 5-OH-U and A, G, or T. In the C: 5-OH-U base-pair the NMR spectral line broadening of the H3-imino hydrogen was rationalized by the inter-conversion between two equally stable, twisted conformations. The two bases were twisted ±40° relative to each other in order to relieve the electrostatic repulsion between the cytosine O2 and 5-OH-U O2 carbonyl oxygen atoms. In this report, using higher theory levels, we contrast the interaction energies of uracil and 5-OH-U. Because NMR data[17] indicated that all of the base-pairs remained in the *anti* conformation, we have not calculated the base-pair energies for the *syn* conformations in which the 5-hydroxy group would be involved in the base-pairing hydrogen bonds. Although the 5-hydroxy substitution would influence base stacking interactions and may form water or cation bridges, these subjects were not investigated in this study.

Results and discussion

The energies and geometries of the base-pairs of 5-hydroxy-uracil (5-OH-U) with the standard DNA bases were calculated at the HF and DFT levels of theory using the 6-311G(d,p) basis set (data not shown) starting with bases attached to ribose and methyl-capped phosphate moieties (*i.e.* methyl-capped nucleotides). These calculations were repeated using bases in which the N1 (pyrimidines) or N9 (purines) nitrogen atoms were capped with a methyl group in place of the C1′ carbon atom. Although removal of the phosphate and sugar moieties significantly changed the absolute energies of the bases and their complexes, the calculations indicated a constant interaction energy change of −0.2 kcal mol^{-1} at the DFT level, presumably due to the large distance of the these atoms from the H-bonding atoms. Therefore, all subsequent calculations were performed using the 6-311G(d,p) basis set and the N-methylated forms of the bases (Fig. 1).

Among the standard DNA bases, 5-OH-U forms the strongest base-pair with guanine (Table 1) at all levels tested, and the bond distances determined (Table 2) in the MP2-BSSE-free calculation, 1.883 Å (O6–H3) and 1.860 (H1–O2)Å, are among the shortest distances observed. The interaction energy of the G: 5-OH-U base-pair is −13.8 and −16.7 kcal mol^{-1} at the HF and DFT(B3LYP) levels of theory, respectively, when calculated using the full nucleotides. This change of −2.9 kcal mol^{-1} between the

Table 1 Hydrogen bond strengths[a] (kcal mol^{-1}) of uracil and 5-hydroxy-uracil bases paired with standard DNA bases and comparison to standard Watson–Crick (WC) GC and AT base-pairs

Base-pair	ΔE^{DFT}	$\Delta E^{DFT,BSSE}$	ΔE^{MP2}	$\Delta E^{MP2,BSSE}$
G · C WC	−28.3	−27.8	−28.0	−24.9
G · 5-OH-U	−16.9	−15.8	−18.0	−14.9
G · U	−15.4	−14.1	−16.5	−13.4
A · 5-OH-U	−15.2	−13.4	−17.5	−13.3
A · U	−15.2	−13.2	−17.2	−12.9
A · T WC	−15.1	−13.1	−17.2	−12.9
C · 5-OH-U	−13.4	−12.1	−15.4	−12.2
C · U	−13.7	−12.4	−15.3	−12.0
T · 5-OH-U(2)	−13.2	−11.5	−14.4	−10.9
T · U(2)	−12.1	−10.2	−13.3	−9.8
T · 5-OH-U(1)	−12.5	−10.6	−13.4	−10.2
T · U(1)	−12.3	−10.5	−13.4	−9.9

[a] All present calculations used the 6-311G(d,p) basis set and the geometries were optimized at the theory level indicated. DFT was performed with B3LYP, and the BSSE geometry optimizations used the counterpoise method as described in experimental methods. The BSSE-free optimized interaction energies include deformation energies, and they are therefore equal to the complexation energies of the dimers. For the BSSE-contaminated calculations (ΔE^{DFT} and ΔE^{EMP2}), neither the geometry nor the energy have been corrected for BSSE. For the BSSE-free calculations ($\Delta E^{DFT,BSSE}$ and $\Delta E^{EMP2,BSSE}$), both the geometries and the energies have been corrected for BSSE. No thermal or zero point energy corrections have been made.

Table 2 Hydrogen bond distances of structures shown in Fig. 1 calculated at different levels of theory

Base-pair	DFT		DFT(BSSE-free)		MP2		MP2(BSSE-free)	
	r_1[Å]	r_2 [Å]	r_1 [Å]	r_2 [Å]	r_1 [Å]	r_2 [Å]	r_1 [Å]	r_2 [Å]
G · C WC[b]	[a]1.77	1.91	[a]1.79	1.94	[a]1.78	1.89	[a]1.86	1.98
G · 5-OH-U	[a]1.79	[a]1.80	[a]1.81	[a]1.83	[a]1.77	[a]1.79	[a]1.86	[a]1.88
G · U	[a]1.81	[a]1.83	[a]1.84	[a]1.86	[a]1.79	[a]1.84	[a]1.87	[a]1.94
A · 5-OH-U	1.80	[a]1.96	1.84	[a]1.98	1.77	[a]1.98	1.87	[a]2.04
A · U	1.82	[a]1.93	1.87	[a]1.94	1.80	[a]1.95	1.90	[a]2.01
A · T WC	1.82	[a]1.93	1.87	[a]1.94	1.80	[a]1.95	1.90	[a]2.01
C · 5-OH-U	[a]1.88	1.94	[a]1.91	1.99	1.85	[a]1.90	1.99	[a]1.99
C · U	[a]1.84	1.95	[a]1.87	2.01	[a]1.87	1.88	[a]1.95	2.01
T · 5-OH-U(2)	[a]1.82	[a]1.85	[a]1.86	[a]1.89	[a]1.82	[a]1.84	[a]1.91	[a]1.94
T · U(2)	[a]1.86	[a]1.87	[a]1.90	[a]1.90	[a]1.86	[a]1.86	[a]1.95	[a]1.95
T · 5-OH-U(1)	[a]1.84	[a]1.89	[a]1.88	[a]1.92	[a]1.83	[a]1.88	[a]1.92	[a]1.98
T · U(1)	[a]1.85	[a]1.87	[a]1.89	[a]1.90	[a]1.84	[a]1.86	[a]1.94	[a]1.96

[a] Indicates a hydrogen bond involving a carbonyl group. [b] $r3 = 1.92$, 1.94, 1.93 and 2.01* at the DFT, DFT(BSSSE-free), MP2 and MP2(BSSE-free) levels, respectively.

HF and DFT calculations is typical of those observed (ranging from −2.9 to −4.1 kcal mol^{-1}) for all of the five complexes calculated using the full nucleotides. The strength of the G : 5-OH-U interaction at the MP2 BSSE-free level, −14.9 kcal mol^{-1}, is 10 kcal mol^{-1} less stable than the interaction in the Watson–Crick (WC) G : C base-pair but 1.5 kcal mol^{-1} more stable than the G : U mismatch. Thus, the 5-hydroxy substitution stabilizes the base-pair by 1.5 kcal mol^{-1}.

It is biologically significant that 5-OH-U forms the most stable base-pair with G, because 5-OH-U arises from the oxidation of a cytosine. Thus, upon DNA strand separation and replication the G formerly paired with 5-OH-U will now pair with a C, and this C will then correctly be incorporated into the daughter strand. Likewise, the 5-OH-U will most often pair with a G, which will also correctly be incorporated into the other daughter DNA strand. In this situation, no mutations will be propagated to the subsequent generations.

The A : 5-OH-U base-pair is the next most stable base-pair involving 5-OH-U at all levels studied, having −13.3 kcal mol^{-1} of interaction energy in the MP2-BSSE-free calculation. One of the hydrogen bonds is formed between the adenine H6 hydrogen and the 5-OH-U O4 oxygen ($r1 = 2.044$ Å). The second hydrogen bond is formed between adenine N1 and the 5-OH-U H3 hydrogen atom ($r2 = 1.869$ Å). The A : 5-OH-U BSSE-free MP2 interaction energy, −13.3 kcal mol^{-1}, is only 0.4 kcal mol^{-1} less than the interaction energies of the A : U and A : T base-pairs (−12.9 kcal mol^{-1}), indicating that the uracil 5-substituent (H, OH or Me) has little effect on the interaction energy. In contrast, the 5-hydroxy substituent makes a markedly larger difference (−1.5 kcal mol^{-1}) when binding with guanine. Because the A : 5-OH-U base-pair is only 1.6 kcal mol^{-1} less stable than the G : 5-OH-U base-pair, some formation of A : 5-OH-U would be expected within the cell. The incorporation of this adenine opposite to the 5-OH-U would lead to the incorporation of a T in place of the 5-OH-U in daughter DNA strands. Thus, formation of the A : 5-OH-U base-pair, if not excised by glycosylases, results in the G : C to A : T transition mutation observed in cells.

At the MP2-BSSE-free level, the C : 5-OH-U base-pairs are less stable than the G : 5-OH-U and A : 5-OH-U base-pairs by 2.7 and 1.1 kcal mol^{-1}, respectively. However, in the two equally energetic C : 5-OH-U base-pairs, the two bases are twisted 38 degrees relative to each other in order to minimize the electrostatic repulsion between the O2 oxygen atoms of C and 5-OH-U. In the context of a DNA duplex, such base twisting would greatly distort surrounding base-pairs making the overall interaction significantly less stable than calculated here. Indeed, both UV-monitored melting experiments and NMR spectroscopy have shown that among the four DNA bases, cytosine makes the least stable complex with 5-OH-U in solution. Because of the twisted nature of these base-pairs, a direct comparison of the cytosine base-pairs *versus* the A, G and T base-pairs with uracil or 5-hydroxy uracil is not appropriate. However, one can compare the MP2-BSSE-free interaction energies of the C : 5-OH-U (−12.2 kcal mol^{-1}) and C : U (−12.0 kcal mol^{-1}) base-pairs. This small energy difference between the 5-substituents, (−0.2 kcal mol^{-1}) is similar to the change observed in the A : 5-OH-U base-pairs (−0.4 kcal mol^{-1}) but contrasts with the larger 1.5 kcal mol^{-1} change observed in the G base-pairs.

Thymine can possibly form two base-pairs with 5-OH-U (Fig. 1). In the most stable arrangement (conformation 2), the thymine H3 hydrogen forms a hydrogen bond with the O2 oxygen of 5-OH-U residue ($r1 = 1.909$ Å), and the H3 hydrogen of 5-OH-U forms a hydrogen bond with the O4 oxygen of T ($r2 = 1.941$ Å). The MP2-BSSE-free interaction energy of this conformation is −10.9 kcal mol^{-1}, which is significantly less stable than the G : 5-OH-U (−14.9 kcal mol^{-1}) and A : 5-OH-U (−13.3) base-pairs.

In conformation 1 (Fig. 1), the H3 hydrogen of T forms a hydrogen bond with the O4 oxygen of 5-OH-U ($r2 = 1.977$) and the H3 hydrogen of 5-OH-U forms a hydrogen-bond with the T O2 oxygen atom ($r1 = 1.923$). The T : 5-OH-U(1) MP2-BSSE-free interaction energy (−10.2 kcal mol^{-1}) is 0.7 kcal mol^{-1} less stable than that for conformation 2. The T : 5-OH-U(1) and TU(1) base-pair energy difference is negligible (0.2 kcal mol^{-1}). This is similar to the small changes observed for the 5-OH-U and U base-pairs with A and C. Interestingly in the other conformation, T : 5-OH-U(2) is 1.1 kcal mol^{-1} more stable than T : U(2).

Of the structures studied here, G : 5-OH-U and conformation 2 of T : 5-OH-U are significantly more stable than the corresponding base-pairs formed with U. In the G-containing base-pairs, the G : 5-OH-U base-pair is 1.5 kcal mol^{-1} more stable than the G : U base-pair. For the T : 5-OH-U(2) base-pair, the 5-hydroxy substituent stabilizes the interaction by 1.1 kcal mol^{-1}. In both

of these structures, hydrogen bonds are formed with the uracil O2 oxygen atom while no hydrogen bonds involve the uracil O4 oxygen atom. In the remaining base-pairs containing A, C or T in conformation 1, the 5-hydroxy substitution only stabilizes the base-pairs by 0.4, 0.2 and 0.3 kcal mol^{-1}, respectively. Each of these structures contain a hydrogen bond to the uracil's O4 oxygen rather than to the O2 oxygen. Thus, the intra-molecular hydrogen bonding in 5-OH-U between the hydroxyl proton and O4 reduces the hydrogen bond strength between O4 and the proton of the pairing base. However, in both G : 5-OH-U and T : 5-OH-U(2) the O4 oxygen of 5-OH-U is not involved in the inter-molecular hydrogen bond, and therefore the 5-OH-U O4 : OH intra-molecular hydrogen bond will further stabilize the molecule. In addition, the hydroxyl oxygen adds an extra repulsive interaction with the hydrogen-bonding oxygen of the other base when the 5-OH-U O4 oxygen atom is involved in the hydrogen bond.

Effects of theory level

Under all methods used, the ordering of base-pair energies remains the same, with the sole exception being the relative energies of the C : 5-OH-U and CU base-pairs, which differ by 0.3 kcal mol^{-1} at most. The DFT calculations without BSSE geometry corrections typically over-estimate the stabilities of the base-pairs by 1.1–2.0 kcal mol^{-1}. However, the interaction energy of the G:C base-pair at the DFT level is over-estimated by only 0.5 kcal mol^{-1}. When using DFT methods, the BSSE geometry and energy corrections make a larger difference for the U or 5-OH-U base-pairs involving either A (1.8 to 2.0 kcal mol^{-1}) or T (1.7 to 1.9 kcal mol^{-1}), than they did for the G (1.1–1.3 kcal mol^{-1}) or C (1.3 kcal mol^{-1}) base-pairs. The DFT-BSSE-free geometry optimizations typically lengthen the hydrogen bonds by about 0.03 Å, but the bond increases do range from 0.01–0.06 Å. Interestingly, the hydrogen bonds that involve a nitrogen atom are lengthened more by the DFT-BSSE correction compared to hydrogen bonds that involve an oxygen atom. Thus, the bonds formed between the H3 amido proton of U or 5-OH-U and the imino nitrogen of either A or C increase by 0.04 to 0.06 Å, while the oxygen-containing hydrogen bonds are increased by 0.02–0.04 Å.

Effects of BSSE correction to MP2 calculations

The effects of the BSSE correction are clearly much larger for the MP2 calculations compared to the DFT calculations. For the BSSE-contaminated calculations (Table 1: energy columns 1 and 3), neither the geometry nor the energies reported have been BSSE-corrected. For these entries, the energy was calculated as the energy of the hydrogen-bonded base-pair minus the energy of the isolated components. Typically, the BSSE-free geometry optimization at the MP2 level lowered the interaction energy by about 3.0 kcal mol^{-1}. The corresponding energy corrections at the DFT level mostly ranged from 1 to 2 kcal mol^{-1}. A greater bond lengthening can also be observed in the MP2-BSSE-free calculations. The BSSE-corrected geometries obtained at the MP2 level have hydrogen bonds lengths that are typically 0.10 Å larger than those observed in the MP2 BSSE-contaminated base-pair structures. At the DFT level, these changes were typically one half as large.

While the BSSE-contaminated-MP2 calculated energies were significantly lower than the BSSE-contaminated-DFT energies (except for GC), the energy changes observed after BSSE-free geometry optimization were significantly larger for the MP2 calculations than for the DFT calculations. Thus, although the uncorrected DFT and MP2 calculated energies differed by up to 2.3 kcal mol^{-1}, after BSSE-free geometry corrections, the DFT and MP2 energies generally differed by about 0.5 kcal mol^{-1} or less. The differences were slightly larger for the GC WC and G : 5-OH-U base-pairs.

Experimental

All calculations were performed with Gaussian98.[24] Calculations at the RHF level were performed on Silicon Graphics Inc. Fuel or Octane 2 computers, calculations at the DFT (B3LYP) level were performed on a 12-node Sun64-SVR-4 computer, and the MP2/BSSE calculations were performed on the NCSA IBM P690 or Sun64-SVR-4 computers. The full nucleotides consisted of the base and a ribose sugar in which the 3′- and 5′ oxygens were each capped by a methyl group. The trimmed nucleotides consisted of the appropriate base in which the N1 or N9 nitrogen was capped with a methyl group. Geometry and energy corrections for basis set superposition error (BSSE) were first calculated at the DFT theory level and subsequently at the MP2 level of theory, using the Counterpoise Correction method of Boys and Bernardi[25] implemented in Gaussian98 as outlined by Simon, Duran and Dannenberg.[26] The 6-311G(d,p) basis set was used for all data reported here. Thermal corrections and zero point energy corrections were not performed. At this level of theory, the use of larger basis sets was deemed to be too expensive, and the dispersion error due to the use of moderately-sized basis sets has been estimated to be −2.0 to −3.0 kcal mol^{-1}.[27]

Conclusions

The oxidative product 5-OH-U forms base-pairs with the natural DNA bases G, A, C and T with MP2-BSSE-free interaction energies of −14.9, −13.3, −12.2 and −10.9 kcal mol^{-1}, respectively. This ordering is similar to that observed in analogous base-pairs with T in place of 5-OH-U,[17] although due to the differing theoretical treatments, the results cannot be directly compared. The effect of the 5-hydroxy substitution, *i.e.* 5-OH-U *vs* U, is to stabilize the base-pairs with G and T (in conformation 2) with energy changes of −1.5 and −1.1 kcal mol^{-1}, respectively, while having a much smaller effect on the energies of base-pairs with A, C or T in conformation 1, where the base-pairs energy changes are only −0.4, −0.2 and −0.3 kcal mol^{-1}, respectively, The BSSE-free DFT and MP2 energies generally agreed within about 0.5 kcal mol^{-1}, with the exceptions of the G : 5-OH-U (0.9 kcal mol^{-1}) and GC WC (2.9 kcal mol^{-1}) base-pairs.

Acknowledgements

This work was supported by grants from the NIEHS (ES06676), the Welch Foundation (H-1296) and the National Center for Supercomputing Applications (MCB050002N). We thank T. David Power for helpful discussions.

References

1 G. Mitra, G. T. Pauly, R. Kumar, G. K. Pei, S. H. Hughes, R. C. Moschel and M. Barbacid, *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 8650.
2 R. Goth and M. F. Rajewsky, *Proc. Natl. Acad. Sci. USA*, 1974, **71**, 639.
3 A. A. Purmal, Y. W. Kow and S. S. Wallace, *Nucleic Acids Res.*, 1994, **22**, 3930–3935.
4 D. I. Feig, L. C. Sowers and L. A. Loeb, *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 6609–6613.
5 D. A. Kreutzer and J. M. Essigmann, *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 3578–3582.
6 R. M. Schaaper, B. N. Danforth and B. W. Glickman, *J. Mol. Biol.*, 1986, **189**, 273–284.
7 E. C. Moraes, S. M. Keyse and R. W. Tyrrell, *Carcinogenesis*, 1990, **11**, 283.
8 L. K. Tkeshelashvili, T. Mcbride, K. Spence and L. A. Loeb, *J. Biol. Chem.*, 1991, **266**, 6401.
9 M. Dizdaroglu, E. Holwitt, M. P. Hagan and W. F. Blakey, *Biochem. J.*, 1986, **235**, 531–536.
10 J. R. Wagner, C. C. Hu and B. N. Ames, *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 3380–3384.
11 T. K. Hazra, T. Izumi, Y. W. Kow and S. Mitra, *Carcinogenesis*, 2003, **24**, 155–157.
12 T. K. Hazra, T. Izumi, I. Boldogh, B. Imhoff, Y. W. Kow, P. Jaruga, M. Dizdaroglu and S. Mitra, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 3523–3528.
13 R. C. Moschel and E. J. Behrman, *J. Org. Chem.*, 1974, **39**, 1983–1989.
14 A. Vaisman and R. Woodgate, *EMBO J.*, 2001, **20**, 6520–6529.
15 P. Hobza and J. Šponer, *Chem. Rev.*, 1999, **99**, 3247–3276.
16 E. T. Kool, J. C. Morales and K. M. Guckian, *Angew. Chem., Int. Ed.*, 2000, **39**, 990–1009.
17 J. Šponer, J. Leszczynski and P. Hobza, *Biopolymers*, 2001/2002, **61**, 3–31.
18 E. S. Kryachko and J. R. Sabin, *Int. J. Quantum Chem.*, 2003, **91**, 695–710.
19 O. V. Shishkin, L. Gorb, A. V. Luzanov, M. Elstner, S. Suhai and J. Leszczynski, *J. Mol. Struct. (Theochem)*, 2003, **625**, 295–303.
20 X. Chen, J. A. McDowell, R. Kierzek, T. R. Krugh and D. H. Turner, *Biochemistry*, 2000, **39**, 8970–8982.
21 Y. Pan, U. D. Priyakumar and A. D. MacKerell, Jr., *Biochemistry*, 2005, **44**, 1433–1443.
22 J. Šponer, P. Jurečka and P. Hobza, *J. Am. Chem. Soc.*, 2004, **126**, 10142–10151.
23 V. Thiviyanathan, A. Somasunderam, D. E. Volk and D. G. Gorenstein, *Chem. Commun.*, 2005, 400–402.
24 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle and J. A. Pople, *GAUSSIAN 98 (Revision A.11)*, Gaussian, Inc., Pittsburgh, PA, 2001.
25 S. F. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553–566.
26 S. Simon, M. Duran and J. J. Dannenberg, *J. Chem. Phys.*, 1996, **105**, 11024–11031.
27 J. Šponer and P. Hobza, *J. Phys. Chem.*, 2000, **104**, 4592–4597.