

# Thermodynamics of 5-bromouracil tautomerisation from first-principles molecular dynamics simulations

*Leo F. Holroyd,<sup>[a]</sup> Michael Bühl,<sup>[a]</sup> Marie-Pierre Gaigeot<sup>[b]</sup> and Tanja van Mourik<sup>\*[a]</sup>*

[a] University of St Andrews, EaStCHEM School of Chemistry, North Haugh, St Andrews, Fife, KY16 9ST, United Kingdom. [b] LAMBE UMR 8587, Université d'Evry val d'Essonne, Université Paris-Saclay, Blvd F. Mitterrand, Bat Maupertuis, 91025 Evry, France.

## **Corresponding Author**

\* email: [tanja.vanmourik@st-andrews.ac.uk](mailto:tanja.vanmourik@st-andrews.ac.uk).

## **Abstract**

We modelled the driving force for aqueous keto-to-enol tautomerisation of 5-bromouracil, a mutagenic thymine analogue, by first-principles molecular dynamics simulations with thermodynamic integration. Using interatomic distance constraints to model the water-assisted (de)protonation of 5-bromouracil in a periodic water box, we show that the free energy for its enolisation is lower than that of the parent compound, uracil, by around 3.0 kcal/mol (BLYP-D2 level), enough to significantly alter the relative tautomeric ratios. Assuming the energetic difference also holds in the cell, this finding is evidence for the “rare tautomer” hypothesis of 5-bromouracil mutagenicity (and, possibly, that of other base analogues).

**KEYWORDS** density functional theory, nucleobases, solvation effects, mutagenesis, CPMD, bromouracil.

## 1. Introduction

5-Bromouracil (BrU) is a structural analogue of thymine in which the methyl group at the 5-position is replaced by bromine (see Figure 1). BrU can be incorporated into DNA in place of thymine,<sup>1</sup> and is a well-known mutagen owing to its ambivalent base-pairing ability with either adenine or guanine.<sup>2</sup> Following a suggestion by Watson and Crick,<sup>3</sup> Freese<sup>4</sup> and Topal and Fresco<sup>5</sup> developed the argument that this behaviour was due to the “rare” enol tautomer of BrU – presumably stabilised by the bromine substituent – which has the appropriate structure to mispair with guanine in pseudo-Watson–Crick geometry via three hydrogen bonds (see Figure 2). Measurements of the tautomeric constants of BrU and its parent compound uracil (U) in solution appeared to validate this view,<sup>6</sup> as did other experiments.<sup>7</sup> We note that both U and BrU have a large number of tautomers, which have been studied computationally before,<sup>8</sup> with the canonical diketo form shown on the left in Figure 1 being found to be the most stable for both U and BrU. In this paper we focus on the particular enol tautomer shown in Figure 1, as this is the tautomer relevant to the mutagenicity of BrU. This is the tautomer referred to as U3 in Ref.<sup>8g</sup>, which is slightly less stable than the U2 enol tautomer.

**[Insert Figure 1 here]**

**Figure 1.** (Br)U (left) and their enol forms (right).

The canonical keto tautomer of BrU can also form a mismatched base pair, in which BrU and guanine adopt a “wobble pair” arrangement stabilised by two hydrogen bonds (Figure 2). This geometry has been experimentally observed in both the crystal phase<sup>9</sup> and solution,<sup>10</sup> where, in the latter case, it is in pH-dependent equilibrium with an ionised mispair that has the same stereochemistry as the enolised mispair. However, notwithstanding their existence in vitro, the

involvement of non-Watson–Crick intermediates in mutagenesis is called into question by the fact that shape complementarity plays a key role in pairing fidelity. In fact, faithful replication can be maintained by base isosteres without any hydrogen-bonding capability at all.<sup>11</sup> Recently, NMR rotation dispersion experiments provided evidence for spontaneous transition from wobble base pairs to Watson–Crick-like mismatches stabilised by tautomerisation and enolisation,<sup>12</sup> further casting doubt upon the involvement of non-Watson–Crick base pairs. Furthermore, it has been demonstrated that the mutagenicity and incorporation specificity of base analogues are related to the abundance of their rare tautomers.<sup>13</sup> The activity of DNA polymerases and repair enzymes is sensitive to base-pair geometry. Thus, replication is stereochemically as well as thermodynamically regulated.

**[Insert Figure 2 here]**

**Figure 2.** Base pairs involving U or BrU: The canonical A-(BrU)(keto) base pair, the G-(BrU)(enol) mispair and the G-(Br)U wobble mispair. X is H (in U) or Br (in BrU). U = uracil; G = guanine; A = adenine.

Despite these arguments, the “rare tautomer” hypothesis for BrU mutagenicity remains unverified. It has been challenged on theoretical grounds by Orozco *et al.*,<sup>14</sup> who calculated that the enol forms of BrU were unstable in the gas and aqueous phases as well as in DNA, and similar arguments were later advanced by Hobza *et al.*<sup>8c</sup> A limitation of these studies is that they employed either a continuum solvent model of water or monohydration, neither of which can fully capture the energetics of BrU in systems approaching the bulk water limit, or described bulk solvation through Monte-Carlo techniques with empirical force fields,<sup>14</sup> with unknown accuracy for the special problem at hand. Van Mourik *et al.*<sup>15</sup> subsequently calculated at the density functional theory (DFT) level that for BrU (but not U) the 4-enol tautomer is absolutely favoured over the keto tautomer in clusters of 50 or 100 explicit water molecules. Other studies

investigating the role of water in tautomer conversion of uracil and uracil derivatives (using either continuum solvent or a microhydrated environment) generally found that water affected the relative stability and lowered the tautomerisation barriers.<sup>8d-h, 16</sup>

It thus appears that water has a highly stabilising effect on the apparently mutagenic “rare” tautomer of BrU. Although water is naturally present in the cellular environment of the DNA helix, this has properties that are significantly different from bulk water.<sup>17</sup> However, DNA bases could be tautomerised somewhere else, such as on a single-stranded region on DNA or somewhere else in water-rich cytosol, and then be incorporated into DNA.<sup>18</sup> Furthermore, Leszczynski *et al.* concluded that inclusion of dynamical effects is essential for studying water-assisted proton transfer of DNA bases.<sup>18</sup> In this paper, to elucidate the thermodynamic driving force for the keto–enol tautomerisation, we report simulations of the water-induced proton transfer in both BrU and U based on constrained Car–Parrinello Molecular Dynamics (CPMD) free energy calculations,<sup>19</sup> a semi-quantum form of molecular dynamics (MD) combining an electronic structure derived from first principles (Kohn–Sham orbitals) with classical nuclei. To the best of our knowledge, no dynamical study of the enolisation of aqueous BrU – that is, proton transfer from N3 to O4 – has previously been performed. In this work we focus only on the keto→enol tautomerisation of BrU. Another aspect of the nature of BrU tautomerisation involves the G-BrU→GBrU(enol) and G-BrU→G(enol)BrU processes, which have been extensively investigated by others.<sup>20</sup>

## 2. Methodology

### 2.1 General methodology

Proton transfers are rare events that can be very difficult to sample over the time-scale of short DFT-based molecular dynamics. Therefore, to model the N3-to-O4 proton transfer in (Br)U, we divided this process into three distinct steps: deprotonation at N3, intra-solvent proton-hopping, and reprotonation at O4 (see Figure 1 for labelling). The first and last of these steps were modelled by a set of constrained MD trajectories, in which the constraints were the N3–H3 or O4–H4 distances, respectively, where H3(4) is the nearest interacting H atom to the respective nucleobase atom (both constraints labelled  $\zeta$ ). Because the free energy is a state function, the driving forces that we compute are in principle independent of the chosen reaction coordinate. However, any kinetic barriers that would appear may well be artefacts of that coordinate; we do not, though, see any such barriers because proton transfers are inherently fast in water (in fact they can occur spontaneously as has happened during our simulations, see Section 2.4 “Progress of the trajectories” below). The choice of a particular reaction coordinate (such as simple bond distances) might introduce a certain bias; however, in context with simulations of  $pK_a$  values it has been shown that using simple X–H distances as constraints usually gives very similar results to more sophisticated coordinates (such as coordination numbers).<sup>21</sup> The simulations for the first steps started from equilibrated solutions of (Br)U in periodic water boxes containing 49 solvent molecules (a typical system size affording a ca. 1 M solution), with consecutive stepwise elongation of the N3–H3 distance (BLYP-D2 level of DFT; this functional was employed because it is still one of the best choices to model liquid water and aqueous solutions – see Section 2.6 “Error evaluation” below for further details and

justification). Analogous simulations were performed for the third steps, starting from equilibrated solutions of the enol forms of (Br)U with stepwise elongation of the O4–H4 distance (see Section 2.2 “Mechanistic approach” for further details). This deprotonation has, in principle, the same endpoint as deprotonation at N3, *i.e.*, a contact ion pair between the deprotonated enolate and a hydronium ion. In practice, there is a free-energy difference between the two resulting species, owing to the different ionic interaction sites (N3 and O4). In order to close the thermodynamic cycles, we have estimated the energy difference between the two contact ion pairs to be just 1.5 kcal/mol<sup>1</sup> at the SCS-MP2/aug-cc-pVTZ/COSMO level (see Section 2.7 below for details), for both U and BrU (with the N•••hydronium ion pair to be the more stable of the two). This error is likely to be partially offset by the fact that the free energies of separation of the two contact ion pairs – which presumably also differ – are already neglected in the free-energy calculations. Assuming the proton-hopping step between solvent molecules to be essentially thermoneutral, and the free energy of reprotonation at O4 to be the negative of the corresponding deprotonation energy of the enol form, this yields the thermodynamics of enolisation (via thermodynamic integration<sup>22</sup>).

## 2.2 Mechanistic approach

Eight values of the constrained distance  $\zeta$  were chosen for each of the two constraints (N3–H3 or O4–H4) in each of the two systems (U and BrU). The  $\zeta$  values were as follows (in Å):

U(N3–H3): 1.05, 1.15, 1.21, 1.27, 1.34, 1.48, 1.61, 1.94;

U(O4–H4): 1.06, 1.14, 1.22, 1.28, 1.38, 1.54, 1.70, 1.86;

BrU(N3–H3): 1.04, 1.15, 1.23, 1.30, 1.37, 1.60, 1.88, 2.17;

---

<sup>1</sup> 1 kcal/mol = 4.184 kJ/mol

BrU(O4–H4): 1.00, 1.10, 1.14, 1.21, 1.28, 1.42, 1.56, 1.71.

The minimum and maximum values of  $\zeta$  for each system are labelled  $\zeta_1$  and  $\zeta_2$ .

The N3–H3 systems were designed to simulate deprotonation of the keto forms. The starting geometries of the setups with the longest N3–H3 constraints were arranged such that a hydronium ion was formed within non-covalent-interaction distance of N3, with one of the hydronium protons being the constrained H atom.

Likewise, the trajectories with O4–H4 constraints were designed to calculate the free energies of reprotonation at O4, forming the enol tautomer. Here, in each starting structure, the base was already fully deprotonated at N3, by simply deleting H3. A new proton (labelled H4) was then inserted between O4 and the nearest water oxygen ( $O_w$ ), forming a linear structure, O4–H4– $O_w$ , with increasing constrained values of the O4–H4 distance. The increasing O4–H4 distances were accompanied by decreasing H4– $O_w$  distances: for BrU, the maximum value of H4– $O_w$  was 1.75 Å (when O4–H4 = 1.00 Å), while the minimum was 1.04 Å (when O4–H4 = 1.71 Å). Likewise, for U, H4– $O_w$  varied between 1.67 and 0.96 Å. Therefore, the smallest values of  $\zeta(O4-H4)$  corresponded to the O4-enol tautomer interacting non-covalently with water at O4, while the largest values corresponded to the deprotonated enolate interacting non-covalently with hydronium at O4 (i.e., a contact ion pair). As examples, Figures 3 and 4 show the starting geometries for BrU with  $\zeta(O4-H4) = 1.00$  and 1.71 Å, respectively.

**[Insert Figure 3 here]**

**Figure 3.** Starting geometry of the BrU(H<sub>2</sub>O)<sub>49</sub> trajectory with  $\zeta(O4-H4) = 1.00$  Å (distance labelled).



**[Insert Figure 4 here]**

**Figure 4.** Starting geometry of the BrU(H<sub>2</sub>O)<sub>49</sub> trajectory with  $\zeta(\text{O4-H4}) = 1.71 \text{ \AA}$  (distance labelled). The resulting hydronium molecule is shown in ball-and-stick representation.

For each  $\zeta$  value, a trajectory of length 6.5–10 ps was simulated (long enough in comparison to the expected proton transfer time-scale). Each trajectory was divided into three stages, termed Equilibration (lasting for either 1 or 2.5 ps), Thermalisation (with durations varying from 1 to 5 ps to ensure force convergence), and Production (the remainder of the trajectory). During Equilibration, the temperature was constrained to  $320 \text{ K} \pm 50 \text{ K}$  by rescaling the velocities when the temperature went outside this range. Rescaling was not performed during either Thermalisation or Production. However, only Production was used for the final calculations of the free energy.

### 2.3 Free energy calculation

In each trajectory, the force needed to maintain the distance constraint at each timestep was recorded. According to Sprik and Ciccotti<sup>22</sup> the relative free energy between states  $\zeta_1$  and  $\zeta_2$  can then be obtained by the following expression:

$$\Delta F = - \int_{\zeta_1}^{\zeta_2} f'_\zeta d\zeta \quad (1)$$

where

$$f'_\zeta = \frac{\langle Z^{-1/2} [\lambda - k_B T G] \rangle_\zeta}{\langle Z^{-1/2} \rangle_\zeta} \quad (2)$$

is the average force of constraint,  $k_B$  is the Boltzmann constant,  $T$  is the temperature, and  $Z$  and  $G$  are factors that compensate for the bias introduced by the constraint (see ref. <sup>22</sup> for the full derivation). In the case of a distance constraint as used here,  $G = 0$ , and

$$f_{\zeta} = \langle \lambda \rangle_{\zeta} \quad (3)$$

is simply the ensemble average of the Lagrange multiplier  $\lambda$  maintaining the constraint. The Lagrange multiplier values,  $\lambda$ , were accumulated over each trajectory. It was confirmed that the running average of  $\lambda$  had reached a steady value by the beginning of the Production phase in each trajectory (see example in Figure 5), with one exception: the BrU(N3–H3) trajectory with  $\zeta = 1.15 \text{ \AA}$ . In this case,  $\langle \lambda \rangle_{\zeta}$  failed to converge to a stable value even after 10 ps of simulation time. This trajectory was therefore discarded.

**[Insert Figure 5 here]**

**Figure 5.** Running average of constraint force  $\lambda$  during 8.0-ps BrU(N3–H3) trajectory with  $\zeta = 1.30 \text{ \AA}$ .

## 2.4 Progress of the trajectories

In eight trajectories (see below) the artificially-formed hydronium ion dissociated mid-trajectory, when one of its unconstrained protons transferred to a neighbouring water molecule and joined the bulk solution, a process known as proton hopping. In six of these cases, dissociation occurred during the Equilibration stage, so these trajectories were discarded from the final calculation of the free energy. In the other two cases (BrU(N3–H3) with  $\zeta = 1.60 \text{ \AA}$  and BrU(O4–H4) with  $\zeta = 1.42 \text{ \AA}$ ), dissociation occurred during the Production stage. Therefore, when calculating the mean force,  $\langle \lambda \rangle_{\zeta}$ , for these two trajectories, we discarded all values of  $\lambda$  taken at timesteps subsequent to dissociation, as those timesteps contained no information about deprotonation of the bases. The timesteps at which dissociation occurred, and the corresponding amounts of “real” simulation time prior to these timesteps, were as follows:

BrU(N3–H3):  $\zeta = 1.60 \text{ \AA}$ , timestep at  $\text{H}_3\text{O}^+$  dissociation = 81180  $\equiv$  5.8 ps

BrU(N3–H3):  $\zeta = 1.88 \text{ \AA}$ , timestep at  $\text{H}_3\text{O}^+$  dissociation = 7159  $\equiv$  0.5 ps (discarded)

BrU(N3–H3):  $\zeta = 2.17 \text{ \AA}$ , timestep at  $\text{H}_3\text{O}^+$  dissociation = 571  $\equiv$  0.04 ps (discarded)

U(N3–H3):  $\zeta = 1.94 \text{ \AA}$ , timestep at  $\text{H}_3\text{O}^+$  dissociation = 2000  $\equiv$  0.2 ps (discarded)

BrU(O4–H4):  $\zeta = 1.42 \text{ \AA}$ , timestep at  $\text{H}_3\text{O}^+$  dissociation = 114175  $\equiv$  8.2 ps

BrU(O4–H4):  $\zeta = 1.71 \text{ \AA}$ , timestep at  $\text{H}_3\text{O}^+$  dissociation = 3450  $\equiv$  0.2 ps (discarded)

U(O4–H4):  $\zeta = 1.70 \text{ \AA}$ , timestep at  $\text{H}_3\text{O}^+$  dissociation = 298  $\equiv$  0.03 ps (discarded)

U(O4–H4):  $\zeta = 1.86 \text{ \AA}$ , timestep at  $\text{H}_3\text{O}^+$  dissociation = 62  $\equiv$  0.01 ps (discarded)

Evidently, in the six cases of dissociation during Equilibration, the  $\zeta$  values (all  $\geq 1.70 \text{ \AA}$ ) were too long for the contact ion interaction between hydronium and base to be maintained.

Additionally, for U(O4–H4) with  $\zeta = 1.28 \text{ \AA}$ , reprotonation of the base at the N3 position (forming the neutral diketo tautomer) occurred after  $\sim 0.4$  ps. This trajectory therefore also had to be discarded.

For each of the four systems, we calculated  $\langle \lambda \rangle_{\zeta}$  for each non-discarded trajectory, and plotted these values against the corresponding constrained distances,  $\zeta$ . As an example, the plot for BrU(N3–H3) is shown in Figure 6.

**[Insert Figure 6 here]**

**Figure 6.** Mean force of constraint  $\langle \lambda \rangle$  as a function of constraint distance  $\zeta$  for valid BrU(N3–H3) trajectories.

For all four sets of trajectories, the mean value of  $\lambda$  was near zero when the shortest N3–H3 or O4–H4 constraint (1.00–1.06  $\text{\AA}$ ) was imposed. This state corresponds to the neutral diketo (or O4-enol) tautomer of each base, featuring a covalent N3–H3 (or O4–H4) bond that requires little force to maintain. The forces of constraint reached their greatest magnitudes with the second- or

third-shortest value of  $\zeta$  (1.14–1.23 Å), indicating an intermediate state between diketo (or enol) tautomer and contact ion pair. The forces then decreased with  $\zeta$ , and came close to zero again when the constrained distance was around 1.4–1.6 Å. This state corresponds to deprotonation at N3 (or O4), forming a contact ion pair. However, we note that in the case of BrU(O4–H4), a second (very shallow) minimum of the force occurred at  $\zeta = 1.28$  Å before the force reached  $\sim 0$ .

## 2.5 CPMD methodology

The DFT-based CPMD simulations followed the general setup of our previous simulations (e.g., ref. <sup>23</sup>), with all molecules (solute and solvent) explicitly represented. All simulations were carried out with the CPMD<sup>24</sup> package. We used the Becke, Lee, Yang, and Parr gradient-corrected functional<sup>25</sup> with Grimme dispersion correction<sup>26</sup> (BLYP-D2) for the exchange and correlation terms. The one-electron orbitals were expanded in a plane-wave basis set with a kinetic energy cutoff of 90 Ry restricted to the gamma point of the Brillouin zone. Medium-soft norm-conserving pseudopotentials of the Martins–Troullier type<sup>27</sup> were used. The core–valence interaction in C, N, and O was treated by *s* and *p* potentials with pseudisation radii of 1.23, 1.12, and 1.05 a.u., respectively (taking the same radius for *s* and *p*), whereas H atoms were treated as an *s* potential with a 0.5 a.u. radius. Energy expectation values were calculated in reciprocal space using the Kleinman–Bylander transformation.<sup>28</sup> Dynamics were performed in the microcanonical ensemble (at constant volume and internal energy) using a fictitious electron mass of 400 a.u. (for U) or 300 a.u. (for BrU) and a time step of 4.0 a.u. (for U) or 3.0 a.u. (for BrU), and periodic boundary conditions were applied to a cubic box of length 11.5 Å. The volume and number of molecules in the box were chosen in order to provide a density of liquid water of 1 g/cm<sup>3</sup> together with sufficient layers of water surrounding the solute. The initial

conformations of all trajectories herein were based on our previous CPMD investigation of U immersed in liquid water.<sup>29</sup>

## 2.6 Error evaluation

Various sources of error need to be considered.

Firstly, the use of a *single* distance constraint in each set of trajectories may have imposed some bias. Here we have run two separate series of constraints ( $\zeta = \text{N3-H3}$  or  $\text{O4-H4}$ ) for each base, generating separate free-energy measurements of the two protonation/deprotonation steps. The energetics of proton transfer from N3 to O4 – a multi-step process, potentially involving several H atoms – cannot be modelled by a single sequence of trajectories if the same single distance constraint is employed in each of them.<sup>30</sup> A seamless set of MD trajectories would require a more sophisticated choice of reaction coordinate. For example, Sprik<sup>21b</sup> modelled the auto-dissociation of water using a coordination number constraint, constraining the number of protons within bonding distance of a chosen oxygen. By applying a weighting function to protons at intermediate O–H distances, the constraint could be varied smoothly toward either 3 (enforcing the creation of a hydronium) or 1 (enforcing the creation of a hydroxyl). This in principle provides a more flexible description of the reaction coordinate than a simple distance constraint. However, in practice, the final energy was found to be independent of the type of constraint. Another choice is the “difference-of-distances,” in which a triatomic bond-breaking/bond-making process,  $\text{A-B} + \text{C} \rightarrow \text{A} + \text{B-C}$ , is studied by constraining the difference between the A–B and B–C distances.<sup>31</sup>

Secondly, uncertainties in the force of constraint,  $\lambda$ , lead to a statistical error,  $\delta\Delta F$ , in each free energy. This error manifests as a fluctuation in the running average of  $\lambda$  even after the

constraining force has reached a reasonably steady value.  $\delta\Delta F$  can hence be estimated by measuring the standard deviation of this running average during the last picosecond of each trajectory, then integrating the standard deviations across the corresponding constraint distances,  $\zeta$ , of each trajectory (a procedure adopted in many of our previous studies, see *e.g.* Ref<sup>32</sup>). Using an in-house program, the standard deviations in the N3–H3 trajectories were found to vary between  $2.7 \times 10^{-5}$  and  $8.7 \times 10^{-4}$  a.u. for U, and between  $2.8 \times 10^{-4}$  and  $1.3 \times 10^{-3}$  a.u. for BrU. The corresponding statistical errors in the free energies are  $\delta\Delta F = \pm 0.22$  kcal/mol for U and  $\pm 0.61$  kcal/mol for BrU. In the O4–H4 trajectories, the standard deviations varied between  $5.7 \times 10^{-5}$  and  $7.1 \times 10^{-4}$  a.u. for U, and  $1.3 \times 10^{-4}$  and  $2.0 \times 10^{-3}$  a.u. for BrU. The corresponding statistical errors are  $\delta\Delta F = \pm 0.12$  kcal/mol for U and  $\pm 0.53$  kcal/mol for BrU. Thus, the errors are 1–2 orders of magnitude smaller than the energies. A more sophisticated method of estimating  $\delta\Delta F$  from the uncertainty in  $\lambda$ , based on Simpson’s rule, was developed by Senn *et al.*,<sup>31</sup> but since it is probably not the largest source of numerical error in the first place, further precision in its estimation is not necessary.

Thirdly, the choice of functional plays a role. The use of BLYP for simulations of small biomolecules immersed in liquid water is well established in our group.<sup>23,30</sup> However, other authors have detected weaknesses in the dynamical description of water by this method. In CPMD simulations of the transport of a hydrated excess proton, Izvekov and Voth noted that BLYP caused an unphysical over-structuring of water, resulting in an attraction between the excess proton and one “special” water oxygen, and thus significantly reducing the self-diffusion coefficient of water compared with experiment.<sup>33</sup> Likewise, Todorova *et al.* observed over-structuring of water by BLYP, and recommended hybrid functionals for the more accurate calculation of self-diffusion coefficients and radial distribution functions.<sup>34</sup> However, our study

was not concerned with the overall water structure, nor with the behavior of the hydronium proton once it had diffused into solution, but only with the constrained base–water interactions at N3 and O4.

## 2.7 Calculation of high-level gas-phase tautomerisation energies and contact ion pairs

The structures of the two tautomeric forms of U and BrU were first optimised at the M06-2X<sup>35</sup>/6-31+G\*\* level using Gaussian.<sup>36</sup> We subsequently performed single-point calculations with the spin-component-scaled MP2 (SCS-MP2)<sup>37</sup> method using the aug-cc-pVQZ basis set<sup>38</sup> using ORCA.<sup>39</sup> The resolution of the identity (RI) approximation was evoked employing the aug-cc-pVQZ/C auxiliary basis set. We note that the computed aug-cc-pVQZ gas-phase tautomerisation energies are nearly identical (differences < 0.02 kcal/mol) to those calculated using the smaller aug-cc-pVDZ and aug-cc-pVTZ basis sets, showing that the results are adequately converged with respect to basis set size. A similar methodology was employed to calculate the energy difference between the two contact ion pairs that are the end points of deprotonation at N3 or O4: optimisation using M06-2X/6-31+G\*\* with PCM (polarizable continuum model)<sup>40</sup> water and subsequent single-point calculation using SCS-MP2/aug-cc-pVTZ employing the aug-cc-pVTZ/C auxiliary basis set with COSMO (conductor-like screening model)<sup>41</sup> water. The two ion pairs considered are the deprotonated (Br)U base interacting with a hydronium ion at either N3 or O4. The O4•••H(hydronium) or N3•••H(hydronium) distances were kept fixed at 1.6 Å. To keep the hydronium ion intact, an additional water molecule was added (see Figure 7). To prevent re-protonation of the base, the water OH pointing to N3 in the O4•••hydronium pair and the water OH pointing to O4 in the N3•••hydronium pair were kept fixed at 0.96885 Å (the H-bonded O–H distance in a water dimer optimised with M06-2X/6-31+G\*\*).

[Insert Figure 7 here]

Figure 7. The O4•••hydronium and N3•••hydronium ion pairs.

### 3. Results and discussion

From the trajectories in which the nascent hydronium ion interacting with N3 or O4 remained stable during the production phase (between five and seven cases for U and BrU, the free energies of deprotonation in water were calculated, and are shown in Table 1.

Table 1. Helmholtz free-energy changes,  $\Delta F$  (in kcal/mol), for deprotonation of uracil and 5-bromouracil at N3, and for deprotonation of the respective enol forms at O4.

System	$\Delta F_{\text{deprot(N3)}}^{\text{a,b}}$	$\Delta F_{\text{deprot(O4)}}^{\text{b,c}}$
U(H <sub>2</sub> O) <sub>49</sub>	14.9(6)	4.6(1)
BrU(H <sub>2</sub> O) <sub>49</sub>	9.4(2)	3.3(5)

<sup>a</sup> Deprotonation of (Br)U at N3. <sup>b</sup> In parentheses: estimated uncertainties from the standard deviations of  $\langle\lambda\rangle$ . <sup>c</sup> Deprotonation of the enol forms of (Br)U at O4 (see Figure 1 for labelling).

Comparison of the  $\Delta F$  values for the N3–H3 and O4–H4 systems shows that, for both bases, the diketo form is more stable against deprotonation than the O4-enol form, in agreement with the literature.<sup>42</sup> With the reasonable assumption that  $\Delta F$  can be equated with the Gibbs free energy  $\Delta G$  (since volume does not vary significantly with pressure), we can use  $\Delta F_{\text{deprot}} (\approx \Delta G_{\text{deprot}})$  to estimate the  $\text{p}K_{\text{a}}$  of each protic site, according to<sup>21a</sup>

$$\Delta G_{\text{deprot}} = 2.303RT \text{p}K_{\text{a}} \quad (4)$$



With  $T = 320$  K, solving for the right-hand side yields  $\text{p}K_a(\text{N3}) = 10.1$  for U and 6.4 for BrU, *i.e.*, BrU is a stronger proton donor. In 1962, Katritzky and Waring estimated these values as 9.5 and 7.8, respectively, by titration at 297 K.<sup>6</sup> More recent experimental estimates for U have been 9.3 and 9.8,<sup>43</sup> while B3LYP/continuum-solvent calculations by Jang *et al.* yielded a value for U at 298 K of 9.3.<sup>44</sup> In 1989, Sowers *et al.* determined that the  $\text{p}K_a$  of the dBrU nucleoside at 296 K was 8.1,<sup>10</sup> while Wood *et al.* more recently measured a value of 8.2 for this nucleoside.<sup>45</sup> We note the good qualitative accord between our computed data and these earlier estimates, although we calculate an appreciably greater acidity for BrU at N3 than the experimental measurements of the nucleoside.

The calculated values of  $\Delta F_{\text{deprot(O4)}} (\approx \Delta G_{\text{deprot(O4)}})$  can likewise be used to calculate the  $\text{p}K_a$  of each enol form at O4. Solving Eq. 1 yields  $\text{p}K_a(\text{O4}) = 3.2$  for U and 2.3 for BrU. The aqueous free energies of tautomerisation,  $\Delta F_{\text{keto} \rightarrow \text{enol}}(\text{aq.})$ , of the two bases can be taken as the difference between the deprotonation free energies at the two sites.<sup>46</sup> The tautomerisation energies estimated in this way are markedly different for the two bases: for U,  $\Delta F_{\text{keto} \rightarrow \text{enol}}(\text{aq.}) \approx 10.2$  kcal/mol, while for BrU,  $\Delta F_{\text{keto} \rightarrow \text{enol}}(\text{aq.}) \approx 6.1$  kcal/mol. Tautomerisation is therefore more favourable for BrU than for U by 4.1 kcal/mol, in qualitative agreement with the “rare tautomer” hypothesis of BrU mutagenicity. Loeb and Kunkel estimated that a free-energy difference of  $1.4n$  kcal/mol should cause a mispairing rate of 1 in  $10^n$ ,<sup>47</sup> which according to the present results implies that BrU should mispair more often than U by a factor of  $10^{(4.1/1.4)} \approx 848$ . (Here we assume the same changes in  $\text{p}K_a$  upon going from the free nucleobases, U and BrU, to the bound nucleotides in helical DNA.)

The difference between the aqueous free energies of tautomerisation derived here at the BLYP-D2 level ( $\Delta F_{\text{keto} \rightarrow \text{enol}}^{\text{BLYP-D2}}(\text{aq.})$ ), and the corresponding gas-phase potential energies of

tautomerisation ( $\Delta E_{\text{keto} \rightarrow \text{enol}}^{\text{BLYP-D2}}(\text{gas})$ ), is due to the effect of solvation on the tautomerism of U and BrU. This difference, which has enthalpic, entropic, and solute–solvent contributions, can be used to estimate the aqueous tautomerisation energies at higher levels of theory, for which only the gas-phase term is known, according to:

$$\Delta E_{\text{keto} \rightarrow \text{enol}}^{\text{high-level}}(\text{aq.}) \approx \Delta E_{\text{keto} \rightarrow \text{enol}}^{\text{high-level}}(\text{gas}) + [\Delta F_{\text{keto} \rightarrow \text{enol}}^{\text{BLYP-D2}}(\text{aq.}) - \Delta E_{\text{keto} \rightarrow \text{enol}}^{\text{BLYP-D2}}(\text{gas})] \quad (5)$$

Here we assume that the effect of solvation (the term in square brackets) is method-invariant, so the BLYP-D2 solvation term can be combined with gas-phase potential energies from higher levels of theory. The gas-phase BLYP-D2/plane-wave tautomerisation energies, calculated using CPMD, are 11.2 and 12.0 kcal/mol for U and BrU, respectively. Thus, the contribution of solvation to the tautomerisation free energy (derived from the  $\Delta F_{\text{keto} \rightarrow \text{enol}}^{\text{BLYP-D2}}(\text{aq.})$  values of 10.2 kcal/mol for U and 6.1 kcal/mol for BrU) is  $-0.9$  kcal/mol for U and  $-5.9$  kcal/mol for BrU. Applying these corrections to the gas-phase SCS-MP2/aug-cc-pVQZ (aVQZ) tautomerisation energies (calculated as 11.9 and 12.5 kcal/mol for U and BrU, respectively, using ORCA<sup>39</sup>) yields  $\Delta F_{\text{keto} \rightarrow \text{enol}}^{\text{SCS-MP2/aVQZ}}(\text{aq.}) = 11.0$  kcal/mol for U and 6.6 kcal/mol for BrU. The difference in tautomerisation energies is therefore 4.4 kcal/mol at this level of theory, which by Loeb and Kunkel’s equation<sup>47</sup> implies a BrU-to-U mispairing ratio of  $10^{(4.4/1.4)} \approx 1389$ , compared with 848 for BLYP-D2/CPMD.

In contrast to the force-field-based evaluation of bulk solvation effects on the tautomerisation equilibria performed in Ref. <sup>14</sup> (through Monte Carlo alchemical mutation studies), our quantum-mechanical-based models indicate that a substantial shift in this equilibrium upon going from U to BrU is preserved in solution. Our results thus reinforce the rare-tautomer hypothesis.

## 4. Conclusion

We have shown through constrained CPMD simulations and thermodynamic integration that the 5-bromine substituent significantly lowers the  $pK_a$  of BrU compared with U, thus increasing the likelihood that the “mutagenic” enol (or ionic) tautomers are formed in water. Corroborating previous suggestions in the literature, our results thus provide new evidence for the previously contested “rare tautomer” hypothesis for BrU mutagenicity. To what extent the propensity for forming this rare tautomer in water, as shown in this work, is maintained in the partially hydrated environment of DNA is an interesting topic for further investigation.

## Acknowledgements

We are grateful for support from the Engineering and Physical Sciences Research Council (EPSRC) UK National Service for Computational Chemistry Software (NSCCS); from GENCI (Grand équipement national de calcul intensif); and from CINES (Centre informatique national de l'enseignement supérieur). LFH and TvM gratefully acknowledge support from the HPC-EUROPA2 project with the support of the European Commission - Capacities Area - Research Infrastructures. LFH is grateful to the EPSRC for studentship support through the Doctoral Training Account scheme (grant code EP/K503162/1). The research data supporting this publication can be accessed at: <http://dx.doi.org/10.17630/XXX>.

## References

1. Dunn, D. B.; Smith, J. D., Incorporation of halogenated pyrimidines into the deoxyribonucleic acids of *Bacterium coli* and its bacteriophages. *Nature* **1954**, *174*, 305-306.
2. Davidson, R. L.; Broeker, P.; Ashman, C. R., DNA base sequence changes and sequence specificity of bromodeoxyuridine-induced mutations in mammalian cells. *Proc. Natl. Acad. Sci. USA* **1988**, *85* (12), 4406-4410.
3. Watson, J. D.; Crick, F. H. C., The structure of DNA. *Cold Spring Harb. Symp. Quant. Biol.* **1953**, *18*, 123-131.

4. Freese, E., The specific mutagenic effect of base analogues on Phage T4. *J. Mol. Biol.* **1959**, *1*, 87-105.
5. Topal, M. D.; Fresco, J. R., Complementary base pairing and the origin of substitution mutations. *Nature* **1976**, *263*, 285-289.
6. Katritzky, A. R.; Waring, A. J., Tautomeric azines. Part I. The tautomerism of 1-methyluracil and 5-bromo-1-methyluracil. *J. Chem. Soc.* **1962**, 1540.
7. Terzaghi, B. E.; Streisinger, G.; Stahl, F. W., The mechanism of 5-bromouracil mutagenesis in the bacteriophage T4. *Proc. Natl. Acad. Sci. USA* **1962**, *48* (9), 1519-1524.
8. (a) Yekeler, H.; Özbakır, D., Concerning the solvent effect in the tautomerism of uracil, 5-fluorouracil, and thymine by density-functional theory and ab initio calculations. *J. Mol. Model.* **2001**, *7* (4), 103-111; (b) Bachorz, R. A.; Rak, J.; Gutowski, M., Stabilization of very rare tautomers of uracil by an excess electron. *Phys. Chem. Chem. Phys.* **2005**, *7* (10), 2116-2125; (c) Hanus, M.; Kabelac, M.; Nachtigallova, D.; Hobza, P., Mutagenic properties of 5-halogenuracils: Correlated quantum chemical ab initio study. *Biochemistry* **2005**, *44* (5), 1701-1707; (d) Hu, X.; Li, H.; Liang, W.; Han, S., Systematic study of the tautomerism of uracil induced by proton transfer. Exploration of water stabilization and mutagenicity. *J. Phys. Chem. B* **2005**, *109* (12), 5935-5944; (e) Markova, N.; Enchev, V.; Timtcheva, I., Oxo-hydroxy tautomerism of 5-fluorouracil: Water-assisted proton transfer. *J. Phys. Chem. A* **2005**, *109* (9), 1981-1988; (f) Hu, X.; Li, H.; Zhang, L.; Han, S., Tautomerism of uracil and 5-bromouracil in a microcosmic environment with water and metal ions. What roles do metal ions play? *J. Phys. Chem. B* **2007**, *111* (31), 9347-9354; (g) Palafox, M. A.; Rastogi, V. K.; Kumar, H.; Kostova, I.; Vats, J. K., Tautomerism in 5-bromouracil: Relationships with other 5-haloderivatives and effect of the microhydration. *Spectrosc. Lett.* **2011**, *44* (4), 300-306; (h) Valadbeigi, Y.; Farrokhpour, H., Effect of hydration on the stability and tautomerisms of different isomers of uracil. *RSC Adv.* **2014**, *4* (106), 61643-61651.
9. Brown, T.; Kneale, G.; Hunter, W. N.; Kennard, O., Structural characterisation of the bromouracil.guanine base pair mismatch in a Z-DNA fragment. *Nucl. Acids Res.* **1986**, *14* (4), 1801-1809.
10. Sowers, L. C.; Goodman, M. F.; Eritja, R.; Kaplan, B.; Fazakerley, G. V., Ionized and wobble base-pairing for bromouracil-guanine in equilibrium under physiological conditions: A nuclear magnetic resonance study on an oligonucleotide containing a bromouracil-guanine base-pair as a function of pH. *J. Mol. Biol.* **1989**, *205* (2), 437-447.
11. Moran, S.; Ren, R. X. F.; Rumney IV, S.; Kool, E. T., Difluorotoluene, a nonpolar isostere for thymine, codes specifically and efficiently for adenine in DNA replication. *J. Am. Chem. Soc.* **1997**, *119* (8), 2056-2057.
12. (a) Kimsey, I. J.; Petzold, K.; Sathyamoorthy, B.; Stein, Z. W.; Al-Hashimi, H. M., Visualizing transient Watson-Crick-like mispairs in DNA and RNA duplexes. *Nature* **2015**, *519*, 315; (b) Szymanski, E. S.; Kimsey, I. J.; Al-Hashimi, H. M., Direct NMR evidence that transient tautomeric and anionic states in dG·dT form Watson-Crick-like base pairs. *J. Am. Chem. Soc.* **2017**, *139* (12), 4326-4329.
13. (a) Harris, V. H.; Smith, C. L.; Cummins, J. W.; Hamilton, A. L.; Adams, H.; Dickman, M.; Hornby, D. P.; Williams, D. M., The effect of tautomeric constant on the specificity of nucleotide incorporation during DNA replication: Support for the rare tautomer Hypothesis of substitution mutagenesis. *J. Mol. Biol.* **2003**, *326* (5), 1389-1401; (b) Suen, W.; Spiro, T. G.; Sowers, L. C.; Fresco, J. R., Identification by UV resonance Raman spectroscopy of an imino tautomer of 5-hydroxy-2'-deoxycytidine, a powerful base analog transition mutagen with a much

higher unfavored tautomer frequency than that of the natural residue 2'-deoxycytidine. *Proc. Natl. Acad. Sci. USA* **1999**, *96* (8), 4500-4505.

14. Orozco, M.; Hernández, A.; Luque, F. J., Tautomerism of 1-methyl derivatives of uracil, thymine, and 5-bromouracil. Is tautomerism the basis for the mutagenicity of 5-bromouridine? *J. Phys. Chem. B* **1998**, *102*, 5228-5233.

15. (a) Danilov, V. I.; van Mourik, T.; Kurita, N.; Wakabayashi, H.; Tsukamoto, T.; Hovorun, D. M., On the mechanism of the mutagenic action of 5-bromouracil: A DFT study of uracil and 5-bromouracil in a water cluster. *J. Phys. Chem. A* **2009**, *113*, 2233-2235; (b) van Mourik, T.; Danilov, V. I.; Dailidonis, V. V.; Kurita, N.; Wakabayashi, H.; Tsukamoto, T., A DFT study of uracil and 5-bromouracil in nanodroplets. *Theor. Chem. Acc.* **2010**, *125*, 233-244.

16. Li, D.; Ai, H., Catalysis effects of water molecules and of charge on intramolecular proton transfer of uracil. *J. Phys. Chem. B* **2009**, *113* (34), 11732-11742.

17. (a) Kubinec, M. G.; Wemmer, D. E., NMR evidence for DNA bound water in solution. *J. Am. Chem. Soc.* **1992**, *114* (22), 8739-8740; (b) Brovarets', O. O.; Kolomiets', I. M.; Hovorun, D. M., Elementary molecular mechanisms of the spontaneous point mutations in DNA: A novel quantum-chemical insight into the classical understanding. In *Quantum Chemistry - Molecules for Innovations*, Tada, T., Ed. InTech: 2012.

18. Furmanchuk, A.; Isayev, O.; Gorb, L.; Shishkin, O. V.; Hovorun, D. M.; Leszczynski, J., Novel view on the mechanism of water-assisted proton transfer in the DNA bases: bulk water hydration. *Phys. Chem. Chem. Phys.* **2011**, *13* (10), 4311-4317.

19. (a) Car, R.; Parrinello, M., Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **1985**, *55* (22), 2471; (b) Marx, D.; Hutter, J., *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*. Cambridge University Press: Cambridge, 2009; (c) Gaigeot, M.-P., Theoretical spectroscopy of floppy peptides at room temperature. A DFTMD perspective: gas and aqueous phase. *Phys. Chem. Chem. Phys.* **2010**, *12* (14), 3336-3359.

20. (a) Brovarets', O. h. O.; Hovorun, D. M., The nature of the transition mismatches with Watson-Crick architecture: the G\*·T or G·T\* DNA base mispair or both? A QM/QTAIM perspective for the biological problem. *J. Biomol. Struct. Dyn.* **2015**, *33* (5), 925-945; (b) Brovarets', O. h. O.; Hovorun, D. M., How many tautomerization pathways connect Watson-Crick-like G\*·T DNA base mispair and wobble mismatches? *J. Biomol. Struct. Dyn.* **2015**, *33* (11), 2297-2315.

21. (a) Ivanov, I.; Klein, M. L., Deprotonation of a Histidine Residue in Aqueous Solution Using Constrained Ab Initio Molecular Dynamics. *J. Am. Chem. Soc.* **2002**, *124* (45), 13380-13381; (b) Sprik, M., Computation of the pK of liquid water using coordination constraints. *Chem. Phys.* **2000**, *258* (2), 139-150.

22. Sprik, M.; Ciccotti, G., Free energy from constrained molecular dynamics. *J. Chem. Phys.* **1998**, *109*, 7737-7744.

23. López-Tarifa, P.; Gaigeot, M. P.; Vuilleumier, R.; Tavernelli, I.; Alcamí, M.; Martín, F.; Penhoat, M. A. H. d.; Politis, M. F., Ultrafast damage following radiation-induced oxidation of uracil in aqueous solution. *Angew. Chem. Int. Ed.* **2013**, *52* (11), 3160-3163.

24. CPMD, copyright International Business Machines Corporation (1990-2008) and Max Planck Institute für Festkörperforschung, Stuttgart (Germany) (1995-2001).

25. (a) Becke, A. D., Density-functional thermochemistry. I. The effect of the exchange-only gradient correction. *J. Chem. Phys.* **1992**, *96* (3), 2155-2160; (b) Becke, A. D., Density

- functional thermochemistry. 3. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648-5652;
- (c) Lee, C.; Yang, W.; Parr, R. G., *Phys. Rev. B* **1988**, *37*, 785-789.
26. Grimme, S., Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463-1473.
27. Troullier, N.; Martins, J. L., Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **1991**, *43* (3), 1993-2006.
28. Kleinman, L.; Bylander, D. M., Efficacious form for model pseudopotentials. *Phys. Rev. Lett.* **1982**, *48* (20), 1425-1428.
29. (a) Gaigeot, M.-P.; Sprik, M., Ab Initio molecular dynamics computation of the infrared spectrum of aqueous uracil. *J. Phys. Chem. B* **2003**, *107* (38), 10344-10358; (b) Gaigeot, M.-P.; Sprik, M., Ab initio Molecular Dynamics study of uracil in aqueous solution. *J. Phys. Chem. B* **2004**, *108* (22), 7458-7467.
30. Gaigeot, M.-P.; Cimas, A.; Seydou, M.; Kim, J.-Y.; Lee, S.; Schermann, J.-P., Proton transfer from the inactive gas-phase nicotine structure to the bioactive aqueous-phase structure. *J. Am. Chem. Soc.* **2010**, *132* (51), 18067-18077.
31. Senn, H. M.; Thiel, S.; Thiel, W., Enzymatic Hydroxylation in p-Hydroxybenzoate Hydroxylase: A Case Study for QM/MM Molecular Dynamics. *J. Chem. Theor. Comput.* **2005**, *1* (3), 494-505.
32. Bühl, M.; Sieffert, N.; Wipff, G., Structure of a uranyl peroxo complex in aqueous solution from first-principles molecular dynamics simulations. *Dalton Trans.* **2014**, *43* (29), 11129-11137.
33. Izvekov, S.; Voth, G. A., Ab initio molecular-dynamics simulation of aqueous proton solvation and transport revisited. *J. Chem. Phys.* **2005**, *123* (4), 044505.
34. Todorova, T.; Seitsonen, A. P.; Hutter, J.; Kuo, I. F. W.; Mundy, C. J., Molecular Dynamics simulation of liquid water: Hybrid density functionals. *J. phys. Chem. B* **2006**, *110* (8), 3685-3691.
35. Zhao, Y.; Truhlar, D. G., The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06 functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120* (1-3), 215-241.
36. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Rev. A.02*, Wallingford CT, 2009.
37. Grimme, F., Improved second-order Møller-Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies *J. Chem. Phys.* **2003**, *118* (20), 9095-9102.

38. (a) Dunning Jr., T. H., Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007; (b) Kendall, R. A.; Dunning Jr., T. H.; Harrison, R. J., Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796.
39. Neese, F., "The ORCA program system", version 3.0.3. *Wiley Interdiscip. Rev: Comput. Mol Sci.* **2012**, *2*, 73-78.
40. Miertus, S.; Scrocco, E.; Tomasi, J., Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **1981**, *55*, 117-129.
41. Klamt, A.; Schuurmann, G., COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, (5), 799-805.
42. Pérez, P.; Contreras, R., A theoretical analysis of the gas-phase protonation of hydroxylamine, methyl-derivatives and aliphatic amino acids. *Chem. Phys. Lett.* **1998**, *293* (3), 239-244.
43. (a) Kimura, E.; Kitamura, H.; Koike, T.; Shiro, M., Facile and Selective Electrostatic Stabilization of Uracil N(1)- Anion by a Proximate Protonated Amine: A Chemical Implication for Why Uracil N(1) Is Chosen for Glycosylation Site. *J. Am. Chem. Soc.* **1997**, *119* (45), 10909-10919; (b) Kurinovich, M. A.; Lee, J. K., The acidity of uracil and uracil analogs in the gas phase: four surprisingly acidic sites and biological implications. *J. Am. Soc. Mass Spectrom.* **2002**, *13* (8), 985-995.
44. Jang, Y. H.; Sowers, L. C.; Çağın, T.; Goddard, W. A., First Principles Calculation of pKa Values for 5-Substituted Uracils. *J. Phys. Chem. A* **2001**, *105* (1), 274-280.
45. Wood, D. O.; Dinsmore, M. J.; Bare, G. A.; Lee, J. S., M-DNA is stabilised in G•C tracts or by incorporation of 5-fluorouracil. *Nuc. Acids Res.* **2002**, *30* (10), 2244-2250.
46. Holmes, F. L., *Meselson, Stahl, and the Replication of DNA: A History of "The Most Beautiful Experiment in Biology"*. Yale University Press: New Haven, 2008.
47. Loeb, L. A.; Kunkel, T. A., Fidelity of DNA synthesis. *Ann. Rev. Biochem.* **1982**, *51* (1), 429-457.