

# Methylation of Zebularine: A Quantum Mechanical Study Incorporating Interactive 3D pdf Graphs

Lalitha Selvam,<sup>†</sup> Vladislav Vasilyev,<sup>‡</sup> and Feng Wang<sup>\*†</sup>

Centre for Molecular Simulation, Swinburne University of Technology, Hawthorn, Melbourne, Victoria 3122, Australia, National Computational Infrastructure, Australian National University, Canberra, ACT 0200, Australia

Received: February 23, 2009; Revised Manuscript Received: May 14, 2009

Methylation of a cytidine deaminase inhibitor, 1-( $\beta$ -D-ribofuranosyl)-2-pyrimidone (i.e., zebularine (zeb)), which produces 1-( $\beta$ -D-ribofuranosyl)-5-methyl-2-pyrimidinone (d5), has been investigated using density functional theory models. The optimized structures of zeb and d5 and the valence orbitals primarily responsible for the methylation in d5 are presented using state-of-the-art interactive (on a computer or online) three-dimensional (3D) graphics in a portable document format (pdf) file, 3D-PDF (<http://www.web3d.org/x3d/vrml/>). The facility to embed 3D molecular structures into pdf documents has been developed jointly at Swinburne University of Technology and the National Computational Infrastructure, the Australian National University. The methyl fragment in the base moiety shows little effect on the sugar puckering but apparently affects anisotropic properties, such as condensed Fukui functions. Binding energy spectra, both valence space and core space, are noticeably affected; in particular, in the outer-valence space (e.g., IP < 20 eV). The methyl fragment delocalizes and diffuses into almost all valence space, but orbitals 8 (57a, IP = 12.57 eV), 18 (47a, IP = 14.70 eV), and 37 (28a, IP = 22.15 eV) are identified as fingerprint for the methyl fragment. In the inner shell, however, the impact of the methyl can be localized and identified by chemical shift. A small, global, red shift is found for the O–K, N–K and sugar C–K spectra, whereas the base C–K spectrum exhibits apparent methyl-related changes.

## 1. Introduction

Cytidine deaminases (CDAs) are widespread in microorganisms,<sup>1,2</sup> plants,<sup>3</sup> and animals.<sup>4,5</sup> It (CDA) plays an important role in the catalysis of hydrolytic deamination of cytidine to uridine. Pyrimidin-2-one nucleosides have been reported as inhibitors of enzyme CDAs.<sup>6,7</sup> Experimental and theoretical methods have been applied to study CDAs as well as inhibitor interactions with enzymes.<sup>8–14</sup> Zebularine (zeb or 1-( $\beta$ -D-ribofuranosyl)-2-pyrimidone), a cytidine analogue, was evaluated for its antitumor properties<sup>15</sup> and also found to be an effective inhibitor of CDAs. It was primarily synthesized as a bacteriostatic agent<sup>16</sup> nearly three decades ago. Later, biochemical investigations established that zeb is a potent inhibitor of CDAs,<sup>7</sup> and it acts as a competitive inhibitor of CDAs with a dissociation constant ( $K_i$ ) of  $1.2 \times 10^{-12}$  M, roughly 8 orders of magnitude lower than the value of the substrate (cytidine) concentration ( $K_m$ ).<sup>17</sup> Pharmacokinetics studies have reported that zeb causes low bioavailability in rodents and extremely low bioavailability in the rhesus monkey.<sup>18</sup> Moreover, zeb has been identified as a novel mechanism-based inhibitor of DNA methylation,<sup>19</sup> forming a covalent complex with DNA methyl transferases,<sup>19</sup> depleting human DNMT1<sup>20</sup> and involving the epigenetic silencing of tumor suppressor genes.<sup>21</sup> Although compared to other antitumor drugs such as 5-azacytidine, zeb has a considerably reduced potency, the favorable chemical stability of zeb facilitates oral administration and makes it such a promising candidate for reversing DNA methylation with minimal toxicity.<sup>22</sup>

Modulation of histone modification (acetylation, phosphorylation, and methylation) and DNA methylation are the principal driving forces behind the phenomenon of epigenetics.<sup>23–25</sup> Although histone modification is restricted to the eukarya, in organisms ranging from bacteriophage to human, differential DNA methylation has been co-opted for the regulation of genetic transactions, including transcription, imprinting, and recombination, and classically provides a barrier to host-specific restriction endonucleases.<sup>26–29</sup>

Chemically, zebularine, or 1-( $\beta$ -D-ribofuranosyl)-2-pyrimidone, is composed of a sugar moiety linked to a pyrimidine base. It differs from its nucleoside analog cytidine only in its lack of the  $-\text{NH}_2$  group at the C(4) position of the pyrimidine ring. Compared to its metabolites, such as uridine, uracil, dihydrouracil,  $\beta$ -ureidipropionic acid, and  $\beta$ -alanine, its large energy gap between the highest occupied molecular orbital and lowest unoccupied molecular orbital (HOMO–LUMO) makes zebularine kinetically inert.<sup>30,31</sup> In addition, zeb is utilized as a cytosine (C) analog while incorporating into DNA strands. It pairs with the opposite guanine (G) and, hence, forms one less hydrogen bond in the G:Z pair in relation to the G:C pair, due to the absence of a  $-\text{NH}_2$  at the C(4) position of the pyrimidine base.<sup>31</sup>

Methylation of 1-( $\beta$ -D-ribofuranosyl)-2-pyrimidone (zebularine) produces 1-( $\beta$ -D-ribofuranosyl)-5-methyl-2-pyrimidinone (d5). It can be considered a modified thymine derivative.<sup>32,33</sup> A study of 5-methyl cytidine derivatives in mammalian enzymes reported a lower rate of deamination, which may be due to the steric consideration.<sup>7</sup> In a DNA decamer, d5 substituted for thymidine was found to affect the overall behavior of DNA strands due to local perturbation.<sup>34</sup> Recently, theoretical investigation of 5-methyl-2-pyrimidinone (5M2P) (base alone) was

\* To whom correspondence should be addressed. E-mail: fwang@swin.edu.au.

<sup>†</sup> Swinburne University of Technology.

<sup>‡</sup> Australian National University.

carried out to analyze its fluorescence behavior.<sup>35</sup> To our knowledge, very little structural and energetic information is available for d5. In this study, the impact of methylation on zeb is investigated with respect to the geometric and electronic properties of d5, which are presented in a state-of-the-art interactive (on computer or online) 3D-PDF. The latter is the outcome of a recent project we developed in the past few months. According to L. Page, 3D-PDF will be a “significant milestone” in 2009, and similar techniques will be employed in Nature.<sup>36</sup>

## 2. Computational Details and Generation of 3D-PDF.

Geometric structures of zeb and d5 in their ground electronic states ( $XA^1$ ) were produced using optimization in isolation, followed by frequency calculations to ensure that the structures are true minima without negative frequencies. Becke’s three-parameter-hybrid (B3LYP) density functional theory (DFT) method was employed in the optimization, together with a correlation-consistent, polarized valence, triple- $\zeta$ , cc-pVTZ, basis set. Appropriate quantum mechanical models, such as B3LYP/aug-cc-pVTZ, OVGf/TZVP, SAOP/et-pVQZ<sup>37</sup> (for valence shell), and LB94/et-pVQZ<sup>38</sup> (for core shell), are employed to produce molecular electrostatic potentials, Hirshfeld charges, Fukui functions and binding energy spectra. Here, the aug-cc-pVTZ and TZVP are Gaussian basis sets, but et-pVQZ is an even-tempered slater-type basis set.<sup>39</sup>

Atomic-site-specific electron charges based on the Hirshfeld scheme<sup>40,41</sup> were calculated on the basis of the LB94/et-pVQZ wave functions. Hirshfeld charges were obtained by subtracting the integral of the density associated with atom A from the corresponding nuclear charge,  $Z_A$ .<sup>42</sup>

$$Q_A^{\text{Hirshfeld}} = Z_A - \int \frac{\rho_A(r)}{\rho_{\text{promolecule}}(r)} \rho_{\text{molecule}}(r) dr \quad (1)$$

The condensed Fukui functions for electrophilic attack ( $f^-$ ) were calculated following a procedure as previously described in Saha et al.<sup>43</sup> and Zhu et al.<sup>44</sup> All calculations were performed using computational chemistry programs, such as Gaussian03,<sup>45</sup> Gamess,<sup>46</sup> and ADF.<sup>47</sup> Sugar puckering of the nucleoside pair, zeb and d5, was performed using the method proposed in ref 48 for pseudorotational analyses.

In this paper, we use a new 3D feature in the ubiquitous portable document format (pdf), 3D-PDF. Using 3D-PDF, it is possible to embed 3D imagery within a pdf file, in which the readers are able to manipulate 3D structures on their computers or online to observe the structures interactively. No additional 3D-rendering software is required, since the 3D information has been already integrated within the pdf document. The 3D-PDF file can be displayed on any computer with up-to-date Adobe Acrobat software (a pdf viewer) installed.

Embedding 3D images into a pdf document could be done using recently released Adobe Acrobat 9 Pro-Extended. It can either convert virtually any computer-aided design (CAD) files into a pdf document or capture a 3D model from any 3D CAD applications that support OpenGL (standard graphics language for 3D models). We used the former approach and saved each molecular scene in the virtual reality modeling language (VRML) format.<sup>49,50</sup> VRML is a text file format that can be visualized by VRML viewers and can be edited in text or VRML editors; that is, the VRML format gives more control over the 3D scene for advanced users. The 3D-PDF technique used in the present study in principle is similar to those recently employed by Barnes and Fluke<sup>51</sup> and Kumar et al.<sup>52</sup> However,

the latter used either s2plot<sup>51</sup> or Adobe Acrobat 3D Toolkit,<sup>52</sup> both of which are technically the same as the one we used, for the creation of 3D-PDF.

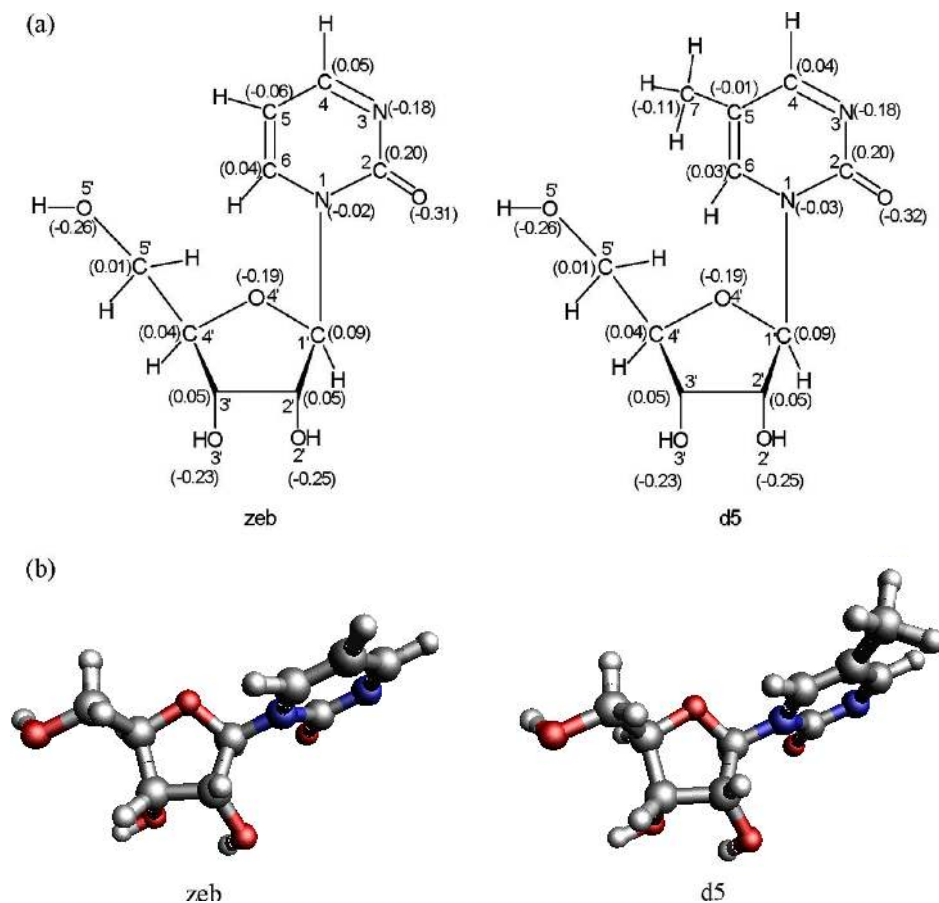
In the present study, VRML files were prepared using Jamberoo,<sup>53</sup> an open-source program for displaying, analyzing, and editing molecular systems, that was developed at the National Computational Infrastructure at the Australian National University. Jamberoo uses Java3d application programming interface (API)<sup>54,55</sup> for high-performance 3D graphics. In Java3d, a collection of objects to be rendered forms hierarchical tree structures or scene graphs. A program module for Jamberoo has been developed that takes Java3d scene graphs as input, explores the scene graph path, and produces graphs that are nearly equivalent to the VRML with the same geometric shapes (spheres, cylinders, surfaces, etc.), geometric transformations, and appearance properties (colors, material attributes, and lighting, etc.).

## 3. Results and Discussion

**3.1. Molecular Structures in 3D-PDF.** Chemical structures and nomenclature of 1-( $\beta$ -D-ribofuranosyl)-2-pyrimidone (zeb) and 1-( $\beta$ -D-ribofuranosyl)-5-methyl-2-pyrimidinone (d5) are given in Figure 1a, together with their state-of-the-art interactive 3D structures in Figure 1b. To view the 3D structures of zeb and d5 interactively on a computer (it requires Adobe Acrobat 8.1 or above), the readers can double click Figure 1b on a personal computer where the pdf file is stored, and the buttons will be displayed once they are activated. Chemically, zeb and d5 differ only by a methyl fragment attached at the C(5) position of the pyrimidine base ring in d5, which shows clearly in their 3D structures.

Optimized geometric and conformational parameters of the nucleoside pair are presented in Table 1. Addition of the methyl ( $\text{CH}_3$ ) group does not impose significant isotropic changes to their geometries. For example, small relaxations in bond lengths are observed in d5 with respect to zeb, which compensates within the aromatic ring so that the perimeters<sup>56</sup> of the hexagon ( $R_6$ ) and the sugar pentagon ( $R_5$ ) rings of the species remain unchanged; that is, 8.25 Å for hexagons and 7.49 Å for pentagons. However, anisotropic properties, such as bond angles, dihedral angles, and dipole moments ( $\mu$ ), exhibit variations between the analogs. For example, the total dipole moment of zeb (6.81 D) has been reduced (6.75 D) in d5 due to the additional methyl fragment, whereas the pseudorotational angle ( $P$ ) reduces  $\sim 6^\circ$ .

Structural details can help us to understand the relationship between conformations and biological activities of a molecule. The glycosyl torsion angle ( $\chi$ ), ( $\text{O}(4')\text{---C}(1')\text{---N}(1)\text{---C}(2)$ ), which determines orientations of the base and sugar in a nucleoside, exhibits an anti orientation for both zeb and d5. It is also seen geometrically that addition of a methyl fragment in the base moiety of zeb causes a certain local perturbation but does not change the orientation of the sugar moiety significantly. For example, the torsional angle ( $\gamma$ ), ( $\text{C}(3')\text{---C}(4')\text{---C}(5')\text{---O}(5')$ ), which describes orientation of the 5'-hydroxyl group relative to the sugar ring, exhibits only a minor change of  $0.71^\circ$ . Moreover, sugar puckering in d5 and zeb also receives only small perturbations from their south orientation, as shown by a small relaxation of the pseudorotational angles and puckering amplitude,  $\nu_m$ . The present results agree well with statistical findings from NMR<sup>57,58</sup> and X-ray<sup>59</sup> for nucleoside analogs;<sup>48</sup> that is, the glycosyl torsion angle in the anti range and the sugar puckering in the south are the necessary orientation for biological activity<sup>60</sup> of nucleosides for the purpose of a strong interaction.



**Figure 1.** Two-dimensional (2D, a) and three-dimensional (3D, b) structures and nomenclature of 1-( $\beta$ -D-ribofuranosyl)-2-pyrimidone (zeb) and 1-( $\beta$ -D-ribofuranosyl)-5-methyl-2-pyrimidinone (d5) with Hirshfeld charges in parentheses of the 2D structures. Double click the 3D structures (part b) on a computer or online to activate the interactive 3D structures.

**3.2. Hirshfeld Charges and Fukui Indices.** Atomic charges are important anisotropic properties to trace behavior of atoms in molecules. Hirshfeld charges of zeb and d5 are presented in parentheses along with the structures given in Figure 1a. The results show that the N sites and the O sites are electron donors (negative charges) because they possess large electron densities, whereas all the C sites are electron acceptors with positive charges except at the C(5) site. Here, site C(5) possesses negative charges because this site connects with the methyl group C(7) in d5. This observation agrees with that of cytidine<sup>40</sup> and other pyrimidine nucleic acids previously studied.<sup>43,61</sup> Within the same species, the oxygen sites attract more intensive electron charges than the nitrogen sites. It is noted that the ketone O(2)=C(2) double bond is strongly polarized in both zeb and d5. Although varied Hirshfeld charges are found, depending on a specific site, the carbon sites on the sugar ring are all positive, whereas the site-specific Hirshfeld charges,  $Q(H)$ , on the base vary significantly, either positively or negatively. Comparing the Hirshfeld charges between zeb and d5, all atomic sites in the nucleoside pair do not change much, except the methyl branch connected at the C(5) site, which is the only negative  $Q(H)$  charged carbon site in d5 and zeb. The negative  $Q(H)$  of zeb (-0.06) at C(5) is largely reduced when the methyl fragment (d5, -0.01) is attached.

Condensed Fukui functions  $f^-$ , which is one of the useful descriptors to monitor electrophilic behavior of a specific site in a molecule, are given in Figure 2. Here, the atomic sites with primes (') are associated with sugar; otherwise, with the base. Distributions of the site-specific Fukui functions of the pair can be divided into three regions: that is, an inactive region with

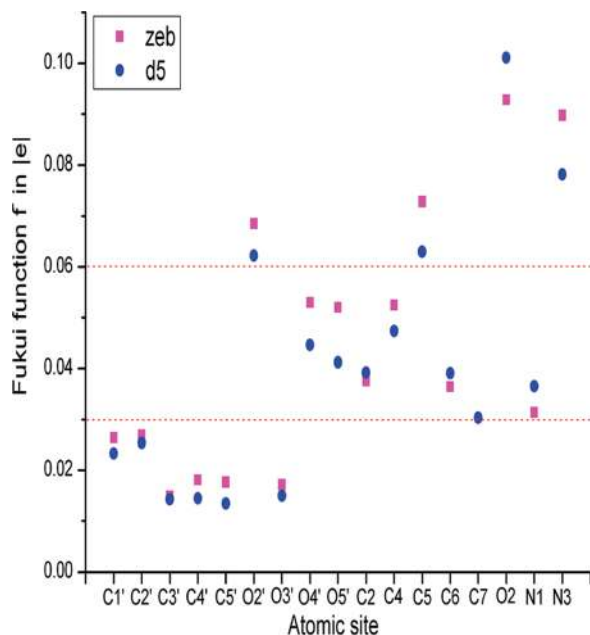
$f^- < 0.03$ , a moderately active region with  $0.03 \leq f^- < 0.06$ , and an active region with  $f^- \geq 0.06$ , as separated by the dashed lines in this figure. In general, the sugar sites are less active than the base sites. The sugar sites are associated with either inactive or moderately active regions, whereas the base sites are dominated by either active or moderately active  $f^-$  regions in this figure. It is noted that only one exception is the sugar O(2'), which is a very active site on the sugar ring. Apart from O(2'), all sites—C(5), O(2), and N(3)—which locate in the active  $f^-$  region either possess large negative  $Q(H)$ s or directly connect with the methyl (C(5)). The oxygen sites on the sugar represent the most diverse Fukui functions in this figure, spanning into all three regions with inactive, moderately active, and active Fukui functions,  $f^-$ . This suggests that the electrophilic reactivities of these oxygen atoms are quite different from one another. In addition, when the methyl group is attached, all carbon sites do not exhibit apparent changes in electrophilic reactivity behavior, except for C(5), and all noncarbon sites split obviously in their Fukui functions, except for O(3'), indicating electrophilic activities of the noncarbon sites in zeb are very different from those in d5.

**Molecular Electrostatic Potentials (MEPs).** Molecular electrostatic potentials (MEPs) of zeb and d5 are presented in Figure 3a–d, with parts a and b being the projection on the “base plane” formed by C(5)–N(3)–N(1), whereas parts c and d are the projection on the “sugar plane” formed by C(1')–O(4')–C(4'). Similarities and differences can be seen clearly in the two-dimensional MEPs. For example, in Figure 3a and b, some local density contributed by methyl is observed. Interestingly, although the MEPs in the base region show similarities, the MEPs

**TABLE 1: Comparison of Geometric and Sugar Puckering Parameters of 1-( $\beta$ -D-Ribofuranosyl)-2-pyrimidone (zeb) and 1-( $\beta$ -D-Ribofuranosyl)-5-methyl-2-pyrimidinone (d5) Using the B3LYP/cc-pVTZ Model**

parameters	zeb	d5
bond length, Å (selected)		
N(1)–C(2)	1.439	1.430
N(3)=C(4)	1.301	1.299
C(4)–C(5)	1.416	1.424
C(6)–N(1)	1.352	1.359
N(1)–C(1')	1.450	1.450
C(2')–C(3')	1.540	1.542
C(3')–C(4')	1.543	1.545
C(4')–O(4')	1.439	1.436
O(4')–C(1')	1.424	1.425
ring perimeters, Å		
R <sub>6</sub> <sup>a</sup>	8.25	8.25
R <sub>5</sub> <sup>a</sup>	7.49	7.49
dihedral angles, deg		
C(2)–N(1)–C(1')–C(2')	114.85	117.05
N(1)–C(1')–O(4')–C(4')	–160.81	–163.25
C(6)–N(1)–C(1')–C(2')	–66.04	–63.44
C(1')–C(2')–C(3')–C(4')	–30.20	–27.69
C(1')–O(4')–C(4')–C(3')	17.33	21.58
dipole moment, D		
$\mu$	6.81	6.75
pseudorotational parameters, deg		
$\chi$ (O(4')–C(1')–N(1)–C(2))	–129.20 (anti)	–127.37 (anti)
$\gamma$ (C(3')–C(4')–C(5')–O(5'))	–65.81	–65.10
pseudorotational angle, $P^b$ , deg		
	137.07 (south)	131.23 (south)
puckering amplitude, $v_m^b$ , deg		
	41.24	42.01
type		
	PYR, C1'-exo	PYR, C1'-exo

<sup>a</sup> Ring perimeters. See ref 56. <sup>b</sup> See ref 48, where  $P = ((v_4 - v_o) - (v_3 - v_1)) / ((2 * v_2 * (\sin 36^\circ + \sin 72^\circ))$ ;  $v_{max} = \text{abs}((v_2) / (\cos P))$ .



**Figure 2.** Site-specific condensed Fukui function variation between 1-( $\beta$ -D-ribofuranosyl)-2-pyrimidone (zeb) and 1-( $\beta$ -D-ribofuranosyl)-5-methyl-2-pyrimidinone (d5) based on the LB94/et-pVQZ model. The variations are the results of methylation.

on the sugar ring are quite different. For example, the C(5')H<sub>2</sub> moiety connecting at the C(4') position in zeb contributes significantly to the “base plane”. Through the “sugar plane” in Figure 3c and d, the methyl group in d5 reduces the pseudorotational angle ( $P$ ) by an apparent 6° so that the MEP closes the loop at C(1') position in the “sugar plane” and builds up at O(4').

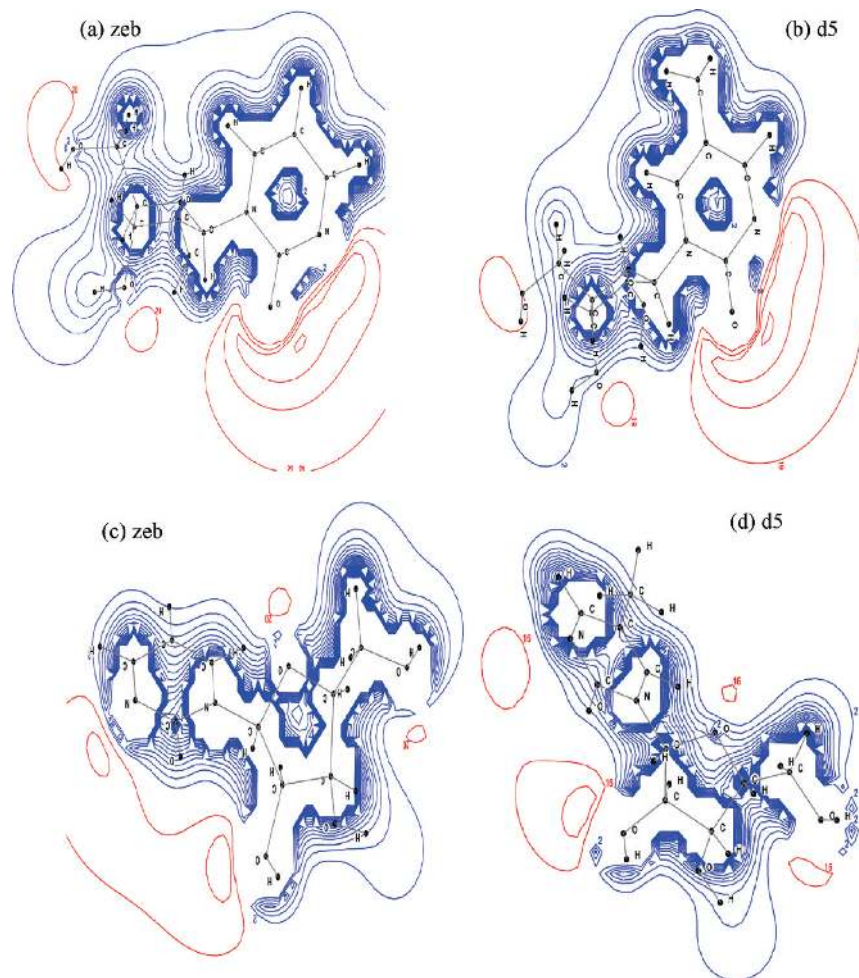
In addition, the methyl opens up the MEP gap along the C(4')–C(5')–O(5') chain.

**Valence Ionization Spectra.** Valence ionization potentials (IPs) directly connect the valence electronic structure of a molecule with the molecular orbital theory. Ionization potentials help the understanding of charge mobility along the DNA double helix. Although experimental photoelectron spectroscopic (PES) measurements for zeb and d5 in the gas phase are not available yet, a number of such measurements for pyrimidine nucleosides and deoxynucleotides have been achieved.<sup>62–64</sup> Accurate theoretical calculations for other IPs of biomolecules the size of nucleosides or nucleotides still present a challenge for other IPs rather than the IP of the HOMO. And most IPs for larger molecules come from semiempirically estimated results rather than from being accurately quantum mechanically calculated. Ab initio methods such as MP2 are hardly applied to calculate IPs of nucleosides with a reasonably large basis set, and MP2 suffers from spin contamination.<sup>64</sup> In addition to the same problem as MP2, the outer valence green function (OVGF) model<sup>65,66</sup> cannot be extended to the inner valence shell. A number of DFT models, such as B3LYP, etc., suffer less from spin contamination than MP2, but self-energy error is significant.<sup>62,64</sup>

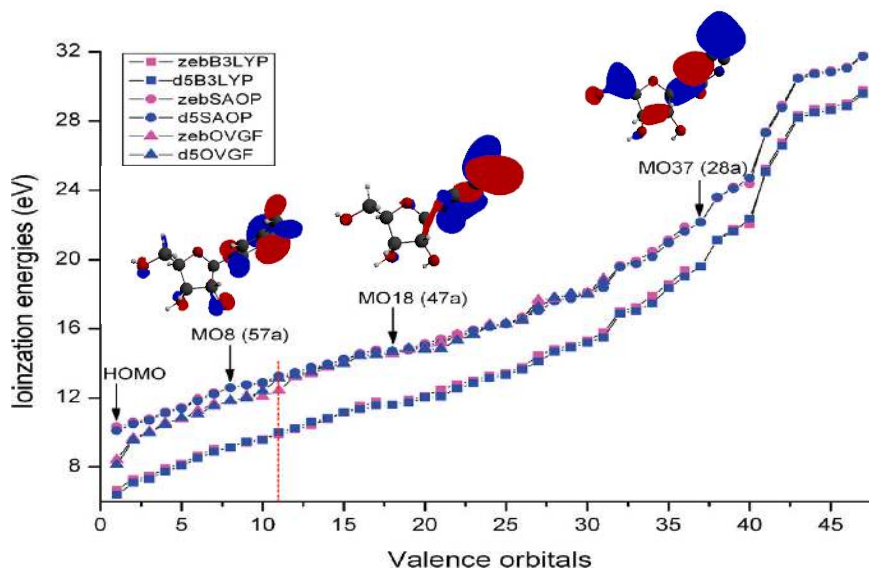
In the present study, valence ionization spectra of zeb and d5 are calculated quantum mechanically using established models, such as the DFT-based statistically averaged orbital potential (SAOP), SAOP/et-pVQZ<sup>37</sup> and OVGF/TZVP.<sup>65,66</sup> The SAOP/et-pVQZ model has been proven<sup>37,64,67</sup> as an accurate DFT-based model for outer valence space as well as inner valence space.<sup>68</sup> Figure 4 compares the vertical IP results for zeb and d5, calculated using the SAOP/et-pVQZ and OVGF/TZVP models. The negative orbital energies ( $-\epsilon_i$ ) produced by the B3LYP/aug-cc-pVTZ model are also given in this figure as a reference<sup>69</sup> to indicate the improvement that the other models have offered.

Figure 4 provides an interesting IP pattern for the nucleoside analogs as well as the quantum mechanical models. The energies increase when the hole moves toward the inner valence shell; that is, the more inward the hole, the larger the energy that is required to ionize the bound electron. After taking consideration of three methyl molecular orbitals (MO, note: HOMO as MO1), that is, MO8, MO18, and MO37 in d5 (as marked in the figure) in the valence space, the IP patterns produced by the SAOP/et-pVQZ and OVGF/TZVP models are consistent except for the HOMOs, which represents a sharp change of vertical IPs compared to other outer valence IPs. It has been a known fact that the SAOP/et-pVQZ model does not produce very accurate first IPs.<sup>67,70,71</sup> In addition, the vertical IPs larger than 20 eV may be subject to satellite shake-ups and shake-offs,<sup>72</sup> which may be partly considered in the SAOP model from the statistical orbital average. It is seen in Figure 4 that the vertical IPs produced by the SAOP/et-pVQZ and OVGF/TZVP models agree well in the outer valence space (IP < 20 eV), except for the outermost valence orbitals of IP < 13 eV, as marked by a vertical dashed line in the figure. The SAOP model continues producing consistent IPs into the inner valence space where the OVGF model ceases (at MO31). Indeed, all three orbitals exhibit methyl-dominant electron density distributions (shown in the insert 3D orbitals), which largely concentrate on the base moiety of the nucleosides.

Valence binding energy spectra of molecules depend not only on their binding energies but also on the distribution or intensities of the ionized states in the region. Figure 5a compares the simulated PES of zeb and d5 on the basis of the SAOP/et-pVQZ model (with a fwhm of 0.40 eV). From a spectral point



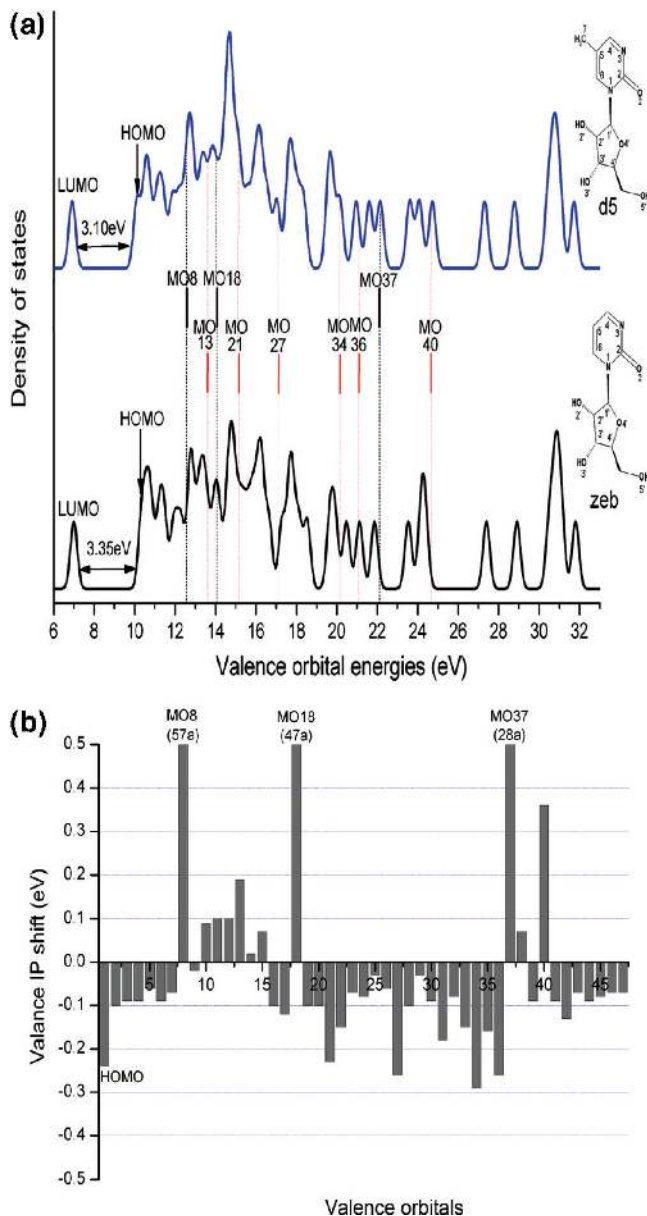
**Figure 3.** MEP of 1-( $\beta$ -D-ribofuranosyl)-2-pyrimidone (zeb) and 1-( $\beta$ -D-ribofuranosyl)-5-methyl-2-pyrimidinone (d5) projected onto the base plane of C(5)–N(3)–N(1) (a, b) and sugar plane of C(1')–O(4')–C(4') (c, d). Calculations are based on the B3LYP/aug-cc-pVTZ model.



**Figure 4.** Valence vertical ionization potentials of 1-( $\beta$ -D-ribofuranosyl)-2-pyrimidone and 1-( $\beta$ -D-ribofuranosyl)-5-methyl-2-pyrimidinone, generated using SAOP/et-pVQZ, OVGf/TZVP, and B3LYP/aug-cc-pVTZ. The nomenclature starts from HOMO as MO1 and inward. The differences indicate species dependency (by color) and model dependency (by symbol), which could be separated by orbitals or energy regions. Double click the insert orbitals, MO8, MO18, and MO37 on a computer or online to activate the interactive 3D structures.

of view, the inner valence shell of the nucleoside pair exhibits certain similarities; in particular, in the region above 25 eV. The outer valence region of the spectra, however, is quite different. The more intensive IP distribution incorporated with

the Gaussian line shape function employed in the simulation contributes to the apparent differences in the PES in the energy region of IP < 25 eV. It is seen that the energy gaps ( $\Delta E_{\text{HOMO-LUMO}}$ ) between the highest occupied molecular orbital



**Figure 5.** (a) Valence electron spectra of zeb (lower) and d5 (upper) based on the SAOP/et-pVQZ model. Inner-valence region (IP > 19 eV) indicates global energies shift, but outer-valence region reveals differences in orbital distribution and density. The vertical bars are primary methyl-dependent orbitals (black) and secondary methyl-dependent orbitals (red). (b) Orbital-based valence ionization energy differential spectrum of zeb and d5 due to the methyl attachment. Calculations are based on SAOP/et-pVQZ. The nomenclature starts from HOMO as MO1. That is, for zeb, HOMO is MO1 or 60a, whereas for d5, HOMO is MO1 or 64a.

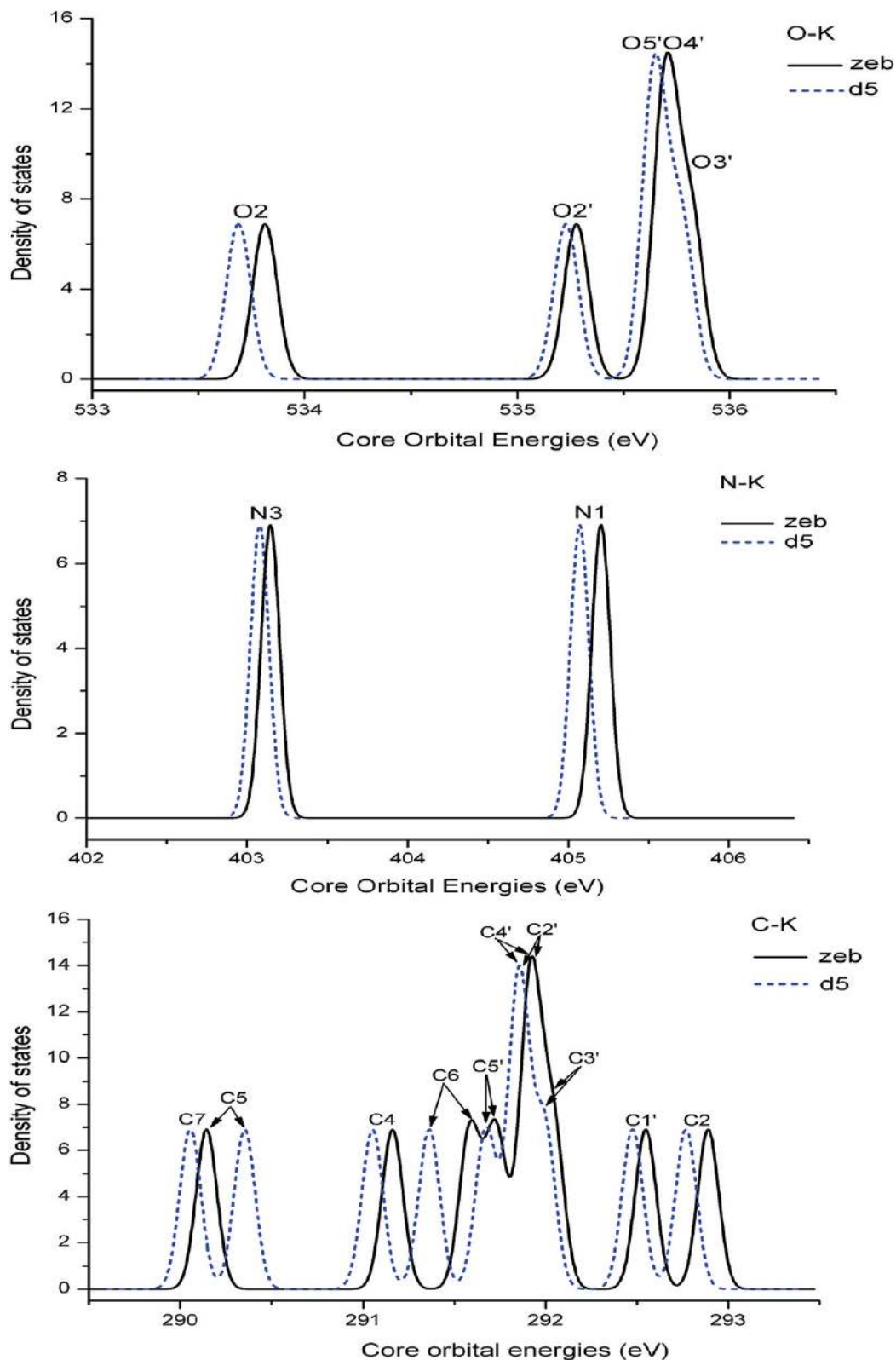
and lowest unoccupied molecular orbital of the molecule pair differ by as much as 0.25 eV. For example, the  $\Delta E_{\text{HOMO-LUMO}}$  is given by 3.35 eV for zeb but reduces to 3.10 eV for d5 in the present calculation. The reduction of the HOMO-LUMO gap in d5 indicates that methyl is capable of activating the nucleoside, which may contribute to the potency if it is used as a drug.

To further monitor the methyl effect in the entire valence shell, the orbital-based differential IP spectrum, which corresponds to the orbital-based vertical IPs of d5 with respect to zeb (i.e.,  $\Delta\text{IP} = \text{IP}(\text{d5}) - \text{IP}(\text{zeb})$ ), are given in Figure 5b. Three large IPs of MO8, MO18, and MO37 in this figure are contributed primarily by the methyl moiety in d5. However,

the IP differential spectrum clearly indicates that methyl effects in the valence space are delocalized: almost all valence orbitals of d5 are more or less affected by the methyl group. The majority of valence orbital IP differences are negative, and therefore, to reduce the total electronic energy of d5. From an energy point of view, the IP differential spectrum between zeb and d5 due to methyl attachment is less different from the IP spectrum between two didehydrodeoxycytidine nucleoside isomer pairs, 1',2'-didehydro-3',4'-deoxycytidine and 3',4'-didehydro-2',4'-deoxycytidine,<sup>67</sup> which differ by only the position of a C=C double bond in the sugar ring. Apart from the methyl-dominant MO8, MO18, and MO37 orbitals, a group of secondary orbitals can be found in this figure as the HOMO, MO13, MO21, MO27, MO34, MO36, and MO40 ( $|\Delta\text{IP}| > 0.2$  eV). The primary and secondary methyl-related orbitals are marked in the valence binding energy spectra of Figure 5a as vertical bars. It is these orbitals that are the dominant sources of variations in the spectrum of d5, unlike small molecules, such as L-alanine,<sup>73</sup> in which the methyl group contributes to four clearly identifiable orbitals in alanine with respect to glycine. However, as shown in Figure 5, contribution from the methyl group in d5 diffuses over the valence space, particularly the outer valence space.

**Core Binding Energy Spectra.** Nucleosides are large molecules in quantum mechanics. As seen in the previous section, the methylation effect is delocalized into a large number of orbitals in the valence space. The complex electronic structural changes brought in by the methyl group affect almost all valence orbitals. However, a more localized picture of methyl effects in the inner shell can be observed. Figure 6 provides the simulated core binding energy spectra of the molecule pair, based on the LB94/et-pVQZ model. An energy resolution (fwhm) of 0.40 eV, which is the same resolution employed in the simulation of the PES for cytidine,<sup>74</sup> is employed in the simulated spectra. Because methyl does not have other non-hydrogen atoms other than C, the O-K and N-K spectra of zeb and d5 are nearly identical, with only subtle global red shifts in d5. The double-bonded O-K site, O(2), of the ketone in the base is well-separated from the single-bonded sugar O-K sites (as labeled by primes). In the sugar O-K sites, the O(2') site is separated from other sugar O-K sites, as noted in their very different condensed Fukui functions in previous sections. The N-K spectra exhibit a simple structure with well-separated N(3) and N(1) peaks, which are associated with the imino ( $-\text{N}=\text{C}$ ) and amino ( $-\text{N}-\text{C}$ ) N-K sites, respectively, as noted in a study of cytidine by Wang.<sup>75</sup> The O-K and N-K spectra indicate that the chemical environment of a double bond is different from a single bond.

Methyl effects on the C-K spectrum of d5 with respect to zeb are also revealed in Figure 6. It indicates that the core shell of the species is different but related. The energy ordering and positions of sugar carbons exhibit small variations. For example, the energy ordering of the sugar C1s sites is given by  $\text{C}(5') < \text{C}(4') < \text{C}(2') < \text{C}(3') < \text{C}(1')$  in the energy region of 291.1–292.5 eV for both zeb and d5. The fact that the sugar C1s sites are more congested than the base C1s sites indicates that the sugar carbons are chemically more similar to one another due to the single bonds in the sugar moiety. The ketone carbon, C(2), always locates as the highest binding energy peak in the spectra, which makes C(2) the most difficult site to ionize. Energy ordering of the C1s sites in the base, which is  $\text{C}(7) < \text{C}(5) < \text{C}(4) < \text{C}(6) < \text{C}(2)$ , remains unchanged in zeb and d5. However, the positions of the C1s sites shift apparently in order to accommodate the methyl carbon, C(7), a



**Figure 6.** Comparison of the O–K, N–K, and C–K binding energy spectra of zeb and d5, simulated using LB94/et-pVQZ calculation with a fwhm of 0.40 eV.

site which directly connects on the C(5) site. The C(5) site and its neighbors C(6) and C(4) shift in opposite directions; that is, C(5) exhibits a blue shift (toward larger energy), whereas C(6) and C(4) result in red shift (toward lower energy). Finally, the addition of methyl splits the peak formed by C(6) and C(5'), indicating that their similar chemical environment in zeb no more exists in d5, as seen in this figure.

## Conclusions

Geometric and electronic impact of the methyl group with respect to two related nucleoside drugs, 1-( $\beta$ -D-ribofuranosyl)-2-pyrimidone (zebularine) and its methylated analog, 1-( $\beta$ -D-ribofuranosyl)-5-methyl-2-pyrimidinone (d5), have been studied quantum mechanically using density functional theory models.

To obtain a thorough insight of the structures and methyl-dominated orbitals in 3D space, a state-of-the-art interactive 3D graph in a portable document format (pdf) file, 3D-PDF, is developed and presented (Figures 1 and 4). It is found that methyl attachment has small effects on the geometries and sugar puckering. However, electronic properties, such as condensed Fukui functions, change more apparently. In the valence space, the methyl fragment affects binding energy spectra more significantly in the outer-valence region rather than in the inner-valence space. The methyl effects, however, are more delocalized and diffused into the valence space, which is quite different from the methyl effects in smaller molecules, such as alanine-glycine.<sup>70</sup> However, the ionization energy differential spectrum between zeb and d5 exhibits less significant changes in valence energy than a didehydrodeoxycytidine nucleoside pair, which differs only by the location of a C=C double bond in the sugar ring. The impact of methyl in the inner shell can be observed in the nucleic base or the sugar. It is found in the present study that the methyl group affects core binding energy spectra of the nucleoside pair with small global red shifts for the O-K, N-K, and sugar C-K spectra, whereas methyl affects the C-K spectrum of the base carbon sites locally with respect to the sites that directly connect to methyl.

**Acknowledgment.** This work is supported by an award under the Merit Allocation Scheme on the National Computational Infrastructure at the ANU. L.S. acknowledges a Swinburne University Postgraduate Research Award (SUPRA). Useful discussions and assistance from Dr. Saumitra Saha and Dr. Quan Zhu must also be acknowledged.

## References and Notes

- (1) Cohen, R. M.; Wolfenden, R. *J. Biol. Chem.* **1971**, *246*, 7561.
- (2) Ipata, P. L.; Cercignani, G.; Balestreri, E. *Biochemistry* **1970**, *9*, 3390.
- (3) Achar, B. S.; Maller, R. K.; Vaidyanathan, C. S. *Indian J. Biochem.* **1966**, *3*, 133.
- (4) Tomchick, R.; Saslaw, L. D.; Waravdekar, V. S. *J. Biol. Chem.* **1968**, *243*, 2534.
- (5) Chabner, B. A.; Johns, D. G.; Coleman, C. N.; Drake, J. C.; Evans, W. H. *J. Clin. Invest.* **1974**, *53*, 922.
- (6) Kim, C. H.; Marquez, V. E.; Mao, D. T.; Haines, D. R.; McCormack, J. J. *J. Med. Chem.* **1986**, *29*, 1374.
- (7) Holy, A.; Ludzisa, A.; Votruba, I.; Sediva, K.; Pischel, H. *Collect. Czech. Chem. Commun.* **1985**, *50*, 393.
- (8) Betts, L.; Xiang, S.; Short, S. A.; Wolfenden, R.; Carter, C. W. *J. Mol. Biol.* **1994**, *235*, 635.
- (9) Xiang, S.; Short, S. A.; Wolfenden, R.; Carter, C. W., Jr. *Biochemistry* **1996**, *35*, 1335.
- (10) Carlow, D.; Wolfenden, R. *Biochemistry* **1998**, *37*, 11873.
- (11) Lewis, J. P.; Carter, C. W.; Hermans, J.; Pan, W.; Lee, T. S.; Yang, W. *J. Am. Chem. Soc.* **1998**, *120*, 5407.
- (12) Matsubara, T.; Ishikura, M.; Aida, M. *J. Chem. Inf. Model.* **2006**, *46*, 1276.
- (13) Matsubara, T.; Dupuis, M.; Aida, M. *Chem. Phys. Lett.* **2007**, *437*, 138.
- (14) Kedzierski, P.; Sokalski, W. A.; Cheng, H.; Mitchell, J.; Leszczynski, J. *Chem. Phys. Lett.* **2003**, *381*, 660.
- (15) Driscoll, J. S.; Marquez, V. E.; Plowman, J.; Liu, P. S.; Kelley, J. A.; Barchi, J. J. *J. Med. Chem.* **1991**, *34*, 3280.
- (16) McCormack, J. J.; Marquez, V. E.; Liu, P. S.; Vistica, D. T.; Driscoll, J. S. *Biochem. Pharmacol.* **1980**, *29*, 830.
- (17) Frick, L.; Yang, C.; Marquez, V. E.; Wolfenden, R. *Biochemistry* **1989**, *28*, 9423.
- (18) Holleran, J. L.; Parise, R. A.; Joseph, E.; Eiseman, J. L.; Covey, J. M.; Glaze, E. R.; Lyubimov, A. V.; Chen, Y.-F.; D'Argenio, D. Z.; Egorin, M. J. *Cancer Ther.: Preclin.* **2005**, *11*, 3862.
- (19) Zhou, L.; Cheng, X.; Connolly, B. A.; Dickman, M. J.; Hurd, P. J.; Hornby, D. P. *J. Mol. Biol.* **2002**, *321*, 591.
- (20) Cheng, J. C.; Weisenberger, D. J.; Gonzales, F. A.; Liang, G.; Xu, G.-L.; Hu, Y.-G.; Marquez, V. E.; Jones, P. A. *Mol. Cell. Biol.* **2004**, *24*, 1270.
- (21) Yoo, C. B.; Cheng, J. C.; Jones, P. A. *Biochem. Soc. Trans.* **2004**, *32*, 910.
- (22) Marquez, V. E.; Kelley, J. A.; Agbaria, R.; Ben-Kasus, T.; Cheng, J. C.; Yoo, C. B.; Jones, P. A. *Nucleosides Nucleotides Nucleic Acids* **2005**, *1058*, 246.
- (23) Bestor, T. H. *Nature* **1998**, *393*, 311.
- (24) Bestor, T. H. *Hum. Mol. Genet.* **2000**, *9*, 2395.
- (25) Jones, P. A.; Takai, D. *Epigenetics* **2001**, *293*, 1068.
- (26) Bestor, T. H.; Verdine, G. L. *Curr. Opin. Cell. Biol.* **1994**, *6*, 380.
- (27) Cheng, X. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 293.
- (28) Redaschi, N.; Bickle, T. A. DNA Restriction and Modification Systems. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd ed.; Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Brooks Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, R., Schaechter, M., Umberger, H. E., Eds.; ASM Press: Washington, DC, 1996; pp 773-781.
- (29) Rountree, M. R.; Bachman, K. E.; Herman, J. G.; Baylin, S. B. *Oncogene* **2001**, *20*, 3156.
- (30) Huq, F. J. *Pharmacol. Toxicol.* **2006**, *1*, 317.
- (31) Dowd, C. L.; Sutch, B. T.; Haworth, I. S.; Eritja, R.; Marquez, V. E.; Yang, A. S. *Nucleosides Nucleotides Nucleic Acids* **2008**, *27*, 131.
- (32) Connolly, B. A.; Newman, P. C. *Nucleic Acids Res.* **1989**, *17*, 4957.
- (33) Gildea, B.; McLaughlin, L. W. *Nucleic Acids Res.* **1989**, *17*, 2261.
- (34) Wu, P.; Nordlund, T. M.; Gildea, B.; McLaughlin, L. W. *Biochemistry* **1990**, *29*, 6508.
- (35) Kistler, K. A.; Matsika, S. *Photochem. Photobiol.* **2007**, *83*, 611.
- (36) Page, L. [http://www.theregister.co.uk/2009/01/06/nature\\_3d\\_dendrogram\\_gas\\_cloud/](http://www.theregister.co.uk/2009/01/06/nature_3d_dendrogram_gas_cloud/); accessed January 6, 2009.
- (37) Schipper, P. R. T.; Gritsenko, O. V.; van Gisbergen, S. J. A.; Baerends, E. J. *J. Chem. Phys.* **2000**, *112*, 1344.
- (38) Van Leeuwen, R.; Baerends, E. J. *Phys. Rev. A* **1994**, *49*, 2421.
- (39) Chong, D. P.; Van Lenthe, E.; van Gisbergen, S.; Baerends, E. J. *J. Comput. Chem.* **2004**, *25*, 1030.
- (40) Hirshfeld, F. L. *Theor. Chim. Acta* **1977**, *44*, 129.
- (41) Ayers, P. *Theor. Chem. Acc.* **2006**, *115*, 370.
- (42) Guerra, C. F.; Handgraaf, J.-W.; Baerends, E. J.; Bickelhaupt, F. M. *J. Comput. Chem.* **2004**, *25*, 189.
- (43) Saha, S.; Wang, F.; Brunger, M. J. *Mol. Simul.* **2006**, *32*, 1261.
- (44) Zhu, Q.; Wang, F.; Ivanova, E. *J. Synchrotron Radiat.* **2009**, *16*, 545.
- (45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Kelna, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, A.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Ak-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian03*; Gaussian, Inc.: Wallingford, CT, 2004.
- (46) Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.
- (47) Baerends, E. J.; Autschbach, J.; Bérces, A.; Bo, C.; Boerrigter, P. M.; Cavallo, L.; Chong, D. P.; Deng, L.; Dickson, R. M.; Ellis, D. E.; van Faassen, M.; Fan, L.; Fischer, T. H.; Guerra, C. F.; van Gisbergen, S. J. A.; Groeneveld, J. A.; Gritsenko, O. V.; Grüning, M.; Harris, F. E.; Van den Hoek, P.; Jacobsen, H.; Jensen, L.; van Kessel, G.; Kootstra, F.; Van Lenthe, E.; McCormack, D.; Michalak, A.; Osinga, V. P.; Patchkovskii, S.; Philipsen, P. H. T.; Post, D.; Pye, C. C.; Ravenek, W.; Ros, P.; Schipper, P. R. T.; Schreckenbach, G.; Snijders, J. G.; Solà, M.; Swart, M.; Swerhone, D.; te Velde, G.; Vernooijs, P.; Versluis, L.; Visser, O.; Wang, F.; van Wezenbeek, E.; Wiesenecker, G.; Wolff, S. K.; Woo, T. K.; Yakovlev, A.; Ziegler, T. *Theoretical Chemistry*; ADF 2006.01, SCM; Vrije Universiteit: Amsterdam, The Netherlands, 2006.
- (48) Sun, G.; Voigt, J. H.; Filippov, I. V.; Marquez, V. E.; Nicklaus, M. C. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1752.
- (49) <http://www.web3d.org/x3d/vrml/>.
- (50) Ames, A. L.; Nadeau, D. R.; Moreland, J. L. *VRML 2.0 Sourcebook*, 2nd ed.; Wiley: New York, 1996.
- (51) Barnes, D. G.; Fluke, C. J. *N. Astron.* **2008**, *13*, 599.
- (52) Kumar, P.; Ziegler, A.; Ziegler, J.; Uchanska-Ziegler, B.; Ziegler, A. *Trends Biochem. Sci.* **2008**, *33*, 408.
- (53) Jamberoo. <http://sf.anu.edu.au/~vvv900/cct/appl/jmoleitor/>.
- (54) <http://java.sun.com/javase/technologies/desktop/java3d/> (accessed February 3, 2009).
- (55) Selman, D. *Java 3D Programming*, 1st ed.; Manning Publications: Greenwich, CT, 2002.



- (56) Wang, F.; Downton, M. T.; Kidwani, N. *J. Theor. Comput. Chem.* **2005**, *4*, 247.
- (57) Painter, G. R.; Aulabaugh, A. E.; Andrews, C. W. *Biochem. Biophys. Res. Commun.* **1993**, *191*, 1166.
- (58) Painter, G. R.; Andrews, C. W.; Furman, P. A. *Nucleosides Nucleotides Nucleic Acids* **2000**, *19*, 13.
- (59) Birnbaum, G. I.; Giziewicz, J.; Gabe, E. J.; Lin, T. S.; Prusoff, W. H. *Can. J. Chem.* **1987**, *65*, 2135.
- (60) Camerman, A.; Mastropaolo, A.; Camerman, N. *Proc. Natl. Acad. Sci., U.S.A.* **1987**, 8239.
- (61) Wang, F.; Zhu, Q.; Ivanova, E. *J. Synchrotron Radiat.* **2008**, *15*, 624.
- (62) Yang, X.; Wang, X.-B.; Vorpapel, E. R.; Wang, L.-S. *Proc. Natl. Acad. Sci., U.S.A.* **2004**, *101*, 17588.
- (63) Novak, I.; Kovač, B. *J. Org. Chem.* **2003**, *68*, 5777.
- (64) Slavicek, P.; Winter, B.; Faubel, M.; Bradforth, S. E.; Jungwirth, P. *J. Am. Chem. Soc.* **2009**, *131*, 6460.
- (65) Cederbaum, L. S. *J. Phys. B: Atom. Molec. Phys.* **1975**, *8*, 290.
- (66) Ortiz, J. V. *J. Chem. Phys.* **1988**, *89*, 6348.
- (67) Wang, F. *J. Phys. Chem. B* **2007**, *111*, 9628.
- (68) Wang, F.; Pang, W. *Mol. Simul.* **2007**, *33*, 1173.
- (69) Tu, G.; Tu, Y.; Vahtras, O.; Agren, H. *Chem. Phys. Lett.* **2009**, *468*, 294.
- (70) Falzon, C. T.; Wang, F. *J. Chem. Phys.* **2005**, *123*, 214307.
- (71) Saha, S.; Wang, F.; Falzon, C. T.; Brunger, M. J. *J. Chem. Phys.* **2005**, *123*, 124315.
- (72) Ning, C. G.; Huang, Y. R.; Zhang, S. F.; Deng, J. K.; Liu, K.; Luo, Z. H.; Wang, F. *J. Phys. Chem. A* **2008**, *112*, 11078.
- (73) Falzon, C. T.; Wang, F.; Pang, W. *J. Phys. Chem. B* **2006**, *110*, 9713.
- (74) Thompson, A.; Saha, S.; Wang, F.; Tsuchimochi, T.; Nakata, A.; Imamura, Y.; Nakai, H. *Bull. Chem. Soc. Jpn.* **2009**, *82*, 187.
- (75) Wang, F. *Micro Nano Lett.* **2006**, *1*, 23.

JP901678G