



X-ray photoelectron spectroscopy of Thymine and 5-Bromouracil studied by Symmetry-Adapted-Cluster Configuration-Interaction (SAC-CI) theory

Yusaku I. Kurokawa^a, Hiroshi Nakatsuji^{a,*}, Misaki Hirato^b, Akinari Yokoya^c

^a Quantum Chemistry Research Institute, Kyoto Technoscience Center 16, 14 Yoshida Kawara-machi, Sakyo-Ku, Kyoto 606-8305, Japan

^b Graduate School of Science and Engineering, Ibaraki University, Mito, Ibaraki 310-8512, Japan

^c Institute for Quantum Life Science, National Institutes for Quantum Science and Technology, 2-4 Shirakata, Tokai, Ibaraki 319-1106, Japan

ARTICLE INFO

This paper is dedicated to the Late Professor Kozo Kuchitsu.

Keywords:

X-ray Photoelectron Spectroscopy

5-bromouracil

Thymine

SAC-CI method

Nature of the inner-core ionization spectra

ABSTRACT

The X-ray Photoelectron Spectra (XPS) of thymine (T) and 5-Bromouracil (BrU) were studied and assigned using Symmetry-Adapted-Cluster Configuration-Interaction (SAC-CI) theory. The differences between theory and experiment were within 1.0 eV when compared with the spectra measured in vacuum conditions. It was clarified that inner-core ionization processes involve many electronic processes like ionizations, accompanying excitations, reorganizations, and electron correlations. Therefore, accurate calculations require both good theory and good basis set. The present SAC-CI calculations with extended basis sets including even Rydberg basis functions may satisfy these requirements.

1. Introduction

X-ray photoelectron spectroscopy (XPS) is an important tool to elucidate the electronic structure of the target molecule: the XPS spectra provide the information of the electronic structure of the molecule not only of the ionizing electron, but also of all the electrons of the molecule. Therefore, to understand the implications of the observed XPS spectra, theoretical examinations of the observed spectra using a reliable theoretical method is necessary. We use here the symmetry-adapted-cluster (SAC)/SAC configuration-interaction (SAC-CI) method, because this is a standard theory that is applicable to all the ground, excited, ionized, and electron attached states of molecules in the same reliable frameworks including electron correlations [1–11]. The SAC/SAC-CI method is convenient also since it is widely distributed as one of the important theoretical methodologies included in Gaussian program package [12]. This method has been applied to many subjects related to the photoelectron spectroscopy of valence, inner-valence, and inner-core electrons [5–11].

Recently, Yokoya and co-workers have measured the XPS spectra of film samples of Thymine (T) and 5-bromouracil (BrU) in the wide range of 0–2000 eV [13]. These molecules are important because halogenation of genomic DNA could raise the sensitivity of the cells to ionizing

radiations, particularly the cells possessing halogenated genomic DNA show enhanced cell-killing effects through the ionization [14]. To clarify the mechanism of this increasing sensitivity to the ionizing radiation, many studies using synchrotron X-rays exposed to various DNA related molecules have been done experimentally [15–20]. In this study, we study the XPS spectra of T and BrU in the single-molecular vacuum condition by calculating the ionization spectra of their inner-shell electrons by the SAC/SAC-CI method. Though the film sample and the single-molecular vacuum condition are much different from the actual conditions in DNAs, we expect that some insights may be obtained from this study for the sensitivity of the cells to ionizing radiations.

2. Computational details

We first optimized the geometry of T and BrU by the density functional theory (DFT) with the B3LYP functional. The geometries of T and BrU were constrained as *Cs* symmetry. The basis sets used for this purpose were the D95* for C, N, and O atoms, Dunning-Hay's DZP basis for H atom, and Huzinaga's (43321/4321/31) basis set augmented with a polarization *d* function ($\alpha = 0.389$) for Br atom [21]. These basis sets were used for geometry optimization and referred to as Basis I. At the optimized geometry sketched in Fig. 1, two hydrogen atoms in the

* Corresponding author.

E-mail address: h.nakatsuji@qcri.or.jp (H. Nakatsuji).

<https://doi.org/10.1016/j.cplett.2022.139957>

Received 30 May 2022; Accepted 8 August 2022

Available online 11 August 2022

0009-2614/© 2022 Elsevier B.V. All rights reserved.

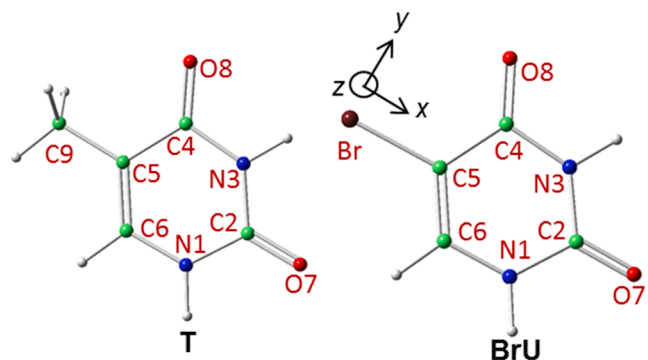


Fig. 1. Optimized geometry of Thymine (T) and 5-bromouracil (BrU), and their atomic labels. The 2p, 3p and 3d orbitals of Br atom are defined along with the local Cartesian coordinates shown in the figure. The Br 3d_{xy} orbital is directing toward the x direction.

methyl group of T were directing toward the neighbouring O atom keeping C_s symmetry, and the other hydrogen atom of the methyl group was in the molecular plane and directing opposite direction. The lowest harmonic frequencies at the optimized geometries were both positive.

With this optimized geometry, we performed the SAC/SAC-CI calculations using relatively large basis sets to describe the orbital reorganization and electron correlations, which will be referred to as Basis II. Namely, for C, N, and O atoms, the outer-most primitive Gaussian functions of D95* were split to (51111/311), the polarization *d* function of the D95* was substituted by the double-zeta *d* functions ($\alpha_d(\text{C}) = 1.3350, 0.28800, \alpha_d(\text{N}) = 1.9860, 0.4120, \alpha_d(\text{O}) = 2.7040, 0.5350$), and further augmented with the double-zeta Rydberg 3s, 3p, and 3d functions ($\alpha_s(\text{C}) = 0.04370, 0.01725, \alpha_p(\text{C}) = 0.03990, 0.01575, \alpha_d(\text{C}) = 0.02850, 0.01125, \alpha_s(\text{N}) = 0.05320, 0.02100, \alpha_p(\text{N}) = 0.04750, 0.01875, \alpha_s(\text{O}) = 0.06080, 0.02400, \alpha_p(\text{O}) = 0.05320, 0.02100, \alpha_d(\text{O}) = 0.02850, 0.01125$). For Br atom, the Huzinaga's (43321/4321/31) basis set was split to be (311212111/3121111/211) and augmented with the polarization *f* functions ($\alpha_f = 0.55150, 0.25800$) and the single-zeta Rydberg 3s, 3p, 3d functions ($\alpha_s = 0.04559, 0.03514, 0.10470$). For H atom, the Dunning-Hay's DZP basis was split to (211/1). These basis sets are referred to as Basis II. In the SAC/SAC-CI calculations, all the molecular orbitals (MOs) were used as the active space. All single and selected double excitation operators concerning the target inner-shell orbitals were included and perturbation selection was carried out with the threshold sets of LevelThree. The ionization cross sections were calculated using the monopole approximation to estimate the relative intensities of the peaks. Both the initial- and final-ionic-state correlation effects were included [22]. All calculations were performed with Gaussian 16 rev. C02.[12].

3. Results and discussion

3.1. Molecular Orbitals

The molecular orbitals obtained by the HF calculation were used in the subsequent SAC/SAC-CI calculations. The atomic labels were shown in Fig. 1. The inner-core orbitals of both T and BrU were localized on one atom and almost isolated from other atomic orbitals, as shown in Figs. S1 and S2 in Supporting Information. The energy orderings of the 1s orbitals of O, N, C of T were the same as those of the corresponding atoms in BrU. The orbital energies of the core orbitals, Br 1s, 2s, 2p, 3s, 3p, 3d, O 1s, N 1s, and C 1s, were separated from others by more than 30 eV. Therefore, we could specify the atomic orbitals from which the electrons were ionized. The O 1s, N 1s and C2 1s orbital energies of BrU were slightly lower (about 0.37–0.50 eV) than those of corresponding atoms in T, but the C5 1s orbital energy was much stabilized by 1.48 eV compared with the C5 1s orbital of T. This is the effect from the bonding

Br atom in BrU, which withdraws the sigma electron from C5 1s orbital, leading to the stabilization. Likewise, the 1s orbitals of the neighbouring atoms, C4 and C6, were somewhat stabilized by 0.53 and 0.72 eV, respectively.

3.2. Inner-core Ionization Spectra of Thymine

The inner-core spectra of T are shown in Table 1 and Fig. 2. The O 1s and N 1s ionized states are described with the SAC-CI method by one-electron ionized configuration with coefficients $|c| > 0.89$, as shown in Table 1 under “Main Configurations”. The C 1s first and second ionized states are described by one-electron ionized configuration from C2 and C4 with coefficients $|c| > 0.88$. The C 1s third ionized state is described as ionizing from C6 with coefficients $|c| \approx 0.88$, and with an additional excitation from MO 33 to MO 88, namely the configuration $(\text{C6 } 1s)^{-1}(\text{MO } 33)^{-1}(\text{MO } 88)$ with $|c| \approx 0.11$, where MOs 33 and 88 are the HOMO and a Rydberg orbital, respectively. The C 1s fourth ionized state is described interestingly by the *mixed* configurations ionizing from C5 and C9 with coefficients of 0.81 and 0.37, respectively, and additional two-electron excited configurations of $(\text{C5 } 1s)^{-1}(\text{MO } 33)^{-1}(\text{MO } 211)$ with coefficients ≈ 0.10 , where MO 33 is HOMO and MO 211 is a Rydberg orbital. These two ionizations are interesting because the HOMO \rightarrow Rydberg excitations accompany with the C 1s ionizations, showing an important role of the Rydberg orbitals assisting the C 1s inner-core ionization. This also shows importance of the Rydberg-type basis function for studying XPS spectra of the inner 1s electron ionizations. The C 1s fifth ionized state is described by the configurations ionizing from C9 and C5 with coefficients $|c| \approx 0.87$ and 0.18, respectively. Thus, the fourth and fifth ionized states are mixed $(1s)^{-1}$ states from C5 and its bonding C9 of methyl group (see Fig. 1). Such a mixing implies that the 1s orbital energy levels of C5 and C9 are close (285.8 and 285.1 eV from film sample data in Table 1) and therefore, their ionization spectra appear overlappingly. This clearly explains why the intensity of the peak at 285.2 eV is about twice stronger than that for 288.5 eV in Fig. 2 for C 1s. Further, this mixing, and such a stronger peak, both do not exist in BrU, because in BrU this methyl group is replaced with Br.

Compared with the inner-core ionization spectra of Thymine measured by Plekan *et al.* [23] and Itälä *et al.* [24] in the single-molecule vacuum condition, the SAC-CI IPs were calculated to be higher by only $\Delta E = 0.8\text{--}1.3$ eV (0.95 eV in average) for the O 1s and N 1s, and only $\Delta E = 0.5\text{--}1.2$ eV (1.0 eV in average) for the C 1s states. It is concluded that the SAC-CI inner-shell (O 1s, N 1s, and C 1s) IP agrees with the spectrum measured in single-molecule vacuum condition, if it is shifted by 0.9–1.0 eV lower. On the other hand, the errors of the IPs calculated by the ADC(4) theory [23] are $\Delta E = 1.6, 1.2$, and 1.0 eV for O 1s, N 1s, and C 1s IPs, respectively, while those by the present SAC-CI calculations are a little better than them as $\Delta E = 1.3, 0.8$, and 0.9 eV for O 1s, N 1s, and C 1s IPs, respectively, in average.

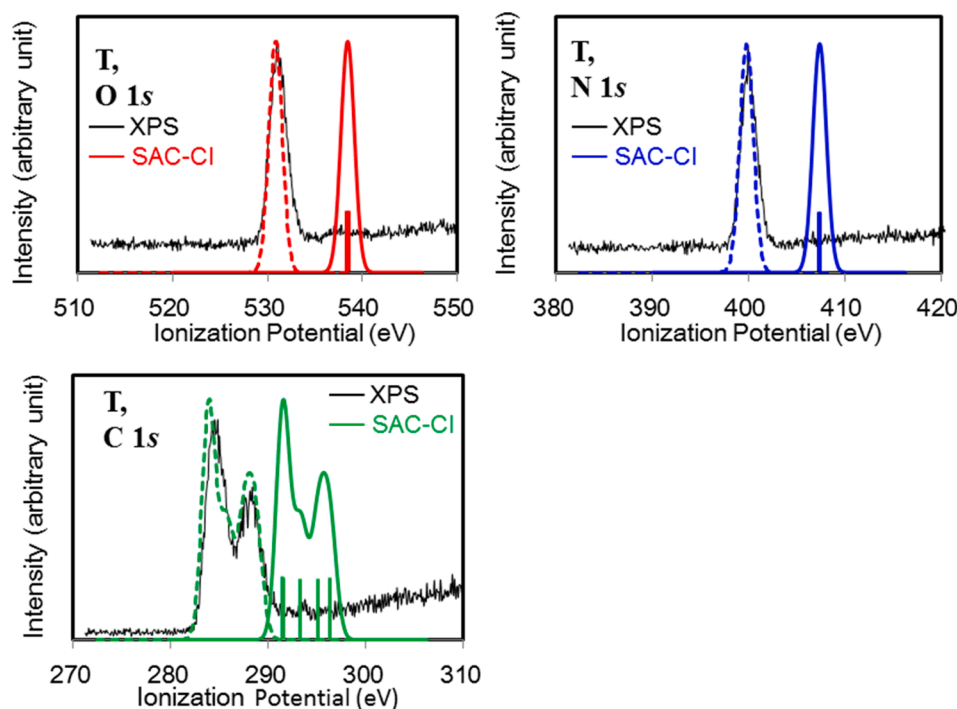
When we compare with the experiments with film sample, due to Yokoya *et al.* [13] and by Ptasińska *et al.* [25], the SAC-CI IPs were calculated higher by $\Delta E = 7.1\text{--}8.1$ eV (7.6 eV in average) and $\Delta E = 6.3\text{--}7.5$ eV (6.8 eV in average), respectively, for O 1s, N 1s, and C 1s ionized states. In the film condition, the interactions between T and surrounding polymer molecules are large, so that the single molecular calculations due to the SAC-CI is not well comparable with experiments.

3.3. Inner-core ionization spectra of 5-Bromouracil

The inner-core ionization potential (IP) of 5-Bromouracil (BrU) calculated by the SAC-CI method were shown in Table 2 and Fig. 3 with experiments. All of the calculated inner-core ionized states are described by one-electron ionized configuration from the original orbital with coefficients $|c| > 0.87$, without two-electron ionized configurations, unlike T. The first and second peaks, calculated at IP = 13269.62 and 1747.10 eV, were assigned to the Br 1s and Br 2s ionizations,

Table 1The inner-core ionization spectra of Thymine (T) calculated by the SAC-CI method, experiments, and other theoretical calculation.^a

Origin	State	SAC-CI with Basis II			Koopmans (eV) ^c	Exptl. in vacuum condition by Plekan et. al. ^d		Exptl. in vacuum condition by Itälä et. al. ^e	Exptl. of film sample by Yokoya et. al. ^f			Exptl. of film sample by Ptasíńska et. al. ^g		Theor. by Plekan et. al. ^d
		IP (eV)	Intensity	Main Configurations ^b		IP (eV)	ΔE (eV)		ΔE (eV)	IP (eV)	ΔE (eV)	IP (eV)	ΔE (eV)	
O7 1s	A'	538.56	0.8074	0.899(O7 1s ⁻¹)	559.10	537.3	1.3			531.5	7.1	532.1	6.5	537.46
O8 1s	A'	538.33	0.8039	0.897(O8 1s ⁻¹)	559.10									537.24
N1 1s	A'	407.49	0.7930	0.891(N1 1s ⁻¹)	425.68	406.7	0.8			400.23	7.3	400.0	7.5	406.80
N3 1s	A'	407.29	0.7947	0.891(N3 1s ⁻¹)	425.36									406.52
C2 1s	A'	296.29	0.7749	0.880(C2 1s ⁻¹)	310.80	295.2	1.1	295.1	1.2	288.5	7.8	290.0	6.3	295.36
C4 1s	A'	295.15	0.7730	0.879(C4 1s ⁻¹)	309.83	294.2	0.9	294.1	1.0			289.0		294.19
C6 1s	A'	293.34	0.7663	0.875(C6 1s ⁻¹)	308.43	292.3	1.0	292.1	1.2	285.2	8.1	285.8	7.5	295.29
				+0.106(C6 1s ⁻¹ (33) ⁻¹ (88))										
C5 1s	A'	291.55	0.7810	0.805(C5 1s ⁻¹)	306.62	291.0	0.5	290.9	0.6			285.1	6.4	290.95
				+0.365(C9 1s ⁻¹)										
				+0.104(C5 1s ⁻¹ (33) ⁻¹ (211))										
C9 1s	A'	291.51	0.7891	0.870(C9 1s ⁻¹)	305.94									290.67
				-0.178(C5 1s ⁻¹)										

^a IP and ΔE are in eV. ΔE is the difference from the SAC-CI IP.^b The main configurations with coefficients |c| > 0.10 are shown "-1" represents ionization or excitation, and the number in parentheses represents the MO number. See Fig. S1 for the MOs.^c The orbital energy of the Hartree Fock (HF) method with a minus sign.^d Ref [23].^e Ref [24].^f Ref [13].^g Ref [25].^h Calculated by the ADC(4) method. The IP has been corrected by 1.60, 1.20 and 1.0 eV for C 1s, N 1s and O 1s ionization, respectively.**Fig. 2.** The binding energy of Thymine (T) by the XPS (black) [13] and the SAC-CI calculation. The SAC-CI results were convoluted with the Gaussian curve of the half-width 1.65 eV. Broken lines are shifted curves by -7.6 eV.

respectively. The third to fifth states are assigned to the Br 2p ionized state: the third and fifth states are both described by a sum of the two one-electron ionized configurations from the Br 2px and Br 2py, but the main configuration is Br 2px⁻¹ for the third states and Br 2py⁻¹ for the fifth state. The fourth state is described by the one-electron ionized configurations from the Br 2pz. There is a discrepancy in the order of the

IPs between SAC-CI and Koopmans's: in Table 2, the main configurations of the fourth and fifth highest ionized states are from the Br 2pz and Br 2py orbitals, respectively, in the SAC-CI, while Br 2py and Br 2pz orbitals, respectively, in the Koopmans's. This discrepancy comes from the orbital relaxation effect and electron-correlation effects that are considered in the SAC-CI method but not in the Koopmans theorem.

Table 2The inner-core ionization spectra of 5-Bromouracil (BrU) calculated by the SAC-CI method and experiments^a.

Origin	State	SAC-CI with Basis II			Koopmans (eV) ^c	Exptl. in vacuum condition by Itälä et. al. ^d		Exptl. of film sample by Yokoya et. al. ^e		Exptl. of film sample by Plekan et. al. ^f	
		IP (eV)	Intensity	Main Configurations ^b		IP (eV)	ΔE (eV)	IP (eV)	ΔE (eV)	IP (eV)	ΔE (eV)
Br 1s	A'	13269.62	0.9000	0.949(Br 1s ⁻¹)	13318.59						
Br 2s	A'	1747.10	0.8861	0.941(Br 2s ⁻¹)	1773.44						
Br 2p	A'	1564.21	0.8877	0.926(Br 2px ⁻¹) +0.173(Br 2py ⁻¹)	1593.35			1596.29	-32.1		
	A''	1564.18	0.8878	0.942(Br 2pz ⁻¹)	1593.311			1550.67	13.5		
	A'	1564.16	0.8877	0.929(Br 2py ⁻¹) +0.157(Br 2px ⁻¹)	1593.310						
O7 1s	A'	538.82	0.8045	0.897(O7 1s ⁻¹)	559.50			531.40	7.4	531.8	7.1
O8 1s	A'	538.66	0.8021	0.896(O8 1s ⁻¹)	559.48						
N1 1s	A'	407.80	0.7890	0.888(N1 1s ⁻¹)	426.17			400.23	7.6	399.9	7.9
N3 1s	A'	407.48	0.7912	0.890(N3 1s ⁻¹)	425.74						
C2 1s	A'	296.66	0.7746	0.880(C2 1s ⁻¹)	311.21	295.6	1.1	288.10	8.6	288.8	7.9
C4 1s	A'	295.62	0.7723	0.879(C4 1s ⁻¹)	310.36	294.7	0.9			287.9	7.8
C6 1s	A'	293.93	0.7601	0.872(C6 1s ⁻¹)	309.15	292.7	1.2			285.6	8.3
C5 1s	A'	292.49	0.7675	0.876(C5 1s ⁻¹)	308.10	292.0	0.5	285.40	7.1	285.1	7.4
Br 3s	A'	259.32	0.8709	0.933(Br 3s ⁻¹)	268.40						
Br 3p	A'	194.28	0.8688	0.932(Br 3px ⁻¹)	203.45			183.94	10.3		
	A''	194.09	0.8710	0.933(Br 3pz ⁻¹)	203.25						
	A'	194.06	0.8707	0.933(Br 3py ⁻¹)	203.24						
Br 3d	A'	78.72	0.8693	0.930(Br 3dσ ⁻¹)	87.38			69.88	8.8	70.8	7.9
	A''	78.65	0.8698	0.931(Br 3dπ ⁻¹)	87.27					69.8	8.9
	A'	78.62	0.8699	0.929(Br 3dπ ⁻¹)	87.27						
	A''	78.42	0.8715	0.932(Br 3dδ ⁻¹)	87.00						
	A'	78.40	0.8715	0.933(Br 3dδ ⁻¹)	87.00						

^a IP and ΔE are in eV. ΔE is the difference from the SAC-CI IP.^b The main configurations with coefficients $|c| > 0.10$ are shown “-1” represents ionization, See Fig. S1 for the MOs.^c orbital energy of the Hartree Fock (HF) method with a minus sign.^d Ref [24].^e Ref [13].^f Ref [26].

Such discrepancy was found for 16 th and 17 th states: the main configurations of the 16 th and 17 th highest ionized states are from the Br 3pz and Br 3py orbitals by the SAC-CI, but Br 3py and Br 3pz orbitals by the Koopmans's. Further, comparing the bottom figures of Figs. 2 and 3, we clearly notice that the intensities of the experimental C 1s peaks from 280 to 290 eV are different between T and BrU, the reason is the absence of the methyl group in BrU: the C 1s peak due to the CH₃ group of T exists at around 285 eV (left peak of the two), but this carbon does not exist in BrU because the methyl group is replaced with Br in BrU: this has been discussed already for T using Fig. 2.

Compared with the experimental results measured in the single-molecule vacuum condition by Itälä et. al. [24] shown in Table 2, the SAC-CI IPs of the C 1s states were calculated a little higher by ΔE = 0.5–1.2 eV (0.92 eV in average). These errors are almost the same as those for T and they are small enough to assign the nature of the experimental spectra.

Next, we compare the SAC-CI results with the experiments by Yokoya et. al. [13] and by Plekan et. al. [26] shown in Table 2. The Br 1s and 2s peaks were not assigned in Fig. 3a by experiments [13]. For the spectra about 1550–1600 eV, Yokoya et. al. assigned them to Br 2p_{1/2} and Br 2p_{3/2} states. With the SAC-CI calculations, they were assigned to three Br 2p states: note that the SAC-CI calculations did not include the relativistic effect. If we consider the relativistic effect, it is expected that the 2s orbital of the bromine atom shrinks, the 2s orbital energy becomes lower, and the higher IP is obtained than the present SAC-CI results. The Br 2p orbitals will split, due to the spin–orbit interaction, to be a lower Br 2p_{1/2} and higher Br 2p_{3/2} orbitals in energy. Actually, the experimental Br 2p_{1/2} IP is larger by 32.1 eV and that for Br 2p_{3/2} IP is smaller by 13.5 eV than the SAC-CI Br 2p IP.

The gas phase experiments by Itälä et al. [24] reported only the C 1s peaks. Others are the film sample experiments. Except for these 2p-

ionization states, the SAC-CI IPs were calculated higher by ΔE = 7.1–10.3 eV (8.3 eV in average) than the experiments by Yokoya et. al. and ΔE = 7.1–8.9 eV (7.9 eV in average) than the experiments by Plekan et. al. These differences are mostly due to the use of the film sample in the experiments.

3.4. Importance of the polarization and Rydberg basis functions

Finally, we show the importance of the polarization and Rydberg basis functions for the studies of the inner-core ionization spectra by showing the results of the SAC-CI calculations of T and BrU with the Basis I that was used here for the geometry optimizations. In contrast to Basis II, in Basis I, the contracted orbitals are used without splitting and the additional polarization and Rydberg functions were not included. Tables 3 and 4 show the SAC-CI results of Basis I and the differences from the results with Basis II for T and Br U, respectively.

For T, the Koopmans's IPs calculated with Basis I were almost the same as those calculated with Basis II, and the largest difference was ΔE = 0.17 eV for C2 1s IP. For BrU, the Koopmans's IPs of Br 1s, 2s, and 2p calculated with Basis I were a little higher in energy than those with Basis II by ΔE = 2.85 eV, 0.49 eV, and 0.72 eV, respectively. But for O 1s, N 1s, C 1s, Br 3s, Br 3p, and Br 3d, the differences between Basis I and II were much smaller, and the largest difference was ΔE = 0.19 eV for C2 1s IP. These results were not surprising because, in general, the energies of the occupied orbitals do not change much by the change of the basis set.

Let us compare the SAC-CI results. The SAC-CI IPs calculated with Basis I and the differences from those calculated with Basis II were shown in Tables 3 and 4 for T and BrU, respectively. For T, the SAC-CI IPs calculated with Basis I were much higher in energy than those calculated with Basis II, and the differences were ΔE = 2.17–3.68 eV

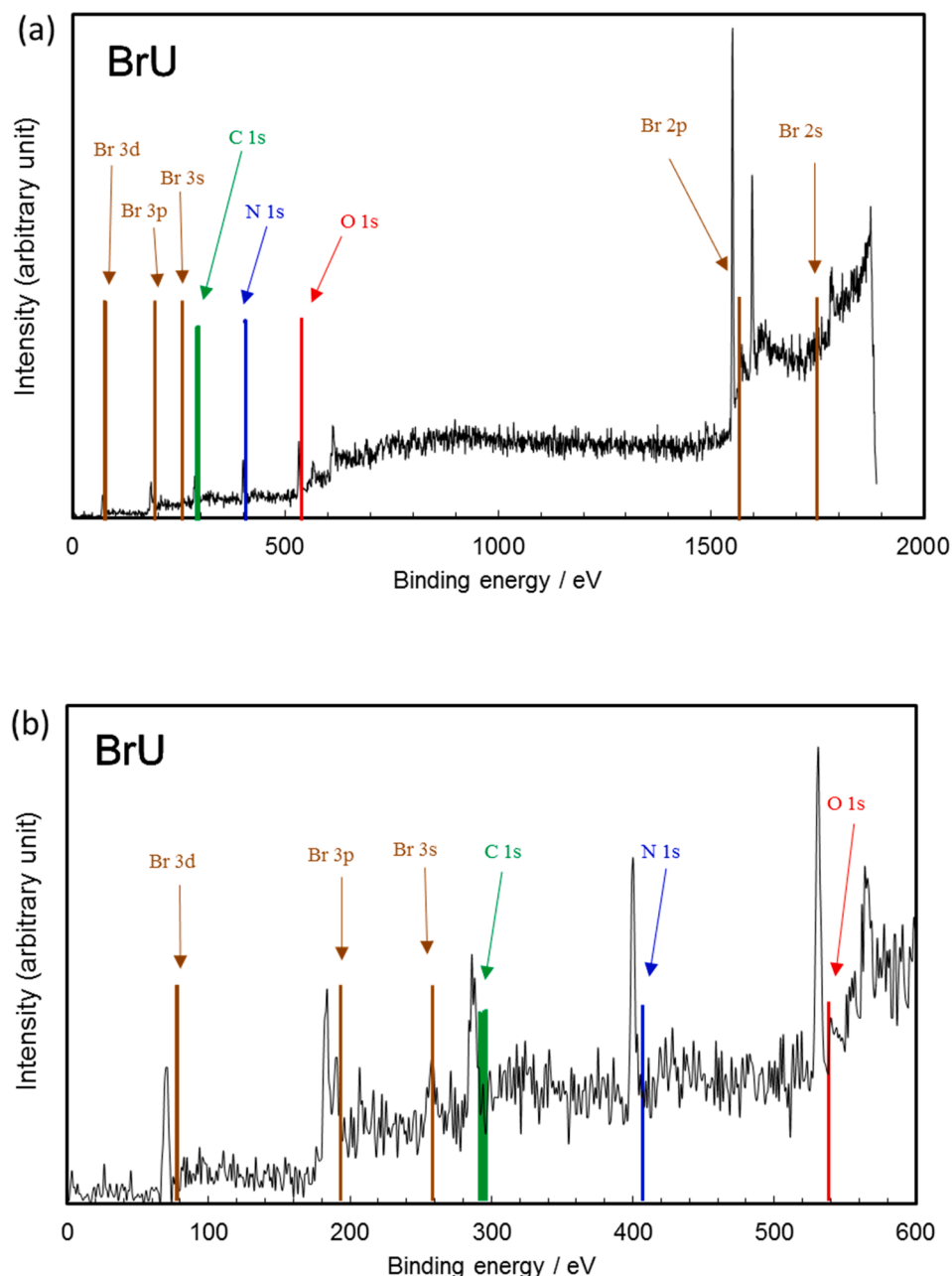


Fig. 3. The binding energy of 5-bromouracil by the XPS (black) [13] and SAC-CI calculation. (a) The whole region (b) The region of 0–600 eV. (c) O 1s, N 1s, and C 1s region. The SAC-CI results were convoluted with the Gaussian curve of the half-width 1.65 eV. Broken lines are shifted curves by -7.6 eV.

(2.64 eV in average) for O 1s, N 1s and C 1s IPs. For BrU, SAC-CI IPs calculated with Basis I were much higher in energy than those calculated with Basis II, and the differences were $\Delta E = 1.21$ – 22.69 eV (3.97 eV in average) for Br 1s, Br 2s, Br 2p, O 1s, N 1s, C 1s, Br 3s, Br 3p, and Br 3d IPs. Thus, the SAC-CI IPs were much improved with the larger and more flexible basis, Basis II, while the Koopmans's IPs were not improved with Basis II. This is interesting because, despite that the orbital energies of inner-core orbitals were almost the same between Basis I and II, the ionization energies were calculated to be much different between Basis I and II. We see from the comparisons of Tables 3 and 4 of Basis I with Tables 1 and 2 of Basis II that the SAC-CI calculations of Basis I used larger (than the case of Basis II) number of two electron processes which are the ionizations accompanied by the excitations from some higher occupied orbitals to some lower unoccupied orbitals. These two-electron excitation configurations were necessary to describe electron

reorganizations and real accompanying ionizations. This clearly shows that the electron re-organization and the changes in the electron correlations caused by the ionizations must be correctly described. For this purpose, we need good theory like SAC-CI that describes not only the direct ionization, but also the accompanied electron reorganization like the accompanying electron excitation from lower occupied orbitals to lower Rydberg orbitals, and at the same time, the good flexible basis set that can describe these complex physics well: without the Rydberg orbital basis set, the ionizations accompanying excitations from HOMO to the Rydberg orbitals cannot be described well.

4. Summary

In this study, we calculated the ionization spectra of Thymine (T) and 5-Bromouracil (BrU) with the SAC-CI method using different basis sets

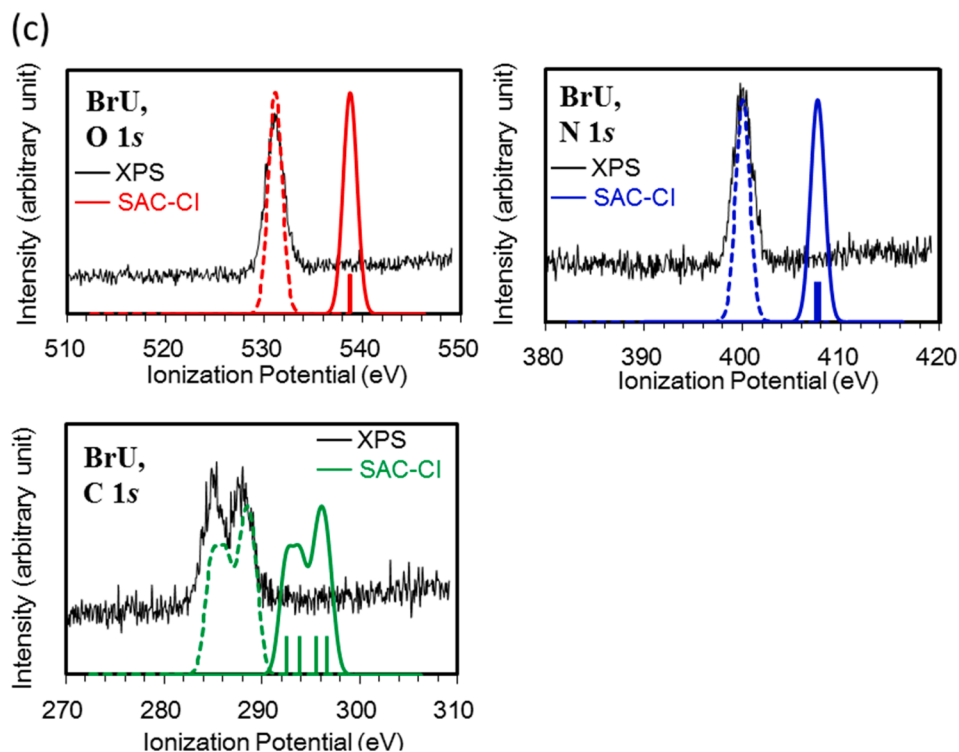


Fig. 3. (continued).

Table 3

The inner-core ionization spectra of Thymine (T) calculated by the SAC-CI method with Basis I.

Origin	State	SAC-CI with Basis I			Koopmans (eV) ^b	Koopmans with Basis II ^c		SAC-CI with Basis II ^d	
		IP (eV)	Intensity	Main Configurations ^a		ΔE (eV)		ΔE (eV)	
O7 1s	A'	542.21	0.8226	0.905 (O7 1s ⁻¹) +0.1061(O7 1s ⁻¹ (32) ⁻¹ (77))	559.05	0.06		3.65	
O8 1s	A'	542.01	0.8193	0.907 (O8 1s ⁻¹) -0.117(O8 1s ⁻¹ (32) ⁻¹ (35))	559.04	0.07		3.68	
N1 1s	A'	410.17	0.8104	0.900 (N3 1s ⁻¹)	425.57	0.10		2.68	
N3 1s	A'	409.83	0.8107	0.900 (N4 1s ⁻¹)	425.22	0.14		2.54	
C2 1s	A'	298.46	0.7930	0.890 (C2 1s ⁻¹) +0.127(C2 1s ⁻¹ (22) ⁻¹ (35)) +0.116(C2 1s ⁻¹ (22) ⁻¹ (49))	310.97	-0.17		2.17	
C4 1s	A'	297.40	0.7920	0.890 (C4 1s ⁻¹) -0.111(C4 1s ⁻¹ (27) ⁻¹ (35)) -0.110(C4 1s ⁻¹ (27) ⁻¹ (49))	309.98	-0.15		2.25	
C6 1s	A'	295.63	0.7864	0.887 (C6 1s ⁻¹) +0.147(C6 1s ⁻¹ (33) ⁻¹ (34)) +0.124(C6 1s ⁻¹ (33) ⁻¹ (55)) +0.121(C6 1s ⁻¹ (29) ⁻¹ (34)) +0.102(C6 1s ⁻¹ (29) ⁻¹ (55))	308.54	-0.11		2.29	
C5 1s	A'	293.81	0.7983	0.876 (C5 1s ⁻¹) +0.173 (C9 1s ⁻¹) -0.174(C5 1s ⁻¹ (33) ⁻¹ (61)) -0.101(C5 1s ⁻¹ (33) ⁻¹ (34))	306.66	-0.04		2.26	
C9 1s	A'	293.74	0.8096	0.894 (C9 1s ⁻¹) -0.104 (C5 1s ⁻¹) +0.135(C9 1s ⁻¹ (25) ⁻¹ (52)) +0.110(C9 1s ⁻¹ (28) ⁻¹ (50))	305.97	-0.03		2.23	

^a The main configurations with coefficients $|c| > 0.10$ are shown. “-1” represents ionization or excitation, and the number in parentheses represents the MO number.^b The orbital energy of the Hartree Fock (HF) method with a minus sign.^c The difference from the Koopmans's IP calculated with Basis II is shown.^d The difference from the SAC-CI IP calculated with Basis II is shown.

and assigned the experimental peaks. We have shown the necessity of using the flexible basis set that includes even the lower Rydberg orbitals. Compared with the spectra measured in the single-molecule vacuum condition, the SAC-CI inner - core Ionization spectra agreed with the

experimental spectra within the errors of $\Delta E \approx 1.0$ eV for both molecules. When we compare with the experiments using the film sample, the SAC-CI spectra must be shifted by $\Delta E = 7-8$ eV.

From the analysis of the results of the SAC-CI calculations using

Table 4

The inner-core ionization spectra of 5-Bromouracil (BrU) calculated by the SAC-CI method with Basis I.

Origin	State	SAC-CI with Basis I			Koopmans (eV) ^b	Koopmans with Basis II ^c		SAC-CI with Basis II ^d	
		IP (eV)	Intensity	Main Configurations ^a		ΔE (eV)	ΔE (eV)		
Br 1s	A'	13292.31	0.8835	0.940(Br 1s ⁻¹)	13315.75	2.85		22.69	
Br 2s	A'	1753.61	0.8875	0.942(Br 2s ⁻¹)	1772.95	0.49		6.51	
Br 2p	A'	1572.25	0.8867	0.942(Br 2px ⁻¹)	1592.64	0.72		8.04	
	A''	1572.22	0.8869	0.942(Br 2pz ⁻¹)	1592.59	0.72		8.04	
O7 1s	A'	1572.21	0.8869	0.942(Br 2py ⁻¹)	1592.59	0.72		8.05	
	A'	542.61	0.8212	0.906(O7 1s ⁻¹)	559.46	0.04		3.79	
O8 1s	A'	542.48	0.8193	+0.105(O7 1s ⁻¹ (44) ⁻¹ (49))	559.45	0.03		3.82	
				0.896(O8 1s ⁻¹)					
				+0.124(O8 1s ⁻¹ (44) ⁻¹ (92)) -0.109(O8 1s ⁻¹ (44) ⁻¹ (49))					
N1 1s	A'	410.65	0.8087	0.899(N1 1s ⁻¹)	426.10	0.07		2.85	
N3 1s	A'	410.20	0.8099	0.900(N3 1s ⁻¹)	425.61	0.12		2.72	
				-0.109(N3 1s ⁻¹ (44) ⁻¹ (82)) -0.101(N3 1s ⁻¹ (44) ⁻¹ (52))					
				0.890(C2 1s ⁻¹)					
C2 1s	A'	298.90	0.7929	0.890(C2 1s ⁻¹)	311.40	-0.19		2.25	
				+0.125(C2 1s ⁻¹ (35) ⁻¹ (49)) -0.116(C2 1s ⁻¹ (35) ⁻¹ (62)) -0.108(C2 1s ⁻¹ (38) ⁻¹ (49))					
				0.890(C4 1s ⁻¹)					
				-0.134(C4 1s ⁻¹ (38) ⁻¹ (49)) +0.103(C4 1s ⁻¹ (38) ⁻¹ (67))					
C4 1s	A'	297.93	0.7915	0.883(C6 1s ⁻¹)	310.54	-0.18		2.31	
				-0.165(C6 1s ⁻¹ (46) ⁻¹ (47)) +0.141(C6 1s ⁻¹ (40) ⁻¹ (47)) +0.138(C6 1s ⁻¹ (46) ⁻¹ (66)) -0.115(C6 1s ⁻¹ (40) ⁻¹ (66))					
				0.889(C5 1s ⁻¹)					
C6 1s	A'	296.32	0.7802	+0.140(C5 1s ⁻¹ (39) ⁻¹ (48)) +0.136(C5 1s ⁻¹ (46) ⁻¹ (71))	309.31	-0.16		2.39	
				0.939(Br 3s ⁻¹)					
				0.938(Br 3px ⁻¹)					
				0.939(Br 3pz ⁻¹)					
C5 1s	A'	294.90	0.7901	0.939(Br 3py ⁻¹)	308.11	-0.01		2.41	
				0.938(Br 3dσ ⁻¹)					
Br 3s	A'	260.53	0.8821	0.938(Br 3dπ ⁻¹)	268.40	0.00		1.21	
Br 3p	A'	195.58	0.8796	0.938(Br 3dπ ⁻¹)	203.51	-0.06		1.29	
	A''	195.30	0.8823	0.938(Br 3dπ ⁻¹)	203.23	0.02		1.22	
Br 3d	A'	195.29	0.8818	0.939(Br 3dπ ⁻¹)	203.23	0.01		1.23	
	A'	80.05	0.8799	0.938(Br 3dσ ⁻¹)	87.42	-0.04		1.34	
	A''	79.95	0.8804	0.938(Br 3dπ ⁻¹)	87.29	-0.02		1.30	
	A'	79.94	0.8802	0.938(Br 3dπ ⁻¹)	87.29	-0.02		1.32	
	A''	79.69	0.8829	0.940(Br 3dδ ⁻¹)	86.99	0.01		1.27	
	A'	79.68	0.8823	0.939(Br 3dδ ⁻¹)	86.99	0.01		1.28	

^a The main configurations with coefficients $|c| > 0.10$ are shown. “-1” represents ionization or excitation, and the number in parentheses represents the MO number.^b The orbital energy of the Hartree Fock (HF) method with a minus sign.^c The difference from the Koopmans's IP calculated with Basis II is shown.^d The difference from the SAC-CI IP calculated with Basis II is shown.

different basis sets, it was clarified that the inner-core ionizations involve many different electronic processes like ionizations, accompanying electron excitations, electron reorganizations, and electron correlations. Therefore, for accurate theoretical descriptions of XPS and inner-core ionizations, we need not only good theory, but also good basis set: we need even the Rydberg orbitals for describing the ionizations accompanying excitations. Therefore, the SAC-CI calculations with the flexible basis set that includes lower Rydberg orbitals could give reliable results that were useful for the study of the XPS spectra of T and BrU.

CRediT authorship contribution statement

Yusaku I. Kurokawa: Software, Investigation. **Hiroshi Nakatsuji:** Methodology, Conceptualization. **Misaki Hirato:** Validation. **Akinari Yokoya:** Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgement

The computation was performed using Research Center for Computational Science, Okazaki, Japan (Project: 22-IMS-C012). This work was supported by JSPS KAKENHI Grants Nos. 20H04338 and 22H02045.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cplett.2022.139957>.

References

- [1] H. Nakatsuji, K. Hirao, J. Chem. Phys. 68 (1978) 2053.
- [2] H. Nakatsuji, Chem. Phys. Lett., 59 (1978) 362; 67 (1979) 329.
- [3] H. Nakatsuji, Acta Chimica Hungarica, Models Chem. 129 (5) (1992) 719.
- [4] H. Nakatsuji, Computational Chemistry, Rev. Curr. Trends 2 (1997) 62.
- [5] H. Nakatsuji, Chem. Phys. 75 (1983) 425.
- [6] M. Ehara, H. Nakatsuji, Chem. Phys. Lett. 282 (1998) 347–354.
- [7] M. Ehara, S. Yasuda, H. Nakatsuji, Z. Phys. Chem. 217 (2003) 161–176.
- [8] M. Ehara, Y. Ohtsuka, H. Nakatsuji, M. Takahashi, Y. Udagawa, J. Chem. Phys. 122 (2005), 234319-1-10.

- [9] M. Ehara, K. Kuramoto, H. Nakatsuji, M. Hoshino, T. Tanaka, M. Kitajima, H. Tanaka, Y. Tamenori, A. De Fanis, K. Ueda, *J. Chem. Phys.* 125 (2006), 114304-1-10.
- [10] M. Ehara, K. Kuramoto, H. Nakatsuji, *Chem. Phys.* 356 (2009) 195–198.
- [11] R. Fukuda, M. Ehara, H. Nakatsuji, N. Kishimoto, K. Ohno, *J. Chem. Phys.* 132 (2010), 084302-1-12.
- [12] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, G.A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B.G. Janesko, R. Gomperts, B. Mennucci, H.P. Hratchian, J. V. Ortiz, A.F. Izmaylov, J.L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. J. Bearpark, J.J. Heyd, E.N. Brothers, K.N. Kudin, V.N. Staroverov, T.A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A.P. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, J.M. Millam, M. Klene, C. Adamo, R. Cammi, J.W. Ochterski, R.L. Martin, K. Morokuma, O. Farkas, J.B. Foresman, D.J. Fox, *Gaussian 16*, Revision C.02, Gaussian, Inc., Wallingford CT, 2019.
- [13] M. Hirato, A. Yokoya, Y. Baba, Y. Kurokawa, H. Nakatsuji, S. Mori, K. Fujii, S.-I. Wada, Y. Izumi, Y. Haga, to be submitted.
- [14] T.L. Morgan, J.L. Redpath, J.F. Ward, *Int. J. Radiat Biol. Relat Stud. Phys. Chem. Med.* 45 (1984) 217–226.
- [15] A. Yokoya, T. Ito, *Int. J. Rad. Bio.* 93 (2017) 743–756.
- [16] M. Hirato, M. Onizawa, Y. Baba, Y. Haga, K. Fujii, S.-I. Wada, A. Yokoya, *Int. J. Rad. Bio.* (2020).
- [17] Y. Wang, G.E. Pantelias, G. Iliakis, *Int. J. Rad. Bio.* 66 (1994) 133–142.
- [18] J.D. Zimbrick, J.F. Ward, L.S. Myers, *Int. J. Radiat Biol. Relat Stud. Phys. Chem. Med.* 16 (1969) 505–523.
- [19] R. Watanabe, H. Nikjoo, *Int. J. Rad. Bio.* 78 (2002) 953–966.
- [20] S. Cecchini, S. Girouard, M.A. Huels, L. Sanche, D.J. Hunting, *Radiat. Res.* 162 (2004) 604–615.
- [21] S. Huzinaga, *Gaussian Basis Sets for Molecular Calculations*, Elsevier Science Ltd, 1984.
- [22] R.I. Martin, D.A. Shirley, *J. Chem. Phys.* 64 (1976) 3685.
- [23] O. Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K.C. Prince, A. B. Trofimov, E.V. Gromov, I.L. Zaytseva, J. Schirmer, *Chem. Phys.* 347 (2008) 360–375.
- [24] E. Itälä, D.T. Ha, K. Kooser, E. Rachlew, M.A. Huels, E. Kukk, *J. Chem. Phys.* 133 (2010), 154316-1-9.
- [25] S. Ptasińska, A. Stypczyńska, T. Nixon, N.J. Mason, D.V. Klyachko, L. Sanche, *J. Chem. Phys.* 129 (2008), 065102.
- [26] O. Plekan, V. Feyer, N. Tsud, M. Vondráček, V. Cháb, V. Matolín, K.C. Prince, *Surf. Sci.* 606 (2012) 435–443.