Journal of Molecular Graphics and Modelling 47 (2014) 62–76

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Investigation of base pairs containing oxidized guanine using *ab initio* method and ABEEM$\sigma\pi$ polarizable force field

Cui Liu*, Yang Wang, Dongxia Zhao, Lidong Gong, Zhongzhi Yang

Liaoning Normal University, Dalian 116029, China

ARTICLE INFO

Article history:
Accepted 21 October 2013
Available online 31 October 2013

Keywords:
Ab initio method
ABEEM$\sigma\pi$ polarizable force field
Mutation
Hydrogen-bonding and stacking interaction
Oxidized guanine base pairs

ABSTRACT

The integrity of the genetic information is constantly threatened by oxidizing agents. Oxidized guanines have all been linked to different types of cancers. Theoretical approaches supplement the assorted experimental techniques, and bring new sight and opportunities to investigate the underlying microscopic mechanics. Unfortunately, there is no specific force field to DNA system including oxidized guanines. Taking high level *ab initio* calculations as benchmark, we developed the ABEEM$\sigma\pi$ fluctuating charge force field, which uses multiple fluctuating charges per atom. And it was applied to study the energies, structures and mutations of base pairs containing oxidized guanines. The geometries were obtained in reference to other studies or using B3LYP/6-31+G* level optimization, which is more rational and time-saving among 24 quantum mechanical methods selected and tested by this work. The energies were determined at MP2/aug-cc-pVDZ level with BSSE corrections. Results show that the constructed potential function can accurately simulate the change of H-bond and the buckled angle formed by two base planes induced by oxidized guanine, and it provides reliable information of hydrogen bonding, stacking interaction and the mutation processes. The performance of ABEEM$\sigma\pi$ polarizable force field in predicting the bond lengths, bond angles, dipole moments *etc.* is generally better than those of the common force fields. And the accuracy of ABEEM$\sigma\pi$ PFF is close to that of the MP2 method. This shows that ABEEM$\sigma\pi$ model is a reliable choice for further research of dynamics behavior of DNA fragment including oxidized guanine.

1. Introduction

Generally, the canonical nucleic acid bases, adenine (A), thymine (T), guanine (G) and cytosine (C) exist as the main form in the double helix of DNA. The formation of specific purinepyrimidine Watson–Crick (WC) hydrogen bonds (H-bonds) is responsible for the maintenance of the genetic code. The integrity of the genetic information is constantly threatened by oxidizing agents [1–5]. Cellular DNA is constantly bombarded by reactive oxygen species (ROS) from both intra- and extracellular sources [5–7]. Of the four DNA nucleobases, guanine has the lowest redox potential [7] and is the most readily oxidized by ROS. Despite nearly a quarter of a century study, and a large number of base- and sugar-derived DNA lesions having been identified, the majority of studies have focused upon the guanine modification [8]. Oxidized guanines have all been linked to different types of cancers. For example, oxidation of guanine to 8-oxoguanine (8-oxo-G) leads to G→T transversions [1–5]. During replication of the flawed strand, oxidized guanines can be misread by polymerases as another kind of base and associated with a wrong purine or pyrimidine in the newly synthesized strand. These may lead to the introduction of a wrong genetic code [9]. The mechanism by which DNA repair enzymes such as formamidopyrimidine glycosylase (FPG/MutM) and human oGG1 enzyme (hoGG1) differentiate between 8-oxo-G and G. H-bonding and stacking interactions contribute to the stability and conformational variability of nucleic acids [10]. A proper description of these properties is useful to understand the basic principles governing the formation of the 3D nucleic acid architectures. Owing to the importance, there are numerous experimental [11–17] and computational [18–25] studies relative to the nucleotide base pairs. However, despite prolonged efforts, it remains a great challenge to understand the H-bonding and stacking interactions in oxidized guanine base pairs, and the mutagenic mechanism induced by oxidized guanine at the molecular level.

Detection and biophysical studies of modified nucleobases are challenging because they are not fluorescent and have broad electronic spectra that overlap with those of normal bases. *Ab initio* calculations supply the assorted experimental techniques, which are already capable of supplying the raw data and facts that are prerequisites for a complete understanding of the H-bonding and stacking interactions, and offered the opportunity to probe into the underlying sources of observable effects [18–21]. But, *ab initio*

* Corresponding author. Tel.: +86 411 82159607; fax: +86 411 82158977.
E-mail address: liuc@lnnu.edu.cn (C. Liu).

1093-3263/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2013.10.008

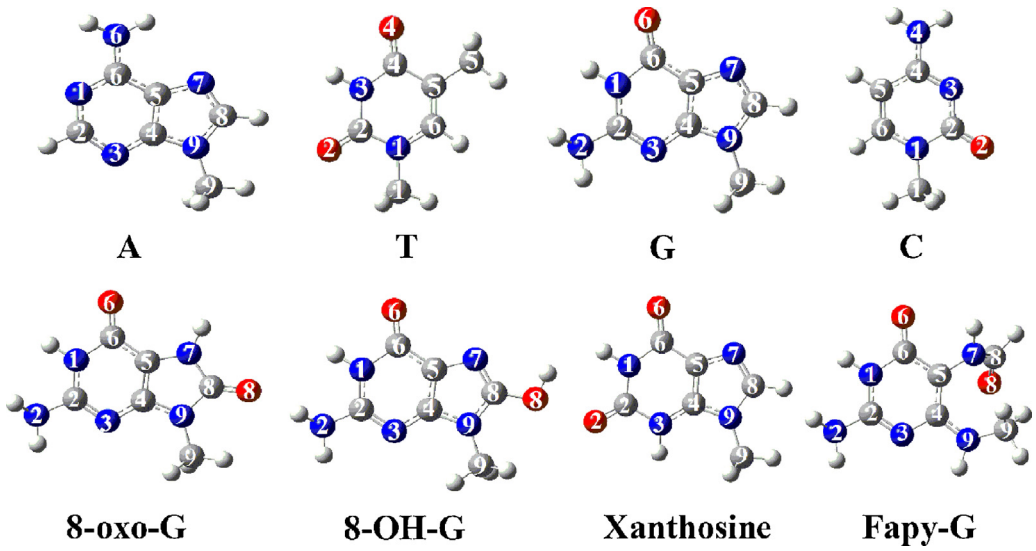

Fig. 1. The geometries and labels of canonical bases and oxidized guanines.

methods need high computational expense. Computer simulations by molecular mechanics have been a powerful tool to investigate properties of biomolecules [26,27]. The electrostatic energy makes significant contributions to the overall interaction energy of base pairs. Polarization and charge transfer effects are certainly not negligible. In view of the significance of the polarization effect, much has been achieved since the 1970s in the development and employment of the polarizable force field (PFF) for biomolecule simulation, such as AMBER [28–32], CHARMM [33–36], OPLS [37–40], NEMO [41], AMOEBA [42,43], QMPFF [44–46] *etc.* PFF mainly includes induced dipole or multipole [41–43], classical Drude oscillator [35,36] and fluctuating charge (FQ) models [33,37,47]. The computational expense of the induced dipole model is relatively high because of the iteration processes to obtain reasonable dipole, quadrupole or multipole moment. Classical Drude oscillator model introduces a massless charged particle attached to the polarizable atomic center. The positions of these "auxiliary" particles are then adjusted self-consistently to take into account the polarization. Classical Drude oscillator and induced dipole models do not consider charge transfer, which is very important for biomolecular simulation. FQ model can take charge transfer into account without iteration processes. It allows the values of partial charges to respond to the electric field of their environment. It can quickly calculate the partial charges of all the sites in the system by performing electronegativity equalization (EE) [48] or chemical potential equalization (CPE) [49] schemes in density functional theory.

ABEEM$\sigma\pi$, *i.e.*, the atom-bond electronegativity equalization method at $\sigma\pi$ level, presented by Yang and coworkers [50–52] introduces polarizability by allowing the values of partial charges to respond to their environment. It divides the single electron density of a molecule into atomic, bond (including σ and π bonds) and lone-pair regions. Recently, it was fused into molecular mechanics (MM) to establish a new generation PFF [52,53], ABEEM$\sigma\pi$ PFF or ABEEM$\sigma\pi$ fluctuating charge FF, which has been successfully applied to gas-phase water clusters and bulk water [53], aqueous ionic solutions [54–56], and biomolecules [52,57]. The results indicate that the ABEEM$\sigma\pi$ PFF is generally better than the common force fields (FFs), and its accuracy can reach or approach that of MP2 method, so it is promising to be broadly applicable. It is a well known that MP2 method coupled with large basis sets owns a much smaller error for H-bonding interactions [20,22,23], but it overestimates the binding energy of stacked dimers. Because developing the ABEEM$\sigma\pi$ fluctuating charge force field is our goal, and it was applied to study the energies, structures and mutations of base pairs containing oxidized guanines, it would be more appropriate to validate ABEEM$\sigma\pi$ results using a single *ab initio* method MP2/aug-cc-pVDZ calculation. Thus, the MP2/aug-cc-pVDZ method was selected as benchmark. In the present work, we performed theoretical investigations by *ab initio* method and ABEEM$\sigma\pi$ PFF on the stacking and H-bonding interactions of base pairs containing oxidized guanines. Then, we will summarize the structural (including geometric and electronic structure) and energetic changes caused by oxidized guanines, and further analyze the mutagenic mechanism induced by them.

2. Model molecules and methods

2.1. Model molecules

In DNA, ribose links to phosphate. In this study, we are interested in the methylated bases, oxidized guanines and their base pairs to investigate the H-bonding interaction, which have the similar chemical-bonding environment with nucleotide system. The PES of H-bonded pairs contains a small number of deep energy minima while the PES of a stacked dimer contains broad and shallow low-energy regions separated by low energy barriers. Further, as revealed by gradient geometry optimization, some stacked dimers appear to not have a minimum corresponding to a stacked arrangement, since there has been a spontaneous transition to a H-bonded structure. The bases were not methylated being kept rigid and planar, assuming MP2-optimised monomer geometries as other references [19,23]. 8-Oxo-G, 8-OH-G, Xanthosine and Fapy-G are the common forms of oxidative guanine. 8-Oxo-G results from the oxidation of guanine C8-N7 bond (an O atom in place of an H atom at C8 and an H atom at N7); 8-OH-G is an OH in place of an H atom at C8; the NH_2 at C2 is substituted with an O atom to form Xanthosine; C8-N9 bond in normal guanine is broken to form Fapy-G. These structures of canonical bases and oxidized guanines studied in this work are shown in Fig. 1. Basically, the double helix of DNA employs a Watson–Crick G:C and A:T H-bonding patterns. In Watson–Crick A:T pattern, A also has additional H-bond donor and acceptor, so it can pair with T to form Hoogsteen or reverse Hoogsteen H-bonding pattern, *i.e.* T:A:T or T:A:T(R) triplet bases, which are the fundamental constitutional units in triple helix. G also can pair with C to form Hoogsteen H-bonding pattern, where C is in the

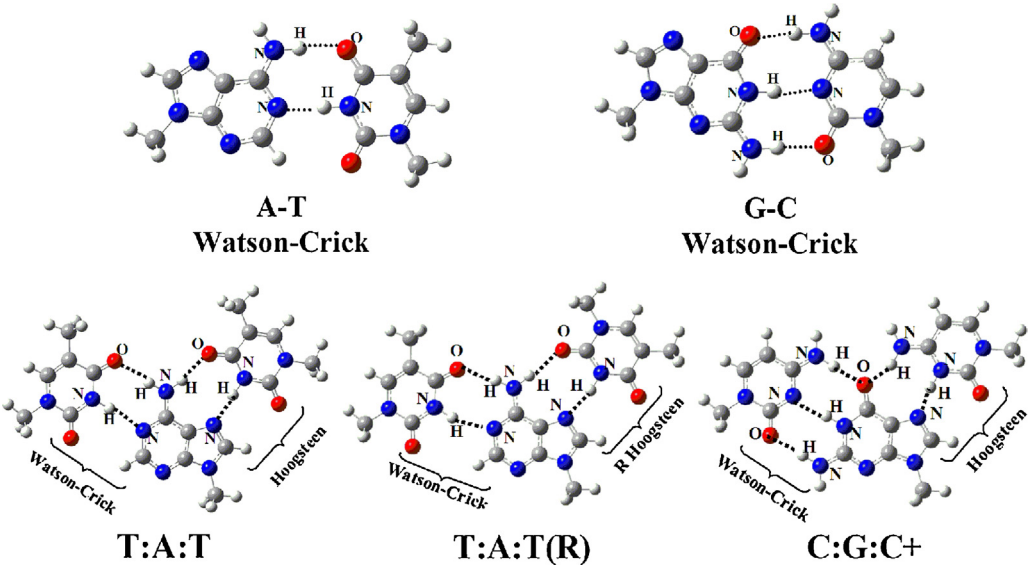

Fig. 2. The geometries of base pairs and triplet bases.

protonated form, and this pattern is not considered in the present study. Base pairs and triplet bases are depicted in Fig. 2.

The geometry of guanine has been changed after oxidation, so there are many types of binding, not only Watson–Crick, Hoogsteen or reverse Hoogsteen patterns but also other special patterns. In order to comprehensively get insight into the pairing behavior of the oxidized guanines, this work investigates 42 modified base pairs according to above introduction. Their geometries are shown in Figs. 3–6.

2.2. Quantum mechanic calculations

The structures of canonical bases and their base pairs have been measured by experiment, while the information of modified bases and their base pairs is insufficient. In order to choose a proper method for optimizing the structures, we use various combinations of four *ab initio* methods (MP2, B3LYP, B97D, wB97XD) and six basis sets (6-31G, 6-31G*, 6-31G**, 6-31+G*, 6-31+G**, 6-311+G**). These structures of A, T, G, C, G:C, A:T from the 24 model chemistries were compared with those of experiment [58]. Tables S1–S3 (Supporting Information) list the root of mean square deviation (RMSD) between the optimized structures and experimental structures. The B3LYP/6-31+G* level is a better compromise in both computational cost and rationality to optimize those structures, which were also characterized as minima by frequency calculation performed using the same level of theory. Subsequently, energetic calculations were performed at the MP2/aug-cc-pVDZ level with basis set superposition error correction by counterpoise method. All calculations have been performed with the Gaussian 03 package [59].

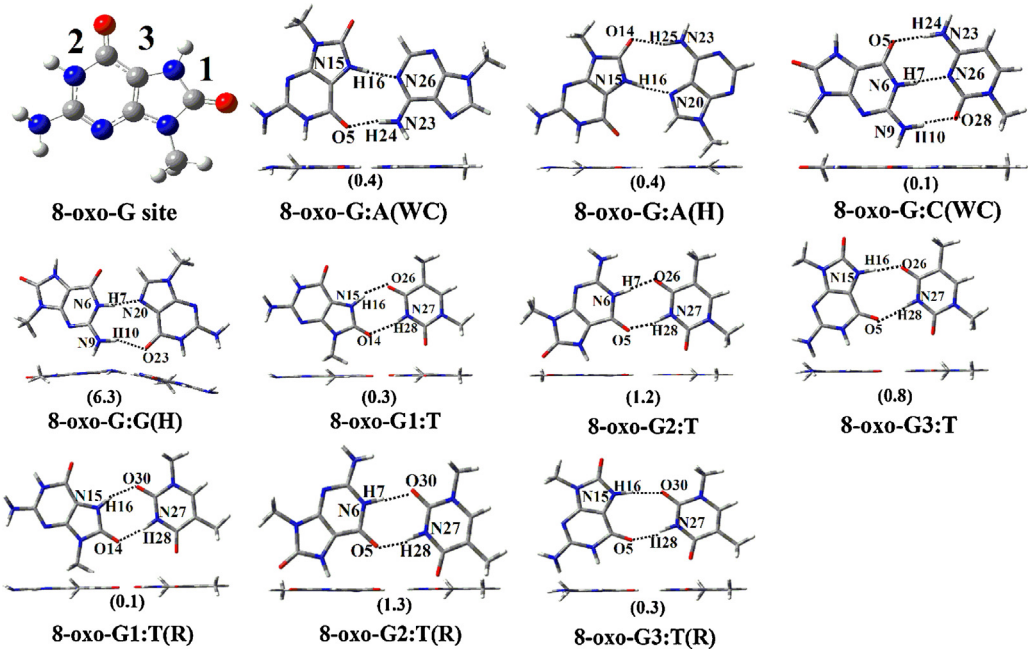

Fig. 3. Oxidized base pairs containing 8-oxo-G. WC represents Watson–Crick H-bonding pattern. H represents Hoogsteen H-bonding pattern. R represents reverse Watson-Crick or reverse Hoogsteen H-bonding pattern. Black numbers 1–3 denote these sites, at which 8-oxo-G pairs with A, T, G, or C. White numbers denote atomic labels. The number (in°) in bracket is the deviation degree defined in following paragraph.

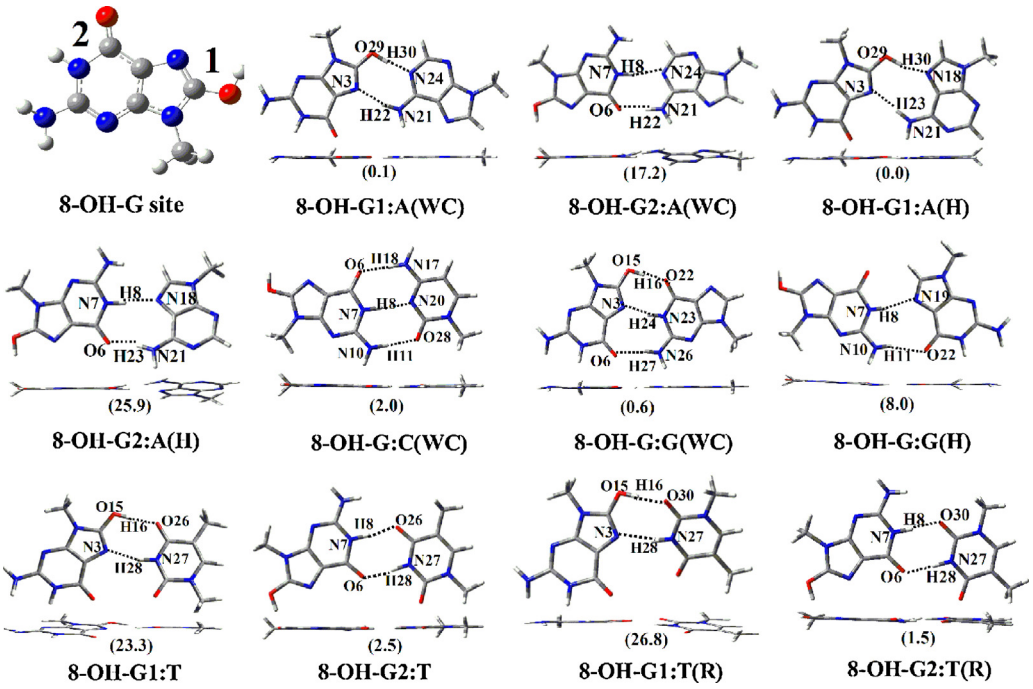

Fig. 4. Oxidized base pairs containing 8-OH-G. These explanations refer to Fig. 3.

2.3. *ABEEMσπ polarizable force field*

The potential of each system was also minimized by ABEEMσπ fluctuating charge FF through the modified Tinker program. The energy function of ABEEMσπ PFF is written as:

$$E_{ABEEM\sigma\pi} = E_b + E_\theta + E_\phi + E_{\text{imptors}} + E_{vdW} + E_{\text{elec}} \quad (1)$$

where E_b is the energy of bond stretching, E_θ represents the energy required for bending an angle, and these two terms is dealt with the usual harmonic oscillator style as follows:

$$E_b(r) = \sum_{\text{bonds}} k_r(r - r_{\text{eq}})^2 \quad (2)$$

$$E_\theta(\theta) = \sum_{\text{angles}} k_\theta(\theta - \theta_{\text{eq}})^2 \quad (3)$$

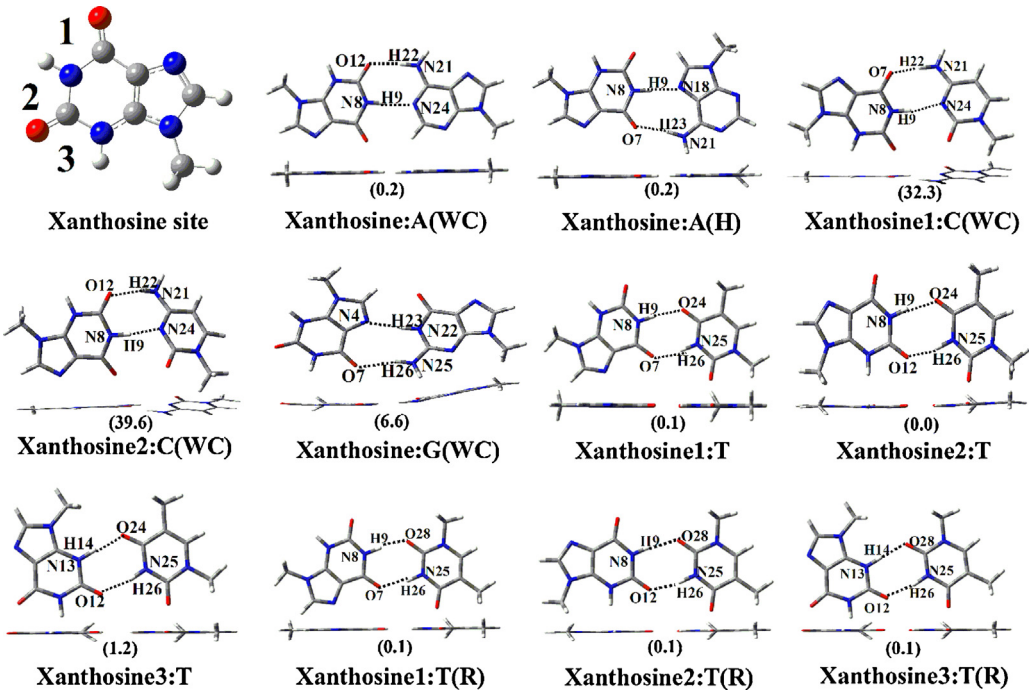

Fig. 5. Oxidized base pairs containing Xanthosine. These explanations refer to Fig. 3.

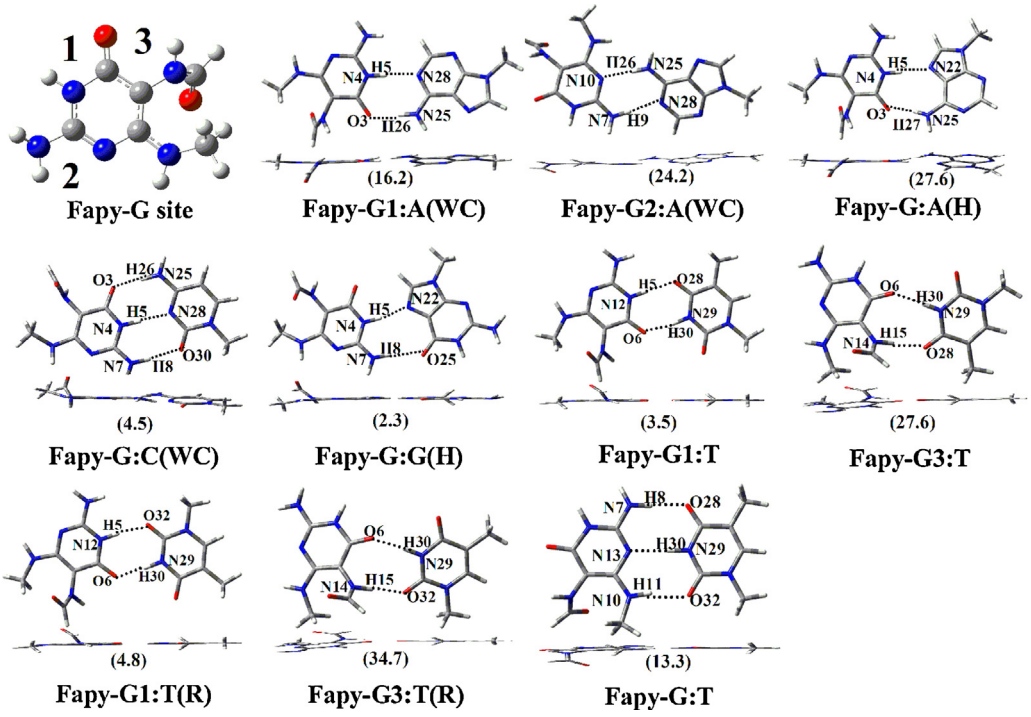

Fig. 6. Oxidized base pairs containing Fapy-G. These explanations refer to Fig. 3.

where k_r and k_θ represent the force constants of the stretching and bending, respectively; r and θ are the actual values of bond length and bond angle, respectively; r_{eq} and θ_{eq} denote the equilibrium values of the bond length and bond angle, respectively.E_φ is the torsional energy for rotation around a bond, and $E_{imptors}$ is the improper dihedral angle term, and they are expressed as:

$$E_\phi(\phi) = \sum_{\text{torsions}} \left[\frac{V_1}{2}(1+\cos\phi) + \frac{V_2}{2}(1-\cos 2\phi) + \frac{V_3}{2}(1+\cos 3\phi)\right] \tag{4}$$

$$E_{\text{imptors}} = \sum_{\text{imptors}} v(1-\cos 2\phi) \tag{5}$$

where V_1, V_2, V_3 and v are the dihedral angle and improper dihedral angle force constants, respectively.

E_{vdW} and E_{elec} are nonbonded interactions. The Lennard–Jones 12-6 function is employed to represent the van der Waals interaction.

$$E_{vdw} = \sum_{i<j} 4 f_{ij}\varepsilon_{ij}\left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}}\right) \tag{6}$$

Geometric combining rules for the Lennard–Jones coefficients are employed: $\sigma_{ij} = (\sigma_{ii}\sigma_{jj})^{1/2}$ and $\varepsilon_{ij} = (\varepsilon_{ii}\varepsilon_{jj})^{1/2}$. The summation runs over all of the pairs of unbonded atoms i and j. If i and j are intramolecular, the coefficient $f_{ij} = 0.0$ for any i–j pair connected by a valence bond (1–2 pairs) or a valence bond angle (1–3 pairs), $f_{ij} = 0.5$ for 1,4 interactions (atoms separated by three bonds), and $f_{ij} = 1.0$ for all other intramolecular and intermolecular cases.

Effective Coulombic form is chosen to calculate the charge–charge interactions.

$$E_{\text{elec}} = \sum_{i<j} \frac{k_{ij} q_i q_j}{r_{ij}} \tag{7}$$

where r_{ij} is the distance between sites i and j. The partial charges q_i and q_j are obtained by the ABEEM$\sigma\pi$ model [50,51]. The ABEEM$\sigma\pi$ regions of 8-oxo-G are shown in Fig. 7. σ bond partial charge is placed on the bond at the point that partitions the bond length according to the ratio of atomic covalent radii of bonding atoms. The two lone pair partial charges of oxygen atom are placed in the base plane as far as the covalent radius (0.74 Å) of O atom. Angles between the two lone pairs and C=O bond are 120°. The lone pair partial charge of nitrogen atom is also placed in the base plane as far as the covalent radius (0.74 Å) of N atom. The line joining N atom and its lone pair is the bisector of angle C=N=C. The π bond regions are perpendicular to the base plane, which exist above and below the corresponding atomic region whose center is located at the covalent radius of the atom, and these π bond regions are omitted in Fig. 7 for brief.

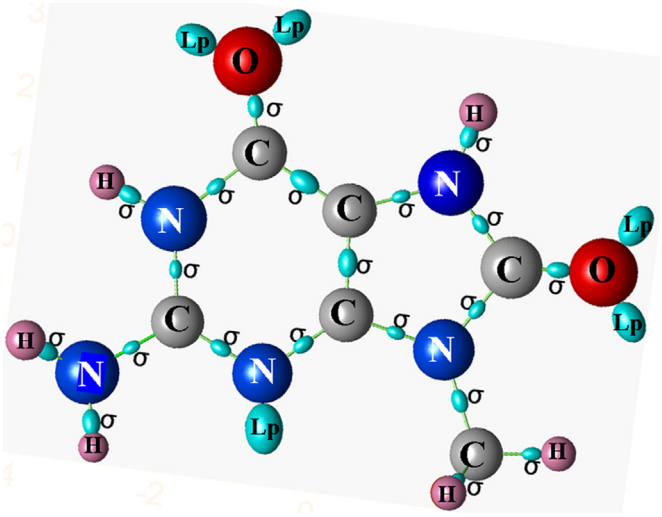

Fig. 7. The sketch of partial charge regions (expect π bond regions) in 8-oxo-G defined in ABEEM$\sigma\pi$ model.

Table 1
Parameters in the function $k_{\text{H-bond}}$.

H-bond type	k'	C	$r_{\text{lp,H}}$	D
lpN_{327}, H—O_{81}	0.64	0.1155	1.1581	0.0010
lpN_{327}, H—N_{347}	0.54	0.1006	1.3223	0.0496
lpN_{327}, H—N_{317}	0.56	0.0800	1.1581	0.0010
lpN_{337}, H—O_{81}	0.67	0.1400	1.1200	0.0011
lpN_{337}, H—N_{347}	0.57	0.1046	1.2956	0.0879
lpN_{337}, H—N_{317}	0.61	0.1100	1.1200	0.0011
lpO_{318}, H—N_{77}	0.61	0.0496	1.3978	0.1034
lpO_{318}, H—O_{81}	0.55	0.0127	1.9482	0.0434
lpO_{318}, H—N_{347}	0.55	0.0931	1.3019	0.0319
lpO_{318}, H—N_{317}	0.55	0.1806	1.2523	0.0536
lpO_{328}, H—N_{77}	0.64	0.0490	1.3978	0.1034
lpO_{328}, H—O_{81}	0.56	0.0107	1.9482	0.0434
lpO_{328}, H—N_{347}	0.59	0.1806	1.2523	0.0536
lpO_{328}, H—N_{317}	0.55	0.0714	1.3019	0.0319
lpO_{338}, H—N_{347}	0.62	0.1806	1.2523	0.0536
lpO_{338}, H—N_{317}	0.54	0.0697	1.3019	0.0319

Table 2
Defined atomic types and nonbonded parameters χ^*, η^*, σ (Å) and ε (kcal/mol) in ABEEM$\sigma\pi$ model.

Type	Element	χ^*[a]	η^*[a]	σ	ε
381	C	2.00	5.23	3.4000	0.0860
338	O	3.95	9.60	2.8600	0.1900
357	N	3.62	6.66	2.7500	0.1700
39	C	2.00	5.23	3.4000	0.0860
187	O_{lp}	4.16	6.81		
188	O_{lp}	4.16	6.81		

[a] The scaling Pauling electronegativity unit is used.

The parameter k_{ij} in equation 7 is a result of considering the exchange, penetration, and shielding effect in the interaction between the two pieces of electron clouds i and j. k_{ij} is set to be 0.57 empirically, except for the H-bond regions. In order to effectively describe the intermolecular H-bonding interaction, especially the electrostatic interaction, we introduce the function, $k_{\text{H-bond}}(R_{\text{lp,H}})$, which is dependent on the distance between the lone-pair electron and the hydrogen atom, $R_{\text{lp,H}}$. They are obtained through the energy fitting between values of *ab initio* method and ABEEM$\sigma\pi$ PFF at different distances of the H atom and its acceptor in small molecule dimers [53]. In the present study, $k_{\text{H-bond}}(R_{\text{lp,H}})$ is expressed as follows:

$$k_{\text{H-bond}}(R_{\text{lp,H}}) = k' - \frac{C}{1 + e^{(R_{\text{lp,H}} - r_{\text{lp,H}}/D)}} \quad (8)$$

k', C, D and $r_{\text{lp,H}}$ are the parameters in the function as listed in Table 1.

We have done a great deal of work to fit those parameters of oxidized guanines. As shown in Fig. 1, C8 in 8-OH-G is a new type relative to the canonical bases, so we have fitted these corresponding parameters. C2 or C8 is substituted with an O atom to form Xanthosine or 8-oxo-G, and the corresponding parameters of angles and torsional angles are obviously different. We can see that C8 locates in five-member ring, but C2 locates in six-member ring, so C8 in 8-oxo-G was defined using a new label. The heavy atoms in the modified bases are in the same plane. To achieve the purpose, in the processing of fitting the parameters, it is necessary to refit the torsional parameters and improper torsional parameters. There are also some similar circumstances. The parameters χ^* and η^* are fitted through a regression and least-squares optimization procedure, where the scaling Pauling electronegativity unit is used. The parameters are fitted not only to reproduce the charges of the *ab initio* calculation but also to obtain good dipole moments, structures, and binding energies compared with those from experimental measurements and quantum mechanic calculations. Parameters for the

Table 3
Bending vibration parameters of bond lengths (Å) and angles (°).

Bond	k_r[a]	r_{eq}[b]	Angle	k_θ[c]	θ[d]	Angle	k_θ	θ
75-23	414.0	1.386	79-71-75	70.00	114.0	75-23-21	50.00	120.9
71-91	418.0	1.381	79-75-71	70.00	130.0	78-75-79	63.00	119.2
73-91	529.0	1.310	71-75-71	70.00	126.2	78-72-78	40.00	116.6
7-91	320.0	1.426	91-71-75	70.00	103.8	75-71-92	70.00	109.7
71-92	418.0	1.381	91-71-83	30.00	115.8	92-71-83	30.00	122.5
81-92	570.0	1.229	71-91-73	70.00	113.0	71-92-81	80.00	126.5
100-92	418.0	1.381	71-91-7	35.00	116.1	92-100-83	30.00	125.2
100-75	436.0	1.376	73-91-7	35.00	128.1	71-92-100	70.00	105.0
100-83	434.0	1.010	91-73-75	70.00	103.8	100-92-81	80.00	129.9
			91-7-8	55.00	109.2	79-75-100	70.00	130.0
			78-75-23	70.00	120.0	100-75-75	70.00	105.2
			79-75-23	70.00	114.5	75-100-92	70.00	109.5
			75-23-24	30.00	119.7	75-100-83	30.00	125.8

[a] Bond length expansion force constant (k_r in kcal/mol/Å^2).
[b] The balance value of bond length.
[c] Angle bending force constant (k_θ in kcal/mol/rad^2).
[d] The balance value of angle.

Table 4
Torsional angles and improper torsional angles parameters (kcal/mol).

Torsion	v	Imptors	v
75-78-74-83	2.250	79-78-75-23	21.0
79-71-75-75	7.250	75-72-79-84	21.0
71-75-71-79	−1.500	71-74-78-72	21.0
23-75-79-71	7.250	74-75-78-72	21.0
23-75-78-72	7.250	75-79-71-83	21.0
78-75-79-71	7.250	71-75-75-71	2.2
72-78-78-71	13.250		

hard degrees of freedom (bond stretching and angle bending) are from AMBER FF without modification, because they can be transferred from one FF to another. We take the torsional, improper torsional and the Lennard–Jones parameters of AMBER as a reference and refit them by fitting quantum mechanic conformational energies, dimer binding energies, dipole moments, and so on, using a regression and least-squares method. The labels of ABEEM$\sigma\pi$ model and force field are listed in Fig. S1. The new labels of the oxidized guanines were added based on our previous work [52], and the newly added parameters are listed in Tables 2–4.

In the ABEEM$\sigma\pi$ method for calculating charge distribution in molecules, σ and π sites, lone pair sites of a molecular system, the effective energy function is expressed as Eq. (9):

where N_{mol} represents the number of interacting molecules in the system. If y stands for an arbitrary partial charge region, then E^*_{iy}, μ^*_{iy}, η^*_{iy} and q_{iy} are the valence state energy, the valence state chemical potential, the valence state hardness and the partial charge of the region y in molecule i, respectively. The y may be a single bond atom a, or double bond atom m, or single bond a—b, or σ bond region of double bond m=n, or π bond region of double bond m=n, or lone-pair lp in molecule i, respectively. The R is the distance between the charge sites.

The effective electronegativity of partial charge site is identified as the negative of the corresponding effective chemical potential:

$$\mu_i = \left(\frac{\partial E}{\partial N_i}\right)_{R,N_j} = -\left(\frac{\partial E}{\partial q_i}\right)_{R,q_j} = -\chi_i \tag{10}$$

$$\begin{aligned}
E = \sum_{i=1}^{N_{mol}} \Bigg\{ & \sum_a (E^*_{ia} - \mu^*_{ia} q_{ia} + \eta^*_{ia} q^2_{ia}) + \sum_m (E^*_{im} - \mu^*_{im} q_{im} + \eta^*_{im} q^2_{im}) + \sum_{a-b} (E^*_{i(a-b)} - \mu^*_{i(a-b)} q_{i(a-b)} + \eta^*_{i(a-b)} q^2_{i(a-b)}) \\
& + \sum_{\sigma m=n} (E^*_{i(\sigma m=n)} - \mu^*_{i(\sigma m=n)} q_{i(\sigma m=n)} + \eta^*_{i(\sigma m=n)} q^2_{i(\sigma m=n)}) + \sum_{\pi m=n} (E^*_{i(\pi m=n)} - \mu^*_{i(\pi m=n)} q_{i(\pi m=n)} + \eta^*_{i(\pi m=n)} q^2_{i(\pi m=n)}) \\
& + \sum_{lp} (E^*_{i(lp)} - \mu^*_{i(lp)} q_{i(lp)} + \eta^*_{i(lp)} q^2_{i(lp)}) \Bigg\} + \sum_{i=1}^{N_{mol}} \Bigg\{ \sum_{j=1}^{N_{mol}} \Bigg[\sum_{\substack{H \in i \\ H,\, lp \text{ in } HBIR}} \sum_{lp \in j} k_{\text{H-bond}}(R_{iH,j(lp)}) \frac{q_{iH} q_{j(lp)}}{R_{iH,j(lp)}} \\
& + k \Bigg(\frac{1}{2} \sum_a \sum_{b \neq a} \sum \frac{q_{ia} q_{jb}}{R_{ia,jb}} + \frac{1}{2} \sum_m \sum_{k \neq m} \sum \frac{q_{im} q_{jk}}{R_{im,jk}} + \frac{1}{2} \sum_{a-b} \sum_{g-h \neq a-b} \frac{q_{i(a-b)} q_{j(g-h)}}{R_{i(a-b),j(g-h)}} + \frac{1}{2} \sum_{\sigma m=n} \sum_{\sigma k=l \neq \sigma m=n} \frac{q_{i(\sigma m=n)} q_{j(\sigma k=l)}}{R_{i(\sigma m=n),j(\sigma k=l)}} \\
& + \frac{1}{2} \sum_{\pi m=n} \sum_{\pi k=l \neq \pi m=n} \frac{q_{i(\pi m=n)} q_{j(\pi k=l)}}{R_{i(\pi m=n),j(\pi k=l)}} + \frac{1}{2} \sum_{lp} \sum_{lp' \neq lp} \frac{q_{i(lp)} q_{j(lp')}}{R_{i(lp),j(lp')}} + \sum_a \sum_m \frac{q_{ia} q_{jm}}{R_{ia,jm}} + \sum_a \sum_{g-h} \frac{q_{ia} q_{j(g-h)}}{R_{ia,j(g-h)}} \\
& + \sum_a \sum_{\sigma m=n} \frac{q_{ia} q_{j(\sigma m=n)}}{R_{ia,j(\sigma m=n)}} + \sum_a \sum_{\pi m=n} \frac{q_{ia} q_{j(\pi m=n)}}{R_{ia,j(\pi m=n)}} + \sum_a \sum_{\pi m=n} \frac{q_{ia} q_{j(\pi m=n)}}{R_{ia,j(\pi m=n)}} + \sum_{a(a \neq H, \text{ and } a=H, H \text{ not in } HBIR)} \sum_{lp} \frac{q_{ia} q_{j(lp)}}{R_{ia,j(lp)}} \\
& + \sum_m \sum_{a-b} \frac{q_{im} q_{j(a-b)}}{R_{im,j(a-b)}} + \sum_m \sum_{\sigma k=l} \frac{q_{im} q_{j(\sigma k=l)}}{R_{im,j(\sigma k=l)}} + \sum_{a-b} \sum_{\pi m=n} \frac{q_{i(a-b)} q_{j(\pi m=n)}}{R_{i(a-b),j(\pi m=n)}} + \sum_{a-b} \sum_{lp} \frac{q_{i(a-b)} q_{j(lp)}}{R_{i(a-b),j(lp)}} + \sum_{\sigma m=n} \sum_{\pi k=l} \frac{q_{i(\sigma m=n)} q_{j(\pi k=l)}}{R_{i(\sigma m=n),j(\pi k=l)}} \\
& + \sum_{\sigma m=n} \sum_{lp} \frac{q_{i(\sigma m=n)} q_{j(lp)}}{R_{i(\sigma m=n),j(lp)}} + \sum_{\pi m=n} \sum_{lp} \frac{q_{i(\pi m=n)} q_{j(lp)}}{R_{i(\pi m=n),j(lp)}} \Bigg) \Bigg] \Bigg\} + \sum_m \sum_{\pi k=l} \frac{q_{im} q_{j(\pi k=l)}}{R_{im,j(\pi k=l)}} \\
& + \sum_m \sum_{lp} \frac{q_{im} q_{j(lp)}}{R_{im,j(lp)}} + \sum_{a-b} \sum_{\sigma m=n} \frac{q_{i(a-b)} q_{j(\sigma m=n)}}{R_{i(a-b),j(\sigma m=n)}} + \sum_{a-b} \sum_{\pi m=n} \frac{q_{i(a-b)} q_{j(\pi m=n)}}{R_{i(a-b),j(\pi m=n)}}
\end{aligned} \tag{9}$$

Thus, based on Eqs. (9) and (10), we obtain Eq. (11) for the effective electronegativities of partial charge site, respectively.

$$
\begin{aligned}
\chi_{ia} &= \chi^{*}_{ia} + 2\eta^{*}_{ia}q_{ia} + \sum_{lp(a=HH,lp\,\text{in}\,HBIR)} k_{lp,H}(R_{i(lp),iH})\frac{q_{i(lp)}}{R_{ia,i(lp)}} + k\left(\sum_{b,b\neq a}\frac{q_{ib}}{R_{ia,ib}} + \sum_{m}\frac{q_{im}}{R_{ia,im}} + \sum_{g-h}\frac{q_{i(g-h)}}{R_{ia,i(g-h)}} + \sum_{\sigma m=n}\frac{q_{i(\sigma m=n)}}{R_{ia,i(\sigma m=n)}}\right.\\
&\left. + \sum_{\pi m=n}\frac{q_{i(\pi m=n)}}{R_{ia,i(\pi m=n)}} + \sum_{lp(lp\,\text{not in}\,HBIR)}\frac{q_{i(lp)}}{R_{ia,i(lp)}}\right) + \sum_{j,j\neq i}\left[\sum_{lp(a=HH,lp\,\text{in}\,HBIR)} k_{lp,H}(R_{iH,j(lp)})\frac{q_{j(lp)}}{R_{ia,j(lp)}} + k\left(\sum_{b}\frac{q_{jb}}{R_{ia,jb}} + \sum_{m}\frac{q_{jm}}{R_{ia,jm}} + \sum_{g-h}\frac{q_{j(g-h)}}{R_{ia,j(g-h)}}\right.\right.\\
&\left.\left. + \sum_{\sigma m=n}\frac{q_{j(\sigma m=n)}}{R_{ia,j(\sigma m=n)}} + \sum_{\pi m=n}\frac{q_{j(\pi m=n)}}{R_{ia,j(\pi m=n)}} + \sum_{lp(lp\,\text{not in}\,HBIR)}\frac{q_{j(lp)}}{R_{ia,j(lp)}}\right)\right]\\
\chi_{im} &= \chi^{*}_{im} + 2\eta^{*}_{im}q_{im} + k\left(\sum_{a}\frac{q_{ia}}{R_{im,ia}} + \sum_{k,k\neq m}\frac{q_{ik}}{R_{im,ik}} + \sum_{a-b}\frac{q_{i(a-b)}}{R_{im,i(a-b)}} + \sum_{\sigma k=l}\frac{q_{i(\sigma k=l)}}{R_{im,i(\sigma k=l)}} + \sum_{\pi k=l}\frac{q_{i(\pi k=l)}}{R_{im,i(\pi k=l)}}\right.\\
&\left. + \sum_{lp}\frac{q_{i(lp)}}{R_{im,i(lp)}}\right) + k\sum_{j,j\neq i}\left(\sum_{a}\frac{q_{ja}}{R_{im,ja}} + \sum_{k}\frac{q_{jk}}{R_{im,jk}} + \sum_{a-b}\frac{q_{j(a-b)}}{R_{im,j(a-b)}} + \sum_{\sigma k=l}\frac{q_{j(\sigma k=l)}}{R_{im,j(\sigma k=l)}} + \sum_{\pi k=l}\frac{q_{j(\pi k=l)}}{R_{im,j(\pi k=l)}} + \sum_{lp}\frac{q_{j(lp)}}{R_{im,j(lp)}}\right)\\
\chi_{i(a-b)} &= \chi^{*}_{i(a-b)} + 2\eta^{*}_{i(a-b)}q_{i(a-b)} + k\left(\sum_{g}\frac{q_{ig}}{R_{i(a-b),ig}} + \sum_{m}\frac{q_{im}}{R_{i(a-b),im}} + \sum_{g-h,g-h\neq a-b}\frac{q_{i(g-h)}}{R_{i(a-b),i(g-h)}} + \sum_{\sigma m=n}\frac{q_{i(\sigma m=n)}}{R_{i(a-b),i(\sigma m=n)}}\right.\\
&\left. + \sum_{\pi m=n}\frac{q_{i(\pi m=n)}}{R_{i(a-b),i(\pi m=n)}} + \sum_{lp}\frac{q_{ilp}}{R_{i(a-b),i(lp)}}\right) + k\sum_{j,j\neq i}\left(\sum_{g}\frac{q_{jg}}{R_{i(a-b),jg}} + \sum_{m}\frac{q_{jm}}{R_{i(a-b),jm}} + \sum_{g-h}\frac{q_{j(g-h)}}{R_{i(a-b),j(g-h)}} + \sum_{\sigma m=n}\frac{q_{j(\sigma m=n)}}{R_{i(a-b),j(\sigma m=n)}}\right.\\
&\left. + \sum_{\pi m=n}\frac{q_{j(\pi m=n)}}{R_{i(a-b),j(\pi m=n)}} + \sum_{lp}\frac{q_{jlp}}{R_{i(a-b),j(lp)}}\right)\\
\chi_{i(\sigma m=n)} &= \chi^{*}_{i(\sigma m=n)} + 2\eta^{*}_{i(\sigma m=n)}q_{i(\sigma m=n)} + k\left(\sum_{a}\frac{q_{ia}}{R_{i(\sigma m=n),ia}} + \sum_{k}\frac{q_{ik}}{R_{i(\sigma m=n),ik}}\right.\\
&\left. + \sum_{a-b}\frac{q_{i(a-b)}}{R_{i(\sigma m=n),i(a-b)}} + \sum_{\sigma k=l,\sigma k=l\neq\sigma m=n}\frac{q_{i(\sigma k=l)}}{R_{i(\sigma m=n),i(\sigma k=l)}} + \sum_{\pi k=l}\frac{q_{i(\pi k=l)}}{R_{i(\sigma m=n),i(\pi k=l)}} + \sum_{lp}\frac{q_{i(lp)}}{R_{i(\sigma m=n),i(lp)}}\right) + k\sum_{j,j\neq i}\left(\sum_{a}\frac{q_{ja}}{R_{i(\sigma m=n),ja}}\right.\\
&\left. + \sum_{k}\frac{q_{jk}}{R_{i(\sigma m=n),jk}} + \sum_{a-b}\frac{q_{j(a-b)}}{R_{i(\sigma m=n),j(a-b)}} + \sum_{\sigma k=l}\frac{q_{j(\sigma k=l)}}{R_{i(\sigma m=n),j(\sigma k=l)}} + \sum_{\pi k=l}\frac{q_{j(\pi k=l)}}{R_{i(\sigma m=n),j(\pi k=l)}} + \sum_{lp}\frac{q_{j(lp)}}{R_{i(\sigma m=n),j(lp)}}\right)\\
\chi_{i(\pi m=n)} &= \chi^{*}_{i(\pi m=n)} + 2\eta^{*}_{i(\pi m=n)}q_{i(\pi m=n)} + k\left(\sum_{a}\frac{q_{ia}}{R_{i(\pi m=n),ia}} + \sum_{k}\frac{q_{ik}}{R_{i(\pi m=n),ik}} + \sum_{a-b}\frac{q_{i(a-b)}}{R_{i(\pi m=n),i(a-b)}} + \sum_{\sigma k=l}\frac{q_{i(\sigma k=l)}}{R_{i(\pi m=n),i(\sigma k=l)}}\right.\\
&\left. + \sum_{\pi k=l,\pi k=l\neq\pi m=n}\frac{q_{i(\pi k=l)}}{R_{i(\pi m=n),i(\pi k=l)}} + \sum_{lp}\frac{q_{i(lp)}}{R_{i(\pi m=n),i(lp)}}\right) + k\sum_{j,j\neq i}\left(\sum_{a}\frac{q_{ja}}{R_{i(\pi m=n),ja}} + \sum_{k}\frac{q_{jk}}{R_{i(\pi m=n),jk}}\right.\\
&\left. + \sum_{a-b}\frac{q_{j(a-b)}}{R_{i(\pi m=n),j(a-b)}} + \sum_{\sigma k=l}\frac{q_{j(\sigma k=l)}}{R_{i(\pi m=n),j(\sigma k=l)}} + \sum_{\pi k=l}\frac{q_{j(\pi k=l)}}{R_{i(\pi m=n),j(\pi k=l)}} + \sum_{lp}\frac{q_{j(lp)}}{R_{i(\pi m=n),j(lp)}}\right)\\
\chi_{i(lp)} &= \chi^{*}_{i(lp)} + 2\eta^{*}_{i(lp)}q_{i(lp)} + \sum_{a\,(a=HH,lp\,\text{in}\,HBIR)} k_{lp,H}(R_{i(lp),ia})\frac{q_{ia}}{R_{i(lp),ia}} + k\left(\sum_{a(a\neq H,\,\text{and}\,a=H,H\,\text{not in}\,HBIR)}\frac{q_{ia}}{R_{i(lp),ia}}\right.\\
&\left. + \sum_{m}\frac{q_{im}}{R_{i(lp),im}} + \sum_{a-b}\frac{q_{i(a-b)}}{R_{i(lp),i(a-b)}} + \sum_{\sigma m=n}\frac{q_{i(\sigma m=n)}}{R_{i(lp),i(\sigma m=n)}} + \sum_{\pi m=n}\frac{q_{i(\pi m=n)}}{R_{i(lp),i(\pi m=n)}} + \sum_{lp'\neq lp}\frac{q_{i(lp')}}{R_{i(lp),i(lp')}}\right)\\
&+ \sum_{j,j\neq i}\left[\sum_{a(a=HH,lp\,\text{in}\,HBIR)} k_{lp,H}(R_{i(lp),ja})\frac{q_{ja}}{R_{i(lp),ja}} + k\left(\sum_{a(a\neq H,\ \text{and}\,a=H,H\,\text{not in}\,HBIR)}\frac{q_{ja}}{R_{i(lp),ja}} + \sum_{m}\frac{q_{jm}}{R_{i(lp),jm}} + \sum_{a-b}\frac{q_{j(a-b)}}{R_{i(lp),j(a-b)}} + \sum_{\sigma m=n}\frac{q_{j(\sigma m=n)}}{R_{i(lp),j(\sigma m=n)}}\right.\right.\\
&\left.\left. + \sum_{\pi m=n}\frac{q_{j(\pi m=n)}}{R_{i(lp),j(\pi m=n)}} + \sum_{lp'}\frac{q_{j(lp')}}{R_{i(lp),j(lp')}}\right)\right]
\end{aligned}
\tag{11}
$$

Table 5
RMSD of bond lengths (Å) and angels (°) of bases relative to that of experimental data or *ab initio* result.

Method	Canonical[a]		Oxidative[b]	
	Bond	Angle	Bond	Angle
B3LYP	0.016	1.0		
ABEEM$\sigma\pi$	0.007	0.7	0.020	1.4
AMBER	0.013	1.2	0.023	2.6

[a] RMSD of A, T, G, and C relative to the experimental data.
[b] RMSD of oxidized bases relative to the results of B3LYP/6-31 + G* level.

For a molecular system with charge Q_i, there is no intermolecular charge transfer from the experience of previous studies [53]. Therefore, there are N_{mol} charge constraint equations.

$$\sum_{a}^{N_a} q_{ia} + \sum_{m}^{N_m} q_{im} + \sum_{a-b}^{N_{a-b}} q_{i(a-b)} + \sum_{\sigma m=n}^{N_{\sigma m=n}} q_{i(\sigma m=n)} + \sum_{\pi m=n}^{N_{\pi m=n}} q_{i(\pi m=n)} + \sum_{lp}^{N_{lp}} q_{i(lp)} = Q_i \quad (i = 1, 2, \ldots N_{mol}) \tag{12}$$

Without intermolecular charge transfer, the effective electronegativity of arbitrary partial charge site will only be equal within a molecule in the electronegativity equalization.

$$\begin{aligned} &\chi_{ia} = \chi_{im} = \chi_{i(a-b)} = \chi_{i(\sigma m=n)} = \chi_{i(\pi m=n)} = \chi_{i(lp)} = \cdots = \tilde{\chi}_i \\ &\chi_{ja} = \chi_{jm} = \chi_{j(a-b)} = \chi_{j(\sigma m=n)} = \chi_{j(\pi m=n)} = \chi_{j(lp)} = \cdots = \tilde{\chi}_j \\ &\vdots \end{aligned}, \tag{13}$$

where $i, j = 1, 2, \ldots, N_{mol}$. $\overline{\chi_j}$, $\overline{\chi_j}$ stand for the effective electronegativity of molecules i and j in the system, respectively. For a molecular system, the region partial charges are not all independent variables since all these effective electronegativities involve the contribution from the partial charges of all the other charge sites of the whole system.

For an arbitrary molecular system containing N_a single bond atoms, N_m double bond atoms, $N_{a\text{-}b}$ single bonds, $N_{\sigma m=n}$ σ bond regions of double bonds, $N_{\pi m=n}$ π bond regions of double bonds, and N_{lp} lone-pair electrons, we yield $N_a + N_m + N_{a\text{-}b} + N_{\sigma m=n} + N_{\pi m=n} + N_{lp}$ simultaneous electronegativity equalization equations of Eq. (13). These equations, along with N_{mol} charge constraint Eq. (12), can be solved to directly give the effective electronegativity of each molecule and the partial charges q on each site, if the parameters in Eq. (10) are known.

Minimizations were performed with the limited memory BFGS quasi-Newton nonlinear optimization routine. The criterion used for convergence was the RMSD energy gradient less than 0.01 kcal/mol/Å. No restrictions were applied, that is, all atoms are allowed to move freely. The binding energy, ΔE, was obtained as the energetic difference between the isolate molecules (E_{A} and E_{B}) and the complex (E_{AB}).

3. Results and discussion

3.1. Geometrical and electronic structures

The geometrical parameters of four canonical bases from ABEEM$\sigma\pi$, AMBER, B3LYP and experiment are listed in Tables S4–S5. The RMSD between experimental data and those of theoretical methods are listed in Table 5. The geometries of four canonical bases calculated by B3LYP/6-31 + G* are the closest to the experimental data among the 24 model chemistries, with the RMSD of 0.016 Å and 1.0° for the bond lengths and the angles, respectively, and so only the results from this model chemistry is given here. Surprisingly, the RMSD of ABEEM$\sigma\pi$ PFF (0.007 Å and 0.7°) is smaller than that of B3LYP/6-31 + G*. Those results show that ABEEM$\sigma\pi$ PFF can accurately simulate the geometries of oxidized guanines. Because of no experimental data available for the modified base, we compare the results from ABEEM$\sigma\pi$ PFF and AMBER FF with those of B3LYP/6-31 + G* method. The RMSD of bond length and angle from ABEEM$\sigma\pi$ PFF are 0.020 Å and 1.4°, respectively, which are smaller than those of AMBER FF (0.023 Å and 2.6°). Electronic structure is also important except for geometric structure. The dipole moment, as a good measure of molecular charge distribution, has been measured by experiment [60] and calculated by MP2 [25,61,62], CHARMM27 [61], ESP [25], *etc.* Table 6 lists the dipole moments of four canonical bases with different methods, in which the results of ABEEM$\sigma\pi$ PFF are fairly consistent with those from experiment.

The geometry of guanine will change after oxidation. In order to investigate the difference between guanine and those oxidized guanines, Table 7 lists their geometry parameters by ABEEM$\sigma\pi$ PFF. Both 8-OH-G and 8-oxo-G are the oxidation products of guanine at C8 position. However, the two functional groups induce obviously different effects to guanine. The most distinct changes in 8-oxo-G are the bond lengths of N7-C8 and C8-N9 increased by 0.064 Å and 0.028 Å; C5-N7-C8 and C8-N9-C4 angles increased by 5.7° and 4.8°; N7-C8-N9 and C8-N9-C9 angles decreased by 8.3° and 4.2°, as compared with the corresponding geometries in G:C(WC). It can be seen that, after C8 position are substituted by C = O, the geometry is obviously changed. But for 8-OH-G, the OH at C8 has slight effect on the structure of normal guanine as shown in Table 7. The introduction of C2 substitution of Xanthosine affects the C2-N3 to increase by 0.048 Å; C2-O2 to decrease by 0.120 Å; C6-N1-C2, C2-N3-C4, N1-C2-N2 to increase by 3.2°, 6.7°, 4.6°, respectively, while N1-C2-N3 to decrease by 6.1°. C2 substitution also obviously affects the geometry. C8-N9 bond in normal guanine is broken to form Fapy-G with nonplanar structure, which leads most of bonds and angles to be changed. Table S6 collects all the geometry information of oxidized guanines from B3LYP/6-31 + G*, ABEEM$\sigma\pi$ PFF and AMBER FF.

3.2. Stacking interaction of base pairs

The interaction of stacked bases plays the same significant role as the interaction of H-bonded bases to the stability of DNA. Here, we investigate the antiparallel stacked structures of A and A (denoted as AAs), C and C (denoted as CCs), G and G (denoted as GGs), and T and T (denoted as TTs) base pairs. The structures are shown in Fig. 8. Single point calculations were done with MP2/aug-cc-pVDZ and ABEEM$\sigma\pi$ PFF methods. We calculated the interaction energies of the antiparallel stacked bases as the vertical separation between two planes formed by two base pairs varies from 2.8 Å to 8.0 Å, potential energy profiles are shown in Fig. 8, and the stacking energies are listed in Table 8. One local minimum is found at the separation of 3.3 Å for each profile. The interaction energies by ABEEM$\sigma\pi$ PFF are in good agreement with the *ab initio* results. The AAD is only 1.3 kcal/mol, which is smaller than that of other methods listed in Table 8. Those π bond regions assigned by ABEEM$\sigma\pi$ PFF contribute to the good result. In addition, the local minima of interaction energies of other methods vary from 3.3 Å to 3.6 Å, which are different from the *ab initio* results.

3.3. Deviation degree of base pair plane

Accurate simulation of the deviation degree of base pair plane is important for a force field. The structure of canonical base is a planar ring. The angle φ formed by the two base planes is close to 180.0°. When the oxidized guanine pairs with canonical base,

Table 6
Dipole moments (D) of A, T, G and C calculated by different methods.

Method	Adenine	Guanine	Cytosine	Thymine
Exp[a]	3.00		7.00	4.10
ABEEM$\sigma\pi$	3.00	6.48	6.59	4.57
MP2/6-311++G(3df,2pd)//MP2/6-31G**[b]	2.55	6.63	6.53	4.32
MP2/aug-cc-pVDZ//MP2/6-31G**[b]	2.56	6.63	6.52	4.32
MP2/aug-cc-pVDZ[c]	2.56	6.49	6.65	4.31
MP2/6-31G*[c]	2.55	6.27	6.45	4.01
CHARMM27[d]	2.91	7.59	7.85	4.50
ESP Charge (MP2/6-31+G*)[b]	2.68	6.82	6.90	4.53

[a] Ref. [60].
[b] Ref. [25].
[c] Ref. [62].
[d] Ref. [61].

Table 7
Structure parameters (Bond: Å; angle:°) of guanine and oxidized guanines from ABEEM$\sigma\pi$ force field.

Parameters	G	8-Oxo-G	8-OH-G	Xanthosine	Fapy-G
C2-N3	1.325	1.329	1.325	1.373	1.317
N7-C8	1.310	1.374	1.307	1.315	1.339
C8-N9	1.376	1.414	1.383	1.375	3.641
C2-N2(O2)(C2-O2)	1.344	1.347	1.342	1.224	1.341
C6-N1-C2	124.5	125.0	124.8	127.7	123.7
N1-C2-N3	123.2	122.6	123.1	117.1	122.6
C2-N3-C4	113.6	113.8	113.5	120.3	119.8
C5-N7-C8	104.2	109.9	103.7	105.4	122.6
N7-C8-N9	113.5	105.2	114.5	113.4	56.3
C8-N9-C4	105.9	110.7	104.8	104.7	64.2
N1-C2-N2	116.9	118.1	116.7	121.5	117.5
C8-N9-C9	128.4	123.2	128.4	129.0	68.2

the angle θ ($\theta = 180.0^\circ - \varphi$) will be buckled to a different value. The buckled angle θ was defined as shown in Fig. 9, considering the random selected heavy atoms 1, 2, 3 on the ring in one base plane and similar atoms 4, 5, 6 on the ring in the other base plane. The average of all these angles is taken as the deviation degree. Therefore, a bigger θ denotes serious deviation, which will impact the stacking effect between bases in double helix of DNA. The θ of canonical base pairs is the range of 0.0–0.5°. Figs. 3–6 display the θ of all model base pairs by ABEEM$\sigma\pi$ PFF, and they are equal to those from B3LYP/6-31 + G*. Results show that base pairs containing 8-oxo-G almost have no effect on the stacking effect, except for 8-oxo-G:G(H), whose deviation angle is only 5.2°. Base pairs containing another three oxidized guanines have larger deviations. From Figs. 4 and 5, we can see that the repulsion of two O atoms or H atoms in oxidized guanine base pair makes large θ. The five-member ring of guanine is broken to form Fapy-G, so those base pairs containing Fapy-G will inevitably affect the stacking in DNA. Our results show 36% of the buckled angles in the oxidized base pairs is larger than 5.0°.

3.4. H-bonding interaction of base pairs

Nucleic acid carries genetic information, and H-bonding interaction plays a unique role for structure and function of nucleic acid. Tables S7–S10 list the distance between proton donor X and proton acceptor Y, and the H-bond angle X-H···Y of several typical base pairs by B3LYP/6-31 + G* method, ABEEM$\sigma\pi$ PFF and AMBER FF. Compared to the B3LYP method, the RMSD of H-bond lengths in those complexes by ABEEM$\sigma\pi$ PFF is only 0.030 Å, which is smaller than that by AMBER FF (0.035 Å). The orientation of the lone pair electron of acceptor directly affects the orientation of H-bond. The lone pair electron sites in ABEEM$\sigma\pi$ PFF fairly control the orientation of H-bond, at the same time, functions $k_{\text{H-bond}}$ and fluctuating charges more flexibly reflect the polarization effect with the changes of environment. The RMSD of H-bond angles in those complexes by ABEEM$\sigma\pi$ PFF is only 1.8°, which is also smaller than that by AMBER FF, 2.5°.

When the guanine has been oxidized, the sites of H-bond donor and acceptor are different, which will change the type and number of H-bond, even cause the mismatch between bases. From Tables S7-S10, the H-bond lengths and angles of 42 modified base pairs rang between 2.620–3.179 Å, and 160.0–179.7°, respectively. Complex 8-OH-G:G forms three H-bonds, while the other three oxidized guanines match with G forming two H-bonds, which obviously decreases the binding energy of OG:G (OG is abbreviation of oxidized guanine) relative to that of G:C. There are two H-bonds in OG:A, and OG:T, except for Fapy-G:T. When 8-oxo-G, 8-OH-G or Fapy-G matches with C, there are three H-bonds, and the H-bond lengths and angles are almost equal to those in G:C. However, there are only two H-bonds in Xanthosine:C complexes.

We have also checked the reliability of binding energies of ABEEM$\sigma\pi$ PFF, which are listed in Table 9. Compared with the

Table 8
The minimum interaction energies (kcal/mol) and the corresponding vertical distance (Å) of the antiparallel stacked base pairs.

Base pair	MP2		ABEEM$\sigma\pi$		AMBER		CHARMM		CFF95		OPLS	
	ΔE	R	ΔE	R	ΔE	R	ΔE	R	ΔE	R	ΔE	R
AA	10.5	3.3	9.7	3.3	9.2	3.4	10.6	3.4	5.5	3.6	7.4	3.6
CC	10.1	3.3	9.0	3.3	11.1	3.4	15.1	3.2	10.3	3.3	11.7	3.4
GG	13.8	3.3	11.8	3.3	7.4	3.5	9.9	3.4	10.0	3.5	7.2	3.6
AAD			1.3	0.0	2.9	0.1	3.0	0.1	3.0	0.2	3.8	0.2

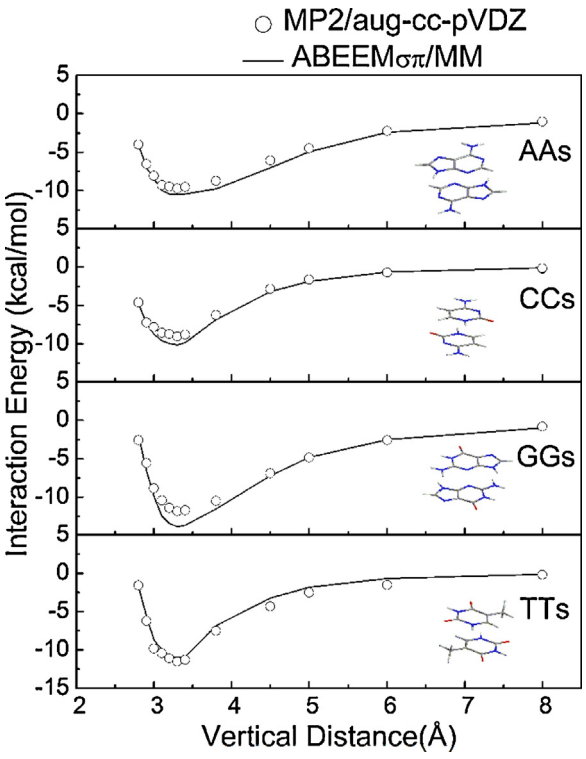

Fig. 8. Potential energy profiles as a function of vertical distance of antiparallel stacked nucleic acid base pairs by MP2/aug-cc-pVDZ and ABEEM$\sigma\pi$ polarizable force field.

Fig. 9. Deviation degree of base pair plane.

Table 9
Binding energies (kcal/mol) of base pairs.

Base pair	MP2/aug-cc-pVDZ	ABEEM$\sigma\pi$	CHARMM[a]	AMBER[b]
G:C(WC)	26.02	26.34	25.69	28.00
A:T(WC)	14.26	12.85	12.66	12.80
A:T(H)	15.38	14.85	13.25	14.50
A:T(R H)	15.17	14.25	12.98	14.50
AAD		0.80	1.56	1.25

[a] Ref. [61].
[b] Ref. [63].

results of CHARMM [61] and AMBER FFs [63], the results of ABEEM$\sigma\pi$ PFF are closed to those of MP2/aug-cc-pVDZ level. The binding energies of oxidized base pairs are shown in Table 10. Compared with the MP2/aug-cc-pVDZ calculation, the RMSD from ABEEM$\sigma\pi$ PFF is only 1.14 kcal/mol, and the linear correlation coefficient is 0.9573.

3.5. Correlation between stacking and H-bonding interactions

With all canonical conventions, the systems in the study reported herein are named CC/GG and CG/CG (base 1 base 3/base 4 base 2 as Fig. 10) [19]. The first case (CC/GG) is when there is an cytosine residue (up) above the other cytosine (down) in one strand and in the other strand a guanine residue (up) above the other guanine (down). The up cytosine is above the down guanine and the up guanine is above the down cytosine, with the two cytosines in the 3′ and the two guanines in the 5′ direction. In our calculations, we did not introduce the two sugar–phosphate backbones between the bases 1 and 3 and the bases 2 and 4, since it is well known that this part of DNA has a minor influence on both the stacking geometry and the hydrogen-atom transfer between the base pairs as well as the mechanism of this transfer [19]. The starting geometries of our two systems are those reported in the supporting information of Ref. [21].

Table 11 shows the distances and angles of the atoms involved in the six hydrogen bonds between the two base pairs, where the subscripts u and d mean “up” and “down” base pairs as shown in Fig. 10, and those of the three hydrogen bonds of the GC monomer system. We can conclude that the two base pairs of the dimers are substantially planar like that of the GC system; the stacking leads to a decrease of the H-bond strength, mainly seeing the H-bond

Table 10
Binding energies (kcal/mol) of oxidized base pairs from MP2/aug-cc-pVDZ and ABEEM$\sigma\pi$ force field.

Base pair	MP2	ABEEM$\sigma\pi$	Base pair	MP2	ABEEM$\sigma\pi$
G:C(WC)	26.02	26.34	8-OH-G1:T(R)	15.67	17.09
A:T(WC)	14.26	12.85	8-OH-G2:T(R)	15.44	16.84
A:T(H)	15.38	14.85	Xanthosine:A(WC)	13.77	13.39
A:T(R H)	15.17	14.25	Xanthosine:A(H)	15.23	15.35
8-Oxo-G:A(WC)	15.46	13.58	Xanthosine1:C(WC)	12.49	14.62
8-Oxo-G:A(H)	16.91	18.31	Xanthosine2:C(WC)	11.49	14.71
8-Oxo-G:C(WC)	26.82	27.65	Xanthosine:G(WC)	18.32	18.01
8-Oxo-G:G(H)	19.71	21.06	Xanthosine1:T	11.90	12.82
8-Oxo-G1:T	13.50	12.44	Xanthosine2:T	11.33	12.89
8-Oxo-G2:T	17.32	17.67	Xanthosine3:T	17.54	14.52
8-Oxo-G3:T	13.46	11.74	Xanthosine1:T(R)	11.23	12.06
8-Oxo-G1:T(R)	13.15	11.92	Xanthosine2:T(R)	11.18	12.02
8-Oxo-G2:T(R)	16.29	16.91	Xanthosine3:T(R)	16.12	13.48
8-Oxo-G3:T(R)	13.17	11.02	Fapy-G1:A(WC)	16.28	16.38
8-OH-G1:A(WC)	19.79	19.40	Fapy-G2:A(WC)	11.89	11.85
8-OH-G2:A(WC)	16.80	17.21	Fapy-G:A(H)	16.28	18.37
8-OH-G1:A(H)	19.88	19.44	Fapy-G:C(WC)	26.21	27.10
8-OH-G2:A(H)	16.42	19.08	Fapy-G:G(H)	18.65	19.28
8-OH-G:C(WC)	25.92	26.47	Fapy-G1:T	17.40	16.88
8-OH-G:G(WC)	29.18	30.07	Fapy-G3:T	13.61	15.12
8-OH-G:G(H)	18.22	19.81	Fapy-G1:T(R)	16.39	15.80
8-OH-G1:T	16.81	18.89	Fapy-G3:T(R)	13.86	13.86
8-OH-G2:T	16.58	17.92	Fapy-G:T	17.59	18.10

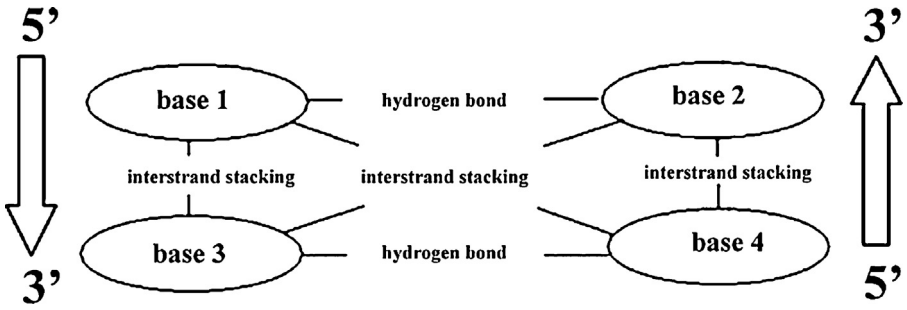

Fig. 10. Numbering scheme of the base-pair dimer. The base pairs 1 and 2 (up) and 3 and 4 (down) are bonded with H-bonds, whereas bases 1 and 3, and 2 and 4, have intra-strand stacking interactions and 1 and 4, and 2 and 3, have inter-strand stacking interactions. The 5′→3′ convention is used.

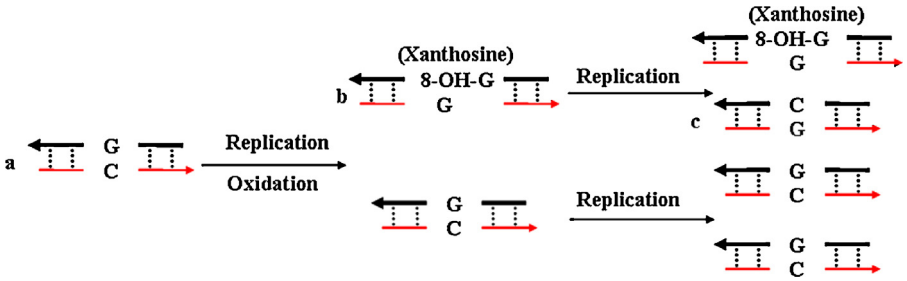

Fig. 11. The mutation of GC→CG. Black and red lines represent sequence of bases. Dotted lines denote H-bond. Sequence of bases is connected to form double helix by H-bond. a–c represent three double helix strands.

lengths and H-bond angles. The H-bond lengths of the three H-bonds in G:C dimer are all increased by about 0.1 Å, and the H-bond angles are decreased by 0.0–4.0°. Really, the strength of a bond can be better discussed in terms of the potential energy curve and of bond index, and this is our choice. ABEEM$\sigma\pi$ PFF gives reasonable geometries of H-bonds compared with the *ab initio* methods. The AAD of H-bond lengths and H-bond angles are 0.0217 Å and 1.0°, respectively.

3.6. Mutation induced by oxidized guanine

According to the results in Table 10, the binding energy of normal base pair G:C(WC) from ABEEM$\sigma\pi$ PFF is 26.34 kcal/mol. The binding energy of 8-oxo-G:C is 27.65 kcal/mol, which is larger than that of G:C(WC) by 1.31 kcal/mol. Similarly, the binding energies of 8-OH-G:C and Fapy-G:C are 26.47 and 27.10 kcal/mol, respectively, which are larger than that of G:C(WC) by 0.13 and 0.76 kcal/mol, respectively. It suggests that the binding energies of these three modified base pairs are nearly the same as that of G:C(WC), indicating that though G is oxidized to 8-oxo-G, Fapy-G or 8-OH-G, and it almost have no effect on the binding ability with C. However, the binding energy of Xanthosine:C is obvious smaller than that of G:C(WC) by 11.31 kcal/mol, indicating the introduction of O atom at C2 substitution affects the stability of the G:C base pair remarkably. At present, ABEEM$\sigma\pi$ PFF and MP2/aug-cc-pVDZ are consistent with the predictions of the mutation. Indeed, the error of some H-bonding interaction energies of ABEEM$\sigma\pi$ PFF with respect of MP2 is more than 20%, but those geometries are not important in prediction of mutations.

When a cell separates into two daughter cells during cell division, the DNA helix begins to unravel. Now, complementary deoxyribonucleotides attach to the original, unravel strands. Complementary chains of DNA is formed by the pairing of bases in the combinations of A with T, or G with C. 8-oxo-G or Fapy-G paired with C is the most preferable pair, so the mismatch between them and other base will not happen. As we concluded, 8-OH-G and Xanthosine are more favorable for paring with G than C, which causes the mutation of GC→CG. As shown in Fig. 11, DNA helix contains normal G:C. If G was oxidated to 8-OH-G (or Xanthosine), it will pair with G in the first-generation progeny (Fig. 11a→b). In the second-generation progeny, G will pair with C (Fig. 11b→c), which will induce the transition from G:C to C:G.

According to our calculation, the binding energies of OG with A are 16.91, 19.88, 15.23, and 16.28 kcal/mol for 8-oxo-G:A(H), 8-OH-G1:A(H), Xanthosine:A(H), and Fapy-G:A(H), respectively, which are larger than that of A:T(WC) (14.26 kcal/mol). It can be concluded that A is more likely to pair with one of the four OG than T. During replication of the flawed strand, OG can be misread by polymerase as T, which can pair with A in the first-generation progeny (Fig. 12a→b or a→b1). According to the above calculation by us, in the second-generation progeny, 8-oxo-G or Fapy-G is much more likely to pair with C (Fig. 12b→c), while 8-OH-G

Table 11
Distances (Å) and angles (°) of atoms involved in the six H-bonds between the two base pairs and those of the three H-bonds of the GC monomer system.[a]

		O(G)···H-N(C)		N(G)-H···N(C)		N(G)-H···O(C)	
		$R_{O(G)-N(C)}$	θ_{OHN}	$R_{N(G)-N(C)}$	θ_{NHN}	$R_{N(G)-O(C)}$	θ_{NHN}
CG		2.8271/2.8703	179.1/178.4	2.9158/2.9627	177.7/178.3	2.9304/2.9324	178.5/177.9
CG/CG	u	2.9209/2.9375	176.3/175.6	3.0364/3.0056	176.1/177.7	3.0174/3.0049	177.2/176.7
CG/CG	d	2.9205/2.9375	176.3/175.6	3.0371/3.0061	176.1/177.7	3.0173/3.0049	177.1/176.8
CC/GG	u	2.9206/3.0035	176.9/174.0	3.0369/3.0407	176.1/175.7	3.0183/3.0513	175.9/176.8
CC/GG	d	2.9205/2.9859	175.4/176.4	3.0371/3.0599	176.1/177.3	3.0172/3.0323	175.3/176.2

[a] The values before "/" are the results from *ab initio*, and those after "/" are the results from ABEEM$\sigma\pi$ PFF. u = up; d = down.

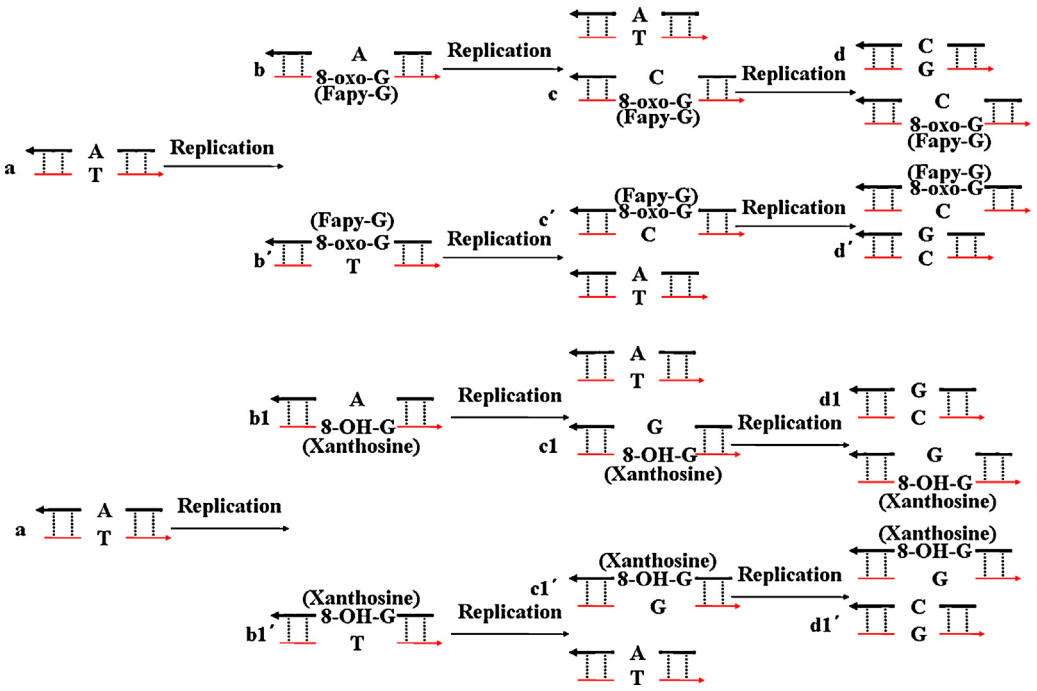

Fig. 12. The mutation of AT→CG and AT→GC.

or Xanthosine is easier to pair with G (Fig. 12b1→c1). In the latter course, C will pair with G (Fig. 12c→d) or G will pair with C (Fig. 12c1→d1) by known modes. It will generate the AT→CG or AT→GC transversion. Moreover, the binding energies of 8-oxo-G2:T (17.32 kcal/mol), 8-OH-G1:T (16.81 kcal/mol), xanthosine3:T (14.52 kcal/mol), and Fapy-G:T (16.14 kcal/mol) are larger than that of A:T(WC) (14.26 kcal/mol). Thus, T is more likely to pair with 8-OH-G, xanthosine, 8-oxo-G or Fapy-G than A. Just like the case shown in Fig. 12a→b→c→d and a→b1→c1→d1. There are also two more routines, a→b′→c′→d′ and a→b1′→c1′→d1′, to lead AT→CG or AT→GC mutation.

In DNA triple helix, the interaction energies of A:T(H) and A:T(RH) are 14.85 and 14.25 kcal/mol, respectively, and smaller than those of the four OGs with A. Thus, the oxidation of G may form Hoogsteen base pair pattern with A (Figs. 3–6). Subsequent to an additional replication cycle, the T:A:T will transverse to G:A:T. In addition, oxidized G will lead to a mutation or genetic information missing. For example, the binding energy of 8-OH-G:G(WC) is 30.07 kcal/mol, which is larger than that of G:C(WC) by 3.73 kcal/mol, and it will make the DNA in natural condition cannot cleave. The binding energy of Xanthosine:G is smaller than that of G:C(WC) by 8.33 kcal/mol, which may inhibit the DNA reunion.

4. Conclusion

Based on the results of *ab initio* method, we developed the ABEEM$\sigma\pi$ PFF to base pairs including oxidized guanine, and this method was applied to study the geometric structures, electronic structures, binding energies, and stacking energies of these model molecules. The results show that the performance of ABEEM$\sigma\pi$ PFF in predicting the geometric parameters (including bond lengths, bond angles, H-bond lengths, H-bond angles), electric structures, and the interaction energies *etc.* is generally better than those of the common force fields, and its accuracy can reach or approach those of the MP2 method. These studies supply a solid basis for further investigations of dynamics behavior of DNA fragment including oxidized guanine, and will provide insight into the mutation of DNA.

The geometries of four oxidized guanines have different extent of change relative to that of guanine, which affects those H-bonding interactive sites. In addition, the planar structure of base is also broken, such as Fapy-G shows the nonplanar structure. When the oxidized guanine couples with canonical base, the type and number of H-bond are different from that of canonical base pair. The angle formed by two base planes is obviously buckled, which will impact the stacking interaction between base pairs. The deviation degree of the planar ring structure of modified base pairs was characterized quantitatively, and our results show that 36% of the buckled angles in the modified base pairs is larger than 5.0°.

Moreover, the pairing behavior of the oxidized guanine was investigated systematically. When guanine is oxidized to 8-oxo-G, Fapy-G or 8-OH-G, the binding ability paired with C dose not obviously change. But, the binding ability of Xanthosine:C is obviously smaller than that of G:C(WC). In the double helix of DNA, the binding energy of 8-oxo-G or Fapy-G paired with C is the largest, so they will not mismatch with other bases. 8-OH-G and Xanthosine will mismatch with G to cause the mutation of GC→CG. The binding ability of 8-OH-G:G is stronger than that of G:C, which will lead the DNA in nature condition cannot cleave. The binding energy of Xanthosine:G is smaller than G:C, which may inhibit the DNA reunion. The binding ability of any of those four OGs paired with A (or T) is stronger than that of A:T (WC), indicating they will mismatch with A (or T) to cause the mutation of AT→CG and AT→GC. In triple helix, these four OGs can all pair with A by Hoogsteen pattern, which causes the mutation of T:A:T to G:A:T.

Acknowledgments

This research has been funded by the grants from the National Natural Science Foundation of China (No. 21133005).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2013.10.008.

References

[1] M.M. Greenberg, The formamidopyrimidines: purine lesions formed in competition with 8-Oxopurines from oxidative stress, Acc. Chem. Res. 45 (2012) 588–597.
[2] C.J. Lord, A. Ashworth, The DNA damage response and cancer therapy, Nature 481 (2012) 287–294.
[3] C.M. Crenshaw, J.E. Wade, H. Arthanari, D. Frueh, B.F. Lane, M.E. Núñez, Hidden in plain sight: subtle effects of the 8-oxoguanine lesion on the structure, dynamics, and thermodynamics of a 15-Base pair oligodeoxynucleotide duplex, Biochemistry 50 (2011) 8463–8477.
[4] A. Naôme, P. Schyman, A. Laaksonen, D.P. Vercauteren, Molecular dynamics simulation of 8-oxoguanine containing DNA fragments reveals altered hydration and ion binding patterns, J. Phys. Chem. B 114 (2010) 4789–4801.
[5] M.L. Hamm, K.A. Crowley, M. Ghio, M.A.M. Lindell, E.J. McFadden, J.S.L. Silberg, A.M. Weaver, Biochemical investigations into the mutagenic potential of 8-Oxo-2′-deoxyguanosine using nucleotide analogues, Chem. Res. Toxicol. 25 (2012) 2577–2588.
[6] T. Lindahl, Instability and decay of the primary structure of DNA, Nature 362 (1993) 709–715.
[7] C.B. Volle, D.A. Jarem, S. Delaney, Trinucleotide repeat DNA alters structure to minimize the thermodynamic impact of 8-Oxo-7, 8-dihydroguanine, Biochemistry 51 (2012) 52–62.
[8] M.D. Evans, M. Dizdaroglu, M.S. Cooke, Oxidative DNA damage and disease: induction, repair and significance, J. Comput. Chem. 567 (2004) 1–61.
[9] X.B. Hu, H. Li, J.Y. Ding, S.J. Han, Mutagenic mechanism of the A-T to G-C transition induced by 5-bromouracil: an ab initio study, Biochemistry 43 (2004) 6361–6369.
[10] N. Jayanth, S. Ramachandran, M. Puranik, Solution structure of the DNA damage lesion 8-oxoguanosine from ultraviolet resonance Raman spectroscopy, J. Phys. Chem. A 113 (2009) 1459–1471.
[11] V. Thiviyanathan, A. Somasunderam, D.E. Volk, T.K. Hazra, S. Mitra, G.D. Gorenstein, Base-pairing properties of the oxidized cytosine derivative, 5-hydroxy uracil, Biochem. Biophys. Res. Commun. 366 (2008) 752–757.
[12] M. Dey, F. Moritz, J. Grotemeyer, E.W. Schlag, Base pair formation of free nucleobases and mononucleosides in the gas phase, J. Am. Chem. Soc. 116 (1994) 9211–9215.
[13] E.S. Manas, Z. Getahun, W.W. Wright, W.F. DeGrado, J.M. Vanderkooi, Infrared spectra of amide groups in R-Helical proteins: evidence for hydrogen bonding between helices and water, J. Am. Chem. Soc. 122 (2000) 9883–9890.
[14] E.D. Isaacs, A. Shukla, P.M. Platzman, D.R. Hamann, B. Barbiellini, C. Tulk, Covalency of the hydrogen bond in ice: a direct X-ray measurement, Phys. Rev. Lett. 82 (1999) 600–603.
[15] A.H. Romero, P.L. Silvestrelli, M. Parinello, Compton scattering and the character of the hydrogen bond in ice I_h, J. Chem. Phys. 115 (2001) 115.
[16] A. Shukla, B. Barbiellini, T. Buslaps, P. Suortti, Compton spectroscopy and chemical bonding, Phys. Chem. (Munich) 215 (2001) 1315–1321.
[17] S. Ragot, J.M. Gillet, P.J. Becker, Interpreting compton anisotropy of ice I_h: a cluster partitioning method, Phys. Rev. B 65 (2002) 235115.
[18] J. Rezac, P. Hobza, Describing noncovalent interactions beyond the common approximations: how accurate is the Gold Standard, CCSD(T) at the complete basis set limit, J. Chem. Theory Comput. 9 (2013) 2151–2155.
[19] G. Villani, Theoretical investigation of the coupling between hydrogen-atom transfer and stacking interaction in adenine–thymine dimmers, Chem. Phys. Chem. 14 (2013) 1256–1263.
[20] K.E. Riley, M. Pitoňák, J. Černý, P. Hobza, On the structure and geometry of biomolecular binding motifs (hydrogen-bonding, stacking X-H$\cdots\pi$), J. Chem. Theory Comput. 6 (2010) 66–80.
[21] D. Svozil, P. Hobza, J. Šponer, Comparison of intrinsic stacking energies of ten unique dinucleotide steps in A-RNA and B-DNA duplexes. Can we determine correct order of stability by quantum-chemical calculations? J. Phys. Chem. B 114 (2010) 1191–1203.
[22] C.D. Sherryll, T. Takatani, E.G. Hohenstein, An assessment of theoretical methods for nonbonded interactions: comparison to complete basis set limit coupled-cluster potential energy curves for the benzene dimer, the methane dimer, benzene-methane, and benzene-H_2S, J. Phys. Chem. A 113 (2009) 10146–10159.
[23] J. Šponer, K.E. Riley, P. Hobza, Nature and magnitude of aromatic stacking of nucleic acid bases, Phys. Chem. Chem. Phys. 10 (2008) 2595–2610.
[24] P.H. obza, Stacking interactions, Phys. Chem. Chem. Phys. 10 (2008) 2581–2583.
[25] S. Nakagawa, Polarizable model potential function for nucleic acid bases, J. Comput. Chem. 28 (2007) 1538–1550.
[26] M. Karplus, J.A. McCammon, Theoretical investigation of the coupling between hydrogen-atom transfer and stacking interaction in adenine-thymine dimmers, Nat. Struct. Biol. 9 (2002) 646–652.
[27] M.W. van der Kamp, K.E. Shaw, C.J. Woods, A.J. Mulholland, Biomolecular simulation and modelling: status, progress and prospects, J. R. Soc. Interface 5 (2008) 173–190.
[28] J. Wang, P. Cieplak, J. Li, Q. Cai, M.J. Hsieh, R. Luo, Y. Duan, Development of polarizable models for molecular mechanical calculations. 4. van der Waals Parametrization, J. Phys. Chem. B 116 (2012) 7088–7101.
[29] J. Wang, P. Cieplak, Q. Cai, M.J. Hsieh, J. Wang, Y. Duan, R. Luo, Development of polarizable models for molecular mechanical calculations. 3. polarizable water models conforming to thole polarization screening schemes, J. Phys. Chem. B 116 (2012) 7999–8008.
[30] J. Wang, P. Cieplak, J. Li, J. Wang, Q. Cai, M.J. Hsieh, H.X. Lei, R. Luo, Y. Duan, Development of polarizable models for molecular mechanical calculations. II. Induced dipole models significantly improve accuracy of intermolecular interaction energies, J. Phys. Chem. B 115 (2011) 3100–3111.
[31] T. Yan, Y.T. Wang, C. Knox, On the structure of ionic liquids: comparisons between electronically polarizable and nonpolarizable models I, J. Phys. Chem. B 114 (2010) 6905–6921.
[32] J. Wang, P. Cieplak, J. Li, T. Hou, R. Luo, Y. Duan, Development of polarizable models for molecular mechanical calculations. I. Parameterization of atomic polarizability, J. Phys. Chem. B 115 (2011) 3091–3099.
[33] S. Patel, A.D. MacKerell Jr., C.L. Brooks III, CHARMM fluctuating charge force field for proteins. II. Protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model, J. Comput. Chem. 25 (2004) 1504–1514.
[34] W. Xie, J. Pu, A.D. MacKerell Jr., J. Gao, Development of a polarizable intermolecular potential function (PIPF) for liquid amides and alkanes, J. Chem. Theory Comput. 3 (2007) 1878–1889.
[35] C.M. Baker, V.M. Anisimov, A.D. MacKerell, Development of CHARMM polarizable force field for nucleic acid bases based on the classical drude oscillator model, J. Phys. Chem. B 115 (2011) 580–596.
[36] E.A. Orabi, G. Lamoureux, Cation–π and π–π interactions in aqueous solution studied using polarizable potential models, J. Chem. Theory Comput. 8 (2012) 182–193.
[37] J.L. Banks, G.A. Kaminski, R. Zhou, D.T. Mainz, B.J. Berne, R.A. Friesner, Parametrizing a polarizable force field from ab initio data. I. The fluctuating point charge model, J. Chem. Phys. 110 (1999) 741–754.
[38] P.J. Winn, G.G. Ferenczy, C.A. Reynolds, Towards improved force fields. III. Polarization through modified atomic charges, J. Comput. Chem. 20 (1999) 704–712.
[39] G.A. Kaminski, H.A. Stern, B.J. Berne, R.A. Friesner, Y.X.X. Cao, R.B. Murphy, R.H. Zhou, T.A. Halgren, Development of a polarizable force field for proteins via *ab initio* quantum chemistry: first generation model and gas phase tests, J. Comput. Chem. 23 (2002) 1515–1531.
[40] W.L. Jorgensen, K.P. Jensen, A.N. Alexandrova, Polarization effects for hydrogen-bonded complexes of substituted phenols with water and chloride ion, J. Chem. Theory Comput. 3 (2007) 1987–1992.
[41] A. Holt, J. Bostrom, G. Karlström, R. Lindh, A NEMO potential that includes the dipole-quadrupole and quadrupole–quadrupole polarizability, J. Comput. Chem. 31 (2010) 1583–1591.
[42] J.W. Ponder, C.J. Wu, P.Y. Ren, V.S. Pande, J.D. Chodera, M.J. Schnieders, I. Haque, D.L. Mobley, D.S. Lambrecht, R.A. DiStasio Jr., M. Head-Gordon, G.N.I. Clark, M.E. Johnson, T. Head-Gordon, Current status of the amoeba polarizable force field, J. Phys. Chem. B 114 (2010) 2549–2564.
[43] P. Ren, C. Wu, J.W. Ponder, Polarizable atomic multipole-based molecular mechanics for organic molecules, J. Chem. Theory Comput. 7 (2011) 3143–3161.
[44] A.G. Donchev, N.G. Galkin, A.A. Illarionov, O.V. Khoruzhii, M.A. Olevanov, V.D. Ozrin, M.V. Subbotin, V.I. Tarasov, Water properties from first principles: simulations by a general-purpose quantum mechanical polarizable force field, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 8613–8617.
[45] O. Khoruzhii, A.G. Donchev, N. Galkin, A. Illarionov, M. Olevanov, V. Ozrin, C. Queen, V.I. Tarasov, Application of a polarizable force field to calculations of relative protein–ligand binding affinities, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 10378–10383.
[46] A.G. Donchev, V.D. Ozrin, M.V. Subbotin, O.V. Tarasov, V.I. Tarasov, A quantum mechanical polarizable force field for biomolecular interactions, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 7829–7834.
[47] B. Bauer, S. Patel, Recent applications and developments of charge equilibration force fields for modeling dynamical charges in classical molecular dynamics simulations, Theory Chem. Acc. 131 (2012) 1153.
[48] W.J. Mortier, S.K. Ghosh, S. Shankar, Electronegativity equalization method for the calculation of atomic charges in molecules, J. Am. Chem. Soc. 108 (1986) 4315–4320.
[49] D.M. York, W. Yang, A chemical potential equalization method for molecular simulations, J. Chem. Phys. 104 (1996) 159–172.
[50] Z.Z. Yang, C.S. Wang, Atom-bond electronegativity equalization method. 1. Calculation of the charge distribution in large molecules, J. Phys. Chem. A 101 (1997) 6315–6321.
[51] Z.Z. Yang, B.Q. Cui, Atomic charge calculation of metallobiomolecules in terms of the ABEEM method, J. Chem. Theory Comput. 3 (2007) 1561–1568.
[52] D.X. Zhao, C. Liu, F.F. Wang, C.Y. Yu, L.D. Gong, S.B. Liu, Z.Z. Yang, Development of a polarizable force field using multiple fluctuating charges per atom, J. Chem. Theory Comput. 6 (2010) 795–804.
[53] Z.Z. Yang, Y. Wu, D.X. Zhao, Atom-bond electronegativity equalization method fused into molecular mechanics. I. A seven-site fluctuating charge and flexible body water potential function for water clusters, J. Chem. Phys. 120 (2004) 2541–2557.
[54] X. Li, Z.Z. Yang, Hydration of Li+-ion in atom-bond electronegativity equalization method-7P water: a molecular dynamics simulation study, J. Chem. Phys. 122 (2005) 084514.
[55] F.F. Wang, D.X. Zhao, Z.Z. Yang, Theoretical studies of uracil–$(H_2O)_n$ (n = 1–7) clusters by ab initio and ABEEM π/MM fluctuating charge model, Chem. Phys. 360 (2009) 141–149.
[56] L.D. Gong, Development and applications of the ABEEM fluctuating charge molecular force field in the ion-containing systems, Sci. China Chem. 55 (2012) 2471–2484.

[57] C.Y. Yu, Z.Z. Yang, Theoretical study of hydrogen peroxide interacting with DNA base and DNA base pair in terms of *ab initio* method and ABEEM π/MM fluctuating charge potential model, Comput. Theory Chem. 967 (2011) 26–36.
[58] L. Clowney, S.C. Jain, A.R. Srinivasan, J. Westbrook, W.K. Olson, H.M. Berman, Geometric parameters in nucleic acids: nitrogenous bases, J. Am. Chem. Soc. 118 (1996) 509–518.
[59] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery Jr., T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J.B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, Gaussian 03, revision B3, Gaussian, Inc, Pittsburgh, PA, 2003.
[60] P.G. Jasien, G. Fitzgerald, Base dipole (from jcc-28-1538), J. Chem. Phys. 93 (1990) 2554–2560.
[61] N. Foloppe, A.D. Mackerell Jr., All-atom empirical force field for nucleic acids. I. Parameter optimization based on small molecule and condensed phase macromolecular target data, J. Comput. Chem. 21 (2000) 86–104.
[62] J. Šponer, J. Leszczyński, P. Hobza, Nature of nucleic acid–base stacking: nonempirical ab initio and empirical potential characterization of 10 stacked base dimers. Comparison of stacked and H-bonded base pairs, J. Phys. Chem. A 100 (1996) 5590–5596.
[63] P. Hobza, M. Kabeláč, J. Šponer, P. Mejzlík, J. Vondrášek, Performance of empirical potentials (AMBER, CFF95, CVFF, CHARMM, OPLS, POLTEV), semiempirical quantum chemical methods (AM1, MNDO/MPM3), and ab initio Hartree–Fock method for interaction of DNA bases: comparison with nonempirical beyond Hartree–Fock results, J. Comput. Chem. 18 (1997) 1136–1150.