

OPEN

# Theoretical analyses and experimental validation of the effects caused by the fluorinated substituent modification of DNA

Jun Koseki<sup>1</sup>, Masamitsu Konno<sup>2</sup>, Ayumu Asai<sup>1,2</sup>, Naohiro Horie<sup>3</sup>, Kenta Tsunekuni<sup>1,2,4</sup>, Koichi Kawamoto<sup>2</sup>, Satoshi Obika<sup>3</sup>, Yuichiro Doki<sup>2</sup>, Masaki Mori<sup>2,5</sup> & Hideshi Ishii<sup>1\*</sup>

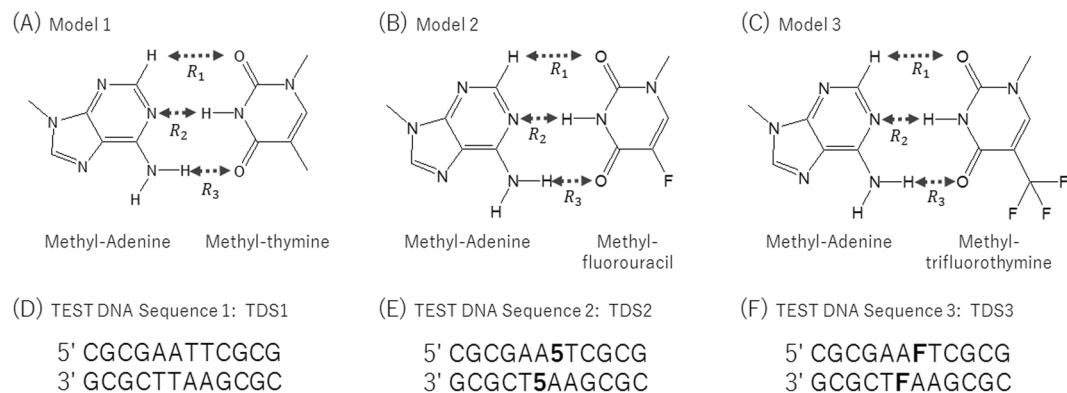
Halogen-modified nucleic acid molecules, such as trifluorothymidine (FTD) and 5-fluorouracil, are widely used in medical science and clinical site. These compounds have a very similar nucleobase structure. It is reported that both of these compounds could be incorporated into DNA. The incorporation of FTD produces highly anti-tumor effect. However, it is not known whether to occur a significant effect by the incorporation of 5-fluorouracil. Nobody knows why such a difference will occur. To understand the reason why there is large differences between trifluorothymidine and 5-fluorouracil, we have performed the molecular dynamics simulations and molecular orbital calculations. Although the active interaction energy between Halogen-modified nucleic acids or and complementary adenine was increased, in only FTD incorporated DNA, more strongly dispersion force interactions with an adjacent base were detected in many thermodynamic DNA conformations. As the results, the conformational changes occur even if it is in internal body temperature. Then the break of hydrogen bonding between FTD and complementary adenine base occur more frequently. The double helix structural destabilization of DNA with FTD is resulted from autoagglutination caused by the bonding via halogen orbitals such as halogen bonding and the general van der Waals interactions such as CH- $\pi$ , lone pair (LP)- $\pi$ , and  $\pi$ - $\pi$  interactions. Therefore, it is strongly speculated that such structural changes caused by trifluoromethyl group is important for the anti-tumor effect of FTD alone.

Some types of nucleic acid modifications such as methylation and acetylation are known to directly affect the functions of DNA and RNA. These modifications play a critical role in DNA manipulation and can considerably affect DNA transcription and replication<sup>1-4</sup>. In addition, artificially modified nucleic acid molecules with halogen atoms, such as 5-fluorouracil (5FU) and trifluorothymidine (FTD), have been used in medical science and for clinical therapies<sup>5-9</sup>. The 5FU acts as an antitumor drug by inhibiting the formation of ribosomal RNA by its incorporation into RNA replacing uracil, or inhibiting DNA synthesis by preventing thymidylate synthesis<sup>10-12</sup>. Recent studies have shown that 5FU also gets incorporated into DNA, however, the association between this incorporation and the antitumor effects remains unknown. Conversely, FTD, which is incorporated with DNA, possesses a highly effective antitumor potency<sup>13-15</sup> although its structure is similar to that of thymidine. The combination of FTD and tipiracil hydrochloride (FTD/TPI) has been approved by some organizations, such as the FDA (Food and Drug Administration of the United States), as a treatment for colorectal cancer<sup>16</sup>. The reason why FTD has antitumor effect is not known yet, although many studies about it have been reported<sup>17-22</sup>. It is said today that the reason is not to form a covalent bond with thymidylate synthase<sup>17</sup> but to lead to DNA dysfunction by the incorporation of FTD into DNA from some experimental results<sup>18-22</sup>. In this way, despite all the different pharmacological mechanisms, the structures of the base moieties of 5FU and FTD are very similar. Why is FTD showing antitumor effect by the incorporation into DNA and is not 5FU showing? How interactions

<sup>1</sup>Department of Medical Data Science, Graduate School of Medicine, Osaka University, Osaka, 565-0871, Japan.

<sup>2</sup>Department of Gastroenterological Surgery, Graduate School of Medicine, Osaka University, Osaka, 565-0871, Japan. <sup>3</sup>Bioorganic Chemistry, Graduate School of Pharmaceutical Sciences, Osaka University, Osaka, 565-0871, Japan. <sup>4</sup>Translational Research Laboratory, Taiho Pharmaceutical Co., Ltd, Tokushima, 771-0194, Japan.

<sup>5</sup>Department of Surgery and Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka, 812-8582, Japan. \*email: [hishii@gesurg.med.osaka-u.ac.jp](mailto:hishii@gesurg.med.osaka-u.ac.jp)



**Figure 1.** Simple base-pair models: methylated adenine and methylated thymine (A), methylated fluorouracil (B), and methylated trifluorothymine pair model (C). Test DNA Sequence 1 (D), 2 (E), and 3 (F).

		$R_1$	$R_2$	$R_3$	Energy
		(Å)	(Å)	(Å)	(kcal/mol)
Model 1	–H...O–/–N...H–	2.72	1.80	1.95	16.3
	–N...N–/–N...O–	—	2.85	2.96	
Model 2	–H...O–/–N...H–	2.67	1.77	1.97	16.7
	–N...N–/–N...O–	—	2.83	2.98	
Model 3	–H...O–/–N...H–	2.68	1.77	1.97	17.1
	–N...N–/–N...O–	—	2.82	2.98	

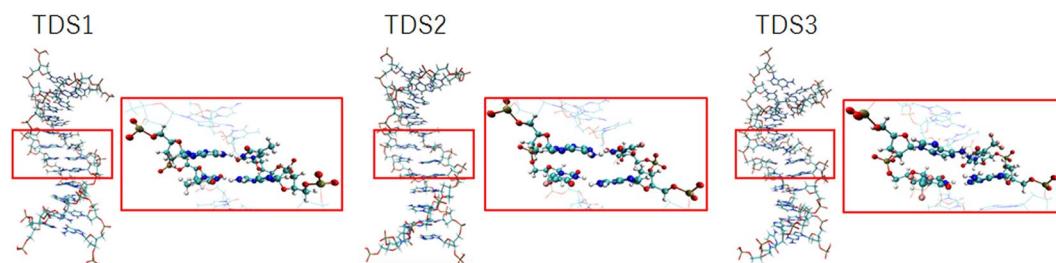
**Table 1.** Distances  $R_1$ ,  $R_2$ , and  $R_3$  in each optimized structure, and the dissociation energies.

with surrounding nucleic acids does these molecules have after incorporation into DNA? To our knowledge, the reasons for these differences remain unclear. Thus, in this study, we performed molecular dynamics (MD) and ab initio molecular orbital (MO) calculations to qualitatively determine the role and effects of the fluorinated groups in 5FU and FTD. We firstly obtained interaction energies and distances between some atoms using the simple model base pairs, and then we compared the stabilization structure of DNA duplexes with each energy minimization. After that, we performed thermal fluctuation analyses of these DNA duplexes with MD calculations, followed by analysis of the interbase interactions. Finally, theoretically predicted results were verified.

## Results and Discussion

**Simple base pair models for dissociation energies analyses.** We constructed simple base pair models, model 1, model 2, and model 3 (Fig. 1A–C), to analyze the difference in the hydrogen bonding intensity according to the presence or absence of the fluorine substituent and dissociation energies between thymidine–adenine, deoxyribose 5FU–adenine, and FTD–adenine pairs, respectively. These simple model structures were optimized by the second-order Møller–Plesset perturbation (MP2) using the 6-31++G\*\* Gaussian basis set, followed by the dissociation energies. In this study, all *ab initio* molecular orbital calculations were performed with Gaussian 09<sup>23</sup>. Table 1 shows the distances,  $R_1$ ,  $R_2$ , and  $R_3$ , in each stabilized structure, and the dissociation energies. For optimized structures of models 1–3, there is no striking difference in each distance. The stabilization energies were 16.3, 16.7, and 17.1 kcal/mol, respectively, which corresponded to the energies required to separate each molecular pair. These results demonstrate that the base pair was slightly stabilized by an increase in fluorine, and suggest that the attractive interaction is high owing to the electronic substituent effect caused by the fluorine group.

**Double-stranded structures of 3 test DNA sequences for structural analyses.** We considered double-stranded structures of 3 test DNA sequences (TDS1, control sequence DNA; TDS2, DNA containing 5FU; and TDS3, DNA containing FTD) to determine the role of the fluorine group in the DNA (Fig. 1D–F). We adopted the short DNA sequence used by Markley *et al.*<sup>24</sup> as a reference for these test DNA constructs. These DNA structures were constructed based on the B-type DNA, which is right-handed with about 10 base pairs per turn. For each the DNA structure in water molecules, energy minimizations were performed using the AMBER 12<sup>25</sup> program package. The AMBER99<sup>26</sup>, general AMBER force field (GAFF)<sup>27</sup> and TIP3P<sup>28</sup> force fields for the complete DNA and water molecules were used respectively. Figure 2 shows the stabilization conformations for TDS1, TDS2, and TDS3, respectively. Disregarding thermal energies, we observed substantial differences between TDS1, TDS2, and TDS3, as shown in Fig. 2. A magnified view of images is shown in Fig. S1. In particular, the structure of TDS3 containing FTD became more distorted compared with TDS1, (the control sequence). Figure S2A shows the superposition of minimization structures with the tube representation for TDS1 (cyan) and

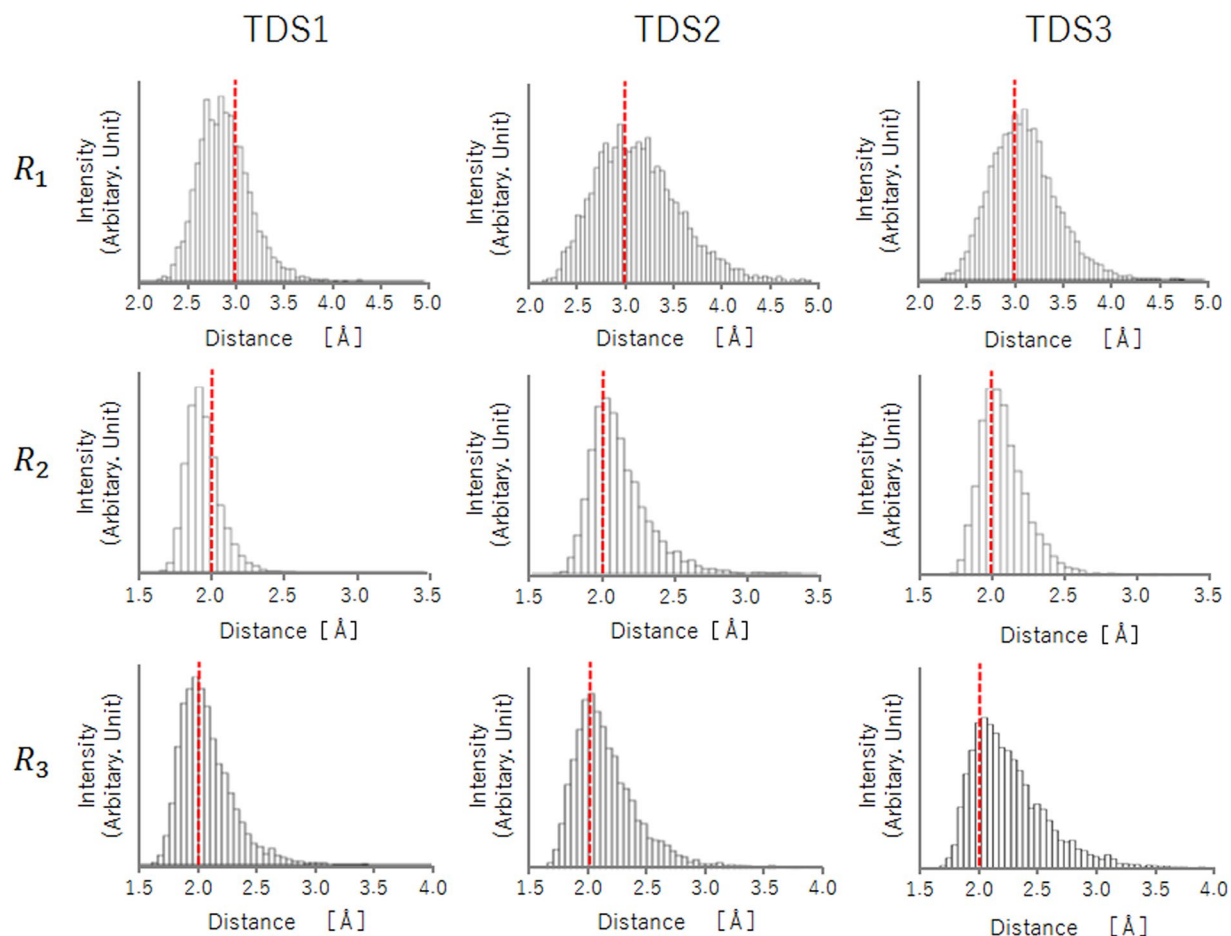


**Figure 2.** Minimization conformations of TDS1, TDS2, and TDS3, in the water phase. A magnified view of images for focused bases (ball and stick representation) are shown in each red frame.

TDS3 (red; the sphere representation implies the position of the fluorine atom). While the double-stranded structure was maintained in TDS1, the hydrogen bonding system between FTD and its complementary adenine broke in TDS3. In TDS2, we did not observe broken hydrogen bonding patterns, but a little distortion occurred in their double-stranded structure. By comparing with results of simple base pair models and test DNA sequences, the destabilization of the double-stranded DNA including FTD, was assumed to not change the hydrogen bonding intensity but the interactions between the trifluoromethyl group and surrounding bases. In previous report, W. H. Gmeiner *et al.*<sup>29</sup> performed the energy minimization simulations for the double helix structure containing FTD. They reported that the trifluoromethyl group does not interfere with base pairing or base stacking, nor does it significantly affect helical dimensions. However, their calculation could not include the solvent effect. In general, it is known that, when the energy minimization of strongly charged molecules such as nucleic acids in gas phase, the hydrogen bonding and electrostatic interaction is overestimated. Therefore, in discussing the structure of the DNA including FTD, it is shown that consideration of solvent effect is indispensable.

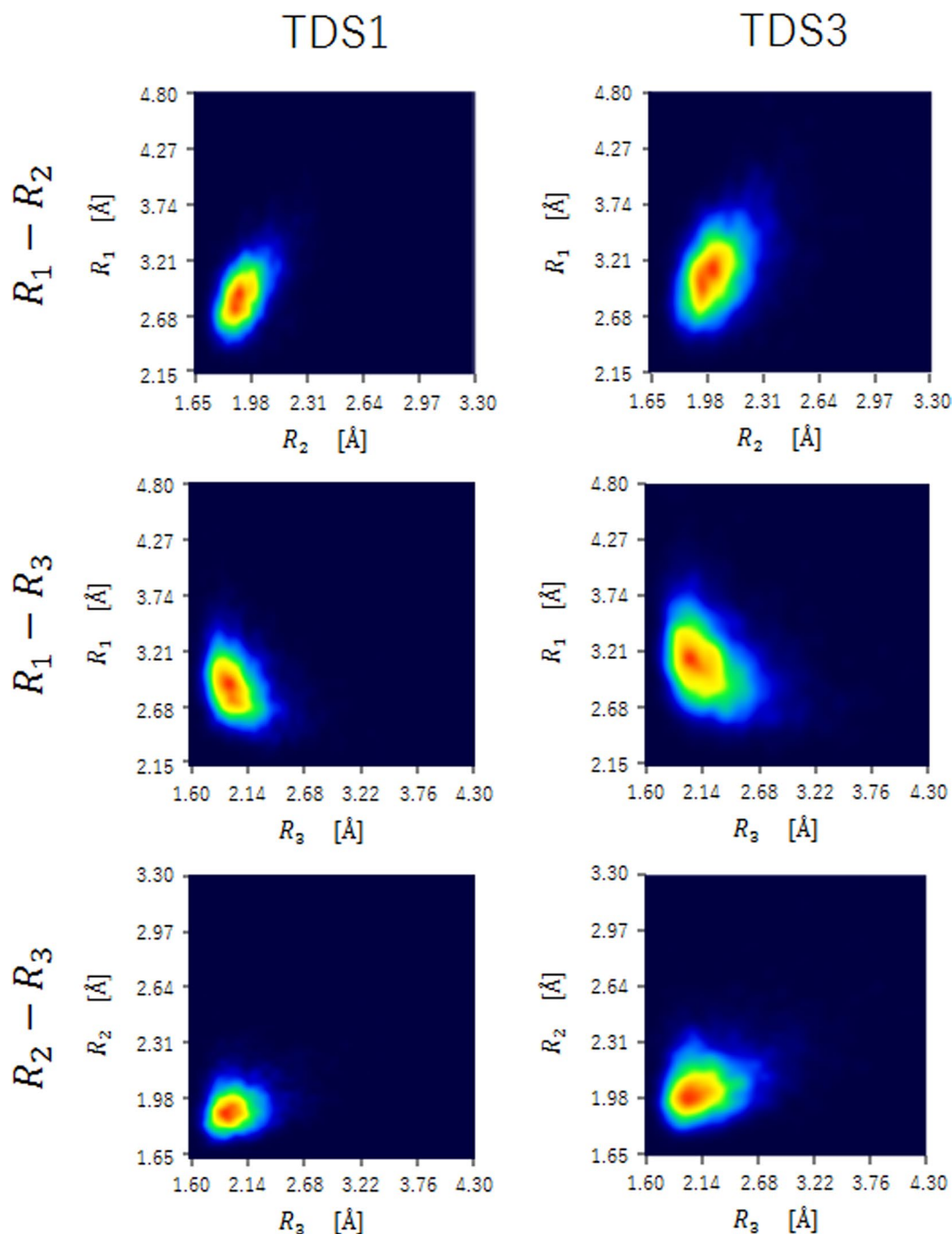
In addition, to elucidate the differences in thermodynamic behaviors in water phase, we performed MD simulations. Supplementary Movies S1–S3 present the thermodynamic behaviors of each TEST DNA. We focused on distances between the complementary base pairs to analyze the difference in these behaviors. Figure 3 shows the distributions of the distances of  $R_1$ ,  $R_2$ , and  $R_3$  in TDS1, TDS2, and TDS3, respectively. In terms of the  $R_1$  distribution, the peak in TDS3 was longer than 3.0 Å, but in TDS1 it was shorter than 3.0 Å. In TDS2, the peak was around 3.0 Å, but the rate located at >3.0 Å was higher than those in other DNAs. The distance of 3.0 Å is the upper limit to form a weak hydrogen bonding<sup>30</sup>. We found that the  $R_2$  distribution had peaks in TDS2 and TDS3 of around 2.0 Å, and longer than the peak in TDS1. In addition, the distribution of TDS3 was quite similar to that of TDS2. The  $R_3$  distribution peaks were similar at around 2.0 Å. Although TDS1 and TDS2 exhibited analogous distributions, the probability of being >2.0 Å in TDS3 was much higher than that in the other test DNA sequences. In this case, it is necessary to maintain a distance  $\leq 2.0$  Å to form a strong hydrogen bonding. The  $R_1$  distance in TDS2 was longer than that in TDS3. However, as weak hydrogen bonding could not be formed at a distance >3.0 Å, no marked differences existed in the thermal stability between TDS2 and TDS3. However, the rate of existence of  $R_3$  at a distance <2.0 Å was higher in TDS2 than in TDS3. In addition, it was easier to take a distance >3.0 Å in TDS3 than in TDS2. Hence, it is more difficult for TDS3 to form a strong hydrogen bonding than for TDS2, and it takes more time to form a weak hydrogen bonding. Of note, TDS3 sometimes does not form hydrogen bonding. In other words, the hydrogen bonding between FTD and its complementary base in TDS3 was weaker than that in TDS2, and the breakage of hydrogen bonding occurs often. In TDS2, although the distance of  $R_1$  which originally has a long interatomic distance distribution is broad, but the  $R_3$  distribution is not significantly different from the distribution in TDS1. On the other hand, in the distance distributions of both  $R_2$  and  $R_3$  of TDS3, many are distributed beyond the distance region forming strong hydrogen bonding (2.0 Å) compare with TDS1. In other word, it means that the double helix structural stabilization of TDS3 is not maintained by forming hydrogen bonding while the stabilization of TDS2 is preserved.

Therefore, in order to understand hydrogen bonding instability in TDS3, molecular oscillation between FTD and complementary Adenine was focused on. The two-dimensional distribution of  $R_1$ - $R_2$ ,  $R_1$ - $R_3$ , and  $R_2$ - $R_3$  in TDS1, and TDS3 were shown in Fig. 4. As shown in  $R_1$ - $R_2$  distributions, the  $R_1$  distance stretched in step with the  $R_2$  distance. The  $R_2$ - $R_3$  distributions exhibited the same tendency as  $R_1$ - $R_2$  distributions. Conversely,  $R_1$ - $R_3$  distributions exhibited a different propensity than these distributions, and the  $R_1$  distance stretched inversely to the  $R_3$  distance. Notably, these tendencies could be observed in common. The difference between each DNA implies the spread of distributions. We observed widespread distributions in TDS3 compared with the distributions in TDS1. In TDS3, the correlation between  $R_2$  and  $R_3$  was comparatively larger than in TDS1 and the distance of  $R_2$  expanded with the extension of the distance of  $R_3$ . As mentioned, the two-dimensional distance distributions in TDS3 exhibited an extensive presence compared with the distributions in TDS1 due to the existence of the interaction by the fluorine substituent. Typically, thermal vibrations of base pair are categorized into normal modes, such as buckle mode, propeller mode, opening mode, shear mode, stretch mode, and stagger mode<sup>31,32</sup>. In the thermal oscillation of the adenine–FTD pair in the MD trajectory of TDS3, many opening modes were observed (Supplementary Movies S4). This could be seen from the following behavior. As the  $R_1$  distance increases in length, the  $R_3$  distance shortens, and the  $R_1$  distance decreases with an increase in the length of the  $R_3$  distance. It might be that this opening mode induces hydrogen bonding fracture of TDS3. In fact, we could observation that water molecules are likely to enter into between two bases in molecular dynamics simulations.



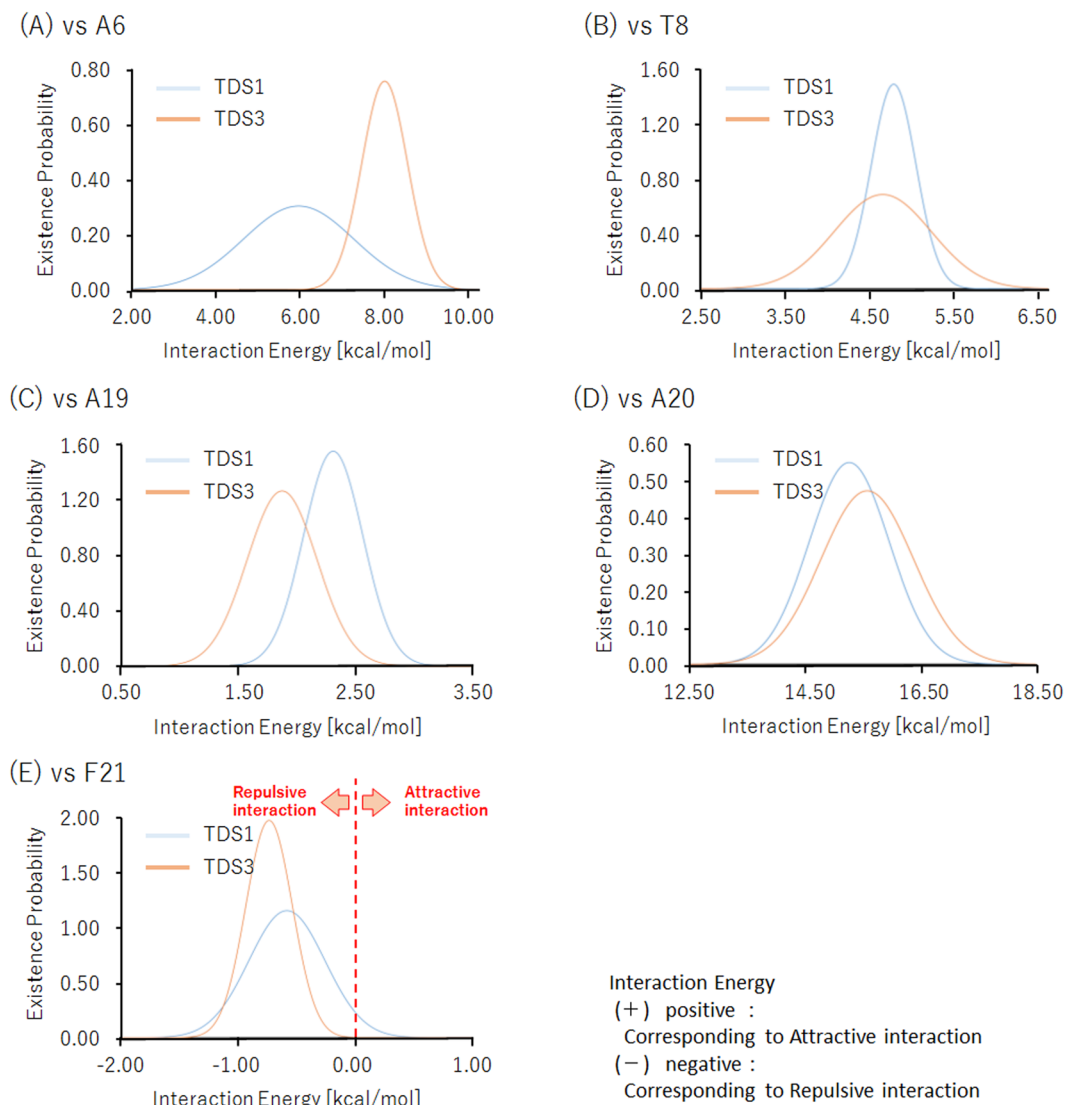
**Figure 3.** Distributions of the distance of  $R_1$ ,  $R_2$ , and  $R_3$  in TDS1, TDS2, and TDS3 in 310 K. The unit is in Å. The distance of 3.0 Å is an upper limiting length to form a weak hydrogen bonding. It is necessary to keep a distance of  $\leq 2.0$  Å for forming a strong hydrogen bond.

**Interbase interaction analyses.** We performed interaction analyses between bases to analyze the factors causing a hydrogen bonding fracture in TDS3. Thus, we introduced the conformational clustering in eight bases, 5th-adenine (A05), 6th-adenine (A06), 7th-thymidine or FTD (T07 or F07), 8th-thymidine (T08), 19th-adenine (A19), 20th-adenine (A20), 21st-thymidine or FTD (T21 or F21), and 22nd-thymidine (T22) to elucidate the possible conformational distribution in MD simulation trajectories (5000 conformations) for TDS1 and TDS3. Eight bases (5A, 6A, “7T or F”, 8T, 19A, 20A “21T or F”, and 22T) were selected from the trajectories of MD simulations for TDS1 and TDS3, respectively, to analyze their interbase interaction energies. Thereafter, we proposed the partial base models comprising these eight bases. Only the hydrogen atoms were reoptimized with MP2/6-31 G\* level. Then, the distribution of interaction energies was evaluated to estimate interaction energies between eight bases, each base energy, and any pair energy using each of the models at the Hartree-Fock (HF) and MP2/6-31 G\* level. It is well known that HF method is unstable for the description of stacking interactions<sup>33</sup> since the method could not restage the dispersion force interactions. However, we venture to estimate the energies with HF method in order to analyze using the difference of the characteristic of MP2. Figure S3 shows the superposition of 5000 conformations in each MD simulation. The variation areas of fluorine atoms in F07 and F21 were more extensive than that of hydrogen atoms in T07 and T21. Table S1A,B presents the interbase interaction energies for TDS1 and TDS3 using five representative conformations selected by conformational clustering. Figure S4A,B shows the structural clustering dendrogram for TDS1 and TDS3. Under these energies, the positive and negative values corresponded to attractive and repulsive interactions, respectively. In TDS1, the sign of interaction energies between T07 and T08 differed from that between HF and MP2. In TDS3, not only the F07-T08 interaction but also the F07-A06 interaction with MP2 were opposite in sign to the interactions with HF. The MP2 calculations can include the effect of dispersion interaction although HF cannot. In other words, it means that given that the estimated energy reflects the contribution of dispersion interaction, the signs of energy levels in HF and MP2 must be different. Figure 5A–E shows the Gaussian approximated distributions of the interaction energy with the MP2 level; we found apparent differences in each interaction energy between TDS1 and TDS3. The A06 interaction in TDS3 was strong and its range shrunk compared to the one in TDS1. The interaction energy peaks for T08 were similar in both TDS1 and TDS3. However, the distribution of TDS3 was more extensive than those in TDS1, implying that TDS3 had some conformations with interaction energies stronger than TDS1. In contrast, for A19 within the complementary strand, the distribution of TDS3 tended to locate in a lower energy region than that of



**Figure 4.** Thermodynamic two-dimensional distributions of the distance Å combination of  $R_1$ ,  $R_2$ , and  $R_3$  in TDS1 and TDS3.

TDS1. In A20, interaction energy distributions were similar to the model system; FTD within TDS3 had a strong interaction energy with the complementary base compared with the thymidine within TDS1. Finally, the interaction energy of F21 was basically negative in TDS3; however, TDS1 exhibited positive interaction energies. In other word, FTD attracts adjacent bases on the same chain strongly, although the interactions with complementary strand bases are weaker than in TDS1 or even repulsive. The DNA duplex instability of TDS3 is induced by these complex factors. In addition, we have also performed natural bond orbital (NBO) analyses<sup>34–36</sup> to determine the difference in their interaction energy intensities. In the A06–F07 and F07–T08 interaction, we found some electronic interactions caused by the halogen bonding and hydrogen bonding via halogen besides the general van der Waals interactions such as CH– $\pi$ , lone pair (LP)– $\pi$ , and  $\pi$ – $\pi$  interactions. The self-cohesive properties due to halogen bonding interactions and hydrogen bonding via halogen orbitals in TDS3 is a unique property not shared by TDS1, which we consider the essential factor causing double-stranded DNA destabilization. Although the conformational change occurs during thermal agitation, the trifluoromethyl group exhibits interactions by the



**Figure 5.** Gaussian approximated distributions of interaction energies between “T or F07” and “(A) A06, (B) T08, (C) A19, (D) A20, or (E) F21”. The positive and negative interactions are corresponding to attractive and repulsive interactions, respectively. The unit is in kcal/mol.

bonding via halogen orbitals to A06 or T08. In fact, when the conformation estimated the minimum interaction energy between F07 and A06, we found the bonding interactions via halogen orbitals to T08.

Although the thermodynamic conformational sampling with molecular dynamics were carried out in the water phase, the above estimations of interaction energies were calculated in the gas phase environment. Therefore, in order to confirm the influence of the solvation effect on the interaction energy estimations, we also tried to estimate the interaction energies with the IEFPCM method for five representative conformations selected by conformational clustering. In a biological environment, it can be assumed that DNA exist in water phase, but individual bases in DNA are sandwiched between other bases having permittivity lower than that of water. Therefore, as a solvent, in addition to the case of water, we considered the case of benzene and DMSO, which have a lower permittivity than water. Figure S5 and Table S2 show the interbase interaction energies between T/F07 and surrounding bases with MP2/6-31 G\* level. These results means that although the interaction energy values generally changes depending on the solvent, there is no qualitative difference in the distribution of interaction energy. Therefore, our results even for gas phase approximation in qualitative discussions at the previous paragraph are considered sufficiently reliable.

**Difference of thermodynamic parameter.** As these results, it can be predicted that DNA containing FTD has a lower melting temperature ( $T_m$ ) of the double helix structure than normal type DNA and DNA containing 5FU. Therefore, we determined the thermodynamic parameters, such as  $T_m$ ,  $-\Delta H$ , and  $-\Delta S$ , of a series of duplexes comprising TDS1, TDS2, and TDS3 by UV melting experiments, to experimentally verify the thermodynamic behavior predicted from the theoretical calculations. Table 2 shows the melting temperature ( $T_m$ ), change

		TDS1	TDS2	TDS3
$T_m$	(°C)	64.3	58.9	56.8
$-\Delta H$	(kcal/mol)	50.7	38.1	31.9
$-\Delta S$	(cal/mol/K)	130	94.4	76.2

**Table 2.** Thermodynamic analyses for TDS1, TDS2, and TDS3.

in enthalpy ( $-\Delta H$ ), and entropy ( $-\Delta S$ ) for each TDS. The values of  $T_m$  for TDS1, TDS2, and TDS3 were 64.2, 58.9, and 56.8 °C, respectively. The  $T_m$  value decreased by approximately 5 °C (or 7 °C, after 5FU (or FTD) incorporation into the DNA. Moreover, the values of enthalpy change and entropy change were altered according to the change in the  $T_m$  value. In other word, DNA duplexes containing 5FU (TDS2) or FTD (TDS3) are unstable compared with the normal DNA (TDS1). In addition, TDS3 is slightly less stable than TDS2. As already mentioned, the interbase interactions substantially change because of the fluorine substituent incorporated into the DNA. We believe these thermodynamic experimental results are due to a change in the hydrogen bonding properties between the complementary strands. Therefore, it might be that the incorporation of FTD into the DNA causes a strong attraction in the nucleic acid chain that breaks the hydrogen bonding to the complementary base, resulting in the loss of thermal DNA stability. It is possible that this difference between the normal DNA and the one incorporating FTD changes the affinity of the DNA binding protein for the molecule, leading to transcription and replication inhibitions similar but stronger to those due to DNA methylation. These physical and biological behaviors may cause the unique pharmaceutical actions of FTD as shown in Fig. S6. Although it is interested in the association between DNA binding protein and FTD-included DNA, experimental verifications to investigate above behaviors are very difficult. Therefore, in order to understand these considerations in detail, we are now developing a novel method for predicting the binding mode between DNA recognition protein and any DNAs with arbitrary sequence.

**Antitumor drugs administration test.** If the thermal instability of the DNA is related to the pharmacological mechanism, FTD should be effective not only for colorectal cancer but also for other various cancer types. Thus, we performed antitumor drug administration tests with a subcutaneous implantation model in nude mice using PANC-1 cells (a human pancreatic adenocarcinoma cell line). Figure S6 shows the results of the time course observation for the relative body weight rate (A) and for the relative tumor volume rate (B). We observed no difference in the relative body weight rate between control, 5FU (S-1), and FTD (trifluorothymidine/tipiracil, TAS-102) groups. As mentioned above, the pharmacological effect of 5FU is not caused by incorporation into DNA like FTD. Then although we could not do a simple comparison with, we found that marked differences in the relative tumor volume rate between these groups. Although we showed antitumor effects in both 5FU and FTD groups, the FTD group exhibited a higher effect than the 5FU group after 24 days. Table S3 presents these numerical data along with the results using gemcitabine (GEM), one of the antitumor drugs for pancreatic cancer; as shown in this table, the FTD group exhibited an antitumor effect potency equivalent to or higher than that in the GEM group. As expected, our results revealed that FTD suppresses the tumor volume increase without causing more weight loss. Since the thermal instability by the FTD incorporation into the DNA causes its antitumor effects, FTD should cause antitumor effects regardless of the cancer type.

## Conclusion

By using theoretical methods such as molecular dynamics and quantum chemistry, we could show the difference of intra-DNA interaction and the change of the 3D conformation by incorporation of nucleic acid having fluorine substituent into DNA. Then, the results that might clarify the pharmacological effect were found out. The incorporation of FTD into DNA destabilizes the conformation of double-helical DNA via halogen bonding and dispersion interactions, therefore, the FDT incorporated DNA will result in alterations in affinity to protein bindings. Thus, the calculation technology is a bona fide useful entity to raise the standard of re-researches in life science.

## Methods

**Energy minimization and Molecular dynamics simulation.** DNA structures of TDS1, TDS2, and TDS3 based on the B-type DNA were constructed, to analyze the conformational change caused by each the fluorine substituent in TDS2 and TDS3, compared with TDS1. The B-type DNA is right-handed with about 10 base pairs per turn. For each the DNA structure in about 6200 water molecules, energy minimizations was performed using the AMBER 12 program package. The AMBER99, GAFF, and TIP3P force fields for the complete DNA and water molecules were used respectively. In particular, a general AMBER force field (GAFF) was used for 5FU and FTD. After the minimization calculations, the NVT MD simulations were performed at 310 K with the periodic boundary condition from minimized structures to obtain a sample for confirming each DNA structure. The time step and total simulation time were 0.2 fs and 10 ns (50,000,000 steps), respectively.

**Ab initio molecular orbital method.** Simple base pair models, such as adenine–thymidine, adenine–fluorouracil, and adenine–FTD, were constructed to analyze the difference in the hydrogen bonding intensity according to the presence or absence of the fluorine substituent. In these models, a methyl group substituted each 1st-carbon in ribose. These simple model structures and simple mode bases were optimized by MP2 using the 6-31++G\*\* Gaussian basis set, followed by the stabilization (dissociation) energies ( $\Delta E_{Dissoc}$ ). In addition, sep-

ated each molecular structure was optimized. We estimated each dissociation energy with the energies of optimized structures ( $\Delta E_{Dimer}$ ,  $\Delta E_{Adenine}$ ,  $\Delta E_{Thymidine\ like}$ ) using the following equation.

$$\Delta E_{Dissoc} = (\Delta E_{Adenine} + \Delta E_{Thymidine\ like}) - \Delta E_{Dimer}$$

In the interbase interaction analyses, eight bases (5A, 6A, “7T or F”, 8T, 19A, 20A “21T or F”, and 22T) were selected from the trajectories of MD simulations for TDS1 and TDS3, respectively. Thereafter, we proposed the partial base models comprising these eight bases. Each 1st-carbon in ribose was the substituted hydrogen atom in these models. Only the hydrogen atoms were reoptimized. Interaction energies were estimated using above the equation for dissociation energy although in these interaction energies calculation the molecular structures were fixed. Then, the distribution of interaction energies was evaluated to estimate interaction energies between eight bases, each base energy, and any pair energy using each of the models at the MP2/6-31 G\* level. In addition, we performed the interaction energies’ estimation with solvation effects using the IEFPCM method. For comprehensively analyzing the electronic interaction, NBO analyses were used. In this study, all ab initio molecular orbital calculations were performed with Gaussian 09.

**Thermal denaturation experiments for intended DNA.** Thermodynamic data, such as  $T_m$ ,  $\Delta H$ , and  $\Delta S$ , were estimated for TDSs with ultraviolet and visible spectrophotometry (SHIMADZU UV-1800 spectrometer) equipped with Tm analysis accessory quartz cuvettes with 0.1 cm optical path length. The samples were prepared in 10 mM sodium phosphate buffer (pH 7.2) containing 100 mM NaCl and 17–70  $\mu$ M of each oligonucleotides at the final concentration, annealed by heating at 100 °C, followed by slow cooling to room temperature. The melting profile was measured at 260 nm at 5–90 °C at the scan rate of 0.5 °C/min. In addition, we used the two-point average method to determine the  $T_m$  values, and the final values were determined by averaging three independent measurements that were accurate within 1 °C.

**Antitumor drugs administration test with subcutaneous implantation model mouse.** This study was approved by the Animal Experiments Committee, Osaka University (approval number: 30-011-008). Our all experiments were performed in accordance with relevant guidelines and regulations. We purchased 6-week-old female Crlj:SHO-PrkdcscidHrhr mice from Charles River Laboratories Japan, and maintained them in a specific pathogen-free room. After 1 week, PANC-1 cells ( $1.0 \times 10^7$  cells/150  $\mu$ L) were subcutaneously transplanted under the skin of their left inguinal part. Once the subcutaneous tumor volume had reached 50 mm<sup>3</sup>, we initiated anticancer drug administrations. Control mice got 10 mL/kg of D-PBS(–) injected into the tail vein once a week. GEM mice got 10 mL/kg of gemcitabine hydrochloride (10 mg/mL) injected into the tail vein once a week. 5FU (S-1) mice got tegafur, gimeracil, and oteracil at a weight ratio of 1:0.29:0.98; 10 mL/kg of this combination (S-1; 0.83 mg/mL) were administered orally once a day for a week. FTD (FTD/TPI, TAS-102) mice got trifluridine and tipiracil hydrochloride at a weight ratio of 1:0.47; 10 mL/kg of this combination (FTD/TPI; 15 mg/mL) were administered orally once a day for 1 week. We evaluated the subcutaneous tumor volume as follows using the following formulas: (greatest diameter)  $\times$  (shortest diameter)<sup>2</sup>/2.

Received: 28 February 2019; Accepted: 3 January 2020;

Published online: 24 January 2020

## References

- Demura, M. & Saijoh, K. The Role of DNA Methylation in Hypertension. *Adv. Exp. Med. Biol.* **956**, 583–598 (2017).
- Zhong, J. *et al.* The Role of DNA Methylation in Cardiovascular Risk and Disease. *Circ. Res.* **118**(1), 119–131 (2016).
- Lizardi, P. M. *et al.* High-Throughput Deep Sequencing for Mapping Mammalian DNA Methylation. *Cold Spring Harb. Protoc.* Dec. **1**, (2017).
- Schuyler, R. P. *et al.* Distinct Trends of DNA Methylation Patterning in the innate and adaptive immune systems. *Cell Rep.* **17**(8), 2101–2111 (2016).
- Gillis, E. P. *et al.* Applications of Fluorine in Medicinal Chemistry. *J. Med. Chem.* **58**(21), 8315–8359 (2015).
- Tachikawa, H. & Kawabata, H. Molecular Design of Ionization-Induced Proton Switching Element Based on Fluorinated DNA Base Pair. *J. Phys. Chem. A* **120**(9), 1529–1535 (2016).
- Liu, P. *et al.* Fluorinated nucleosides: Synthesis and biological implication. *J. Fluorine Chem.* **129**(9), 743–766 (2008).
- Qiu, X.-L. *et al.* Recent advances in the synthesis of fluorinated nucleosides. *Tetrahedron* **66**(4), 789–843 (2010).
- Wójtowicz-Rajchel, H. Synthesis and applications of fluorinated nucleoside analogues. *J. Fluorine Chem.* **143**, 11–48 (2012).
- Hartmann, K.-U. & Heidelberger, C. Studies on fluorinated pyrimidines. *J. Biol. Chem.* **236**(11), 3006–3013 (1961).
- Spiegelman, S. *et al.* Potentiation of the anti-tumor activity 5FU by thymidine and its correlation with the formation of (5FU) RNA. *Cancer* **45**(5), 1129–1134 (1980).
- Morgan, R. G. Leucovorin enhancement of the effects of the fluoropyrimidines on thymidylate synthase. *Cancer* **63**(S6), 1008–1012 (1989).
- Heidelberger, C. *et al.* Syntheses of 5-trifluoromethyluracil and 5-trifluoromethyl-2'-deoxyuridine. *J. Med. Chem.* **7**(1), 1–5 (1964).
- Ryan, K. J. *et al.* Chemical Synthesis of 2'-Deoxy-5-(trifluoromethyl)uridine and the  $\alpha$ -Anomer. *J. Org. Chem.* **31**(4), 1181–1184 (1966).
- Fujiwara, Y. *et al.* Fluorinated Pyrimidines XXXVII. Effects of 5-Trifluoromethyl-2'-deoxyuridine on the Synthesis of Deoxyribonucleic Acid of Mammalian Cells in Culture. *Mol. Pharmacol.* **6**(3), 273–280 (1970).
- Mayer, R. J. *et al.* Randomized Trial of TAS-102 for Refractory Metastatic Colorectal Cancer. *N. Engl. J. Med.* **372**(20), 1909–1919 (2015).
- Santi, D. V. & Sakai, T. T. Thymidylate synthetase. Model studies of inhibition by 5-trifluoromethyl-2'-deoxyuridylic acid. *Biochemistry* **10**(19), 3598–3607 (1971).
- Fukushima, M. *et al.* Structure and activity of specific inhibitors of thymidine phosphorylase to potentiate the function of antitumor 2'-deoxyribonucleosides. *Biochem. Pharmacol.* **59**(10), 1227–1236 (2000).
- Murakami, Y. *et al.* Different mechanisms of acquired resistance to fluorinated pyrimidines in human colorectal cancer cells. *Int. J. Oncol.* **17**(2), 277–283 (2000).



20. Emura, T. *et al.* An optimal dosing schedule for a novel combination antimetabolite, TAS-102, based on its intracellular metabolism and its incorporation into DNA. *Int. J. Oncol.* **13**(2), 249–255 (2004).
21. Emura, T. *et al.* Potentiation of the antitumor activity of  $\alpha$ ,  $\alpha$ ,  $\alpha$ -trifluorothymidine by the co-administration of an inhibitor of thymidine phosphorylase at a suitable molar ratio *in vivo*. *Int. J. Oncol.* **27**(2), 449–455 (2005).
22. Suzuki, N. *et al.* Mode of action of trifluorothymidine (TFT) against DNA replication and repair enzymes. *Int. J. Oncol.* **39**(1), 263–270 (2011).
23. Frisch, M. J. *et al.* *Gaussian 09*, revision A.01; Gaussian Inc.: Wallingford, CT. (2009).
24. Markley, J. C. *et al.* Incorporation of 2'-deoxy-5-(trifluoromethyl)uridine and 5-Cyano-2'-deoxyuridine into DNA. *Bioorg. Med. Chem. Lett.* **11**(18), 2453–2455 (2001).
25. Case, D. A. *et al.* *AMBER 12* 12th ed., University of California, San Francisco (2012).
26. Wang, J. *et al.* How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comp. Chem.* **21**(12), 1049–1074 (2000).
27. Wang, J. *et al.* Development and testing of a general amber force field. *J. Comput. Chem.* **25**(9), 1157–1174 (2004).
28. Jorgensen, W. L. *et al.* Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**(2), 926–935 (1983).
29. Gmeiner, W. H. *et al.* Synthesis, annealing properties, fluorine-19 NMR characterization, and detection limits of a trifluorothymidine-labeled antisense oligodeoxyribonucleotide 21 mer'. *J. Org. Chem.* **56**(11), 3602–3608 (1991).
30. Desiraju, G. & Steiner, T. *The Weak Hydrogen Bond – In Structural Chemistry and Biology*. Oxford Science Publications (1999).
31. Srinivasan, A. R. *et al.* Properties of the Nucleic-acid Bases in Free and Watson-Crick Hydrogen-bonded States: Computational Insights into the Sequence-dependent Features of Double-helical DNA. *Biophys. Rev.* **1**(1), 13–20 (2009).
32. Daido, M. *et al.* Nuclear Quantum Effect and Temperature Dependency on the Hydrogen-Bonded Structure of Base Pairs. *J. Comput. Chem.* **34**(28), 2403–2411 (2013).
33. Elstner, M. *et al.* Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density-functional-theory based treatment. *J. Chem. Phys.* **114**(12), 5149–5155 (2001).
34. Reed, A. E. *et al.* Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **88**(6), 899–926 (1988).
35. Foster, J. P. & Weinhold, F. Natural hybrid orbitals. *J. Am. Chem. Soc.* **102**(24), 7211–7218 (1980).
36. Reed, A. E. & Weinhold, F. Natural localized molecular orbitals. *J. Chem. Phys.* **83**(4), 1736–1740 (1983).

## Acknowledgements

This work received the following financial support: grants-in-aid for Scientific Research from the Japan Agency for Medical Research and Development and the Ministry of Education, Culture, Sports, Science, and Technology (Grant Nos. 19K22658 (H. Ishii), 19K07688 (M. Konno), 18K16356 (J. Koseki), 18K16355 (A. Ayumu), 18KK0251 (H. Ishii), 17H04282 (H. Ishii), and 17K19698 (H. Ishii); and Taiho Pharmaceutical Co., Ltd.

## Author contributions

J.K. performed some computational analyses. J.K., A.A., N.H. and S.O. performed thermodynamic experiment. M.K., K.T., K.K. and H.I. performed some animal experiments. J.K., M.K., A.A., Y.D., M.M. and H.I. designed the analyses. J.K., M.K., A.A. and H.I. wrote the article.

## Competing interests

This study was funded by Taiho Pharmaceutical Co., Ltd. (Tokyo, Japan), Evidence Based Medical (EBM) Research Center INC. (Osaka, Japan), UNITECH Co., Ltd. (Chiba, Japan), ID EA Consultants Inc. (Tokyo, Japan), Kinshu-kai Medical Corporation (Osaka, Japan). These funders had no role in the main experimental equipment, supplies expenses, study design, data collection and analysis, decision to publish, or preparation of this manuscript. K.T. is employed of Taiho Pharmaceutical Co., Ltd. (Tokyo Japan).

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-57899-7>.

**Correspondence** and requests for materials should be addressed to H.I.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020