

**Quantum Chemical Calculations on Locked Nucleic Acid based Antisense Modifications:  
A Density Functional Theory (DFT) Study at Monomer Level**

Mallikarjunachari V. N. Uppuladinne<sup>1</sup>, Dikshita Dowerah<sup>2</sup>, Uddhaves B. Sonavane<sup>1\*</sup>,  
Suvendra Kumar Ray<sup>3</sup>, Ramesh C. Deka<sup>2\*</sup>, Rajendra R. Joshi<sup>1\*</sup>.

*<sup>1</sup>HPC – Medical & Bioinformatics Applications Group, Centre for Development of Advanced Computing (C-DAC), Pune, India.*

*<sup>2</sup>Department of Chemical Sciences, Tezpur University, Assam, India.*

*<sup>3</sup>Department of Molecular Biology and Biotechnology, Tezpur University, Assam, India*

E-mail address: [rajendra@cdac.in](mailto:rajendra@cdac.in), [uddhaveshs@cdac.in](mailto:uddhaveshs@cdac.in), [ramesh@tezu.ernet.in](mailto:ramesh@tezu.ernet.in)

## Abstract

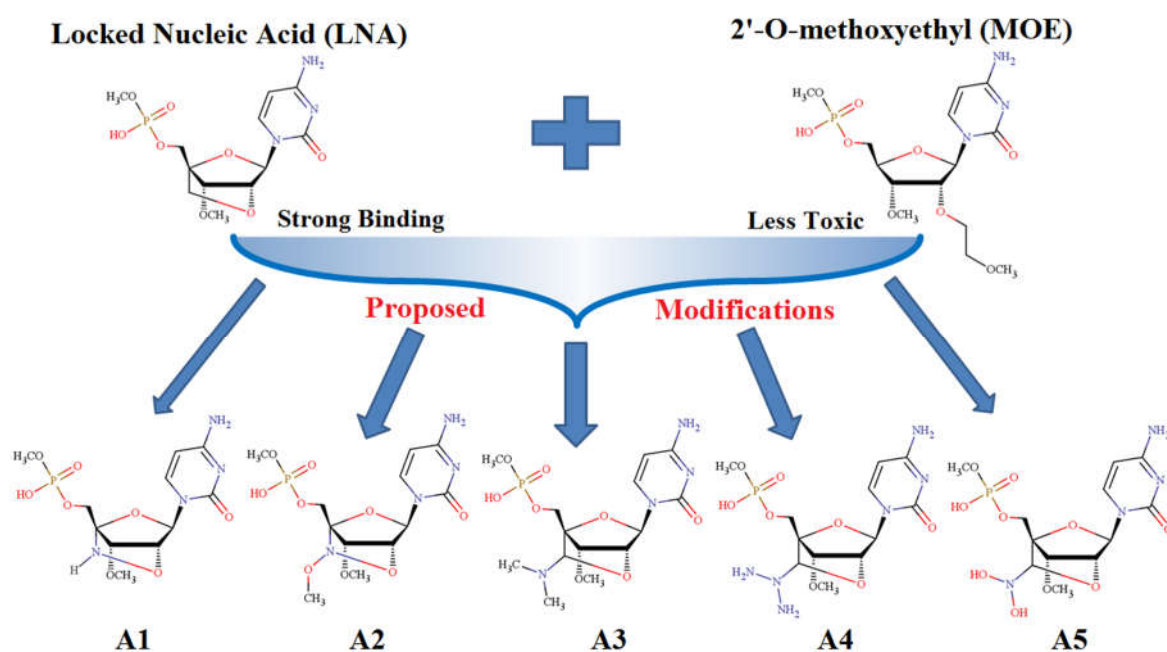
Antisense technology has been developed as the next generation drug discovery methodology by which unwanted gene expression can be inhibited by targeting mRNA specifically with antisense oligonucleotides. It has been observed that a good number of these molecules entered into clinical trials at a faster rate and some of them got approved. The computational studies of antisense modifications based on phosphorothioate (PS), methoxyethyl (MOE), locked nucleic acids (LNA) may help to design better novel modifications. In the present study, newer LNA based modifications have been proposed. The conformational search and density functional theory (DFT) calculations have been used to investigate the quantum chemical parameters of PS, LNA, MOE, and novel LNA based proposed modifications. The conformational search has been done to identify the most and alternative stable conformations. The geometry optimization followed by single point energy calculation has been done at B3LYP/6-31G(d,p) level for gas phase and B3LYP/6-311G(d,p) level for the solvent phase of all modifications. The electronic properties and the quantum chemical descriptors for the frontier molecular orbitals of all the antisense modifications were derived and compared. The local and global reactivity descriptors, such as hardness, chemical potential, electronegativity, electrophilicity index, Fukui function calculated at DFT level for the optimized geometries. These are used for understanding the reactive nature and reactive sites of the modifications. A comparison of global reactivity descriptors confirmed that LNA based modifications are the most reactive modifications and prone to the chemical reactions. It may form stable duplex when it is bound to complementary nucleotides, compared to other modifications. Therefore, we are proposing that one of our proposed antisense modification (A3) may show strong binding to the complementary nucleotide as LNA and may also show reduced toxic effects like MOE.

## Keywords

Antisense modifications, quantum chemical studies, conformational search, LNA, MOE, TANGO, MOPAC, global descriptors, Fukui-function.

## Graphical Abstract

Antisense technology is one of the best methods to regulate gene expression and acts as a therapeutic platform. Structural investigations of the LNA based novel antisense modifications have been carried out using quantum chemical parameters. Semi-empirical based conformational search has been carried out to identify the most stable conformation of these antisense modifications. Comparison has been done of DFT based quantum chemical descriptors of all antisense modifications like phosphorothioate (PS), methoxyethyl (MOE), locked nucleic acids (LNA), and novel proposed LNA based modifications.



## **Introduction**

Antisense technology is one of the therapeutic platforms to regulate gene expression in vivo [1]. The synthetic antisense oligonucleotides would influence gene expression and inhibit protein synthesis. Proteins are having a crucial role in all the cellular processes of human metabolism. The majority of human diseases are the result of inappropriate protein production or disordered protein performance [2]. To inhibit the production of disease-causing proteins, antisense technology based drugs would be designed to bind sequence specifically to concerned mRNA. A wide spectrum of diseases including infectious, inflammatory, cancer, and cardiovascular diseases can be treated by specially designed and synthesized antisense drugs. Antisense oligonucleotides (ASOs) are short chemically modified oligonucleotides that bind to their complementary mRNA by Watson-Crick base-pairing and modulate its function [3]. The ASOs can be more effective and less toxic in targeting any disease because they can have the capacity to bind sequence specifically to the target mRNA. Single-stranded DNA based oligonucleotides activate the RNase H antisense mechanism [4]. The RNase H1 enzyme is ubiquitously expressed which cleaves the RNA strand selectively from the RNA/DNA hetero duplex [5]. The investigation of therapeutic applications of antisense oligonucleotides against various diseases is moving in a fast pace [6]. There are six antisense technology based drugs namely Fomivirsen [7,8], Pegaptanib [6], Mipomersen [9], Eteplirsen [9], Defibrotide [9] and Nusinersen [9] that have been approved by the FDA from 1998 to 2016. It has been observed in recent years that there is a rapid increase in the number of antisense molecules entering into phase III clinical trials [8]. In 1970, first time Zamecnik and Stephenson proposed ASOs as therapeutic agents [10]. The standard unmodified nucleic acids have confined stability in biological media and undergo rapid degradation by nucleases [11,12]. Chemical modifications would be required to protect the oligonucleotides from the cellular nucleases, enhance the stability, improve binding affinity for the target RNA, and improve pharmacokinetic properties

in animals to elicit a functional antisense response. The protective modifications could be introduced at three different sites on the nucleotide [13]. The nitrogen base can be altered or changes in phosphate backbone can be made for DNA and RNA nucleotides. Apart from these two, in RNA nucleotides, the 2' hydroxyl group also can be modified. These modifications are categorized into three generations.

First generation antisense modifications are majorly backbone based modifications like phosphorothioates, methyl phosphonates, and phosphoramidates modifications. One of the non-bridged oxygen attached to the phosphate is replaced by sulfur in phosphodiester backbone of nucleotide (Phosphorothioates), a methyl group (methylphosphonates) and amines (phosphoramidates). In all these backbone modifications, phosphorothioates (PS) are successful and are widely used for gene-silencing because of their resistance against nucleases and the ability to induce the RNase H functions [14]. However, the binding affinity to the target sequences, specificity and cellular uptake profiles of phosphorothioates are less satisfactory [15]. The issues raised with first generation antisense modifications are solved up to some extent with second generation modifications that are majorly based on sugar based modifications. The 2'-O-methyl (OMe) and 2'-O-methoxyethyl (MOE) are well explored and important members of second generation modifications. These sugar based modifications OMe and MOE can be further combined with the phosphorothioate backbone linkage [16]. It was reported that the antisense oligonucleotides having the 2'-O-methoxyethyl modification are less toxic than phosphorothioate antisense oligonucleotides and also shows enhanced affinity towards their complementary RNAs and improved pharmacokinetic properties [17,18]. To improve the thermal stabilities of the antisense oligomer bound to either complementary DNA or RNA, several nucleic acid analogs have been studied and developed. The third generation modifications are peptide nucleic acids (PNAs), locked nucleic acids (LNAs) or bridged nucleic acids, hexitol nucleic acids, and morpholino oligonucleotides to name a few [19-22].

PNAs are oligonucleotide analogues in which the sugar-phosphate backbone has been completely replaced by pseudo peptide linkages. This modification offers increased stability and favourable hybridization kinetics. However, these constructs have problems of solubility and delivery difficulties and they can't activate the RNase H cleavage mechanism [23]. LNAs are the most promising third generation modifications. LNA nucleotides are a class of nucleic acid analogues in which the ribose ring is locked by a methylene bridge connecting the 2'-O atom and the 4'-C atom. The LNAs show increased thermodynamic stability and improved nucleic acid recognition [24,25]. Apart from these generations of modifications, several nucleobase based modifications, backbone based modifications, furanose sugar based modifications, six-membered ring analogues, bicyclo and tricyclo modifications, constrained nucleic acids, etc. have been designed by medicinal chemists [26]. The novel "chimera" modifications in which more than one modification has been made to improve the nuclease resistance and target binding affinity as well. For example, single nucleotide modification can have phosphorothioate backbone modification and MOE sugar based modification. Another advanced strategy being explored recently is the concept of gapmer design. Gapmer is a designed antisense oligomer strand in which both the ends (2-5nt) have particular modification for example MOE or LNA to increase binding affinity, improve pharmacokinetic properties and in the middle region phosphorothioate modification to increase nuclease resistance and activate RNase H activity [27-31].

The importance of antisense technology and other oligonucleotide based therapeutics is continuously increasing. Several studies are being carried out by various labs to design novel drugs against some critical diseases. There is a strong need to identify or design novel modifications to improve the performance of existing antisense oligonucleotides [32]. Various modified antisense oligonucleotides are used in targeting different regions of the mRNA of disease-causing genes. However, the structural information about these antisense modifications

is limited. The structural parameters calculated through computational methods at the quantum mechanical level or molecular dynamics level of the modifications are very useful for understanding the mechanism of action of these modifications and also help in designing novel modifications. The quantum chemical and molecular dynamics simulations studies of some of the antisense modifications have been reported in the literature [33-36]. Many antisense molecules are available in the market as FDA approved drugs, where PS, MOE, LNA modified molecules are being used. There is tremendous scope to develop new modifications that can have a strong binding with DNA/RNA and less toxicity. Modification of nucleic acids can affect chemical stability, pairing, conformation. In the present study, five new antisense modifications have been proposed. The structures of these proposed modifications are designed based on the LNA structure. LNA is having strong binding property but higher toxicity compared to MOE [37-40]. The proposed modifications are designed by changing various atoms or by adding different oxy or nitro groups to the LNA. The chemical modification is key to improving pairing affinity, metabolic stability, and cellular uptake of RNA and absolutely essential in the discovery and development of highly active oligonucleotide / RNA based therapeutic candidates. To enhance the potency and reduce the potential toxicity of antisense oligomers, numerous chemical modification geometries have been established and tested. The basic idea behind proposing these modifications is to increase the binding affinity as well as to reduce the toxic effects and improve the pharmacokinetic properties. The proposed modifications are designed in such a way that they share the structural components from LNA, MOE and other electronegative groups / bulky groups attached to methylene bridge carbon of LNA. In the proposed modifications, the methylene bridge carbon is replaced with nitrogen in A1 and A2 modifications. In A2, extra methoxy group is added to the bridged nitrogen. For A3, A4 and A5 modifications, dimethyl amine group, amine group and dihydroxy amine group are added to the bridged carbon respectively. For proposed modifications, the methoxy group and

other amine groups are added because they are already well established as 2'-O sugar modifications and incorporation of these proved to show good pharmacokinetic properties [17,18,41]. In all proposed modifications, the electronegative groups and other small chemical groups were added to the basic LNA structure to improve the pharmacokinetics and reduce the toxicity without disturbing the strong binding nature of LNA. All these proposed modifications are studied thoroughly through quantum calculations. The 2D structural representation of all the modifications used in the current study is given in Figure 1. The study also aims to identify the stable conformations of the antisense molecules selected from the literature and the proposed novel modifications. Comparison has been made in terms of quantum chemical descriptors which helps to evaluate the different standard antisense modifications and the proposed new modifications.

The conformational search and optimization of all modifications at the monomer level was carried out and identified the most stable conformation for each modification by using in-house developed conformation generation and optimization tool TANGO [42]. The quantum chemical calculations have been carried out for the most stable conformations of all the modifications and derived various quantum chemical descriptors. This study may help in understanding the structural and functional significance of these novel antisense modifications in exhibiting lower toxicity with higher binding affinity and increased potency. Structural insight and information regarding various quantum chemical properties could be useful in their functional understanding, which may guide in the design of better modifications.



## Methods

The starting structures of all the modifications were generated by considering the crystal structures of respective standard modifications. Other modifications were built using molecular building and visualization software Gaussview [43]. The 2D structures of all the modifications are given in Figure 1 and the modification details are given in Table 1. The proposed novel modifications have been labelled as A1 to A5 in Figure 1. The complete methodology can be divided into three sections, conformational search and identifying most stable conformation, running quantum mechanical calculations in various levels and quantum chemical descriptor derivation.

### **Conformational search and identification of most stable conformation for each modification:**

The conformational search was done using an in-house developed TANGO tool [42]. The TANGO tool performs well on parallel HPC clusters with respect to scalability that reduces the time taken for calculations. The information about non-cyclic torsion angles of molecules was provided to TANGO along with the mol2 file and torsion rotation value of our choice and necessary keywords. For the modifications DC, DCS, RC, RCS, LCC, LCS, five torsion angles were identified for conformation generation. For the modifications CME & CMS, eight torsion angles were identified. For A1, five torsion angles, and for A2, A3, A4, A5 six torsion angles were identified for conformation generation. TANGO generates all the possible conformations based on the torsion rotation value given and the number of torsion angles selected for that molecule. The rotational angle was given as 30°, so that single torsion rotation can generate 12 conformers. The conformers generated for any modification will increase exponentially with an increase in the number of torsion angles. For example, DC modification is having 5 torsion angles and the angle rotation is 30°, so it generates  $12^5$  (248832) conformations. Once conformation generation is over, it calculates the MOPAC [44] based energies for all the

conformations based on semi-empirical methods like PM6. All conformations were sorted based on electronic energy. All the conformations were segregated based on their energy vs RMSD of the conformation compared to the lowest energy conformation. The most stable and alternate stable conformational ensembles were plotted. The sugar puckering calculations have also been carried out for the most stable conformations of each modification. The same multidimensional conformational search was carried out for all the modifications listed in Table 1 and the most stable conformations were identified for each modification and selected for further quantum calculations.

### **Running quantum mechanical calculations:**

The most stable conformations of all the modifications were taken for quantum calculations. Full geometry optimization and frequency calculations were done on these structures using density functional theory (DFT) method using the B3LYP/6-31G(d,p) basis set [45] in gas phase. Again, the DFT single point calculations were carried out using B3LYP/6-311G(d,p) basis set [46,47] for the most stable conformations of all the modifications by adding water as a solvent using polarizable continuum solvation model of Gaussian. For the optimization and frequency calculations Gaussian03 [48] was used and for single point calculation with solvent models, Gaussian09 [49] was used.

### **Derivation of quantum chemical reactivity descriptors:**

The global reactivity descriptors [50,51] like highest occupied molecular orbital (HOMO)-lowest unoccupied molecular orbital (LUMO) gap [52], ionization potential (I), electron affinity (A), global hardness ( $\eta$ ), global softness (S), chemical potential ( $\mu$ ), electronegativity ( $\chi$ ), electrophilicity index ( $\omega$ ) were calculated and compared for the optimized geometries of most stable conformations of all modifications [53-57]. The local descriptors like local softness, local electrophilicity, and local electrophilic attack were also derived for all these geometries.

$$\text{Ionization potential (I)} = -E_{\text{HOMO}} \quad (1)$$

$$\text{Electron affinity (A)} = -E_{\text{LUMO}} \quad (2)$$

$$\text{Global hardness } (\eta) = (I-A)/2 \quad (3)$$

$$\text{Global softness (S)} = 1/2\eta \quad (4)$$

$$\text{Chemical potential } (\mu) = - (I+A)/2 \quad (5)$$

$$\text{Electronegativity } (\chi) = (I+A)/2 \quad (6)$$

$$\text{Electrophilicity } (\omega) = \mu^2/2\eta \quad (7)$$

Local reactivity descriptors

Fukui function (FF) [58-60] provides information about the local reactivity site within the molecule and it provides a way for understanding of chemical reactions. These values represent the qualitative descriptors of reactivity of various atoms in the molecule. Fukui functions have been done with the basis of B3LYP/6-311G(d,p) level of theory for electrophilic and nucleophilic attacks. With the help of Mulliken atomic charges of cationic and anionic states, local Fukui functions ( $f_k^+$ ,  $f_k^-$ ) [58-60], local softness values ( $s_k^+$ ,  $s_k^-$ ) [61,62], and local electrophilicity indices ( $\omega_k^+$ ,  $\omega_k^-$ ) [55-57] have been calculated using the following equation:

$$f_k^+ = [q(N+1) - q(N)] \text{ for nucleophilic attack} \quad (8)$$

$$f_k^- = [q(N) - q(N-1)] \text{ for electrophilic attack} \quad (9)$$

$$f_k^0 = 1/2 [q(N+1) + q(N-1)] \text{ for radical attack} \quad (10)$$

Local softness values and electrophilicity indices were calculated using

$$s_k^+ = S f_k^+, \quad s_k^- = S f_k^-$$

$$\omega_k^+ = \omega f_k^+, \quad \omega_k^- = \omega f_k^-,$$

The + and – signs show nucleophilic and electrophilic attack, respectively. Where  $q(N)$  is the charge on the  $k$ th atom for neutral molecule while  $q(N+1)$  and  $q(N-1)$  are the same for its anionic and cationic species, respectively.

## Results and Discussions

### Most Stable Conformation and allowed flexibility of each modification:

All the conformations generated from the TANGO tool were sorted based on the MOPAC energies. The most stable conformation which has the lowest energy compared to the other conformations was identified and all the torsion angles measured for the same. The MOPAC energies and the torsion angles of most stable conformations for all the selected antisense modifications are given in Table 2. The marking of torsion angles ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ , and  $\zeta$ ) are shown in Figure 2. The start structure and the most stable conformation of each modification are shown in Figure 3a and 3b. The number of torsion angles rotated, the number of conformations generated for each modification and the number of conformations found within the range of 10 Kcal from the lowest conformation of respective modifications are given in Supplementary Table S1. The ensembles of conformations having MOPAC energy below 10 Kcal/mol were plotted by considering the energy and RMSD as shown in Figure 4. This plot captures the stable regions and flexibility allowed with each modification. It was observed that the highly populated cluster was not close to the lowest energy for the majority of modifications. Some modifications have not even shown any highly populated clusters. CME modification has shown the highest population of clusters among all the modifications due to the presence of an extra 2-O methoxyethyl free chain and thus a huge number of conformations were generated having energy within 10 Kcal/mol from the lowest energy conformation. The antisense modifications are chosen from literature like the DCS has been showing higher values of energy whereas CME is showing lower values of energy in their respective clusters. The modification A1 & A5 were exhibiting populated clusters but A2, A3, and A4 exhibiting restricted flexibility of the modifications. It was also observed from the sugar puckering of each modification that the DC and DCS were showing C2' endo puckering which is a property of B-form helix generally observed in DNA. The modifications RC, RCS, LCC, LCS, CME,

and CMS were showing C3' endo puckering, a property of A-form helix in nucleic acids. All the proposed novel modifications namely A1, A2, A3, A4 and A5 were preferring C3' endo puckering.

### **Geometry Optimization:**

The structure of a molecule contains the features responsible for its physical, chemical and biological properties. The variations in similar kinds of structures can be correlated with changes in descriptors that reflect their molecular properties. The computational analysis of the molecular geometry of a chemical molecule plays a crucial role in determining the structure-activity relationship. The molecular geometry generally described by the position of atoms in space, bond lengths of two joined atoms, bond angles of three connected atoms, and torsion angles (dihedral angles) of three consecutive bonds. The molecular geometries can be determined by the quantum mechanical behaviour of the electrons computed by ab-initio quantum chemistry methods to high accuracy. Molecular geometry represents the three-dimensional arrangement of the atoms that determines several properties of a substance including its reactivity, polarity, phase of matter, colour, magnetism, and biological activity. The optimization of geometry for the selected molecules has been achieved by energy minimization using DFT at the B3LYP level, employing the basis set 6-31G(d,p). The parameters like electronic energy, dipole moment, polarizability, thermal energy, and heat capacity have been given in Table 3.

The electronic energy trend observed for the standard modifications is DC>RC>LCC>CME. It was also observed that the addition of phosphorothioate to any modified nucleotide lowered the electronic energy. In the proposed modifications the electronic energy trend is A1>A2>A3>A4>A5. Some of the proposed modifications were also showing lower energy values than the standard modifications. A dipole moment is a measurement of the separation of two opposite electrical charges. Even though the total charge on a molecule is zero, the

nature of chemical bonds is such that the positive and negative charges do not completely overlap in most molecules. Such molecules are said to be polar because they possess a permanent dipole moment. The magnitude of the dipole moment induced is a measure of the polarizability of the molecules. The highest dipole moment can be seen in CMS modification compared to all others. Polarizability allows us to better understand the interactions between nonpolar atoms and molecules and other electrically charged species, such as ions or polar molecules with dipole moments. Neutral nonpolar species have spherically symmetric arrangements of electrons in their electron clouds. When in the presence of an electric field, their electron clouds can be distorted. The ease of this distortion is defined as the polarizability of the atom or molecule. The created distortion of the electron cloud causes the originally nonpolar molecule or atom to acquire a dipole moment. It has been suggested that polarizability is related to reactivity [63]. The polarizability is a measure for the change in the charge distribution within a molecule induced by an electric field. Thus, a system with low polarizability is supposed to be more stable, *i.e.* less reactive [64]. According to our calculations, DC, RC, LCC, and A1 were the molecules with the smallest values for the polarizability, with values 173, 174, 181 and 178 a.u., respectively. The addition of phosphorothioate slightly increased the polarizability of the molecules. Thermal energy is also known as random or internal kinetic energy, due to the random motion of molecules in a system. Kinetic energy is seen in three forms namely vibrational, rotational and translational. Thermal energy is directly proportional to the temperature within a given system. The heat capacity is the measurable physical quantity that characterizes the amount of heat required to change a molecule's temperature by a unit amount. Thermal energy is decreased and heat capacity is slightly increased with the addition of phosphorothioate for the modified molecules.

#### **Global reactivity descriptors:**

The frontier orbitals, HOMO and LUMO are the most important in a molecule. These orbitals determine the way how the molecule interacts with other species and gives information about reactivity or stability of specific regions of the molecule. The energy of HOMO characterizes electron donating ability of a molecule while LUMO energy determines the ability to accept an electron. Therefore, higher values of  $E_{\text{HOMO}}$  indicate a better tendency towards the donation of an electron. As can be seen in Table 4 and Table 5, the molecules CME, CMS, RCS & A2 have high LUMO energies, hence they can accept electrons while molecules DC, DCS, LCC & A3 have the highest HOMO energies that allow them to be the best electron donors. Frontier molecular orbitals (FMOs), HOMO and LUMO plot for all the modifications were shown in Figure 5. The values of the calculated quantum chemical parameters such as the energy of highest occupied molecular orbital ( $E_{\text{HOMO}}$ ), the energy of the lowest unoccupied molecular orbital ( $E_{\text{LUMO}}$ ), energy gap ( $\Delta E_{\text{Gap}}$ ), ionization potential (I), electron affinity (A), global hardness ( $\eta$ ), global softness (S), chemical potential ( $\mu$ ), and electrophilicity index ( $\omega$ ) were presented in Table 4 and Table 5. The highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) are very popular quantum chemical parameters. The similar kind of quantum calculations and reactivity descriptor derivation approach has been used by several groups for different kind of molecules like nucleobases, small chemical compounds and their derivatives [65-69]. We have used this strategy for studying antisense modifications at the monomer level. The FMOs are important in determining molecular reactivity and the ability of a molecule to absorb light. The vicinal orbitals of HOMO and LUMO play the same role as electron donor and electron acceptor respectively. The energies of HOMO and LUMO and their neighbouring orbitals were all negative, which indicates that the corresponding molecule is stable. The HOMO-LUMO energy gap ( $\Delta E_{\text{Gap}}$ ) is an important stability index. The HOMO-LUMO energy gap of any molecule reflects the chemical stability of the molecule. Through Koopman's theorem, the HOMO and LUMO energy values are

related to the ionization potential ( $I = -E_{\text{HOMO}}$ ) and electron affinities ( $A = -E_{\text{LUMO}}$ ).  $I$  and  $A$  is calculated as the negative of energy Eigenvalues of HOMO and LUMO respectively.

The energies of the HOMO and LUMO as well as the energy gap separating those are useful descriptors for the reactivity of nucleotide bases and base pairs. Due to the delocalization of the mobile electrons, these energies provide information on the stabilization of the molecules. The HOMO-LUMO energy gap is very important in determining the chemical reactivity of the molecule. The high value of the energy gap indicates that the molecule shows high chemical stability, while a small HOMO-LUMO gap means small excitation energies and hence easily reactive. Higher values of HOMO-LUMO gap was observed in RCS, RC, CMS, and A2 modifications. Ionization potential ( $I$ ), which is defined as the amount of energy needed to remove an electron from a molecule. High ionization energy indicates high stability and chemical inertness and small ionization energy indicates high reactivity of the atoms and molecules. Modifications DC, DCS, LCC, and A3 were showing lower ionization potential values, which indicate better electron donors. The electronic affinity ( $A$ ) is defined as the energy released when an electron is added to a neutral molecule. A molecule with high electron affinity ( $A$ ) values tends to take electrons easily. From Table 4 and Table 5, RCS, CME, CMS and A2 have been the reactive modifications.

This kind of information may be contained within the orbital energies of the HOMO and LUMO but, instead, it may be more useful to study the global reactivity descriptors. The global chemical reactivity descriptors, chemical potential ( $\mu$ ), absolute electronegativity ( $\chi$ ) and chemical hardness ( $\eta$ ), global softness ( $S$ ) and electrophilicity ( $\omega$ ) which were calculated from HOMO and LUMO energies obtained at the level of theory B3LYP/6-31G(d,p) for gas phase as given in Table 4 and B3LYP/6-311G(d,p) for solvent phase as given in Table 5. According to these parameters, the chemical reactivity varies with the structural configuration of molecules. The chemical potential  $\mu$  (eV) measures the escaping tendency of an electron and it



can be associated with the molecular electronegativity then, as  $\mu$  becomes more negative, it is more difficult to lose an electron but easier to gain one. As shown in Figure 6, Tables 4 and 5, modifications DC, DCS, LCC, LCS, and A3 have the least stability and more reactivity among all the modifications. Electronegativity ( $\chi$ ), representing the ability of molecules to attract electrons and is the negative of the chemical potential ( $\mu$ ) in Mulliken sense. The ( $\chi$ ) values displayed in Tables 4 and 5 show that modifications RCS and A2 have higher electronegativity values compared to other modifications. The global hardness ( $\eta$ ) and softness ( $S$ ) are useful concepts for understanding the behaviour of the chemical system. Softness ( $S$ ) is a property of molecules that measures the extent of chemical reactivity; it is the reciprocal of hardness. A hard molecule has a large energy gap and a soft molecule has a small energy gap. Therefore, soft molecules will be more polarizable than hard molecules. Theoretical calculations established that the modifications RCS and A5 have the highest hardness values, which indicates the hardest molecules. The modifications LCS and A3 have the highest softness values or can be stated as chemically more reactive molecules. Parr defined the electrophilicity index ( $\omega$ ), as a numerical value which is related to the stabilization of energy when the system acquires an electronic charge and serves as an indicator of the reactivity of a system towards nucleophiles. Electrophilicity gives an idea of the stabilization energy when the system gets saturated by electrons, which come from the external environment. This reactivity information shows if a molecule is capable of donating electrons. The lower values of ( $\omega$ ) indicate the presence of the nucleophilic character, while higher values indicate the presence of a good electrophilic character. Our results indicate that modifications DC, DCS, LCC, and A5 have lower values of ( $\omega$ ) so that these are good nucleophiles. However, modifications RCS, CMS, and A2 are good electrophiles.

#### **Local reactivity descriptors:**

In a chemical reaction, a change in the number of electrons involves the addition or subtraction of at least one electron in the frontier orbitals. Local reactivity descriptors are used to decide the relative reactivity of different atoms in the molecule. It is established that molecules tend to react where the value of descriptor is largest when attacked by soft reagent and where the value is smaller when attacked by the hard reagent. The use of descriptors for the site selectivity of the molecule for the nucleophilic and electrophilic attack has been made. Calculating Fukui functions (FF) helps us determine the active sites of a molecule, based on the electronic density changes experienced by it during a reaction. Fukui function provides information on the local site reactivity within the molecule and as such it provides a system for the understanding of chemical reactions. These values correspond to the qualitative descriptors of reactivity of different atoms in the molecule. Fukui functions for electrophilic and nucleophilic attacks have been made with the basis of B3LYP/6-311G (d,p) level of theory. With the help of Mulliken atomic charges of cationic and anionic states, local Fukui function ( $f_k^+$ ,  $f_k^-$ ), local softness values ( $s_k^+$ ,  $s_k^-$ ), and local electrophilicity indices ( $\omega_k^+$ ,  $\omega_k^-$ ) have been calculated.

Based on the Fukui function calculations at the DFT level, the most susceptible sites for nucleophilic, electrophilic and free radical attack for all the modifications are shown in Figure 7. Fukui functions, local softness values and local electrophilicity indices for selected atomic sites in different antisense modifications have been listed in Supplementary Table S2. Supplementary Table S2 shows that at the DFT level the most susceptible site to a nucleophilic, electrophilic and free radical attack for modification DC is the same and located on 10C. For modification DCS, the more susceptible site to nucleophilic attack is 6O, while 10C is the most susceptible site for electrophilic and free radical attacks. The most susceptible site to a nucleophilic, electrophilic, and free radical attack for modification RC is the same and located on 12C (Supplementary Table S3). For modification RCS, the more susceptible site to nucleophilic attack is 10C, while 12C is the most susceptible site for electrophilic and free

radical attacks. Modifications LCC and LCS (Supplementary Table S4) show that the most susceptible site to the nucleophilic, electrophilic, and free radical attacks. Supplementary Table S5 shows that the most susceptible active centers to a nucleophilic and free radical attack for CME modification are located on 10C while 11C is the most susceptible site for electrophilic attack. For the modification CMS, the more susceptible site to nucleophilic attack is 21O, while 11C is the most susceptible site for electrophilic and free radical attacks. The proposed novel modifications A1, A2, A4, and A5 (Supplementary Tables S6, S7, S8) show that the most susceptible site to nucleophilic, electrophilic and free radical attacks is 5C. Whereas for the A3 modification the more susceptible site to nucleophilic attack is 5C, while C18 is the most susceptible site for electrophilic and free radical attacks. It has been observed from the Fukui function results that the modifications DC, RC, LCC, LCS, A1, A2, A4, and A5 are showing single susceptible site (10C, 12C, 5C, 5C, 5C, 5C, 5C, 5C respectively) for all the three (nucleophilic, electrophilic and free radical) kinds of attacks. The modifications DCS, RCS, CMS, and A3 are showing single susceptible sites (10C, 12C, 11C, 18C respectively) for electrophilic and free radical attacks. Whereas CME modification showing a single susceptible site (10C) for the nucleophilic and free radical attack.

## Conclusions

In the present work, we have used DFT based reactivity descriptors to study the structure, stability and reactivity of selected antisense molecules along with the proposed novel antisense modifications. We have calculated the geometrical parameters and frontier orbital energies for different antisense modifications both in the gas phase and the solvent phase. One goal of the study was to identify the global reactivity descriptors that could be used in recognizing the chemically reactive and stable molecules, the second was to identify the local reactivity descriptors that could be used in determining the chemically more reactive sites in each modification. From the whole study and the results presented in this contribution, it has been demonstrated that the reactive sites of interaction of all the antisense modifications can be predicted by using DFT-based global reactivity descriptors as well as Fukui-function calculations. A comparison of global reactivity descriptors confirmed that LNA based modifications LCC, LCS, and A3 are most reactive modifications and prone to chemical reactions and may form stable duplexes when bound to complementary nucleotides, compared to other modifications. Theoretical results from reactivity descriptors show that 5C is a more reactive site for nucleophilic, electrophilic and free radical attacks in all LNA based modifications except A3 modification. In the novel proposed modifications A3 was showing similar kinds of properties with LCC and LCS. The A3 modification was also showing closer values of properties with CME and CMS which are claimed to be less toxic modifications. Therefore, A3 antisense modification may also strongly bind to the complementary nucleotides as LNA and may show reduced toxic effects. The molecular dynamics simulations study by incorporating these modifications in a duplex can throw more light and that will be the continuation of this work.

## **Conflict of Interest Statement**

The authors declare no conflict of interest.

## **Acknowledgments**

The authors acknowledges the Bioinformatics Resources and Applications Facility (BRAAF) of C-DAC, for providing the supercomputing facility to carryout quantum calculations of these molecules. Mallikarjunachari Uppuladinne acknowledges his colleague Ms. Shruti Koulgi and other team members for helping in graphics and technical discussions. Dikshita Dowerah acknowledges Tezpur University supercomputing facility PARAMTEZ for carrying out the related quantum chemical calculations. Uddhavesh Sonavane and Ramesh Deka acknowledges the Department of Biotechnology, Government of India, for financial support of this work (DBT Project No. BT/PR16182/NER/95/92/2015). Authors also acknowledges National Supercomputing Mission (NSM) & Ministry of Electronics and Information Technology (MeitY) for the support to carrying out R&D work at C-DAC.

## References

1. S. T. Crooke, *Antisense Drug Technology: Principles, Strategies and Applications*, 2nd ed.; CRC Press: Boca Raton, FL, 2007; p 1–799.
2. J. H. P. Chan, S. Lim, W. S. F. Wong, *Clin. Exp. Pharmacol. Physiol.* 2006, 33, 533.
3. C. F. Bennett, E. E. Swayze, *Annu. Rev. Pharmacol. Toxicol.* 2010, 50, 259–293.
4. W. F. Lima, H. Wu, S. T. Crooke, *Methods Enzymol.* 2001, 341, 430–440.
5. S. M. Cerritelli, R. J. Crouch, *FEBS J.* 2009, 276, 1494–1505.
6. U. Galderisi, A. Cascino, A. Giordano, *J. Cell. Physiol.* 1999, 181, 251.
7. Vitravene Study Group, *American Journal of Ophthalmology*. 2002, 133 (4), 467–474.
8. H. L. Calvez, M. Yu, F. Fand, *Virol. J.* 2004, 1, 12.
9. Stein CA, Castanotto D, *Mol Ther.* 2017, 25(5), 1069–1075.
10. P. C. Zamecnik, M. L. Stephenson, *Proc. Natl. Acad. Sci. USA* 1978, 75, 280.
11. B. Tavitian, S. Terrazzino, B. Kuhnast, S. Marzabal, O. Stettler, F. Dolle, J. R. Deverre, A. Jobert, F. Hinnen, B. Bendriem, C. Crouzel, L. Di Giamberardino, *Nat. Med.* 1998, 4, 467–471.
12. R. S. Geary, *Expert Opin. Drug Metab. Toxicol.* 2009, 5, 381–391.
13. J. Kurreck, *Eur. J. Biochem.* 2003, 270, 1628.
14. J. M. Campbell, T. A. Bacon, E. Wickstrom, *J. Biochem. Biophys. Methods* 1990, 20, 259.
15. X. Chen, N. Dudgeon, L. Shen, J. H. Wang, *Drug Discov. Today* 2005, 10, 587.
16. M. Manoharan, *Biochim. Biophys. Acta* 1999, 1489, 117.
17. Bennett CF, Baker BF, et al. (2017) Pharmacology of antisense drugs. *Annu Rev Pharmacol Toxicol*, 57:81–105.

18. Shen, W., De Hoyos, C.L., Migawa, M.T. *et al. Nat Biotechnol* **37**, 640–650 (2019).  
<https://doi.org/10.1038/s41587-019-0106-2>.
19. B. P. Monia, E. A. Lesnik, C. Gonzalez, W. F. Lima, D. McGee, C. J. Guinasso, A. M. Kawasaki, P. D. Cook, S. M. Freier, *J. Biol. Chem.* 1993, 268, 14514–14522.
20. B. Hyrup, P. E. Nielson, *Bioorg. Med. Chem.* 1996, 4, 5.
21. A. N. Elayadi, D. R. Corey, *Curr. Opin. Investig. Drugs.* 2001, 2, 558.
22. R. Declercq, A. Van Aerschot, J. R. Read, P. Herdewijn, L. van Meervelt, *J. Am. Chem. Soc.* 2002, 124, 928.
23. S. Pensato, M. Saviano, A. Romanelli, *Expert Opin. Biol. Ther.* 2007, 7, 1219.
24. R. N. Veedu, J. Wengel, *RNA Biol.* 2009, 6, 321.
25. C. A. Stein, J. B. Hansen, J. Lai, S. Wu, A. Voskresenskiy, A. Hqg, J. Worm, M. Hedtjarn, N. Souleimanian, P. Miller, H. S. Soifer, D. Castanotto, L. Benimetskaya, H. Qrum, T. Koch, *Nucleic Acids Res.* 2010, 38, e3.
26. W. B. Wan and P. P. Seth, *J. Med Chem.* 2016, 59, 9645-9667.
27. Shibahara, S.; Mukai, S.; Nishihara, T.; Inoue, H.; Ohtsuka, E.; Morisawa, H. Site-Directed Cleavage of RNA. *Nucleic Acids Res.* 1987, 15, 4403–4415.
28. Monia, B. P.; Lesnik, E. A.; Gonzalez, C.; Lima, W. F.; McGee, D.; Guinasso, C. J.; Kawasaki, A. M.; Cook, P. D.; Freier, S. M. *J. Biol. Chem.* 1993, 268, 14514–14522.
29. Henry, S. P.; Geary, R. S.; Yu, R.; Levin, A. A. *Curr. Opin. Investig. Drugs.* 2001, 2, 1444–1449.
30. Yamamoto, T.; Harada-Shiba, M.; Nakatani, M.; Wada, S.; Yasuhara, H.; Narukawa, K.; Sasaki, K.; Shibata, M.-A.; Torigoe, H.; Yamaoka, T.; Imanishi, T.; Obika, S. *Mol. Ther.–Nucleic Acids.* 2012, 1, e22.

31. Pedersen, L.; Hagedorn, P. H.; Lindholm, M. W.; Lindow, M. A. *Mol. Ther.–Nucleic Acids*. 2014, 3, e149.
32. Lundin, K. E., Gissberg, O. & Smith, C. I. *Hum Gene Ther*. 2015, 26, 475–485.
33. S. Sharma, U. B. Sonavane, R. R. Joshi, *Int. J. Quantum Chem*. 2008, 109, 890.
34. J. P. Malrieu, In *Semiempirical Methods of Electronic Structure Calculations, Part A: techniques*; G. A. Segal, Ed.; Plenum: New York, 1977; p 69.
35. M. V. Uppuladinne, V. Jani, U. B. Sonavane, R. R. Joshi, *Int. J. Quantum Chem*. 2013, 113(23), 2523-2533.
36. M. V. Uppuladinne, U. B. Sonavane, R.C. Deck, R. R. Joshi, *J. Biomol Str and Dyn*. 2019, 37(11), 2823-2836.
37. A. A. Koshkin, P. Nielsen, M. Meldgaard, V. K. Rajwanshi, S. K. Singh, J. Wengel, *J. Am. Chem. Soc*. 1998, 120, 13252–13253.
38. A. A. Koshkin, S. K. Singh, P. Nielsen, V. K. Rajwanshi, R. Kumar, M. Meldgaard, C. E. Olsen, J. Wengel, *Tetrahedron*. 1998, 54, 3607–3630.
39. S. Obika, D. Nanbu, Y. Hari, K. I. Morio, Y. In, T. Ishida, T. Imanishi, *Tetrahedron Lett*. 1997, 38, 8735–8738.
40. S. Obika, D. Nanbu, Y. Hari, J. I. Andoh, K. I. Morio, T. Doi, T. Imanishi, *Tetrahedron Lett*. 1998, 39, 5401–5404.
41. Khvorova A, Watts JK. *Nat Biotechnol*. 2017;35(3):238-248. doi:10.1038/nbt.3765.
42. V. Gavane, S. Koulgi, V. Jani, M. V. Uppuladinne, U. Sonavane, and R. Joshi, *J. Comp Chem*. 2019, 40(7), 900-909.
43. GaussView, Version 6.1, R. Dennington, T. A. Keith, and J. M. Millam, Semichem Inc., Shawnee Mission, KS, 2016.



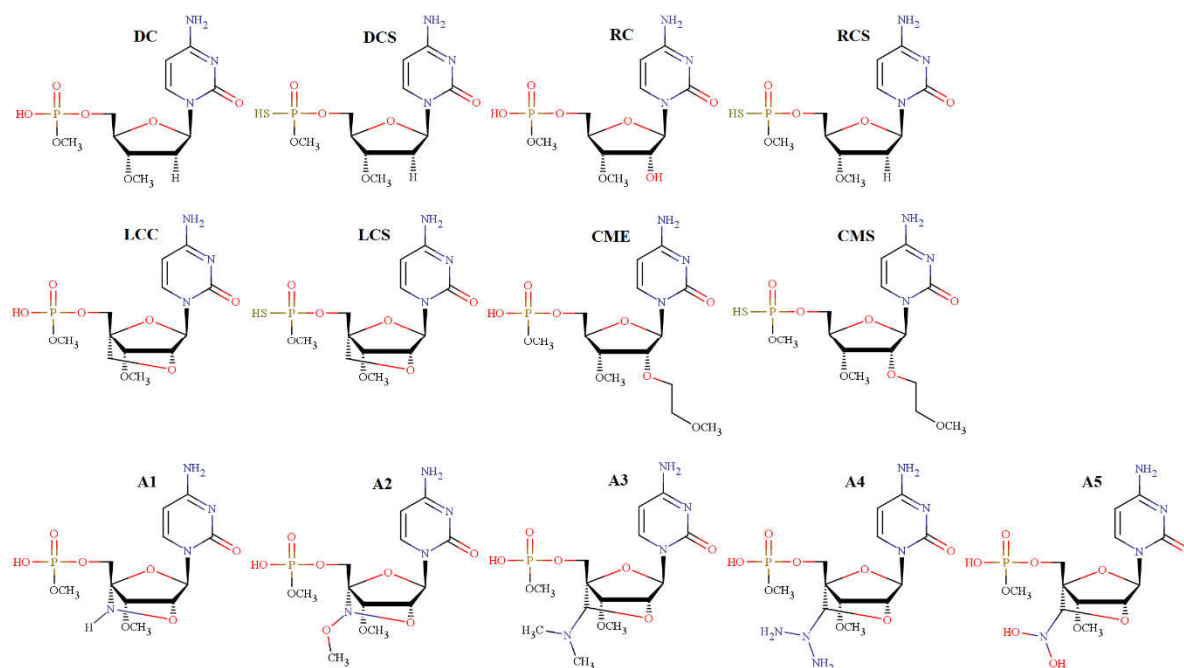
44. MOPAC2009, J. J. P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA, [HTTP://OpenMOPAC.net](http://OpenMOPAC.net), 2008.
45. R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.*, 1971. 54, 724.
46. G. A. Petersson, A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley, and J. Mantzaris, *J. Chem. Phys.*, 1988. 89, 2193-218.
47. K. Raghavachari, J. S. Binkley, R. Seeger, and J. A. Pople, *J. Chem. Phys.*, 1980. 72, 650-54.
48. 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. Millan, 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, V. Bakken, 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, 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 C. 02; Gaussian, Inc.: Wallingford, CT, 2004.
49. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F.

Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian 09, Revision A.02, Gaussian, Inc., Wallingford CT, 2016.

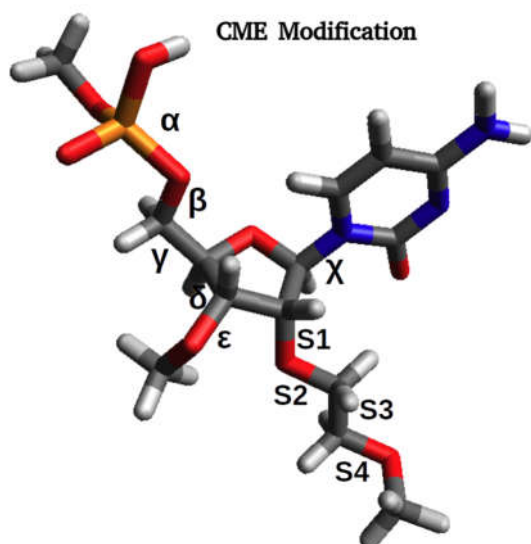
50. P. K. Chattaraj, U. Sarkar, D. R. Roy, *Chem. Rev.* 2006, 106, 2065.
51. K. K. Hazarika, N. C. Baruah, R. C. Deka, *Struct. Chem.* 2009, 20, 1079.
52. A. J. Cohen, P. Mori-Sanchez, W. T. Yang, *Phys Rev B.* 2008, 77, 115123.
53. R. G. Parr, L. V. Szentpály, and S. Liu, *J. Am. Chem. Soc.* 1999, 121(9), 1922–1924.
54. P. K. Chattaraj and S. Giri, *J. Phys Chem A*, 2007, 111(43), 11116–11121.
55. J. Padmanabhan, R. Parthasarathi, V. Subramanian, and P. K. Chattaraj, *J. Phys Chem A*, 2007, 111(7), 1358–1361.
56. P. W. Ayers and R. G. Parr, *J. Am. Chem. Soc.* 2000, 122(9), 2010–2018.
57. Liu S (2009) Chemical reactivity theory a density functional view, Chapt 13. CRC Press, Boca raton.
58. W. T. Yang, R. G. Parr, R. Pucci, *J Chem Phys*, 1984, 81, 2862–2863.
59. R. G. Parr, W. Yang, *J Am Chem Soc*, 1984, 106, 4049.
60. P. W. Ayers, M. Levy, *Theor Chem Accounts*, 2000, 103, 353–360.
61. W. Yang, R. G. Parr, *Proc Natl Acad Sci*, 1985, 82, 6723.

62. A. K. Chandra, M. T. Nguyen, (2008) Fukui function and local softness. In: Chattaraj PK (ed) Chemical reactivity theory: a density-functional view. Taylor and Francis, New York, pp 163–178.
63. P. Senet, in Chemical Reactivity Theory: A Density Functional (Ed: P. K. Chattaraj), CRC Press, Boca Raton, Florida 2009, Ch. 24, pp. 331–361.
64. J. Jerbi, M. Springborg, *J. Comput. Chem.* 2017, 38, 1049.
65. J. Jerbi, M. Springborg, *Int. J. Quantum Chem.* 2018, 118(11), e25538.
66. A. S. Gidado, A. Salihu, M. A. Shariff, *Bayero Journal of Pure and Applied Sciences.* 2017, 10 (1), 115-127.
67. A. Dwivedi, V. Baboo, A. Bajpai, *Journal of Theoretical Chemistry*, 2015.
68. S. Rajkhowa, I. Hussain, K. K. Hazarika, P. Sarmah, R. Chandra Deka, *Combinatorial chemistry & high throughput screening*, 2013. 16(8), 590-602.
69. K. K. Hazarika, N. C. Baruah, R. C. Deka, *Struct Chem*, 2009, 20(6), 1079-1085.

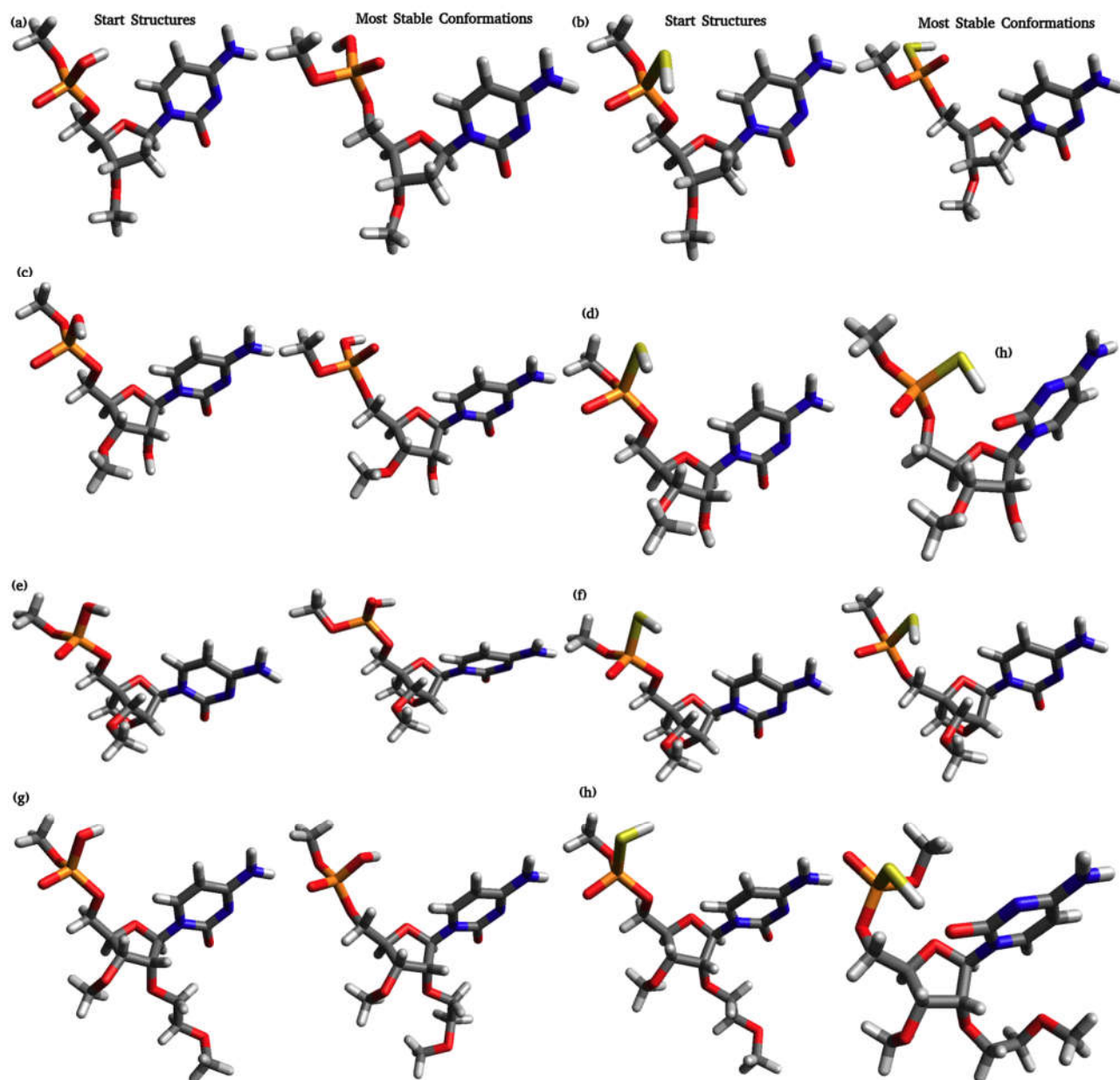
## Figures



**Figure 1:** The 2D structures of all the modifications selected for the study. Structures of DC, DCS, RC, RCS, LCC, LCS, CME, CMS, A1, A2, A3, A4 and A5 are given.

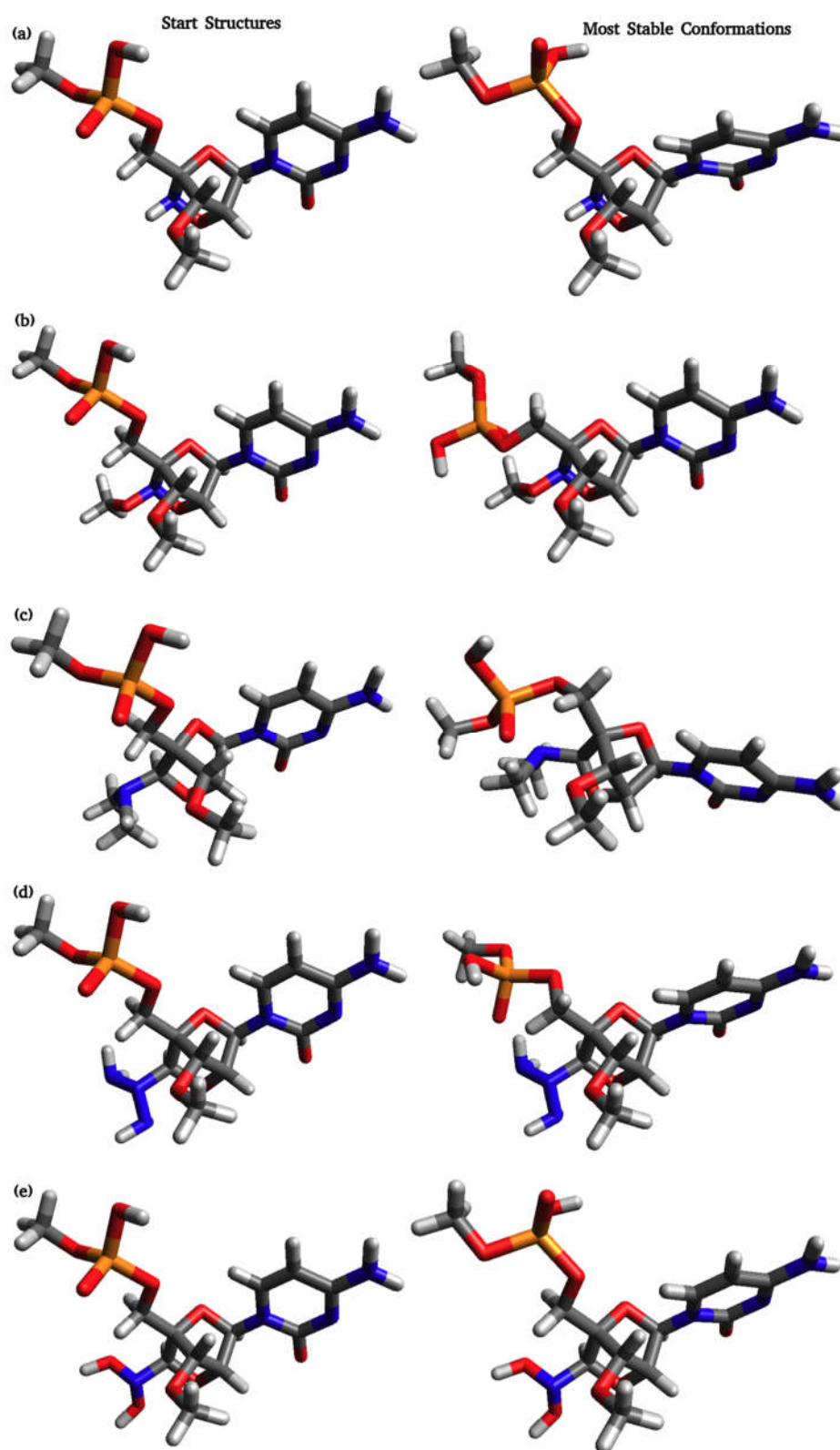


**Figure 2:** The CME modification structure marked with all the torsion angles.

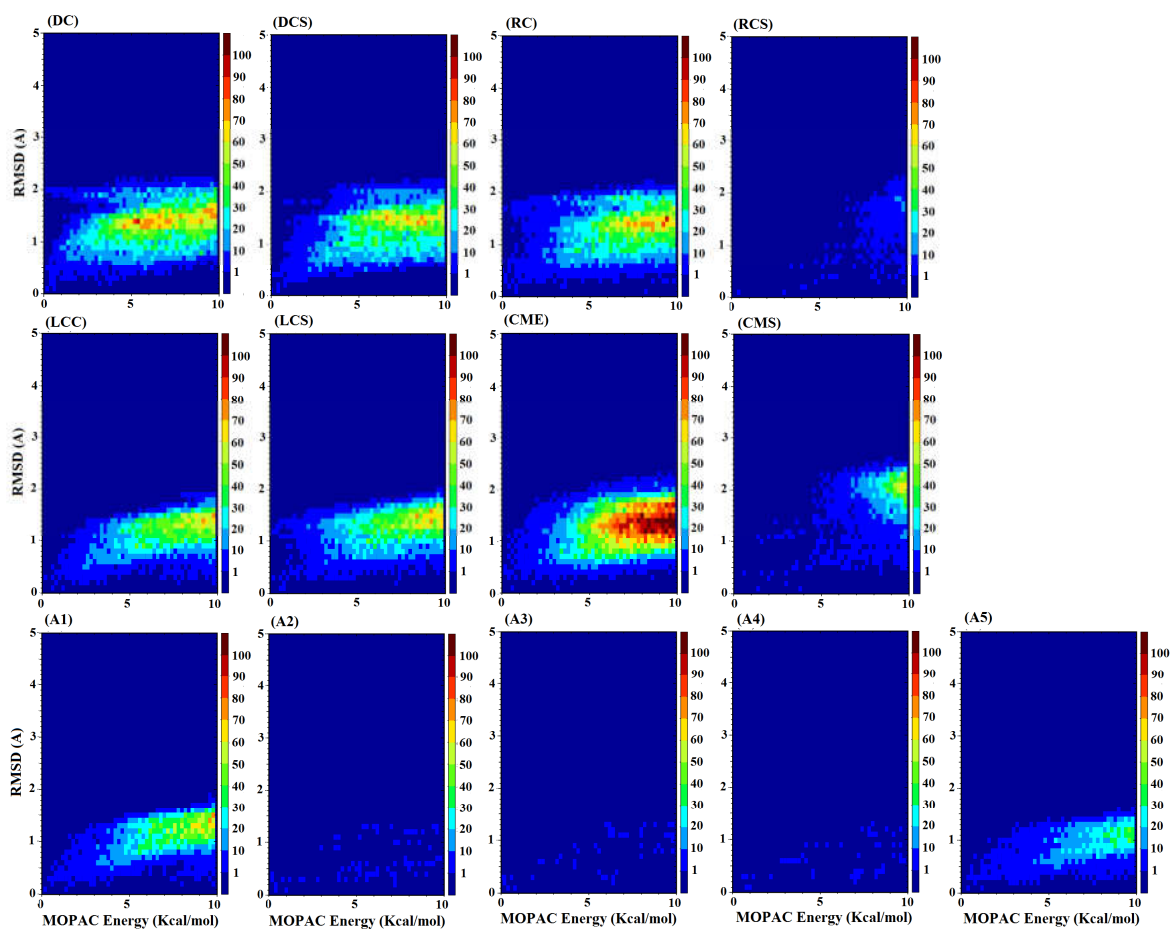


**Figure 3a:** The start structure and most stable conformation of respective structure are given.

(a) DC (b) DCS (c) RC (d) RCS (e) LCC (f) LCS (g) CME (h) CMS.

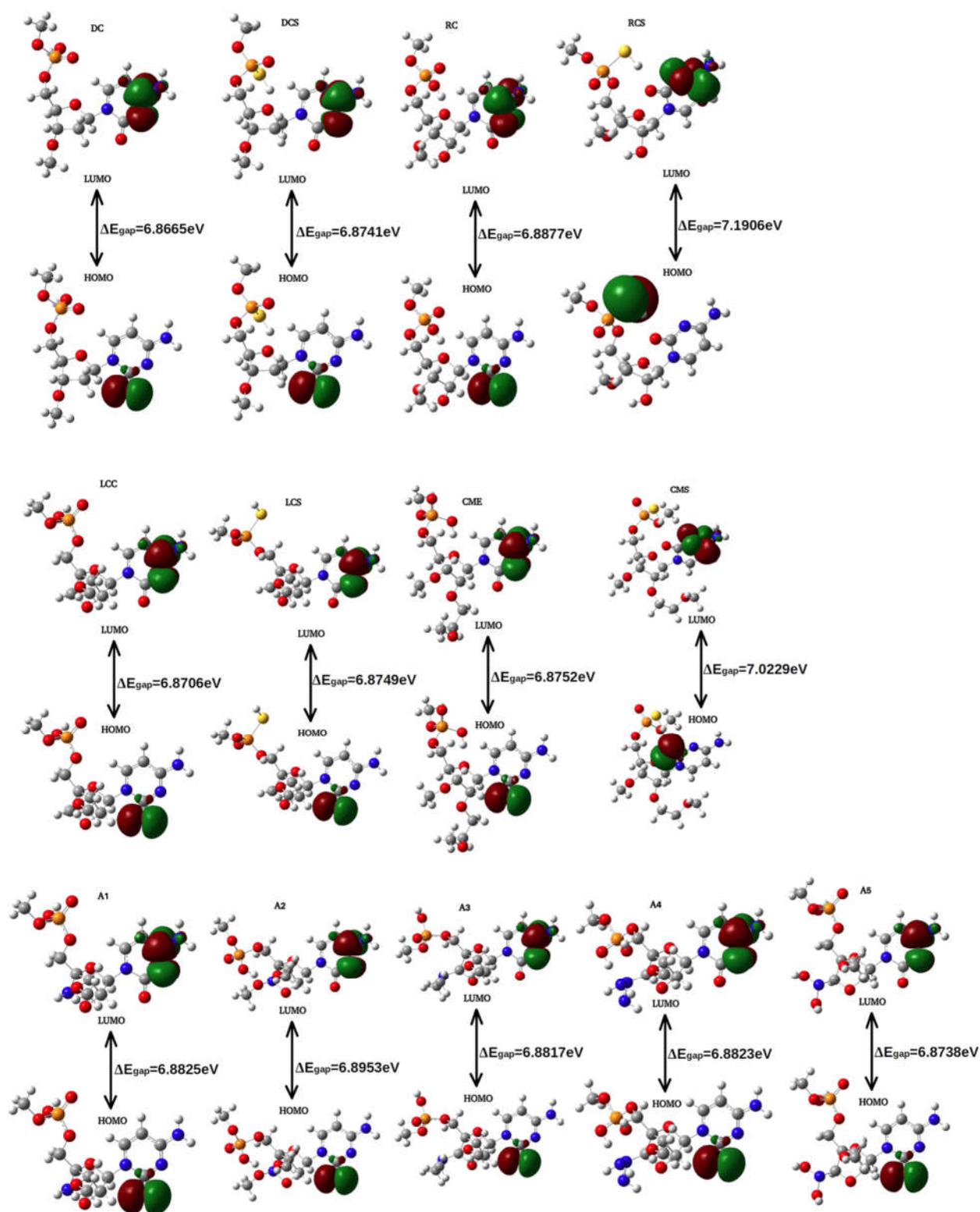


**Figure 3b:** The start structure and most stable conformation of respective structure are given.  
 (a) A1 (b) A2 (c) A3 (d) A4 (e) A5.

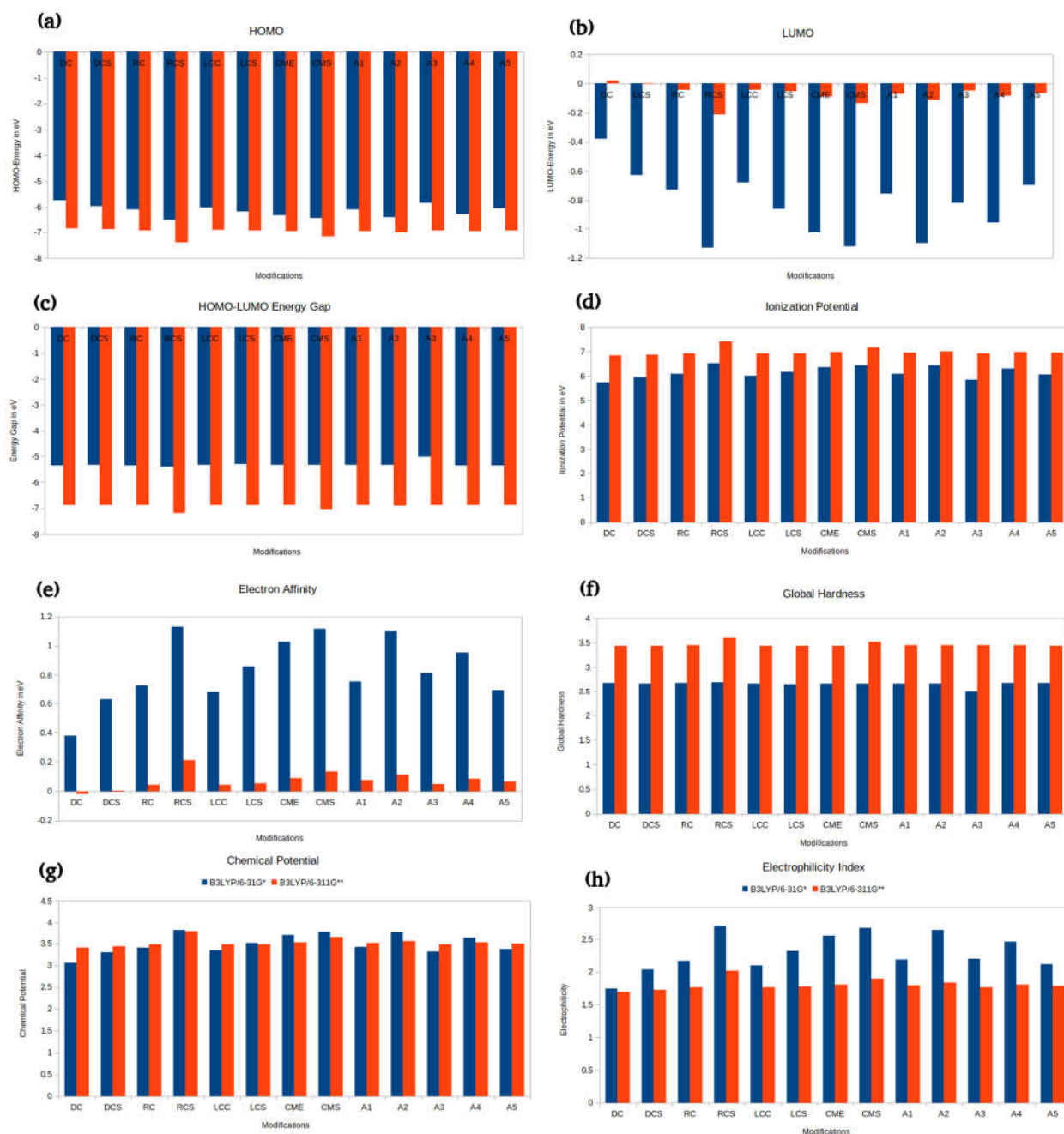


**Figure 4:** The ensembles of conformations generated through TANGO tool plotted considering MOPAC energy vs RMSD for all molecules.

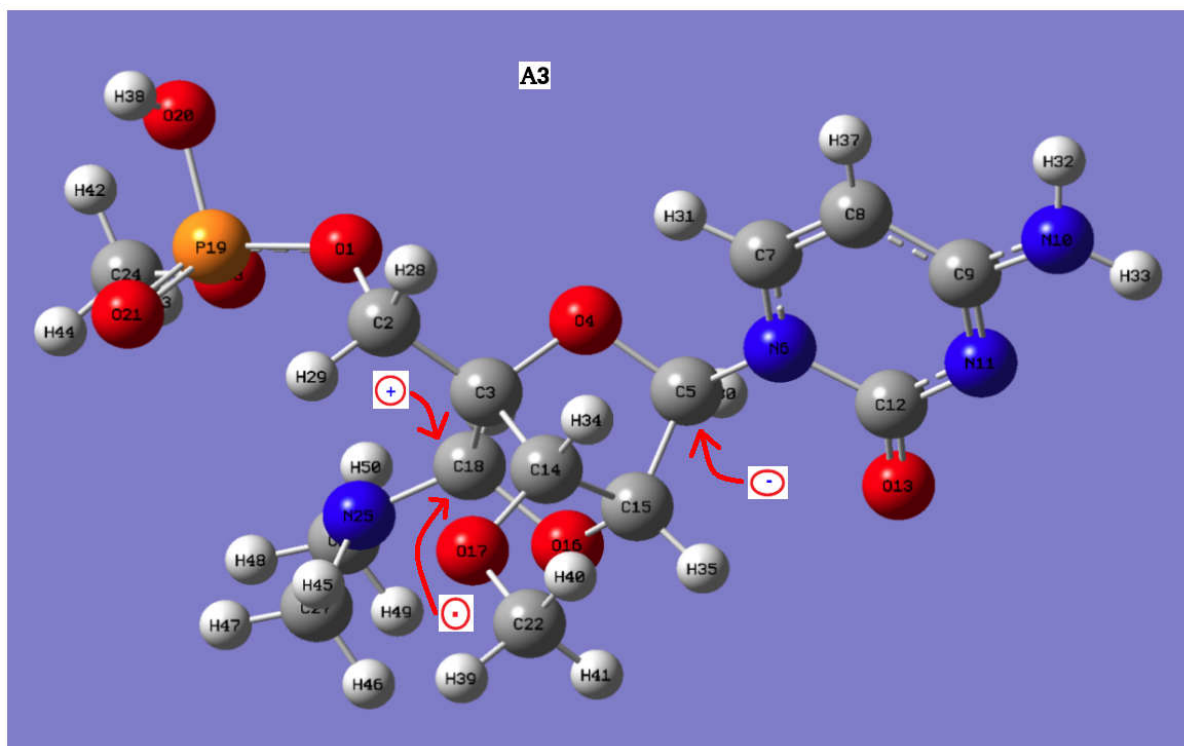




**Figure 5:** The HOMO and LUMO structures and HOMO-LUMO energy gap for all the modifications are given.



**Figure 6:** The calculated quantum chemical parameters for all the antisense modifications for gas phase using B3LYP/6-31G(d,p) method (blue) and for solvent phase using B3LYP/6-311G(d,p) method (red). (a) HOMO, (b) LUMO, (c) HOMO-LUMO energy gap, (d) Ionization potential, (e) Electron affinity, (f) Global hardness, (g) Chemical potential, (h) Electrophilicity index.



**Figure 7:** The most susceptible reactive sites for nucleophilic, electrophilic and free radical attacks marked for A3 modification.

## Tables

**Table 1:** The details of names and coding terms used of all modifications used in this work.

S. No.	Name Code	Base	Modification	Description
1	DC	Cytosine	No modification,	Deoxyribo Cytidine
2	DCS	Cytosine	Phosphorothioate (PS) modification	Deoxyribo Cytidine with PS modification
3	RC	Cytosine	No Modification,	Ribo Cytidine
4	RCS	Cytosine	Phosphorothioate (PS) modification	Ribo Cytidine with PS modification
5	LCC	Cytosine	Locked nucleic acid (LNA) modification	Cytidine with LNA modification
6	LCS	Cytosine	Locked nucleic acid (LNA) and phosphorothioate (PS) modification	Cytidine with LNA and PS modifications
7	CME	Cytosine	2'-O- methoxyethyl (MOE) modification	Ribo Cytidine with MOE modification
8	CMS	Cytosine	2'-O- methoxyethyl (MOE) and phosphorothioate (PS) modification	Ribo Cytidine with MOE and PS modifications
9	A1	Cytosine	N-linked LNA modification	Cytidine with N-linked LNA modifications
10	A2	Cytosine	Methoxy-N-linked LNA modification	Cytidine with methoxy-N-linked LNA modifications
11	A3	Cytosine	LNA linked to dimethyl amine modification	Cytidine with LNA linked to dimethyl amine modification
12	A4	Cytosine	LNA linked to N-diamide modification	Cytidine with LNA linked to N-diamide modification
13	A5	Cytosine	LNA linked to N-dihydroxy modification	Cytidine with LNA linked to N-dihydroxy modification

**Table 2:** The dihedral angles and MOPAC energies of most stable conformations of all modifications.

S. No.	Name	Alpha ( $\alpha$ )	Beta ( $\beta$ )	Gamma ( $\gamma$ )	Delta ( $\delta$ )	Epsilon ( $\epsilon$ )	Chi ( $\chi$ )	S1	S2	S3	S4	Energy (Kcal/mol)
1	DC	-56.87	-156.42	178.38	137.98	-151.78	-150.65					-97875.25
2	DCS	-56.37	-157.50	-151.91	138.40	-151.74	-151.05					-95371.76
3	RC	-65.66	-153.85	-176.09	84.17	-96.11	-143.53					-104619.98
4	RCS	-123.60	-124.81	124.33	84.39	-125.50	66.13					-104527.23
5	LCC	51.24	157.87	63.78	63.78	-166.75	-138.61					-107372.17
6	LCS	171.50	-142.58	-176.10	63.97	-167.18	-168.22					-107272.59
7	CME	-172.82	-152.69	-174.92	79.18	29.24	-134.54	93.72	-172.15	-93.97	97.51	-121758.93
8	CMS	98.05	-92.53	-174.55	79.48	30.46	47.06	124.96	-111.62	86.26	96.63	-121638.14
9	A1	50.08	157.40	63.57	63.75	-168.07	-139.36					-106966.52
10	A2	-43.76	-114.92	-86.64	64.54	-168.17	-170.02					-109241.72
11	A3	-131.55	96.15	-147.47	61.64	-166.85	-139.08					-109887.25
12	A4	-161.11	126.63	150.44	65.84	-166.30	-138.33					-112631.50
13	A5	46.17	155.53	63.62	65.63	-165.62	-138.70					-124924.12

**Table 3:** The electronic energy, dipole moment, polarizability, thermal energy and heat capacity of all the modifications.

S. No.	Name	Electronic Energy (hartree)	Dipole moment (Debye)	Polarizability (a.u)	Thermal Energy (Kcal/mol)	Heat Capacity cal/mol-kelvin
1	DC	-1462.2522	6.561359	173.311667	212.828	83.289
2	DCS	-1785.2066	4.707803	187.250333	210.318	84.867
3	RC	-1537.4683	4.780571	174.736667	216.552	86.529
4	RCS	-1860.4220	4.629668	192.669667	214.053	87.869
5	LCC	-1575.5542	6.479880	181.739000	220.427	88.037
6	LCS	-1898.5040	6.349720	197.898333	217.775	89.822
7	CME	-1730.5920	1.959740	213.966667	276.589	104.445
8	CMS	-2053.5371	11.858483	224.859000	273.966	106.130
9	A1	-1591.5430	6.655189	178.414333	212.655	87.657
10	A2	-1706.0300	7.307183	196.298667	233.822	96.679
11	A3	-1709.5141	5.487077	212.611000	269.087	103.306
12	A4	-1741.545033	4.862393	206.580667	254.958	100.474
13	A5	-1781.237451	6.384333	198.818000	237.584	101.538

**Table 4:** The global quantum chemical descriptors HOMO, LUMO, HOMO-LUMO Gap, Ionization Potential, Electron Affinity, Global Hardness, Global Softness, Chemical Potential and Electrophilicity Index for gas phase calculations by B3LYP/6-31G\*.

S. No.	Name	HOMO (eV)	LUMO (eV)	HOMO-LUMO Gap (eV)	Ionization Potential $I = -E_{\text{HOMO}}$	Electron Affinity $A = -E_{\text{LUMO}}$	Global Hardness $\eta = (I-A)/2$	Global Softness $S = 1/2\eta$	Chemical Potential $\mu = -(I+A)/2$	Electro-negativity $\chi = -\mu$	Electrophilicity Index $\omega = \mu^2/2\eta$
1	DC	-5.7386	-0.3785	<b>5.3601</b>	5.7386	0.3785	2.6800	0.1865	<b>-3.0585</b>	3.0585	<b>1.7452</b>
2	DCS	-5.9622	-0.6274	<b>5.3348</b>	5.9622	0.6274	2.6674	0.1874	<b>-3.2948</b>	3.2948	2.0348
3	RC	-6.0934	-0.7273	<b>5.3661</b>	6.0934	0.7273	2.6830	0.1863	-3.4103	3.4103	2.1673
4	RCS	-6.5263	-1.1311	<b>5.3952</b>	6.5263	1.1311	<b>2.6976</b>	0.1853	-3.8287	<b>3.8287</b>	<b>2.717</b>
5	LCC	-6.0063	-0.6767	<b>5.3296</b>	6.0063	0.6767	2.6648	0.1876	<b>-3.3415</b>	3.3415	2.095
6	LCS	-6.1715	-0.8587	<b>5.3128</b>	6.1715	0.8587	2.6564	<b>0.1882</b>	-3.5151	3.5151	2.3256
7	CME	-6.3484	-1.0269	<b>5.3215</b>	6.3484	1.0269	2.6607	0.1879	-3.6876	3.6876	2.5553
8	CMS	-6.4461	-1.1189	<b>5.3272</b>	6.4461	1.1189	2.6636	0.1877	-3.7825	3.7825	2.6857
9	A1	-6.0765	-0.7551	<b>5.3214</b>	6.0765	0.7551	2.6607	0.1879	-3.4158	3.4158	2.1925
10	A2	-6.4256	-1.0990	<b>5.3266</b>	6.4256	1.0990	2.6633	0.1877	-3.7623	<b>3.7623</b>	<b>2.6573</b>
11	A3	-5.8294	-0.8157	<b>5.0137</b>	5.8294	0.8157	2.5068	<b>0.1994</b>	<b>-3.3225</b>	3.3225	2.2018
12	A4	-6.2999	-0.9556	<b>5.3443</b>	6.2999	0.9556	2.6721	0.1871	-3.6277	3.6277	2.4625
13	A5	-6.0433	-0.6968	<b>5.3465</b>	6.0433	0.6968	<b>2.6732</b>	0.1870	-3.3700	3.3700	<b>2.1242</b>

**Table 5:** The global quantum chemical descriptors HOMO, LUMO, HOMO-LUMO Gap, Ionization Potential, Electron Affinity, Global Hardness, Global Softness, Chemical Potential and Electrophilicity Index for solvent phase calculations by B3LYP/6-311G\*\*.

S. No.	Name	HOMO (eV)	LUMO (eV)	HOMO-LUMO Gap (eV)	Ionization Potential $I = -E_{\text{HOMO}}$	Electron Affinity $A = -E_{\text{LUMO}}$	Global Hardness $\eta = (I-A)/2$	Global Softness $S = 1/2\eta$	Chemical Potential $\mu = -(I+A)/2$	Electro-negativity $\chi = -\mu$	Electrophilicity Index $\omega = \mu^2/2\eta$
1	DC	-6.8466	0.0198	<b>6.8665</b>	6.8466	-0.0198	3.4332	<b>0.1456</b>	<b>-3.4133</b>	3.4133	<b>1.6968</b>
2	DCS	-6.8784	-0.0043	<b>6.8741</b>	6.8784	0.0043	3.4370	0.1454	<b>-3.4414</b>	3.4414	1.7228
3	RC	-6.9293	-0.0416	<b>6.8877</b>	6.9293	0.0416	3.4438	0.1451	-3.4855	3.4855	1.7638
4	RCS	-7.4004	-0.2097	<b>7.1906</b>	7.4004	0.2097	<b>3.5953</b>	0.1390	-3.8051	<b>3.8051</b>	<b>2.0135</b>
5	LCC	-6.9152	-0.0446	<b>6.8706</b>	6.9152	0.0446	3.4353	<b>0.1455</b>	<b>-3.4799</b>	3.4799	1.7625
6	LCS	-6.9274	-0.0525	<b>6.8749</b>	6.9274	0.0525	3.4374	0.1454	<b>-3.4899</b>	3.4899	1.7716
7	CME	-6.9636	-0.0884	<b>6.8752</b>	6.9636	0.0884	3.4376	0.1454	-3.5260	3.5260	1.8083
8	CMS	-7.1576	-0.1346	<b>7.0229</b>	7.1576	0.1346	3.5114	0.1423	-3.6461	3.6461	1.8930
9	A1	-6.9555	-0.0729	<b>6.8825</b>	6.9555	0.0729	3.4412	0.1452	-3.5142	3.5142	1.7943
10	A2	-7.0053	-0.1099	<b>6.8953</b>	7.0053	0.1099	3.4476	0.1450	-3.5576	<b>3.5576</b>	<b>1.8355</b>
11	A3	-6.9277	-0.0459	<b>6.8817</b>	6.9277	0.0459	3.4408	0.1453	<b>-3.4868</b>	3.4868	1.7667
12	A4	-6.9644	-0.0821	<b>6.8823</b>	6.9644	0.0821	3.4411	0.1453	-3.5233	3.5233	1.8037
13	A5	-6.9410	-0.0672	<b>6.8738</b>	6.9410	0.0672	3.4369	0.1454	-3.5041	3.5041	1.7863