

# PCCP

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

## The R- and S-diastereoisomeric effects on the Guanidinohydantoin-Induced Mutations in DNA

N.R. Jena<sup>1\*</sup>, Vivek Gaur<sup>2</sup>, P.C. Mishra<sup>3</sup>

### Abstract

Direct and indirect oxidation of guanine in DNA produces guanidinohydantoin (Gh), which is capable of inhibiting replication and inducing mutation during cellular activities. Although, some biochemical studies have proposed that Gh may induce exclusively G to C mutation in DNA, other studies have predicted occurrence of both G to C and G to T mutations. However, the exact reasons for these mutations and the dubious character of Gh in this context are not yet understood. Further, due to insufficient structural data, the electronic structure of Gh that can participate in the formation of different base pair complexes in DNA is also not known. Here, density functional theory (DFT) is used to find the most stable tautomers of Gh at the base level out of a total 112 possible tautomers and their involvement in mutagenesis is investigated by computing structures, energies and electronic properties of different base pair complexes formed between the *syn*- and *anti*-conformations of the most stable tautomer of Gh (aGh) and all the bases of DNA. It is found that aGh can coexist in R- and S- diastereoisomeric configurations. Due to the flexible guanidinium group, it can rotate about the N3-C4 bond in each of the above diastereoisomers to form two different stable conformations (aGh1 and aGh2). It is further shown that among the different base pair complexes involving aGh1, *syn*-aGh1:G is the most stable. It indicates that G would be easily incorporated against *syn*-aGh1 giving rise to G to C mutations in DNA. However, in the case of aGh2, G is the preferred base pair partner of *syn*-aGh2 and T is the preferred base pair partner of *anti*-aGh2. It implies that in addition to G to C mutations, the occurrence of aGh2 in DNA may also induce G to A mutation. Further, due to similarity between base pairing patterns and binding energies of *syn*-aGh1:A and *syn*-aGh2:A complexes with those of the T:A complex, DNA polymerases may mistakenly insert A opposite aGh1 or aGh2 by misrecognizing the latter as T. This may ultimately induce G to T mutations in DNA. However, as the constraints imposed by the DNA backbones and stacking interactions were not considered here, the possibilities of aGh2:T and aGh2:A base pairs need to be investigated experimentally. It is further found that the mutagenic character of aGh in the R- and S-diastereoisomeric forms is similar.

Keywords: Guanidinohydantoin (Gh), DNA damage, Base modification, Mutation, Cancer, DFT

<sup>1</sup>Discipline of Natural Sciences, Indian Institute of Information Technology, Design and Manufacturing, Khamaria, Jabalpur-482005, India

<sup>2</sup>Discipline of Mechanical Engineering, Indian Institute of Information Technology, Design and Manufacturing, Khamaria, Jabalpur-482005, India

<sup>3</sup>NASI Senior Scientist, Department of Physics, Banaras Hindu University, Varanasi-221005, India

---

\*Corresponding Author's Email Address: [nrjena@iiitdmj.ac.in](mailto:nrjena@iiitdmj.ac.in)

## 1. Introduction

DNA damage due to the oxidation of guanine (G) [1-3] is deleterious as it promotes mutagenesis [4-6], creates replication errors [7] and induces pathological conditions [8,9] including cancer and several neurodegenerative diseases. It was initially believed that the formation of 8-oxoguanine (8-oxoG) due to the oxidative damage of guanine is the main source of mutagenesis [10,11]. However, recent studies have revealed that the secondary oxidation of 8-oxoG leading to the formation of guanidinohydantoin (Gh), spiroiminodihydantoin (Sp), 2-2-diamino-oxazolone (Oz), etc are much more stable, mutagenic and lethal as compared to 8-oxoG [12-17]. These lesions can also be directly produced from guanine by exposure of DNA to several oxidants such as singlet oxygen [18], superoxide radical anion [19], peroxyxynitrite [20] etc [21]. Further, depending upon the reaction environment e.g. pH and temperature, yields of Gh and Sp lesions can be different. For example, in acidic conditions, formation of Gh would be favored over that of Sp [19,22] while in basic conditions, yield of Sp would be dominant [18,21,22]. Similarly, at the nucleotide level, yield of Gh is more than that of Sp while at the nucleoside level, the reverse is true [18,21,23].

Recent observations regarding the formation of Sp in the liver and colon of mice [24] have shown that Sp and Gh are of great biological relevance. Formation of Sp was also detected in the Neibase excision repair enzyme deficient *Escherichia coli* (*E. coli*) after the cells were treated with chromate [25]. Both *in vitro* and *in vivo* studies have shown that Sp can coexist in R- and S-diastereoisomeric configurations, which are stable and distinct [26-29]. Further, the formation of these lesions can induce different mutations characteristic of the stereochemical configurations [30-33]. For example, by inserting oligonucleotides containing Sp into bacteriophage DNA having a sequence of 5'-GXA-3' (X=Sp), it was observed that one of the isomers of Sp (Sp1) can yield 72% of G to C and 27% of G to T mutations, while the other isomer (Sp2) can induce 57% of G to C and 41% of G to T mutations [31]. Similarly, in another study, taking the 5'-TXG-3' (X=Sp) sequence, Sp1 was found to yield 19% of G to C mutations and 78% of G to T mutations, while Sp2 was observed to yield 48% of G to C mutations and 49% of G to T mutations [32]. Although, the absolute configuration of Sp1 and Sp2 was not known in these studies [31,32], in latter studies these lesions were identified to be S- and R-diastereoisomers of Sp respectively [26-29]. Similar studies taking Gh in the 5'-TXA-3' (X=Gh) sequence, yielded 98% of G to C and 2% of G to T mutations [31], while in the 5'-TXG-3' sequence, it yielded 57% of G to C, 40% of G to T and 3% of G to A mutations [32]. Interestingly, in a recent steady-state kinetic study, the incorporation of A opposite Gh in the presence of RB69 gp43 *exo*<sup>-</sup> DNA polymerase was found to be more favored as compared to that of G [34]. Although, these studies pointed out about the insertion of a nucleotide opposite Gh, the actual diastereoisomeric configuration of Gh that was mainly involved in inducing the above mutations is not known. This could be due to difficulty in isolating R- and S-diastereoisomers of Gh owing to their quick conversion from one form to the other [18,19]. It was proposed that Gh diastereoisomers are

interconvertible via the enolization of the C4 carbonyl that has a very fast rate constant [35,36]. Beside this, the formation of R- and S-diastereoisomers of Gh was observed in an NMR-study in which the two  $^{15}\text{N}$ -labelled nitrogen atoms of the guanidinium group were equivalent and coupled to the adjacent  $^{13}\text{C}$  [18].

It was proposed that the occurrence of Gh in DNA would affect its stability and distort it. In order to understand this destabilizing effect, optical melting analysis and differential scanning calorimetry (DSC) experiments were performed by considering a 15-mer DNA duplex, where C was base paired opposite Gh [37]. The results obtained were compared with the same duplex containing G and 8-oxoG. It was indeed found that stability of the Gh:C pair was the least as compared to those of the G:C and 8-oxoG:C pairs [37]. In spite of these results, no structural data are available for base pairs of Gh which are greatly valuable with regard to elucidation of involvement of Gh in mutagenesis. Recently, the crystal structure of a replicative DNA polymerase (RB69) bound to Gh containing DNA has been obtained (pdb 3L8B) [34]. In this structure, Gh was observed to be in the R-diastereoisomeric configuration (Gh-R) and extrahelical due to its rotation toward the major groove. As a result, the insertion of a complementary nucleotide opposite Gh-R was impossible. Although, this study indicated about the ability of Gh-R to inhibit replication, no structural information regarding its mutagenic potential was obtained [34]. It was thus proposed that instead of Gh-R, Gh-S i.e. the S-diastereoisomeric configuration of Gh may be involved in promoting mutagenesis [34].

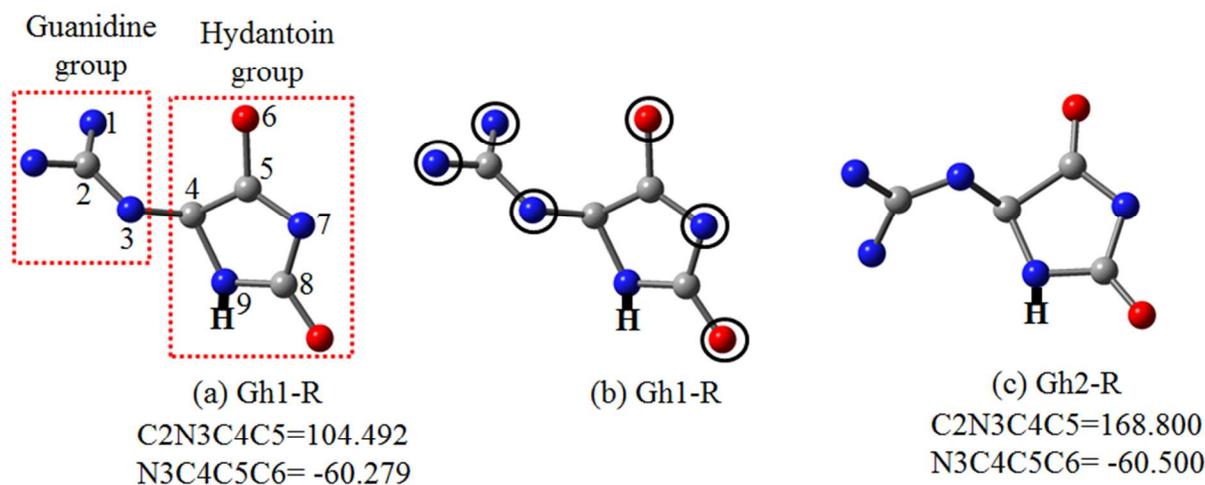
To rationalize the actual cause of Gh mediated mutagenesis, structures, binding energies and electronic properties such as electron density and electrostatic potential distributions in isolated Gh and different base pairs involving its R- and S-diastereoisomeric configurations were studied here by employing density functional theory (DFT). As during different base pairings, the

oxidatively damaged products of guanine can undergo *N*-glycosidic bond rotation, roles of *anti*- and *syn*- conformations of Gh on the formation of different base pairs were also investigated. It should be noted that the typical  $\chi$  value (O4'-C1'-N9-C4 dihedral angle) for the 2'-deoxyguanosine in the *anti*-conformation lies in the range -120 to -180 deg., while in the *syn*-conformation, it lies in the range 0 to 90 deg. [29]. Further, in the *anti*-conformation, the six-membered ring of 2'-deoxyguanosine (Watson-Crick face) is involved in hydrogen bonded interactions with the neighboring base, while in the *syn*-conformation, the five-membered ring (Hoogsteen face) participates in the base pairing interactions. As in the present study we have not considered any *N*-glycosidic bond and the guanidine and hydantoin groups of Gh are created by bond breaking and subsequent rearrangements of the six- and five-membered rings of 2'-deoxyguanosine respectively, these groups are considered to represent the *anti*- and *syn*-confirmations of Gh (Scheme 1).

## 2. Computational Methodology

The coordinates of Gh-R taken from the earlier X-ray crystal structure of DNA containing Gh complexed with the RB69 DNA-polymerase [34] were used to build the starting structure of Gh-R after removing the phosphate and sugar groups for the present study (Scheme 1a). In order to find stable tautomers of Gh-R, hydrogen atoms were placed at different suitable positions as shown in Scheme 1b. This created 28 different conformations of Gh-R. During geometry optimization in aqueous medium, it was found that the guanidine group can freely rotate about the N3-C4 bond which connects it with the hydantoin group, thereby forming another conformation (Scheme 1c). For convenience of description, the former (Scheme 1a) and latter (Scheme 1c) structures would be referred to as Gh1-R and Gh2-R respectively. The main difference between these structures arises due to the dihedral angle C2N3C4C5 which is close to

104 and 169 deg. in the Gh1-R and Gh2-R conformations respectively. It was also found that as in the case of Gh1-R, 28 different tautomers of Gh2-R are possible. The same procedure was adopted to find all the 56 tautomers of Gh1 and Gh2 in the S-diastereoisomeric configuration.



Scheme 1. Structures of Gh-R obtained from the X-ray crystallographic study [34]. (a) Gh1-R, (b) Oxygen and nitrogen atoms (encircled) to which hydrogen atoms would be bonded that can be used to generate different tautomeric conformations of Gh1-R, and (c) Gh2-R as obtained by geometry optimization in aqueous medium. The atomic numbering scheme of Gh used here is shown in (a).

By using the most stable tautomers of Gh1 and Gh2 in the R- and S-diastereoisomeric configurations, structures and stabilities of different base pairs were computed. The B3LYP/6-31+G\* [38,39] and  $\omega$ B97XD/AUG-cc-pVDZ levels of theories [40,41] were used for geometry optimization and single point energy calculations respectively. All calculations were initially performed in the gas phase and subsequently in aqueous medium by using the integral equation formalism of the polarizable continuum model (IEF-PCM) of the self-consistent reaction field theory (SCRF) [42,43]. As geometry optimization was not performed at the  $\omega$ B97XD/AUG-cc-pVDZ level of theory, zero-point energy (ZPE) corrections obtained at the B3LYP/AUG-cc-

pVDZ level were considered to be valid at the former level also. All the calculations were performed using the Gaussian suite of program (G09) [44] and structures were visualized employing the GaussView program (version 5.0) [45]. As the  $\omega$ B97XD functional includes dispersion interaction and the AUG-cc-pVDZ basis set is much larger than the 6-31+G\* basis set, the results obtained at the  $\omega$ B97XD/AUG-cc-pVDZ level of theory are expected to be more accurate than those obtained at the B3LYP/6-31+G\* level. Further, as the aqueous medium may be considered to represent the biological system, we would mainly discuss the results obtained at the  $\omega$ B97XD/AUG-cc-pVDZ level of theory in aqueous medium. To calculate the ZPE-corrected binding energies of the different base-pair complexes, Equation (1) given below was used:

$$AB_{BE} = AB_{TE} - [A_{TE} + B_{TE}] \quad \text{-----(1)}$$

where A and B are any two DNA bases and AB is a base pair between A and B. The subscripts BE and TE stand for ZPE-corrected binding energy and ZPE-corrected total energy, respectively.

### 3.1 Stability of different tautomers of Gh

Geometry optimization in aqueous medium for 28 possible tautomeric structures of each of Gh1-R and Gh2-R yielded their 21 and 28 tautomers respectively. The optimized structures of these tautomers are presented in the Supporting Information (Figs.S1-S3). It may be noted that during geometry optimization, 7 input tautomeric structures of Gh1-R got converted to Gh2-R. If we compare the ZPE-corrected total energies of these tautomers, it becomes evident that the amino tautomers of Gh1-R (aGh1-R) and Gh2-R (aGh2-R) as shown in Figs. 1a,b are the most stable and aGh1-R is about 2.3 kcal/mol more stable than aGh2-R. Interestingly, aGh1-R is found to be about 6.00 kcal/mol more stable than the previously observed imino tautomers of Gh at the nucleoside level by  $^1\text{H}$  NMR spectroscopy (Fig. S1d,f) [46,47] and hence would be likely to be

observed in DNA. This is consistent with an earlier computational study, where it was found that in aqueous solution the population of aGh1-R would be about 76% [48]. It should be noted that like Gh, the amino tautomer of Sp was also found to be more stable than the imino tautomer [30]. Similarly, ZPE-corrected total energies of different tautomers of Gh-S in aqueous medium (Figs. S4-S6) showed the amino tautomers of aGh1-S and aGh2-S to be more stable than the imino tautomers (Fig. 1 c,d). Among these tautomers, aGh1-S is 1.75 kcal/mol more stable than aGh2-S. If we compare the stabilities of aGh1-R and aGh1-S, it is clear that the former is only 0.56 kcal/mol more stable than the latter (Fig. 1). These results indicate that the formation of both R- and S-diastereoisomers in single-stranded DNA would be nearly equally probable. We also noticed that although aGh2 is stable in aqueous medium, it gets converted to aGh1 in gas phase. This and the fact that stabilities of aGh1 and aGh2 in aqueous medium are comparable reveal that interconversion between aGh1 to aGh2 may occur in single-stranded DNA.

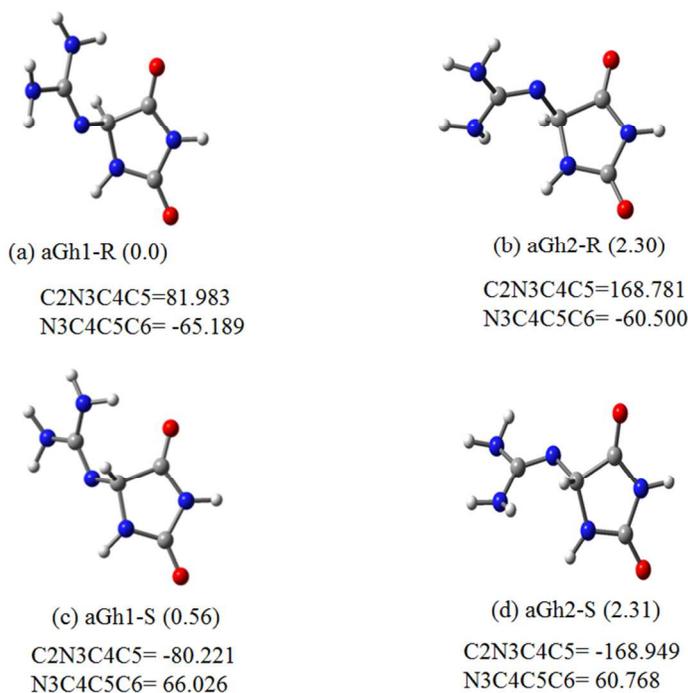


Fig.1. Optimized structures of the most stable amino tautomers aGh1-R, aGh2-R, aGh1-S and aGh2-S in the aqueous medium. ZPE-corrected relative energies (kcal/mol) of the different structures calculated with respect to aGh1-R are given in parentheses and a few important geometrical parameters are also shown.

## 3.2 Structures and binding energies of different base pair complexes involving the R-diastereoisomer of aGh

### 3.2.1 aGh-R:G complexes

The geometry optimization of aGh-R:G complexes in aqueous medium yielded five different possible complex structures. Out of these, four important and most stable base pair complexes are shown in Fig. 2. The ZPE-corrected relative binding energies (kcal/mol) of these complexes are presented in Table 1. The detailed structures and binding energies of all optimized aGh-R:G complexes are provided in the supporting Information (Fig S7, Tables S1, S2). As can be seen from Fig. 2, aGh1-R can bind with G either by engaging its Watson-Crick face (*anti*-conformation) or Hoogsteen face (*syn*-conformation). Irrespective of the *anti*- and *syn*-conformations, aGh1-R makes two strong hydrogen bonds with G. In the *anti*-conformation, the amino group of the guanidine moiety is mainly involved in the formation of hydrogen bonds, while in the *syn*-conformation, the hydantoin group participates in the formation of hydrogen bonds with G. We note that in going from the *anti*-aGh1-R to *syn*-aGh1-R (Fig. 2a,b), G is shifted downward, making the latter complex about 4 kcal/mol more stable than the former (Table 1). This wobble-type hydrogen bonding pattern is similar to that of the T:G complex, where downward shifting of the incoming G is required to make stable interactions with the template T [49]. In order to compare the base pair pattern and binding energy of *syn*-aGh1-R:G with those of the T:G complex, the structure of T:G complex (Fig. 2a) was optimized in the

aqueous medium by employing the B3LYP/6-31+G\* level of theory followed by single-point energy calculation at the  $\omega$ B97XD/AUG-cc-pVDZ level (Table 1). It is thus revealed that not only the *syn*-aGh1-R:G (Fig. 2c) and T:G (Fig. 2a) complexes are structurally similar but also these are energetically comparable (Table 1). This suggests that the insertion of G opposite *syn*-aGh1-R is likely to occur in DNA. However, in the case of normal DNA, base-pair complexes involving the *anti*-conformation are more stable than those involving the *syn*-conformation. However, different oxidatively damaged products of guanine such as 8-oxoG and FapyG were predicted to base pair with complementary bases of DNA in both the *anti*- and *syn*-conformations [50-53]. For example, the incorporation of C opposite *anti*-8-oxoG was predicted to be non-mutagenic [50], while base pairing of A opposite *syn*-8-oxoG was suggested to induce G to T mutations [50]. Similarly, both *anti*-FapyG:A and *syn*-FapyG:A complexes have been recently proposed to yield G to T mutations in DNA [51-53]. These results imply that aGh1-R can also adopt the *syn*-conformation in DNA to make a stable base pair complex with G.

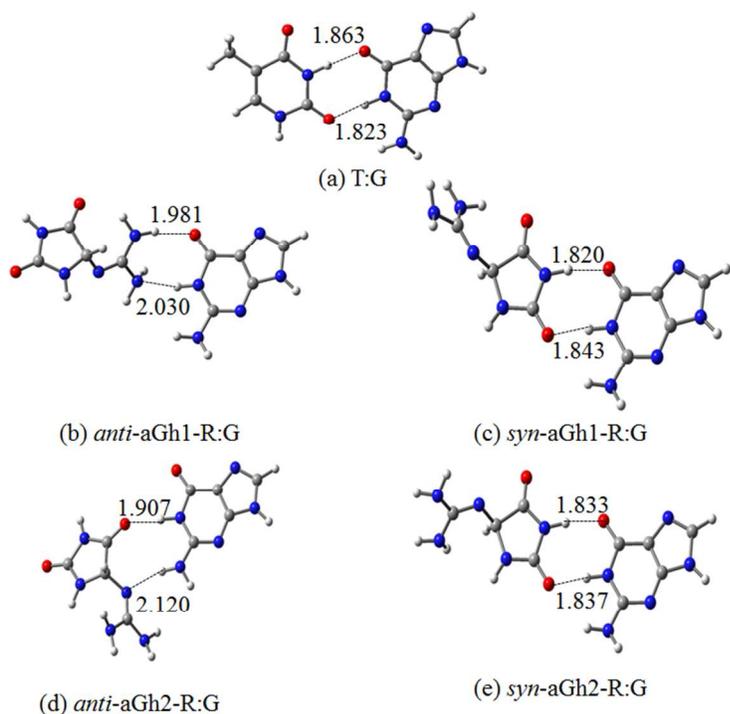


Fig. 2: The optimized structures of (a) T:G and the most stable (b,c) aGh1-R:G and (d,e) aGh2-R:G complexes as obtained in the aqueous medium.

Like aGh1-R, aGh2-R can also make two strong hydrogen bonds with G in both the *anti*- and *syn*-conformations. In the *anti*-aGh2-R:G complex, the N3 and O6 atoms of aGh2-R (Scheme 1a) are involved in the formation of hydrogen bonds, while in the *syn*-aGh2-R:G complex, the N7 and O8 atoms are involved in the formation of hydrogen bonds. In order to make strong hydrogen bonds with the hydantoin group of *syn*-aGh2-R, G moves down relative to its position in the *anti*-aGh2-R:G complex. As a result, the binding pattern of *syn*-aGh2-R:G complex becomes similar to that of the *syn*-aGh1-R:G complex. The only difference arises due to the reorientation of the guanidine group in the former complex. This binding pattern resulted in *syn*-aGh2-R:G complex to be  $\sim 2$  kcal/mol more stable than the *anti*-aGh2-R:G complex, which is energetically comparable with the *syn*-aGh1-R:G and T:G complexes (Table 1). It implies that both aGh1-R and aGh2-R can make equally stable complexes with G only in the *syn*-conformation and during this interaction its guanidine group may switch between aGh1-R and aGh2-R.

Table 1. ZPE-corrected relative binding energies (kcal/mol) of different base pair complexes involving aGh1 and aGh2-R in aqueous medium. To compare the stability of different aGh1-R:G complexes, the relative binding energies of these complexes were calculated with respect to the *anti*-aGh1-R:G complex. Similarly, the relative binding energies of other complexes were calculated with respect to the corresponding *anti*-aGh1-R:X (X=C or A or T) complex. For comparison, the ZPE-corrected relative binding energies of T:G, G:C, and A:T base pair complexes computed with respect to the *anti*-aGh1-R:G, *anti*-aGh1-R:C and *anti*-aGh1-R:A complexes respectively in aqueous medium are given in parentheses.

Base Pair	Method	aGh1-R		aGh2-R	
		<i>anti</i>	<i>syn</i>	<i>anti</i>	<i>syn</i>
Gh:G ( <b>T:G</b> )	$\omega$ B97XD	0.00 <b>(-3.41)</b>	-3.85	-1.77	-3.64
Gh:C ( <b>G:C</b> ) <sup>a</sup>	$\omega$ B97XD	0.00 <b>(-7.95)</b>	-0.89	-1.93	-1.03
Gh:A ( <b>T:A</b> ) <sup>a</sup>	$\omega$ B97XD	0.00 <b>(-4.03)</b>	-3.34	-3.60	-3.44
Gh:T	$\omega$ B97XD	0.00	-2.43	-4.01	-2.55

<sup>a</sup> Ref [51,54].

### 3.2.2 aGh-R:C complexes

Geometry optimization in aqueous medium yielded eight possible base pair aGh-R:C complexes (Fig. S8). Out of these, four most stable base pair complexes involving aGh1-R and aGh2-R are illustrated in Fig.3. As can be seen from this figure, C can bind with aGh1-R and aGh2-R in both *anti*- and *syn*-conformations via two strong hydrogen bonds. However, in the Watson-Crick G:C base pair, it binds with G only in the *anti*-conformation via three strong hydrogen bonds. It is also found that the conversion of G to aGh-R reduces the binding energy of G:C complex by about 6-8 kcal/mol (Table 1) [51,54]. This is in accordance with the earlier melting studies which have shown a reduced stability by ~6-7 kcal/mol for the Gh:C complex [37].

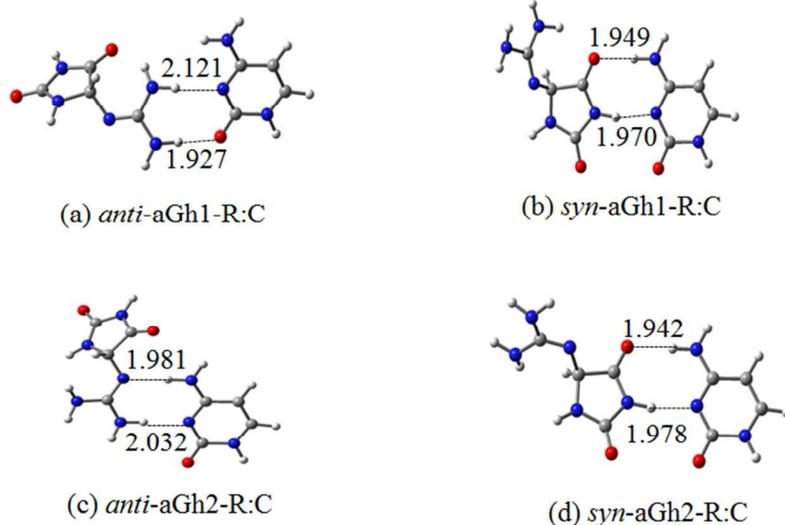


Fig. 3: The optimized structures of the most stable (a,b) aGh1-R:C and (c,d) aGh2-R:C complexes as obtained in the aqueous medium.

### 3.2.3 aGh-R:A complexes

It was found that aGh-R can bind with A in seven possible ways (Fig. S8). Out of these, the most stable conformations of aGh1-R:A and aGh2-R:A in the *anti*- and *syn*-conformations are shown in Fig. 4. To compare these structures with that of the T:A base pair, optimized structure of the latter complex is also shown in this figure (Fig. 4a). It is clear from this figure that the binding modes of *syn*-aGh1-R:A and *syn*-aGh2-R:A complexes are similar to that of the T:A complex. Further, it is found that the *syn*-aGh1-R:A complex is appreciably (about 3 kcal/mol) more stable than the *anti*-aGh1-R:A complex, while the *anti*-aGh2-R:A and *syn*-aGh2-R:A complexes are equally stable. It indicates that during the incorporation of A, aGh1-R may adopt only the *syn*-conformation, while aGh2-R can adopt both the *anti*- and *syn*-conformations. However, from the binding mode of the *anti*-aGh2-R:A complex, it is evident that the hydantoin group is highly non-planar and may interfere with the neighboring 5'-nucleotide, thereby distorting DNA

locally. It should be noted that recently the S-diastereoisomeric configuration of Sp has been crystallized in duplex DNA [55]. As Sp-S contains two rings nearly perpendicular to each other, it was expected to create a large distortion in DNA and probably would not fit at the template position in DNA. However, the successful placement of Sp-S in DNA suggests that it would not be impossible for aGh2-R to adopt the *anti*-conformation to form the *anti*-aGh2-R:A complex in DNA.

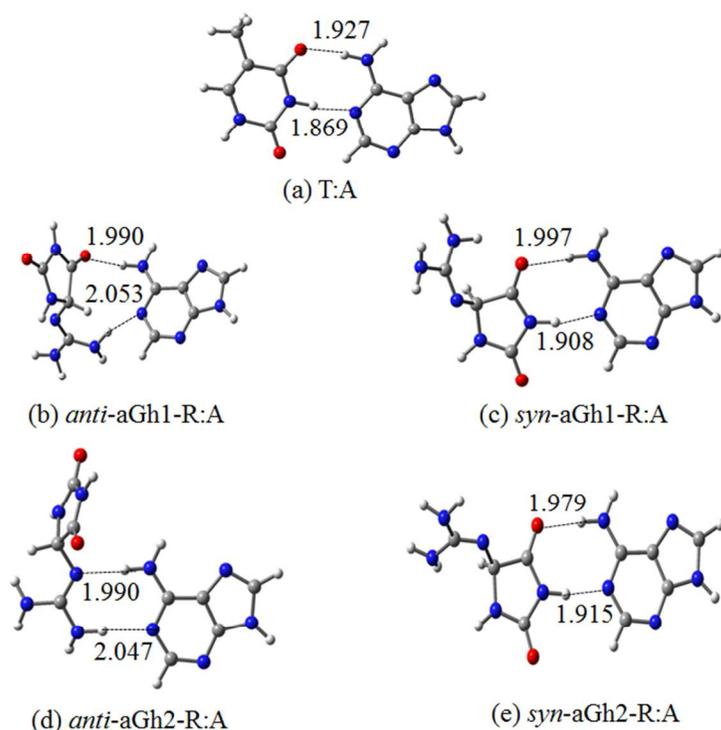


Fig. 4: The optimized structures of (a) T:A and the most stable (b,c) aGh1-R:A and (d,e) aGh2-R:A complexes as obtained in the aqueous medium.

If the formation of these complexes occurs in DNA, they would remain unrecognized by DNA polymerases as their binding energies and binding patterns are similar to that of the normal T:A complex. For this reason, DNA polymerases may mistakenly recognize both aGh1-R and aGh2-R as T and incorporate A opposite it, thereby inducing G to T mutation in DNA. However, the

insertion of A opposite aGh1-R has not been observed by any wild type polymerase. An attempt made by the earlier X-ray crystallographic study in the presence of RB69 polymerase failed to yield any useful base paired structure [34]. However, in a subsequent X-ray crystallographic study, the active site residue Tyr567 of the RB69 polymerase was mutated to Ala567 [56]. This mutation indeed helped to position Gh1-R within the DNA double helix, thereby forming a Gh1-R:A complex. It was argued that the Tyr567-Ala567 mutation mainly assisted in the displacement of G568 from Gh1-R, where it was initially hydrogen bonded with the O8 atom of Gh1-R toward the minor groove. This structural organization helped the polymerase to fully accommodate Gh1-R in the high *syn*-conformation ( $\chi=105$  deg) within the DNA double helix [56]. Notably, the observed *syn*-Gh1-R:A complex is similar to the *syn*-aGh1-R:A complex obtained here (Fig. 4c). However, in the experimentally observed complex, the N7 atom of Gh1-R was only directly hydrogen bonded with the N1 of A and the O6 atom of Gh1-R was indirectly hydrogen bonded with the N6 of A through a water molecule. However, we find that in the *syn*-aGh1-R:A complex, both the N7 and O6 atoms of Gh1-R make strong hydrogen bonds with A (Fig. 4b).

Although, this mutagenesis study [56] suggests that the insertion of A opposite *syn*-Gh1-R is possible, it cannot explain why the wild-type RB69 polymerase could not produce *syn*-Gh1-R:A complex [30]. We stress that as observed recently [56], if the *syn*-aGh1-R:A complex can be formed in DNA, the formation of *syn*-aGh2-R:A complex would also be possible in DNA and these complexes might generate rigidity in DNA when bound to some DNA polymerases. However, the conversion of *syn*-aGh2-R to *anti*-aGh2-R would reduce this rigidity by substantially displacing the hydantoin moiety away from the G568 residue of RB69 polymerase due to which the insertion of A opposite *anti*-aGh2-R may become possible. Thus, it can be

proposed that depending on the fidelity of DNA polymerases, incorporation of A opposite *syn*-aGh1-R, *syn*-aGh2-R and *anti*-aGh2-R may be possible in DNA.

### 3.2.4 aGh-R:T complexes

As obtained here, aGh1-R and aGh2-R can bind with T in five and three possible ways respectively (Fig. S10). The most stable *anti*- and *syn*-conformations of aGh1-R and aGh2-R can make base pair complexes with T by making two strong hydrogen bonds each (Fig. 5). As evident from Table 1, the binding of T with *anti*-aGh2-R is comparatively more stable than those with *syn*-aGh2-R and *syn*-aGh1-R.

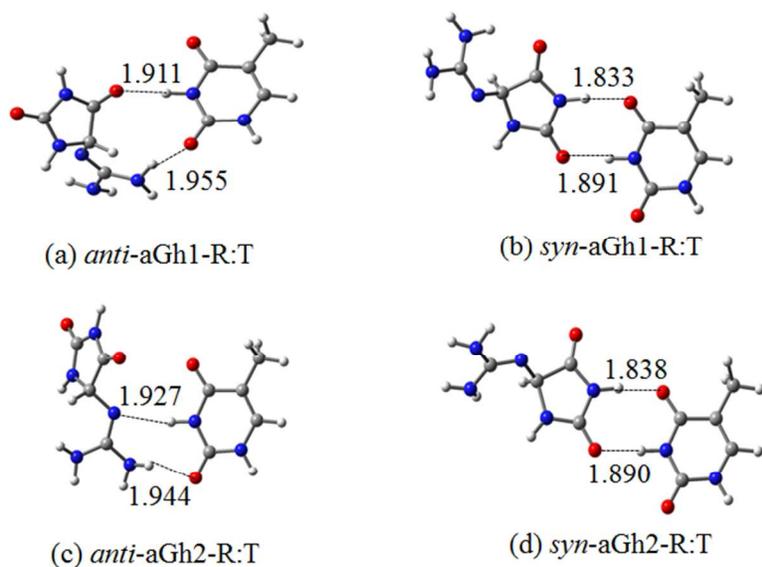


Fig. 5: The optimized structures of the most stable (a,b) aGh1-R:T and (c,d) aGh2-R:T complexes as obtained in the aqueous medium.

In order to evaluate the stabilities of different complexes, the relative ZPE-corrected binding energies of these complexes were calculated with respect to the *anti*-aGh1-R:G as presented in

Table 2. From this Table it is clear that different complexes involving aGh1-R follow the order:  $syn\text{-aGh1-R}_1\text{:G} > syn\text{-aGh1-R}_1\text{:A} \geq syn\text{-aGh1-R}_1\text{:T} > syn\text{-aGh1-R}_1\text{:C}$ . Similarly, binding energies of different complexes involving aGh2-R follow the order:  $anti\text{-aGh2-R}_1\text{:T} \geq syn\text{-aGh2-R}_1\text{:G} > anti\text{-aGh2-R}_1\text{:C} \geq anti\text{-aGh2-R}_1\text{:A} \geq syn\text{-aGh2-R}_1\text{:A} \approx syn\text{-aGh2-R}_1\text{:T} > syn\text{-aGh2-R}_1\text{:C} > anti\text{-Gh2-R}_1\text{:G}$ . It also indicates that the stability of  $anti\text{-aGh2-R}_1\text{:T}$  complex is comparable with those of the  $syn\text{-aGh1-R}_1\text{:G}$  and  $syn\text{-aGh2-R}_1\text{:G}$  complexes (Table 2).

Table 2: ZPE-corrected relative binding energies (kcal/mol) of different base pair complexes involving aGh1 and aGh2-R in aqueous medium calculated with respect to the  $anti\text{-aGh1-R}_1\text{:G}$  complex.

Base Pair	Method	aGh1-R		aGh2-R	
		<i>anti</i>	<i>syn</i>	<i>anti</i>	<i>syn</i>
Gh:G	$\omega$ B97XD	0.00	-3.85	-1.77	-3.64
Gh:C	$\omega$ B97XD	-0.90	-1.79	-2.83	-1.93
Gh:A	$\omega$ B97XD	1.14	-2.20	-2.46	-2.30
Gh:T	$\omega$ B97XD	-0.25	-2.18	-3.76	-2.30

It is worthwhile to mention that although, the DFT methods used in the present study can produce reliable energetic data, use of more sophisticated electron correlation based methods such as MP4, CCSD etc. may further improve these results. As these methods are quite expensive, these were not used here. However, based on the results obtained here, it can be proposed that the insertion of G opposite  $syn\text{-aGh1-R}$  and  $syn\text{-aGh2-R}$  in DNA can yield very tight base pair complexes, prolonged persistence of which may induce G to C mutations in DNA. Similarly, the incorporation of T opposite  $anti\text{-aGh2-R}$  may also yield equally stable base pair complex similar to  $syn\text{-aGh1-R}_1\text{:G}$  and  $syn\text{-aGh2-R}_1\text{:G}$ . However, it is interesting to investigate

whether due to small size of T, it can be inserted opposite aGh2 by overcoming constraints imposed by DNA backbone, base stacking and steric interactions. Similarly, although, the structural and energetic data also indicate that the insertion of A opposite aGh1 (only in the *syn*-conformation) and aGh2 (in both *anti*- and *syn*-conformations) may also be possible in DNA, it remained to be experimentally verified.

### 3.3 Structures and binding energies of different base pair complexes involving the S-diastereoisomer of aGh

#### 3.3.1 aGh-S:G complexes

Structures and relative binding energies of the most stable aGh1-S:G and aGh2-S:G complexes in the *anti*- and *syn*-conformations are presented in Fig.6 and Table 3, respectively. The structures and binding energies of all possible optimized structures of aGh-S:G complexes are presented in Fig. S11, Tables S3,S4 respectively but will not be discussed here. The binding modes depicted in Fig. 6 suggests that G also binds with aGh-S by making two strong hydrogen bonds like those of the aGh-R:G complexes (Fig. 2). It is also found that in the *syn*-conformation, aGh1-S and aGh2-S make more stable base pair complexes with G compared to the *anti*-conformation (Table 3). This implies that in DNA, aGh1-S and aGh2-S would base pair with G only in the *syn*-conformation. Further, a comparison of the binding patterns and binding energies of the *syn*-aGh1-S:G and *syn*-aGh2-S:G complexes (Fig. 6, Table 3) with those of the *syn*-aGh1-R:G, *syn*-aGh2-R:G and T:G complexes (Fig. 2, Table 1) shows that these complexes are quite similar. The only difference between these complexes arises due to positioning of the guanidine group relative to the hydantoin moiety. In the complexes involving the R-diastereoisomeric configuration, the guanidine group lies above the hydantoin group, while in the

complexes involving the S-diastereoisomeric configuration, it lies below the hydantoin group. As in the *syn*-conformation, the hydantoin group of aGh is involved in making the base pair interactions with G, the details of positioning of guanidine moiety would not affect the overall binding modes and binding energies of aGh-S:G and aGh-R:G complexes. This might be the reason why biochemical studies could not quantify the mutagenic potential of the Gh-R and Gh-S conformations separately.

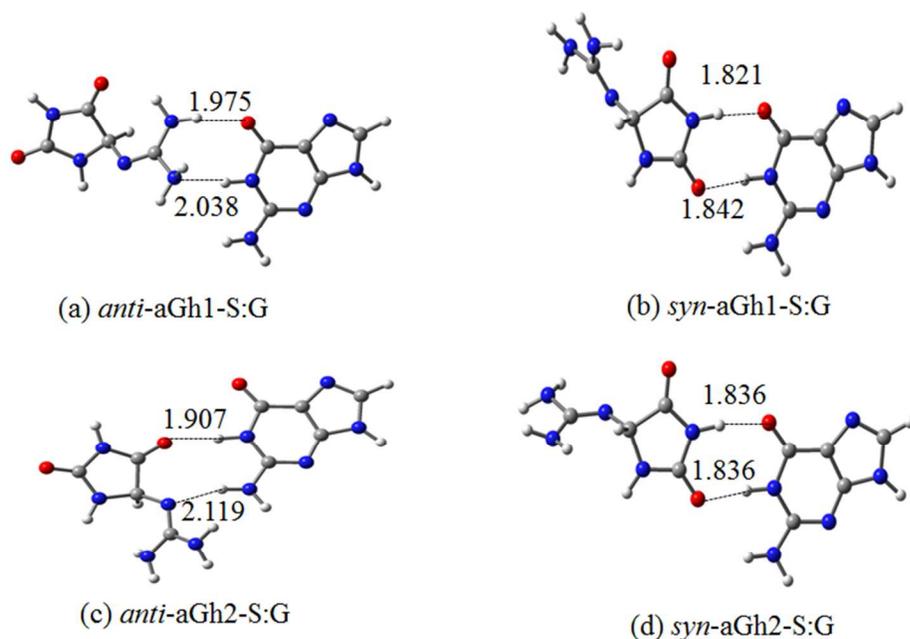


Fig. 6: The optimized structures of the most stable (a,b) aGh1-S:G and (c,d) aGh2-S:G complexes as obtained in the aqueous medium.

Table 3. ZPE-corrected relative binding energies (kcal/mol) of different base pair complexes involving aGh1-S and aGh2-S in aqueous medium. To compare the stability of different aGh1-S:S:G complexes, the relative binding energies of these complexes were calculated with respect to the *anti*-aGh1-S:G complex. Similarly, the relative binding energies of other complexes were calculated with respect to the corresponding *anti*-aGh1-S:X (X=C or A or T) complex.

Base Pair	Method	aGh1-S		aGh2-S	
		<i>anti</i> 1	<i>syn</i> 1	<i>anti</i> 1	<i>syn</i> 1
Gh:G	$\omega$ B97XD	0.00	-3.36	-1.28	-3.31
Gh:C	$\omega$ B97XD	0.00	-0.75	-2.21	-0.88
Gh:A	$\omega$ B97XD	0.00	-1.24	-3.32	-2.93
Gh:T	$\omega$ B97XD	0.00	-2.49	-6.98	-2.66

### 3.3.2 aGh-S:C complexes

All optimized aGh-S:C complexes are shown in Fig. S12 and the most stable aGh-S:C complexes in the *anti*- and *syn*-conformations are illustrated in Fig. 7. The patterns of binding modes and trends of stabilities of these complexes are similar to those of the corresponding complexes involving the R-diastereoisomer of aGh.

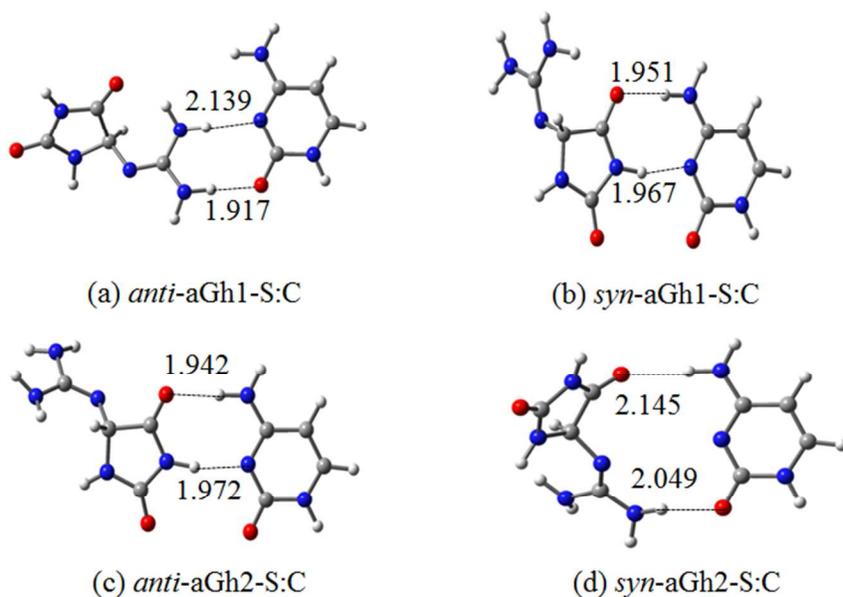


Fig. 7: The optimized structures of the most stable (a,b) aGh1-S:C and (c,d) aGh2-S:C complexes as obtained in the aqueous medium.

### 3.3.3 aGh-S:A complexes

Out of the seven possible conformations of aGh-S:A complexes (Fig. S13), the structures of the most stable aGh-S:A complexes in the *anti*- and *syn*-conformations are presented in Fig. 8. It is found that the insertion of A opposite the *anti*- and *syn*-conformations of aGh2-S would produce relatively tightly bound complexes compared to those of aGh1-S (Table 3). The binding patterns and binding energies of *anti*-aGh2-S:A and *syn*-aGh2-S:A complexes with those of the *anti*-aGh2-R:A and *syn*-aGh2-R:A complexes respectively, are similar, and the similarity also extends to the T:A complex (Fig. 4a). It is further revealed that the mutagenic character of *syn*-aGh1 or *syn*-aGh2 in both the R- and S-diastereoisomeric configurations would be similar, unlike the earlier presumption [34]. This is because in the *syn*-conformation, the chirality of the C4 atom does not influence the base pairing interactions of aGh with A.

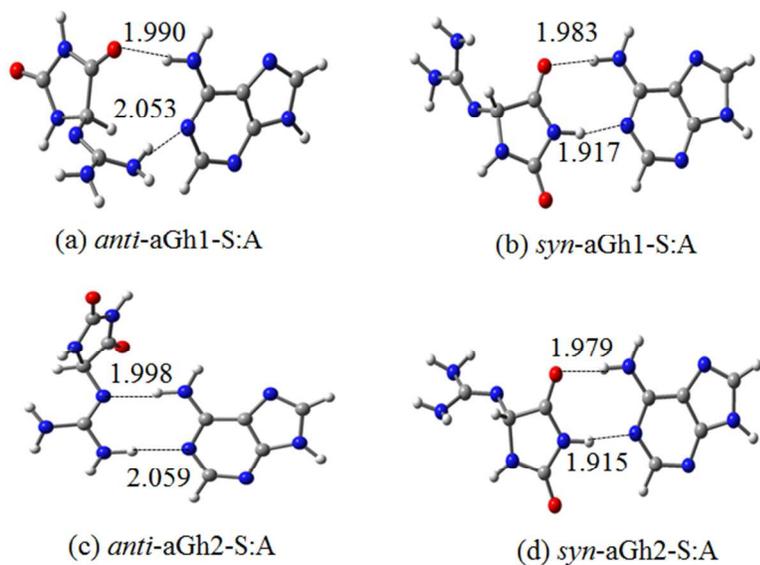


Fig. 8: The optimized structures of the most stable (a,b) aGh1-S:A and (c,d) aGh2-S:A complexes as obtained in the aqueous medium.

In order to verify this, the *syn*-aGh1-R:A and *syn*-aGh1-S:A complexes were aligned as shown in Fig. 9a. The aligned structures clarify that the nature of base pair interactions in these complexes are similar. Further, as discussed earlier, due to the bulky non-planar hydantoin group, formation of the *anti*-aGh2-S:A complex might distort DNA like the *anti*-aGh2-R:A complex. In order to test if the occurrence of *anti*-aGh2-S:A and *anti*-aGh2-R:A complexes would support the insertion of A opposite *anti*-aGh2-R and *anti*-aGh2-S, these complexes were superposed by considering A as the reference base (Fig. 9b). As shown in Fig. 9b, it appears that the formation of these complexes would not be difficult in DNA and it may not drastically affect the stacking interactions with the neighboring bases. However, the position of the hydantoin group may interfere with the neighboring bases. This structure has a distinctive feature in the sense that it can be easily identified and processed by repair enzymes, thereby diminishing its mutagenic potential.

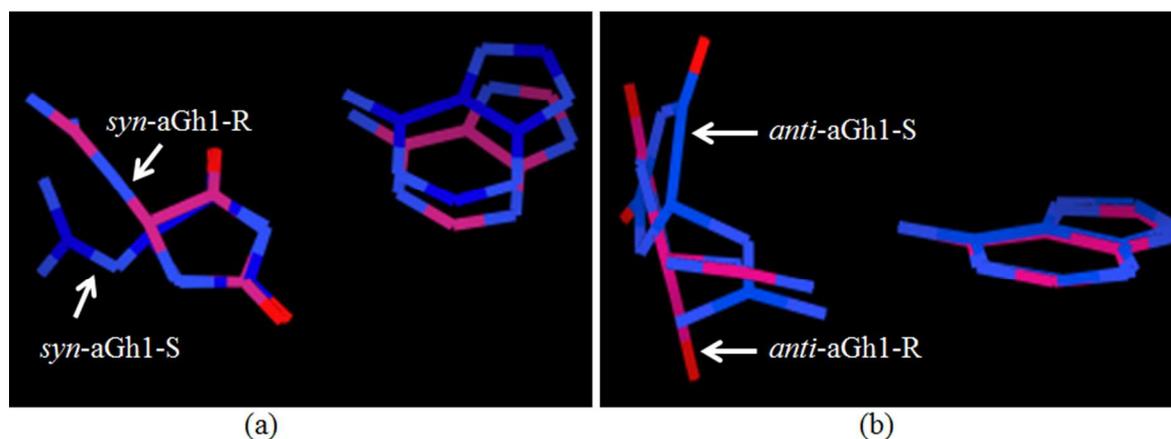


Fig.9. Alignment of complexes: (a) *syn*-aGh1-R:A and *syn*-aGh1-S:A complexes and (b) *anti*-aGh1-R:A and *anti*-aGh1-S:A complexes.

### 3.3.1 aGh-S:T complexes

Out of the ten possible aGh-S:T complexes (Fig. S14), four most stable complexes are shown in Fig. 10. The binding patterns of these complexes are similar to those of the aGh-R:T complexes. It is found that the *syn*-aGh1-S:T complex is  $\sim 2.5$  kcal/mol more stable than the *anti*-aGh1-S:T complex and hence T can bind with aGh1-S only in the *syn*-conformation. Interestingly, the *anti*-aGh2-S:T complex is  $\sim 4$  kcal/mol more stable than the *syn*-aGh2-S:T and *syn*-aGh1-S:T complexes (Table 3). This difference in stability indicates that the binding of T with aGh2-S would only occur in the *anti*-conformation.

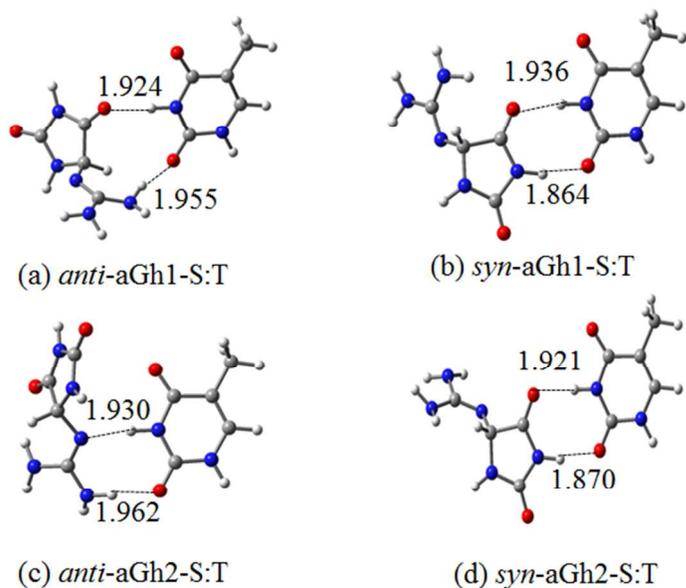


Fig. 10: The optimized structures of the most stable (a,b) aGh1-S:T and (c,d) aGh2-S:T complexes as obtained in the aqueous medium.

In order to compare the stabilities of all possible DNA base pair complexes involving the S-diastereoisomer of aGh, the relative binding energies of these complexes were evaluated with

respect to the *anti*-aGh1-S:G complex (Table 4). As evident from Table 4, G and T are the preferred base pair partners of aGh-S. The adoption of the *syn*-conformation of aGh1-S or aGh2-S would facilitate its binding with G, while the *anti*-conformation of aGh2-S will be preferred for the incorporation of T. However, it remains to be verified whether constraints imposed by DNA backbone, base stacking and steric interactions can allow T to be inserted opposite *anti*-aGh2-S. Even if this insertion is possible, due to the presence of bulky hydantoin group, recognition and repair of *anti*-aGh2-S:T complex by repair enzymes could be easily processed, which will ultimately diminish the level of G to A mutations in DNA.

Table 4. ZPE-corrected relative binding energies (kcal/mol) of different base pair complexes involving aGh1-S and aGh2-S in aqueous medium calculated with respect to *anti*-aGh1-S:G complex.

Base Pair	Method	aGh1-S		aGh2-S	
		<i>anti</i> 1	<i>syn</i> 1	<i>anti</i> 1	<i>syn</i> 1
Gh:G	$\omega$ B97XD	0.00	-3.36	-1.28	-3.31
Gh:C	$\omega$ B97XD	-0.66	-1.41	-2.87	-1.54
Gh:A	$\omega$ B97XD	1.11	-0.13	-2.21	-1.82
Gh:T	$\omega$ B97XD	1.69	-0.80	-5.29	-0.97

It is possible that the occurrence of either the *anti*- or *syn*-conformation of aGh may depend upon the sequence of stacking and steric interactions of DNA [31,32,57]. However, conversion from aGh1 to aGh2 may be spontaneous and independent of the sequence. In this light, the experimental observation of 98% G to C mutation in the 5'-GGhA-3' sequence [31] can be considered to arise due to the adoption of the *syn*-conformation by both aGh1 and aGh2 followed by the insertion of G opposite *syn*-aGh1 and *syn*-aGh2. Similarly, the detection of mixture of G to C, G to T and G to A mutations in the 5'-GGhG-3' sequence [32] may presumably have arisen

due to the insertion of G opposite *syn*-aGh2 followed by its quick conversion to *anti*-aGh2 and subsequent incorporation of A and T opposite *anti*-aGh2. This trend of mutagenicity would be followed in both R- and S-diastereoisomeric configurations. Due to their similar stabilities, it would be quite difficult to distinguish between these diastereoisomers in single- and double-stranded DNA.

### 3.4 Electronic properties of different base pairs involving aGh-R and aGh-S

Shape complementarity is an important factor that plays a crucial role in biomolecular recognition. Therefore, the determination of electron density and electrostatic potential around each molecule is vital in deciding the incorporation of actual base opposite aGh [52]. Calculated electrostatic potentials mapped onto the electron densities ( $0.0004 \text{ e/Bohr}^3$ ) of different vital base pair complexes involving aGh-R and aGh-S are presented in Fig.11 and Fig. S15 respectively. The XYZ-coordinates of these complexes are presented in the Supporting Information (Table S5-S20). For comparison, the distributions of electrostatic potentials mapped onto electron densities of G:C, T:A and T:G base pairs are also shown in these figures. It is obvious from these figures that distributions of electrostatic potentials and densities of *syn*-aGh1:G and *syn*-aGh2:G complexes are similar to those of the T:G complex. It also appears that the bulky guanidine group in *syn*-aGh1:G and *syn*-aGh2:G complexes would play a similar role as the methyl group in the T:G complex. These results indicate that G can be easily inserted opposite aGh in double-stranded DNA by DNA polymerases.

Similarly, the distributions of electrostatic potentials and electron densities of *syn*-aGh1:A and *syn*-aGh2:A complexes are similar to those of the T:A complex. Further, the *anti*-aGh2:A complex has a compact distribution of electron densities, which are generally found in the case

of normal base pair complexes. The distributions of electrostatic potentials around the *anti*-aGh2:A complex indicate that due to the presence of O6 and O8 atoms, the hydantoin group can favorably interact with the environment, thereby providing additional stability to DNA. These results imply that the incorporation of A opposite *syn*-aGh1, *syn*-aGh2, and *anti*-aGh2 may also be possible in DNA. It is also revealed that the distributions of electrostatic potentials and electron densities of *syn*-aGh1:T and *syn*-aGh2:T complexes are similar to those of the *syn*-aGh1:G and *syn*-aGh2:G complexes and electronic properties of *anti*-aGh2:T and *anti*-aGh2:A complexes are similar.

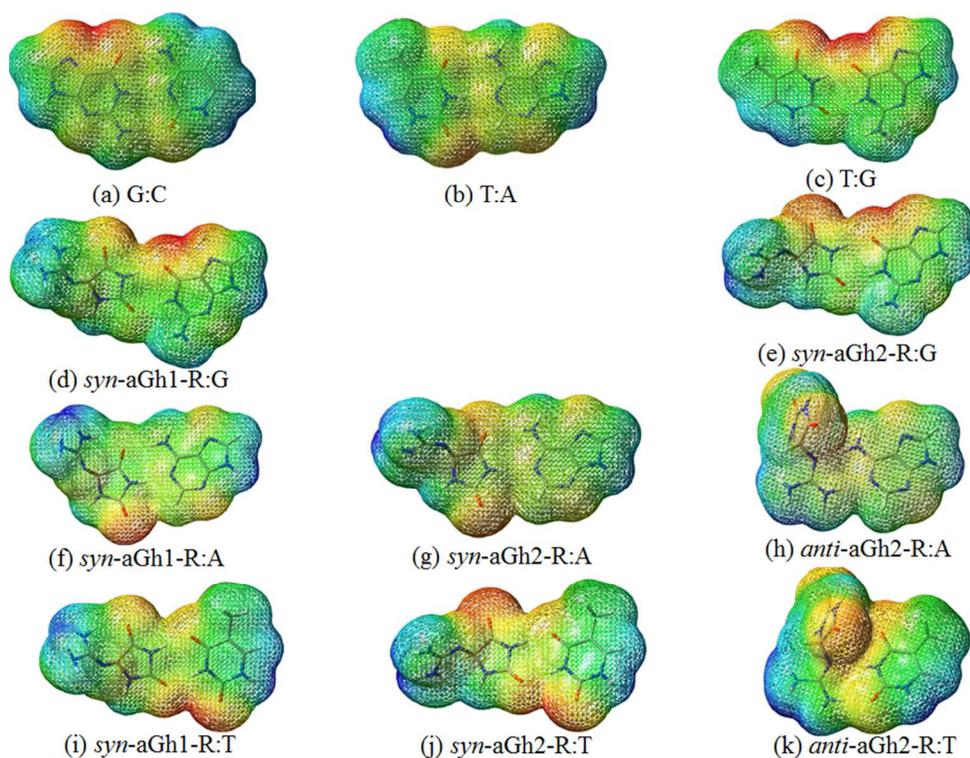


Fig. 11: Electrostatic potentials mapped onto the electronic densities ( $0.0004 \text{ e/Bohr}^3$ ) of different vital base pair complexes involving aGh-R as obtained at the B3LYP/6-31+G\* level in aqueous medium.

#### 4. Conclusion

It is revealed that both the R- and S-diastereoisomeric configurations of Gh are equally stable and difficult to energetically distinguish. Gh in both the stereochemical configurations can adopt two different amino tautomeric conformations namely aGh1 and aGh2. While aGh1 can make stable base pair complexes with the different bases in DNA only in the *syn*-conformation, aGh2 can adopt both *anti*- and *syn*-conformations. Among the different complexes formed in DNA involving aGh1, *syn*-aGh1:G and *syn*-aGh1:T complexes are the most stable. Similarly, among the different base pair complexes involving aGh2, *syn*-aGh2:G and *anti*-aGh2:T complexes are the most stable. These results indicate that if G and T can be inserted opposite aGh by overcoming constraints imposed by DNA backbones, base stacking and steric interactions, it would lead to the formation of highly stable base pair complexes, which might have serious mutagenic implications. For example, the insertion of G opposite aGh may lead to G to C mutations, while the incorporation of T opposite aGh may induce G to A mutations. In addition to this it is further found that the binding patterns and binding energies of *syn*-aGh1:A, *syn*-aGh2:A, and T:A complexes are similar. As a result, *syn*-aGh1 and *syn*-aGh2 can be mistaken as T by DNA polymerases, due to which A may be inserted opposite *syn*-aGh1 and *syn*-aGh2. It would eventually give rise to G to T mutations in DNA. It is further found that the insertion of A opposite *anti*-aGh2 by some DNA polymerases may also be possible. The overall results indicate that there is no preference for the incorporation of any base opposite R- or S-diastereoisomeric configurations of aGh and the trend of mutagenicity is similar in both the cases. Although, these results are remarkable to understand different base pair possibilities and hence Gh-induced mutations, the consideration of DNA backbone, base stacking and steric interactions and complexation of DNA with high fidelity DNA polymerases can yield somewhat different results.

Hence the results obtained here should be considered as a starting point to perform structural experiments to understand Gh-induced mutations in detail.

### Supporting Information

The optimized structures of various tautomers of Gh1 and Gh2, different possible base pair complexes involving Gh1-R, Gh2-R, Gh1-S and Gh2-S, electrostatic potentials mapped onto the electron densities of different complexes involving aGh-S, binding energies of different complexes, and XYZ-coordinates of important base pairs are given in this part.

### Acknowledgement

NRJ is thankful to the IITDM Jabalpur for research initiation grant. PCM is thankful to the National Academy of Sciences, India (NASI) for a Senior Scientist Fellowship.

### References

1. J. Cadet, J.R. Wagner, V. Shafirovich and N.E. Geacintov, *Int. J. Radiat. Biol.* 2014, **90**, 423-432.
2. N.R. Jena and P.C. Mishra *J. Phys. Chem. B.* 2005, **109**, 14205-14218.
3. N.R. Jena and P.C. Mishra, *J. Comput. Chem.* 2007, **28**, 1321-1335.
4. E.C. Friedberg, L.D. McDaniel and R.A. Schultz, *Curr. Opin. Gen. Dev.* 2004, **14**, 5-10.
5. N.R. Jena and P.C. Mishra, *Free Radic. Biol. Med.* 2012, **53**, 81-94.
6. N.R. Jena *J. Biosci.* 2012, **37**, 503-517.
7. V. Duarte, J.G. Muller and C.J. Burrows *Nucleic Acid. Res.* 1999, **27**, 496-502.
8. M.S. Cooke, M.D. Evans, M. Dizdaroglu and J. Lunec, *The FASEB J.* 2003, **17**, 1195-1214.
9. M. Dizdaroglu, *Mutat. Res. Rev. Mutat. Res.* 2015, **763**, 212-245.

10. H.J. Helbock, K.B. Beckman, M.K. Shigenaga, P.B. Walter, A.A. Woodall, H.C. Yeo and B.N. Ames *Proc. Natl. Acad. Sci. USA* 1998, **95**, 288-293.
11. K.C. Cheng, D.S. Cahill, H. Kasai, S. Nishimura and L.A. Loeb *J. Biol. Chem.* 1992, **267**, 166-172.
12. O. Korniyushyna, A.M. Berges, J.G. Muller and C.J. Burrows, *Biochemistry*, 2002, **41**, 15304-15314.
13. V. Duarte, D. Gasparutto, M. Jaquinod and J. Cadet *Nucleic Acid Res.* 2000, **28**, 1555-1563.
14. A. Yadav and P.C. Mishra, *Chem. Phys. Lett.* 2014, **592**, 232-237.
15. B.H. Munk, C.J. Burrows, and H.B. Schlegel *J. Am. Chem. Soc.* 2008, **130**, 5245-5256.
16. B.T. Psciuk and H.B. Schlegel *J. Phys. Chem. B* 2013, **117**, 9518-9531.
17. L.I. Shukla, A. Adhikary, R. Pazdro, D. Becker and M.D. Sevilla, *Nucleic Acid Res.* 2004, **32**, 6565-6574.
18. Y. Ye, J.G. Muller, W. Luo, C.L. Mayne, A.J. Shallop, R.A. Jones, and C.J. Burrows, *J. Am. Chem. Soc.* 2003, **125**, 13926-13927
19. W. Luo, J.G. Muller and C.J. Burrows *Org. Lett.* 2001, **3**, 2801-2804.
20. J.C. Niles, J.S. Wishnok, and S.R. Tannenbaum *Org. Lett.* 2001, **3**, 963-966.
21. C.J. Burrows, J.G. Muller, O. Korniyushyna, W. Luo, V. Duarte, M.D. Leipold and S.S. David, *Environ. Health Perspect.* 2002, **110**, 713-717.
22. Y. Ye, B.H. Munk, J.G. Muller, A. Cogbill, C.J. Burrows and H.B. Schlegel *Chem. Res. Toxicol.* 2009, **22**, 526-535.
23. A.M. Fleming, J.G. Muller, A.C. Dlouhy, and C.J. Burrows, *J. Am. Chem. Soc.* 2012, **134**, 15091-15102.
24. A. Mangerich, C.G. Knutson, N.M. Parry, S. Muthupalani, W. Ye, E. Prestwich, L. Cui, J.L. McFaline, M. Mobley, Z. Ge, K. Taghizadeh, J.S. Wishnok, G.N. Wogan, J.G. Fox, S.R. Tannenbaum, and P.C. Dedon, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, E1820-E1829.

25. M.K. Hailer, P.G. Slade, B.D. Martin, and K.D. Sugden, *Chem. Res. Toxicol.* 2005, **18**, 1378-1383.
26. A. Durandin, L. Jia, C. Crean, A. Kolbanovskiy, S. Ding, V. Shafirovich, S. Broyde, and N.E. Geacintov, *Chem. Res. Toxicol.* 2006, **19**, 908-913.
27. S. Ding, L. Jia, A. Durandin, C. Crean, A. Kolbanovskiy, and N.E. Geacintov, *Chem. Res. Toxicol.* 2009, **22**, 1189-93.
28. A.M. Fleming, A.M. Orendt, Y. He, J. Zhu, R.K. Dukor, and C.J. *J. Am. Chem. Soc.* 2013, **135**, 18191-18204.
29. B. Karwowski, F. Dupeyrat, M. Bardet, M., J.-L. Ravanat, P. Krajewski, and J. Cadet, *J. Chem. Res. Toxicol.* 2006, **19**, 1357-1365.
30. L. Jia, V. Shafirovich, R. Shapiro, N.E. Geacintov, S. Broyde, *Biochemistry*, 2005, **44**, 6043-6051.
31. P.T. Henderson, J.C. Delaney, J.G. Muller, W.L. Leeley, S.R. Tannenbaum, C.J. Burrows, and J.M. Essigmann, *Biochemistry*, 2003, **42**, 9257-9262.
32. S. Delaney, W.L. Neeley, W.L. Deaney, and J.M. Essigmann, *Biochemistry*, 2007, **46**, 1448-1455.
33. P.K. Shukla P.K. and P.C. Mishra, P.C. *Int. J. Quantum. Chem.* 2013, **113**, 2600-2604.
34. P. Aller, Y. Ye, S.S. Wallace, C.J. Burrows and S. Double *Biochemistry*, 2010, **49**, 2502-2509.
35. A. Chworos, Y. Coppel, Y., I. Dubey, G. Pratviel, G. and B. Meunier, *J. Am. Chem. Soc.*, 2001, **123**, 5867-5877.
36. A. Chworos, C. Seguy, G. Pratviel, and B. Meunier, *Chem. Res. Toxicol.* 2002, **15**, 1643-1651.
37. C.J. Yennie and S. Delaney, *Chem. Res. Toxicol.* 2012, **25**, 1732-1739.
38. C. Lee, W. Yang, and R.G. Parr *Phys. Rev. B*, 1988, **37**,785-789.
39. B. Miehlich, A. Savin, H. Stoll, and H. Preuss, *Chem. Phys. Lett.* 1989, **157**, 200-206.
40. J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.* 2008, **10**, 6615-6620.
41. J.-D. Chai and M.J. Head-Gordon, *Chem. Phys.* 2008, **128**, 084106.
42. J. Tomasi, B. Mennucci, and R. Cammi. *Chem. Rev.* 2005, **105**, 2999-3093.
43. G. Scalmani, G. and M.J. Frissh, *J. Chem. Phys.* 2010, **132**, 114110.

44. Gaussian09, Revision A.1, , M.J. Frisch., G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, V. Nakai, T. Vreven, J. Jr J.A. Montgomery, J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, C. J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
45. GaussView, Version 5, R. Dennington, T. Keith and J. Millam, Inc. *Semichem*, KS Shawnee Mission 2009.
46. A. Chworos, C. Seguy, G. Pratviel, G. and B. Meunier, *Chem. Res. Toxicol.* 2002, **15**, 1643-1651.
47. W. Luo, J.G. Muller, E.M. Rachlin, and C.J. Burrows, *Chem. Res. Toxicol.* 2001, **14**, 927-938.
48. V. Verdolino, R. Cammi, R., B.H. Munk, and H.B. Schlegel, *J. Phys. Chem. B*, 2008, **112**, 16860-16873.
49. D.T. Minnick, L. Liu, N.D. Grindley, T.A. Kunkel and C.M. Joyce, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 1194-119.
50. F. Le Page, A. Sarasin, A. Gentil, A., A. Guy, and J. Cadet, *Nucleic Acid Res.* 1988 **26**, 1276-1281.
51. N.R. Jena, and P.C. Mishra, *ChemPhysChem*, 2013, **14**, 3263-3270.
52. N.R. Jena and P.C. Mishra, *ChemPhysChem* 2014, **15**, 1779-1784.
53. T.H. Gehrke, U. Lischke, U., K.L. Gasteiger, S. Schneider, S., Arnold, H.C. Muller, D.S. Stephenson, H. Zipse, and T. Carell, *Nat. Chem. Biol.* 2013, **9**, 455-461.

54. N.R. Jena, and M. Bansal, *Phys. Biol.* 2011, **8**, 046007.
55. B.E. Eckenroth, A.M. Fleming, J.B. Sweasy, C.J. Burrows, and S. Double S. *Biochemistry*, 2014, **53**, 2075-2077.
56. J. Beckman, M. Wang, G. Blaha, J. Wang, and W.H. Konigsberg *Biochemistry*, 2010, **49**, 8554-8563.
57. E.T. Kool, *Annu. Rev. Biophys. Biomol. Struct.* 2001, **30**, 1-22.