

Residue Interactions Affecting The Deprotonation of Internal Guanine Moieties In Oligodeoxyribonucleotides, Calculated By FMO Methods

Absalom Zamorano-Carrillo

Instituto Politécnico Nacional, 07360, Zacatenco-CDMX

Gerardo Arreola-Jardón

Universidad Politécnica de Santa Rosa Jáuregui

Reynaldo C. Pless

Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada

Julio C. González-Olvera (✉ jcgonzalez@upsrj.edu.mx)

Universidad Politécnica de Santa Rosa Jáuregui <https://orcid.org/0000-0001-7652-0617>

Research Article

Keywords: Intrinsic acidity, Deprotonated guanine, Quantum descriptors, DNA oligonucleotides, PIEDA

Posted Date: November 3rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-954859/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Residue interactions affecting the deprotonation of internal guanine moieties in oligodeoxyribonucleotides, calculated by FMO methods

Absalom Zamorano-Carrillo¹ · Gerardo Arreola-Jardón² · Reynaldo C. Pless³ · and Julio C. González-Olvera²

Correspondence to: Julio C. González-Olvera (email: jcgonzalez@upsrj.edu.mx)

¹ Laboratorio de Biofísica Computacional, SEPI-ENMH, Instituto Politécnico Nacional, 07360, Zacatenco-CDMX, Mexico

² Universidad Politécnica de Santa Rosa Jáuregui, 76220, Querétaro, Mexico

³ Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada, Instituto Politécnico Nacional, 76090, Querétaro, Mexico

Abstract

The effect of vicinal molecular groups on the intrinsic acidity of a central guanine residue in short single-stranded DNA models, and the potentials exerted by the backbone and the nucleobases on the leaving proton were determined by the Fragment Molecular Orbital (FMO) method, in terms of quantum descriptors (QD) and pair interaction interfragment decomposition analysis (PIEDA). The acidity of the central guanine moiety decreased with increasing oligonucleotide length, in response to changes by less than 1 eV in the ionization potential, global softness, electrophilicity index and electronegativity descriptors. The differences in these descriptors were majorly interpreted in terms of the electrostatic influence of the negative charges residing on the backbone of the molecule. Additionally, this electric-field effect was determined explicitly for the displacement of the test hydronium ion to a distance of 250 Å from its original position, resulting in good agreement with calculations of the variation in Gibbs free energies, obtained from physical experiments conducted on the identical oligonucleotide sequences. The reported results are useful for biophysical applications of deoxyriboligonucleotides containing guanine residues in order to induce local negative charges at specific positions in the DNA chain.

Keywords Intrinsic acidity · Deprotonated guanine · Quantum descriptors · DNA oligonucleotides · PIEDA

Introduction

Under physiological conditions, DNA molecules are long polyanionic chains, due to the negative charges residing on the internucleotidic phosphate groups. The charges' strong potential affects the properties of the heterocyclic bases, as they are located close to the phosphate groups. Regarding acid/base ionization, electrostatic considerations dictate that the formation of a new local negative charge on a nucleobase should be more difficult in the polynucleotidic context than in the simple nucleoside, whereas for the installation of a new positive charge the reverse should be true. This results in a shift of the pK_a values of nucleobases in the polynucleotidic context to higher values [1]. The acidity and basicity of the heterocyclic bases in polynucleotides are also affected by their neighboring nucleobases through π -orbital coupling [2].

The structure and properties of DNA have been widely studied over the past decades, due to its central role in biology. In this regard, semi- and empirical methods and molecular mechanics (MM) approaches have been widely used to study the molecular interactions between nucleic acids and other molecules [3, 4]. Computational quantum mechanical (QM) methods represent an unsurpassed strategy for the study of areas where experimental determinations and classical computational methods are not possible, i.e., charge transfer and polarization effects for binding interaction energies. A comprehensive review on structures, properties, and functions of DNA fragments analyzed by QM methods has been published by Šponer et al. [5]. In particular, the electrostatic influence of the negative phosphate groups on base properties in oligomeric ADN/ARN models has been the subject of several computational studies [6, 7]. In RNA trinucleotides modeled by the density functional theory (DFT) algorithm, it was found that electrostatic interactions between the sugar-phosphate backbone and the base have the largest effect, rather than the traditionally studied interbase stacking [6]. Hartree-Fock (HF) calculations conducted at the level of 6-31G* on model trinucleotides showed that an exchange of the central base can strongly affect the shape of the electrostatic potential of the molecule [7]. As to nucleobase stacking interactions in similar contexts, it was found to significantly affect the ionization properties of the adenines in the tetramer d-5'-AATT-3', as calculated by various *ab initio* methods [8]. In the case of base-to-base interactions in single chains of DNA of different sizes, the delocalization of orbitals arises in spite of strong stacking interactions [9].

To reduce the limitation imposed by the high computational costs associated with QM calculations applied to large biomolecules, approximative methods have been proposed as an alternative, such as the fragment molecular orbital (FMO) method. This algorithm was designed to compute energies of large fragmented molecules solving the HF equation for each fragment and considering the electrostatic field from the other fragments [10]. Some nucleic acids and other biomolecules have been modeled using this approach: 1) the stability of a DNA duplex bound to a cAMP receptor protein was found to change in dependence of both electrostatic and charge-transfer inter- and intramolecular interactions, in comparison with the uncomplexed situation, as determined by the FMO method, with second order Møller-Plesset (MP2) level calculations, using different basis sets [11, 12]; 2) the binding affinity of the influenza virus hemagglutinin protein to host-cell receptors was analyzed by the MP2/6-31G method, finding that the intramolecular interactions of some protein residues play key roles in the binding process [13]; 3) the interaction energies between cyclin-dependent kinase 2 and a set of 28 inhibitors were computed by the FMO method, at an MP2/6-31G* theory level, contributing to our understanding of these interactions [14]. In summary, FMO calculations can be used to study not only the interactions between large biomolecules, but also those occurring within each molecule, providing total energy and electron density for the whole system along with the interactions at the fragment and orbital levels.

In the present study, we examine single-stranded deoxyribonucleotides, from three to seven nucleotide units long, containing a single guanine base in the central position, and calculate the interactions of the various constituent oligomer fragments with the central guanine group, to assess the energetics governing the

deprotonation of this guanine (the removal of the guanine-N¹ hydrogen as a hydronium ion). Also, these calculations assessed the additional work required to distance the resulting hydronium ion from the negative charges located on the internucleotidic phosphates, on the various nucleobase residues and on the now anionic guanine group. The oligonucleotide structures were initially optimized by an FMO interfaced to a polarizable continuum method algorithm (FMO/PCM), using a level of theory of MP2/6-31G. The ionization potential, molecular softness, electrophilicity index, and electronegativity quantum descriptors (QDs) were then used to describe acidity shifts of the central guanine residue contained in different-sized oligomers. The energy interactions between a hydronium ion and each of the corresponding phosphate groups and nucleobases were computed as a function of distance to obtain the aforementioned potential, using point calculations by the FMO/PCM method combined with pair interaction energy decomposition analysis (FMO/PCM/PIEDA), at the MP2/6-31G* level of theory. The values thus obtained are in good agreement with the p*K*_a values experimentally determined in this work, in aqueous solution, for the model oligodeoxyribonucleotides. The results of this theoretical study describe the base-stacking effects and electrostatic contribution of the phosphate groups to the free energy of deprotonation of a guanine moiety in the different modeled oligonucleotides. A better understanding of these physicochemical properties of short single-stranded DNA oligonucleotides will contribute to the design and development of recently reported clinical applications for these biopolymers, namely oligonucleotide therapeutics [15], and DNA-based nanotechnology, such as nanophotonics, nanoscale lithography, nanometrology, etc [16].

Methods

Molecular models and optimizations

The molecular structure of 2'-deoxyguanosine (2'-dGuo) was obtained from the PDBeChem database, with entry GNG [17]. The oligonucleotides containing a central guanine group were built using the 3D-DART DNA modeling server [18], applying crystallographic dinucleotide step parameters (roll, twist, and slide) reported previously [19], and the complementary strand was removed from the double-stranded polynucleotide which had been generated. This strategy was used because structures of single-stranded oligonucleotides modeled in this work were not found in databases for structures of nucleic acids. Additionally, the marked structure instability of DNA oligonucleotides in this conformation, when they are dissolved in water, makes difficult the selection of a fixed initial structure. The structures thus obtained were fragmented as reported for single-stranded RNA by Kurisaki et al. [10], shown in Fig. 1, and then optimized using the FMO/PCM method [20], at a restricted HF (RHF) level and 6-31G basis set, using the open-access GAMESS software [21], version 11-32. The algorithm used for the FMO/PCM optimization was conjugate gradients for all the cases. All the scripts for computations were elaborated using the Facio program version 18.6.2 [22]. The structures of the oligonucleotides were rendered by the Jmol program, version 13.0.8 [23].

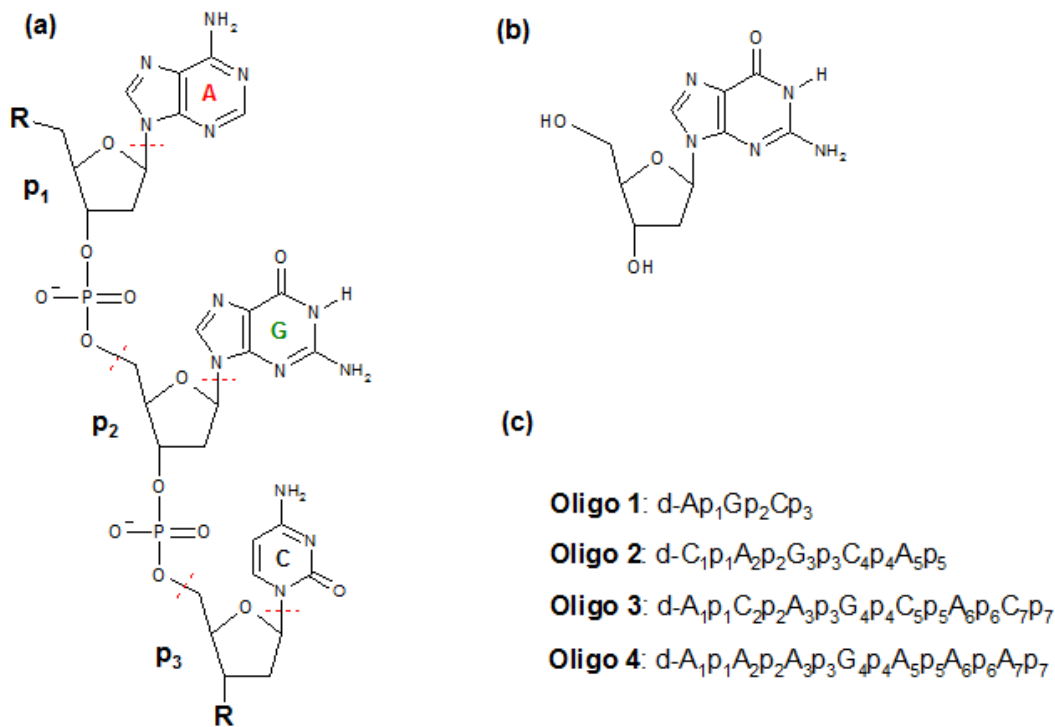


Fig. 1 Fragmentation of the oligonucleotides. (a) The fragmentation for the trimer d-AGC is shown in detail: R denotes the group HO-, at both the 5'- and 3'-ends of the oligomer. For the longer oligonucleotides, R represents the corresponding additional nucleotide units. The dashed lines show the position where the bonds were broken. (b) Chemical structure of 2'-dGuo. (c) Abbreviations for the oligonucleotides modeled in this work

FMO calculations

The FMO method divides a molecule into fragments and assigns the pair of electrons in any covalent bond equally divided between the two adjoining residues. The molecular orbitals for the electrons are calculated for each fragment by the HF method until the densities are self-consistent. Then, the molecular orbitals are computed for these fragments and fragment pairs, applying methods for achieving accurate dispersion energies, such as MP2, in order to obtain the total energy of the whole molecule [24].

The total energy of the molecule, E , for the total number of fragments (N) is calculated as follows:

$$E = \sum_{l>j}^N E'_{lj} - (N-2) \sum_l^N E'_l + \sum_{s>t}^{all\ atoms} \frac{Z_s Z_t}{|\mathbf{r}_s - \mathbf{r}_t|} \quad (1)$$

$$= \sum_{l>j}^N E'_{lj} - (N-2) \sum_l^N E'_l \quad (2)$$

$$= \sum_I^N E'_I + \sum_{I>J}^N (E'_{IJ} - E'_I - E'_J) \quad (3)$$

where I and J are the fragment numbers, E'_I and E'_{IJ} accounts for the RHF energies of the fragment and fragment pair, respectively; the last term in Equation 1 is the nuclear repulsion energy [24, 25].

The FMO calculations can provide interaction energies between fragment pairs, called pair interaction energies (PIEs) or interfragmentary interaction energies (IFIEs) [26], which can be used for residues of interest in a molecule, for example, phosphate and base groups in DNA or aminoacid residues in proteins. These PIEs, ΔE_{IJ} , are calculated by the following equation:

$$\Delta E_{IJ} = (E'_{IJ} - E'_I - E'_J) + \text{Tr}(\Delta \mathbf{D}^{IJ} \mathbf{V}^{IJ}) \quad (4)$$

where E'_I and E'_{IJ} are the electronic energies defined for Equations (1) and (2); $\Delta \mathbf{D}^{IJ}$ is a difference density matrix, and \mathbf{V}^{IJ} is an environmental electrostatic potential for fragment dimer IJ from other fragments [11, 26].

The PIEs can be further decomposed into the energy contributors by the PIEDA method [26], namely electrostatic (ΔE_{IJ}^{es}), exchange (ΔE_{IJ}^{ex}), charge transfer with higher-order mixed terms ($\Delta E_{IJ}^{(ct+mix)}$), and dispersion (ΔE_{IJ}^{disp}) energies, according to Equation 5:

$$\Delta E_{IJ} = \Delta E_{IJ}^{es} + \Delta E_{IJ}^{ex} + \Delta E_{IJ}^{ct+mix} + \Delta E_{IJ}^{disp} \quad (5)$$

The PCM is generally employed to represent the solvation effects in *ab initio* computations. For these calculations the environment is considered as a polarizable continuum, instead of as explicit solvent molecules, and is represented as a cavity surrounding the solute [27]. The solute exerts an electrostatic potential on the cavity surface generating the apparent surface charges, which are calculated as follows:

$$\mathbf{C} \mathbf{q} = \mathbf{g} \quad (6)$$

where \mathbf{C} is a geometric matrix, \mathbf{q} are the apparent surface charges; and \mathbf{g} is a function of the electrostatic potential vector. The interaction between these surface charges and the nuclei of all solute atoms is added to the total energy of the molecule [28].

For the FMO interfaced to PCM, the cavity is constructed around the entire system and is used unchanged in all individual fragment calculations, but the Coulomb interactions due to the apparent surface charges are added to each fragment calculation. For more detailed information on FMO/PCM method see Fedorov et al. [26]. Recently, the PIEDA algorithm was implemented in the FMO/PCM method for treatment of the solvation energies of the fragments of the molecule [29].

For the calculation of interaction energies between the departing N^1 proton of the central guanine residue, represented here as a hydronium ion, and either the backbone fragments or the nucleobase residues, two

different steps of the molecules were employed. In the first set of molecules, the leaving proton was displaced by 2 Å from its original position (tagged as Point A) in the plane of the central guanine ring (see Fig. 4). Then, the hydronium ion was displaced to a distance of 248 Å (Point B). Single-point calculations were performed for each step, using the FMO/PCM/PIEDA algorithm, at a level of theory of MP2/6-31G*, by an updated GAMESS version (2018-R1). Initially, the difference between the two steps was obtained for the total energies of the molecule, and then compared to experimental results obtained by potentiometric titrations of oligonucleotides with identical sequences. Then, the corresponding differences for total PIEs and energy contributors, again for the two steps mentioned above, were obtained by simple subtraction of total PIEs, or contributors, in Point A from those in Point B. The variation between these energies of the two arrays accounts for the additional work required to distance the hydronium cation from the negatively charged backbone and the base groups of the oligonucleotide.

Proton affinity calculation

The intrinsic acidity of the neutral central guanine residue was evaluated by an approach similar to those previously used for the proton affinity (PA) of small molecules [30]. Accordingly, the ionization potential (IP), global softness (GS), electrophilicity index (EI), and electronegativity (χ) QDs were obtained from the last round of the FMO optimization for each modeled molecule. These values were used to compute the correlation coefficients in Equation 7, by iterative linear regression using the Origin2020b program [31].

$$PA = C + C_1(-IP) + C_2GS + C_3\left(\frac{1}{\chi}\right) + C_4\left(\frac{1}{EI}\right) \quad (7)$$

The PA was considered as the dependent variable for the regression, and for the calculations the experimental pK_a values were employed, while the four above-mentioned QDs were taken as independent variables. Then, the PAs were calculated for the monomer 2'-dGuo and the oligomers 1–4, using the QDs obtained at the FMO optimizations for each molecular model.

Potentiometric titrations of oligonucleotides

The experimental pK_a values of oligonucleotides used here to be compared with the theoretical values were determined by a similar procedure to those reported in our previous work [32]. However, the NaCl concentration or ionic strength employed in the experiments was 3 mM, instead of 100 mM. This low ionic strength was selected to be more comparable with the theoretical calculations, where salt molecules were not added.

Results and discussion

Molecular models and optimizations

The FMO/PCM-optimized structures of Oligo 1, Oligo 2, Oligo 3, and Oligo 4 are shown in Fig. 2, superimposed with the initial crystallographic-restrained structures. The sequence selection for these oligonucleotides was carried out to analyze the influence of the molecule length on the acidity of a central

guanine residue in short single-stranded DNA, Oligo 1 to Oligo 3, and the examination of the effect of different flanking sequences on this central base, namely the heptamers Oligo 3 and Oligo 4. The RMS (Root Mean Square) values for gradient energy during optimizations for the corresponding structures, after five iterations, were between 0.0178 and 0.0163 hartree bohr⁻¹, with the exception of Oligo 1, which was optimized until the seventh round to approach these values.

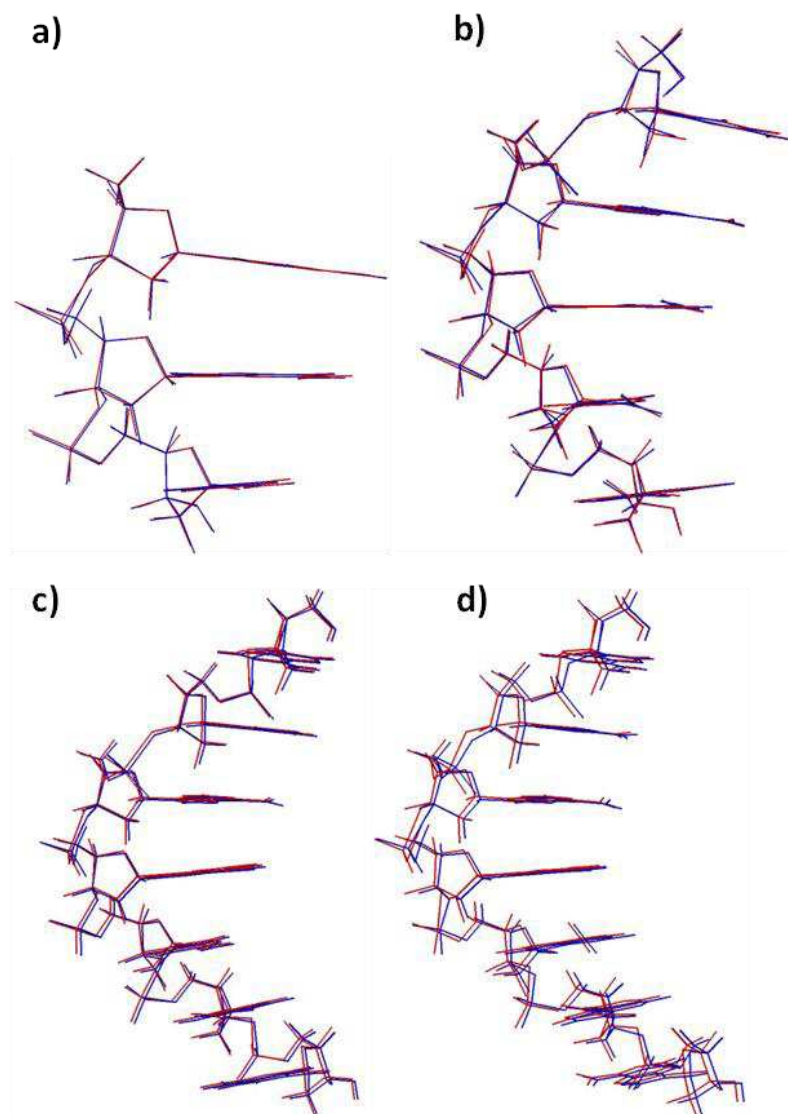


Fig. 2 Superimpositions of initial (red) and FMO/PCM-optimized (blue) structures of short oligomers: a) Oligo 1 (d-AGC), b) Oligo 2 (d-CAGCA), c) Oligo 3 (d-ACAGCAC), and d) Oligo 4 (d-AAAGAAA)

In general, the structures under comparison are similar, and only the heptamer Oligo 4 showed distinct differences between the initial and optimized structure, particularly in the backbone of the molecule. The trimer Oligo 1 showed major changes in the RMSs during optimization, whereas for the other molecules this

value follows a monotonous reduction as the optimization proceeds. For Oligo 4 (the adenine-rich heptamer), the exocyclic amino groups in the tri-adenilyl run which abuts the central guanine on the 3' side, are not coplanar with their heterocycle, whereas for Oligo 3 all nucleobases are essentially coplanar with their exocyclic amino groups (Fig. 2). The rotation of amino groups could be caused by the proximity of the negative π -surface of one pyrimidine ring to the hydrogens attached to nitrogen in these groups, as was suggested by Isaksson and coworkers [33].

Interfragment distances between vicinal bases were determined as twice the van der Waals radii, computed in the last round of FMO optimization for each of the four oligonucleotides. The strongest base stacking in the modeled oligonucleotides was observed for Oligo 4 with an averaged interfragment distance of 2.24 Å, showing the shortest distances in the A-tract that abuts the central guanine on the 3' side, particularly in the steps G4-A5 and A5-A6, with interplanar distances of 2.14 and 2.20 Å, respectively. In contrast, the interfragment distances at the identical positions in Oligo 3 are larger by 0.2 and 0.1 Å for G4-C5 and C5-A6, respectively. In fact, the averaged interfragment distance for this last oligomer was 2.32 Å, which was the highest recorded for all the oligonucleotides analyzed. As a consequence, the distance between internucleotidic phosphates is reduced in Oligo 4, shrinking the length of the molecule in comparison with Oligo 3. This geometrical difference leads to a higher linear density of negative charges in the backbone of the A-rich heptanucleotide, strengthening the influence of the electric field produced by the phosphate groups on the N¹-H of the central guanine residue, thus making the withdrawal of the proton to the solvent more difficult. As to the other oligomers, Oligo 1 showed an averaged interbase distance of 2.31 Å, while for Oligo 2 this same distance was calculated at 2.27 Å, showing an unexpectedly low value of 2.16 Å between the bases C4 and A5, since the shortest distance was expected to be for the stacking between purines base groups.

This strong stacking of purines in the chain should affect the deprotonation of central aglycone at the N¹ position, whether decreasing the intrinsic acidity of the central guanine moiety by redistribution of charge to the vicinal stacked bases [33] or through the greater proximity of the negative charges in the backbone of the molecule [32], both effects hampering the approach of the hydroxide ions that is to abstract the proton. These structural differences relative to the other heptamer, affect the propensity of a guanine residue to dissociate, causing a shift in the corresponding pK_a value. Similar physicochemical behavior was also observed for the progressive lowering of the ionization energies or potentials as the number of consecutive adenines increase, in DNA oligomers modeled at the DFT level [34]. In fact, the strong stacking observed in this work for adenine-rich Oligo 4, with approximately 2.28 Å for the rise parameter, also was reported for poly(dA) chains where helical stacks were identified using a coarse-grained model [35]. Similarly, stacked adenines in short single-stranded DNA oligomers, modeled by the DFT method using the B3LYP-D functionals, were reported to adopt more regular arrangements [36], though in these DFT calculations the modeling of the short oligomers was carried out from an initial B-DNA conformation, whereas in the present work we started from crystallographic-restrained structures in the gaseous state, and then optimized in the polarizable-continuum model. Similar ordered structures for the homopolynucleotide dA₃₀ with strongly-stacked adenines, in

aqueous solution, were obtained by small angle X-ray scattering (SAXS) paired to ensemble fitting methods [37].

Physical descriptors as indicators of acidity for the central guanine residue

To determine the propensity of the central guanine moiety for proton dissociation in the oligomeric context, the QDs such as ionization potential (IP), global softness (GS), electrophilicity index (EI) and electronegativity (χ) of 2'-dGuo and Oligos 1-4, at the one-body level, were used to estimate the intrinsic acidity of the guanine group. For this purpose, we employed a strategy similar to that reported by Rajak et al. [30] for the computation of proton affinity (PA) for several sets of small single molecules using the aforementioned four descriptors to evaluate the PA of different molecules. Since the deprotonation is the inverse process to protonation, the residue acidity in a molecule should exhibit an inverse relationship to the PA, correlating a lower acidity or deprotonation propensity with a higher proton affinity. The calculated QDs for the guanine fragment in the modeled molecules, obtained in the last round of optimization, are presented in Table 1. Electronegativity is the inverse additive of the chemical potential descriptor obtained by FMO optimizations.

Table 1 Quantum descriptors for the central guanine group in the optimized 2'-dGuo and DNA oligomers. The QDs correspond to the fifth round of optimization (seventh round for Oligo 1), and all are given in eV

Model	IP	GS	EI	χ
2'-dGuo	8.3652	0.0839	0.2424	2.404
Oligo 1	7.9659	0.0845	0.1772	2.0482
Oligo 2	7.8615	0.0831	0.141	1.8423
Oligo 3	7.7	0.0833	0.1197	1.6957
Oligo 4	7.8155	0.0833	0.1374	1.8161

Based on the mathematical model for the determination of PA [30], the values for the correlation coefficients obtained after the multiple linear regression were 3108.2606, 186.0409, -10402.3178, 60.054, and -2205.8263 for C , C_1 , C_2 , C_3 and C_4 , respectively. Then, Equation 7 was used to calculate the PA for each of the modeled oligomers and 2'-dGuo, and the inverse of these theoretically obtained values were plotted against the corresponding experimental pK_a values determined in this work, to assess the acidity of the central guanine in the different oligomeric contexts (Fig. 3).

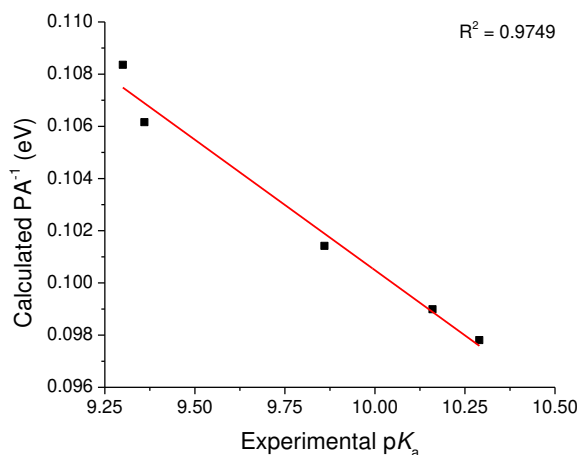


Fig. 3 Plot of the inverse of calculated proton affinity (PA^{-1}) vs. the experimental pK_a values

As expected, the guanine residue in the monomeric unit, 2'-dGuo, presents the highest PA^{-1} value, indicating a higher acidity in comparison to this base group in the oligonucleotidic context. Moreover, as the central-guanine-containing oligonucleotides become longer the PA^{-1} is reduced. This decrease denotes a higher acidity for the guanine group in the trimer in comparison to the longer oligonucleotides, which is in line with experimental results, where UV-monitored alkaline titration of this same set of oligonucleotides, at an ionic strength of 0.1 M, presented the following series of decreasing acidity of the central guanine group: 2'-dGuo \approx d-AGC > d-CAGCA > d-ACAGCAC > d-AAAGAAA [32].

Acidity shifts observed for the central guanine residue could be attributed to the redistribution of negative charges residing on the backbone's phosphate groups. According to our results, the central guanine moiety experiences different physical situations in dependence on the oligomeric context, namely the size of the chain and the sequence flanking the central nucleotide, thus affecting the propensity for dissociation of the N^1 proton of that residue. The physical descriptors obtained indicate a lower acidity of the central guanine base as the size of the oligomer increases, which is in agreement with the experimental results of the potentiometric titration of similar oligonucleotides [32]. The increase in pK_a value for a residue in the oligomeric situation suggests that essentially no negatively-charged guanine groups would be present under physiological conditions, low salt concentration, neutral pH and temperatures around 37°C.

However, in the comparison of the heptamers with a different sequence, namely d-ACAGCAC (Oligo 3) and d-AAAGAAA (Oligo 4), the present results show a shift in the acidity of the guanine moiety, but in the opposite sense to that reported in the experimental titrations, where the central guanine had a higher pK_a in the latter oligomer, in comparison to the former [32]. Locally, the charge transfer in the central guanine, promoted by the redistribution of charges due to both the adjacent phosphate groups and the vicinal bases should contribute to the shifts in the acidity of this central group. Accordingly, future studies may be conducted to calculate this charge transfer employing modeling strategies different from those used in this work, such as DFTB2 method with the TZVP basis set [38] or multilayer-FMO-based excited-state method, employing a 6-31G* basis set [39], along with a wider basis set.

The energy contribution on the free energy of deprotonation for the central guanine residue

The energy through-space effect of the residues constituting the different oligomers on the proton departing the central base, considered as a positively charged hydronium ion, was determined by point calculations using the FMO/PCM/PIEDA algorithm. For this purpose, the computed total energies for the molecules, after MP2 correction, were used to calculate the total work (W) to separate that positive point charge from an initial position at 2 Å from the original site of the N¹ proton, in the deprotonated guanine residue (Point A, Fig. 4), to a linear distance of approximately 250 Å (Point B, Fig. 4), explicitly moved in a straight-line displacement in the plane of the guanine moiety, as shown in Fig. 4. This total work was obtained according to the following model:

$$W = \Delta E^{tot} = E_A^{tot} - E_B^{tot} \quad (8)$$

where E_A^{tot} and E_B^{tot} accounts for the total energy of the molecular configurations with the departing positive charge at point A and at point B, respectively; ΔE^{tot} stands for the interaction force between the residues of the oligomer and the test hydronium cation for the corresponding modeled molecules.

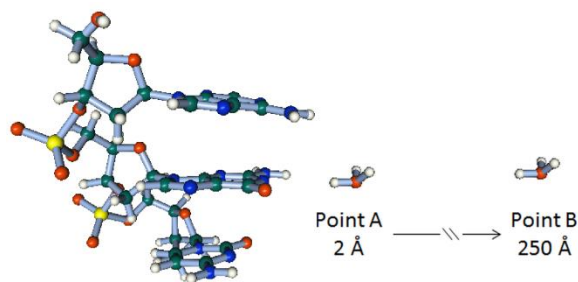


Fig. 4 Graphic representation for the displacement of the hydronium ion (positive charge) from the initial (Point A) to final (Point B) positions for the trimer Oligo 1, as an example. Distances are indicated relative to the initial position of the proton bound at the N¹ site of the central guanine

The proton in the test hydronium cation was located at 2 Å apart from its original position on the atom N¹ of the guanine moiety to avoid unfavorable contacts [40, 41], as is shown in Point A of Fig. 4. The displacement distance of 250 Å (Point B, Fig. 4) corresponds to about half the average distance to the next-nearest oligonucleotide, for the 3 mM concentration used in the experimental assays, which are used for comparison to the theoretical values.

From a thermodynamic point of view, the additional effect of the oligonucleotide backbone and base residues on the removal of the proton from the central guanine moiety must be overcome by a higher chemical potential of the hydroxide ions in the bulk solvent, namely, by a higher hydroxide concentration in the titration of DNA oligomers compared to the case of the monomeric molecule 2'-dGuo. The corresponding free-energy differential is determined by:

$$\Delta\Delta F_{\text{experimental}} = \mu_1 - \mu_2 = RT * \ln\left(\frac{a_2}{a_1}\right) \cong RT * \ln\left(\frac{c_2}{c_1}\right)$$

where R is the universal gas constant ($1.987 \text{ cal mol}^{-1} \text{ K}^{-1}$), T is the temperature of the experiment (298.15 K), and μ_2 , a_2 , and c_2 are the chemical potential, activity, and molar concentration of the hydroxide ions at the midpoint of the deprotonation transition of the corresponding oligodeoxyribonucleotide, while μ_1 , a_1 , and c_1 stand for those values at the midpoint of the deprotonation of the 2'-dGuo model. The NaCl concentration for all the experiments was 3 mM. The experimental pK_a values for 2'-dGuo and the oligomers, as well as the variation of the potential ($\Delta\Delta F$), the computed work (W) for the oligomers, and the variation of W for different oligomers in comparison to 2'-dGuo (ΔW^*), are shown in Table 2.

Table 2 Physical values determined for 2'-dGuo and the different oligonucleotides. The pK_a values were obtained at 25°C and 3 mM of NaCl

Model	pK_a	W (kcal/mol)	$\Delta\Delta F$ (kcal/mol)	ΔW^* (kcal/mol)
2'-dGuo	9.36	-0.126	----	----
Oligo 1	9.30	-0.228	-0.082	0.102
Oligo 2	9.86	-0.312	0.682	0.186
Oligo 3	10.16	-0.394	1.091	0.268
Oligo 4	10.29	-0.375	1.269	0.249

The additional potential that must be overcome for the hydroxide ions to remove the departing proton from the central guanine moiety is virtually zero in the case of the trimer Oligo 1 ($0.102 \text{ kcal mol}^{-1}$), and could be comparable with the situation of the monomer 2'-dGuo, where no internucleotide negative-charged phosphate groups are present. As was stated previously [32], this result was unexpected since the electrostatic interactions present in the trimeric case should be reflected in a higher potential to be required to remove the proton, in comparison to the case of the simple nucleoside, where there are no vicinal heterocyclic bases and phosphate groups to interact electrostatically with the departing positive charge.

For the rest of the modeled oligomers, both the experimentally determined additional potential ($\Delta\Delta F$) and the calculated additional work (ΔW^*) increase as the molecule is lengthened, being comparable in the case of the heptamers. This could be mainly attributed to the addition of internucleotidic negative-charged phosphate groups which affect the separation of the departing proton after the dissociation from the central guanine group.

For the cases of the heptamers Oligo 3 and Oligo 4 differences of about $1.0 \text{ kcal mol}^{-1}$ were found between experimental and theoretical potentials. The higher deviations for the penta- and heptanucleotides could be attributed to the treatment of stacking interactions of nucleobases by the computational algorithms applied in this work. However, differences of a few kilocalories per mole are acceptable for comparison between theoretical calculations and experimental determinations.

PIEs for hydronium ion and residues in the molecular models

To investigate the variation of the energies for the modeled molecules, a pair interaction energies decomposition analysis (PIEDA) was carried out, and the results are shown in Fig. 5. The variation of the non-bonding interaction energies ($\Delta\Delta PIEs$) from the hydronium ion at Point A to Point B, and the different residues of the tested molecules are shown in the first set of bars in Fig. 5. The contributions to $\Delta\Delta PIEs$ are grouped according to the type of energy, namely electrostatic energy (E^{es}), exchange energy (E^{ex}), charge transfer plus mixed energy (E^{ct+mix}) and dispersion energy (E^{disp}). As expected, the largest contribution to the interaction energies is the electrostatic component, due to the attraction of the hydronium ion to both the negatively-charged internucleotidic phosphate groups, which is absent in the case of the simple nucleoside, and the deprotonated guanine group. In all these computations, the repulsive effect of the exchange energies only accounts for about 0.2% of the total effect; this is the lowest contribution to the total pair interaction energy for all the modeled molecules. Other minor contributions to the total energy differentials are the charge transfer plus mixed and dispersion energies with about 10% and 5%, respectively.

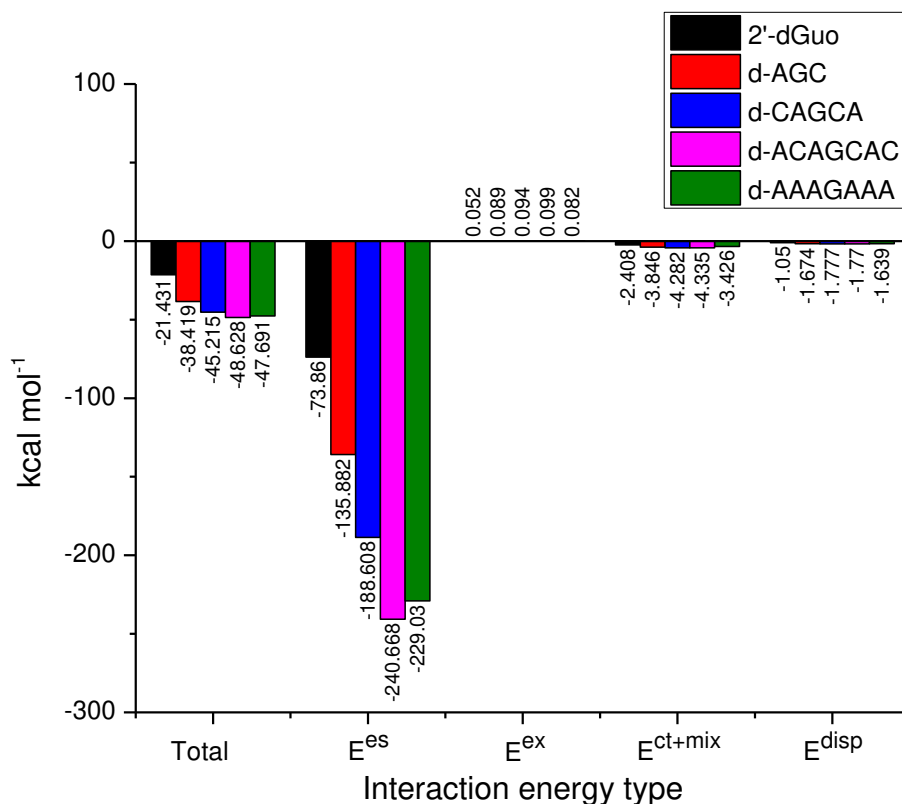


Fig. 5 Variations of non-bonding interaction energies ($\Delta\Delta PIEs$) calculated for 2'-dGuo and the four modeled oligomers. From left to right: $\Delta\Delta PIEs$, $\Delta\Delta E^{es}$, $\Delta\Delta E^{ex}$, $\Delta\Delta E^{ct+mix}$, and $\Delta\Delta E^{disp}$. All the values are expressed in kcal mol⁻¹

The difference of the electrostatic interactions, $\Delta\Delta E^{es}$, between each of the fragments in the guanine-deprotonated chain and the hydronium ion for the polyfragment systems or ssDNA models at both steps,

namely Point A and Point B, are shown in the Figure 6. For all the cases, the major electrostatic potential exerted on the hydronium ion corresponds to the negative-charged guanine moiety, with higher energies of -82.162 and -80.257 kcal/mol for Oligo 2 and Oligo 3, respectively, and lower energies for Oligo 1 and Oligo 4, namely -79.787 and -76.073 kcal/mol, respectively. Similarly, the electrostatic potentials of the phosphate groups to the hydronium ion are lower for Oligo 1 and Oligo 4, between 25 kcal/mol and 26 kcal/mol, in comparison to Oligo 2 and Oligo 3, where they are around 27 kcal/mol (Fig. 6). Although these residues, namely anionic guanine and phosphate groups (bonded to 2'-deoxyribose group), carry a point negative electrical charge at the beginning of the computations, differences of some few calories were observed for their electrostatic interactions with the hydronium cation. These variations can be attributed to the solvation screening exerted by the atoms surrounding each fragment at the initial step (Point A), which are higher for Oligo 1 and Oligo 4, thus affecting the mentioned potentials. Regarding the nucleobases in the modeled oligomers, the 3'-vicinal cytosine nucleobases interact strongly in comparison to the other bases in Oligos 1-3, being comparable for the 3'-vicinal adenine to the other similar bases in the Oligo 4. For all the cases of the bases, the screening is due to positive charges induced on the corresponding surrounding atoms, probably by the strong density of negative charges in the molecules, residing in the deprotonated guanine and phosphate groups.

For the oligonucleotides discussed here, the contribution of nucleobases, other than the anionic central guanine group, to the total pair interaction energies mostly have positive values, counteracting the predominant electrostatic interactions. The highest positive total variation energies between nucleobases and the hydronium ion were obtained for the adenine and cytosine residues, and in just one case did a cytosine group have a negative value, namely the cytosine which is 3'-vicinal to the central anionic guanine in the pentamer d-CAGCA, with a value of -0.079 kcal mol⁻¹.

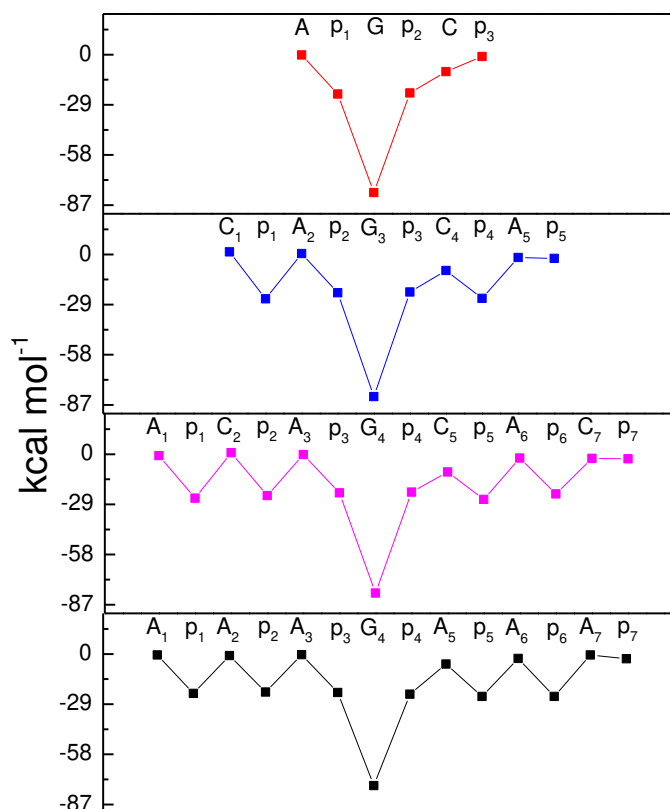


Fig. 6 Difference of the electrostatic interactions, $\Delta\Delta E^{es}$, for each of the fragments in the modeled oligonucleotides: a) Oligo1, b) Oligo 2, c) Oligo 3, and d) Oligo 4

The observed general trends are comparable to those previously reported for these same molecules [32]. After the deprotonation of the central aglycone in oligomers, the interactions with the backbone and the nucleobases affect the withdrawal of the departing proton to the bulk solvent. In that respect, the stabilization produced by the backbone in the transfer RNA conformation has been suggested by theoretical studies by DFT at the M06-2X/6-31+G(d,p) level [6] and for several DNA models [42]. In order to determine the additional potential mentioned above, we considered the negative phosphate as point charges residing on the backbone of oligonucleotides, as well as the deprotonated single guanine residue. As mentioned before, the FMO method is a useful strategy to calculate the electrostatic interactions of fragments in large molecules. For these calculations, we assume that proton separation is not affected by a structural rearrangement of the guanine-deprotonated chain. For this reason, the electrostatic interactions between the atomic groups involved in the deprotonation transition were treated as point charges.

Conclusions

The computational methods and procedures employed in this work to determine both the intra- and interfragment interactions which influence the deprotonation of a central guanine residue in the oligomeric situation, both before and after the N^1 proton dissociation, are in good agreement with reported experimental

results. Although the initial structures were built using parameters of crystallographic dinucleotide steps in the gas phase, the optimization of structures in water solvent shows the expected parameters, particularly in the stacking of adjacent nucleobases, i.e., the interfragment distance between adenines in the A-rich heptamer d-AAAGAAA are comparable to previous reports.

The intrinsic acidity for the central guanine residue in the oligonucleotide context was analyzed by quantum descriptors obtained from the optimization results. The general trend for the decreasing of acidities as the oligonucleotides are lengthened is comparable to that for the corresponding pK_a values. The effect of the sugar-phosphate backbone on the structure, properties and dynamics of the nucleic acids is of interest for a better understanding of these biopolymers. The action of the electric field set up by the negative phosphate groups on the departing proton make the deprotonation of the nucleobase more difficult. The general agreement between theoretical calculations and experimental results on the variation of this potential as a function of linear distance gives a good picture of these interfragment interactions, contributing to our physical understanding of single-stranded DNA.

References

1. Chatterjee S, Pathmasiri W, Plashkevych O, Honcharenko D, Varghese OP, Maiti M, Chattopadhyaya J (2006) The chemical nature of the 2'-substituent in the pentose-sugar dictates the pseudoaromatic character of the nucleobase (pK_a) in DNA/RNA. *Org Biomol Chem* 4:1675–1686.
2. Acharya S, Barman J, Cheruku P, Chatterjee S, Acharya P, Isaksson J, Chattopadhyaya J (2004) Significant pK_a Perturbation of Nucleobases Is an Intrinsic Property of the Sequence Context in DNA and RNA. *J Am Chem Soc* 126:8674-8681.
3. Galindo-Murillo R, Bergonzo C, Cheatham III TE (2014) Molecular Modeling of Nucleic Acid Structure: Energy and Sampling. *Curr Protoc Nucleic Acid Chem* 56:7.10.1. doi:10.1002/0471142700.nc0708s04
4. Šponer J, Banáš P, Jurečka P, Zgarbová M, Kührová P, Havrila M, Krepl M, Stadlbauer P, Otyepka M (2014) Molecular Dynamics Simulations of Nucleic Acids. From Tetranucleotides to the Ribosome. *J Phys Chem Lett* 5:1771-1782.
5. Šponer J, Shukla MK, Wang J, Leszczynski J (2017). Computational Modeling of DNA and RNA Fragments In: *Handbook of Computational Chemistry*. Springer, p 1803-1826.
6. Witts RN, Hopson EC, Koballa DE, Van Boening TA, Hopkins NH, Patterson EV, Nagan MC (2013) Backbone–Base Interactions Critical to Quantum Stabilization of Transfer RNA Anticodon Structure. *J Phys Chem* 117:7489–7497.
7. Koch T, Shim I, Lindow M, Ørum H, Bohr HG (2014) Quantum Mechanical Studies of DNA and LNA. *Nucleic Acid Ther* 24:139–148.
8. Bravaya KB, Epifanovsky E, Krylov AI (2012) Four Bases Score a Run: Ab Initio Calculations Quantify a Cooperative Effect of H-Bonding and π -Stacking on the Ionization Energy of Adenine in the AATT Tetramer. *J Phys Chem Lett* 3:2726–2732.
9. Tonzani S, Schatz GC (2008) Electronic Excitations and Spectra in Single-Stranded DNA. *J Am Chem Soc* 130:7607–7612.

10. Kurisaki I, Fukuzawa K, Nakano T, Mochizuki Y, Watanabe H, Tanaka S (2010) Fragment molecular orbital (FMO) study on stabilization mechanism of neuro-oncological ventral antigen (NOVA)-RNA complex system. *J Mol Struct* 962:45-55.
11. Fukuzawa K, Komeiji Y, Mochizuki Y, Kato A, Nakano T, Tanaka S (2006) Intra- and Intermolecular Interactions between Cyclic-AMP Receptor Protein and DNA: *Ab initio* Fragment Molecular Orbital Study. *J Comput Chem* 27:948–960.
12. Kurisaki I, Fukuzawa K, Komeiji Y, Mochizuki Y, Nakano T, Imada J, Chmielewski A, Rothstein SM, Watanabe H, Tanaka S (2007) Visualization analysis of inter-fragment interaction energies of CRP–cAMP–DNA complex based on the fragment molecular orbital method. *Biophys Chem* 130: 1–9.
13. Iwata T, Fukuzawa K, Nakajima K, Aida-Hyugaji S, Mochizuki Y, Watanabe H, Tanaka S (2008) Theoretical analysis of binding specificity of influenza viral hemagglutinin to avian and human receptors based on the fragment molecular orbital method. *Comput Biol Chem* 32:198-211.
14. Mazanetz MP, Ichihara O, Law RJ, Whittaker M (2011) Prediction of cyclin-dependent kinase 2 inhibitor potency using the fragment molecular orbital method. *J Cheminf* 3:2.
15. Scharner J, Aznarez I (2020). Clinical Applications of Single-Stranded Oligonucleotides: Current Landscape of Approved and In-Development Therapeutics. *Mol Ther* 29:540-554.
16. Bathe M, Rothmund PWK (2017) DNA Nanotechnology: A foundation for Programmable Nanoscale Materials. *MRS Bulletin* 42:882–888.
17. Gutmanas A, Alhroub Y, Battle GM, Berrisford JM, Bochet E, Conroy MJ, Dana JM, Fernández Montecelo MA, van Ginkel G, Gore SP, Haslam P, Hatherley R, Hendrickx PMS, Hirshberg M, Lagerstedt I, Mir S, Mukhopadhyay A, Oldfield TJ, Patwardhan A, Rinaldi L, Sahni G, Sanz-García E, Sen S, Slowley RA, Velankar S, Wainwright ME, Kleywegt GJ (2014) PDBe: Protein Data Bank in Europe. *Nucleic Acids Res* 42:D285.
18. van Dijk M, Bonvin AM (2009) 3D-DART: a DNA structure modelling server. *Nucleic Acids Res.* 37(suppl 2):W235-W239.
19. El Hassan MA, Calladine CR (1997) Conformational characteristics of DNA: empirical classifications and a hypothesis for the conformational behaviour of dinucleotide steps. *Philos T Roy Soc A* 335:43-100.
20. Li H, Fedorov DG, Nagata T, Kitaura K, Jensen JH, Gordon MS (2010) Energy gradients in combined fragment molecular orbital and polarizable continuum model (FMO/PCM) calculation. *J Comput Chem* 31:778-790.
21. Schmidt MW, Baldrige KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery Jr JA (1993) General atomic and molecular electronic structure system. *J Comput Chem* 14:1347-1363.
22. Suenaga M (2005) Facio: New Computational Chemistry Environment for PC GAMESS. *J Comput Chem* 4:25-32.
23. Jmol: an open-source Java viewer for chemical structures in 3D. <http://www.jmol.org/>
24. Kitaura K, Ikeo E, Asada T, Nakano T, Uebayasi M (1999) Fragment molecular orbital method: an approximate computational method for large molecules. *Chem Phys Lett* 313:701–706.
25. Nakano T, Kaminuma T, Sato T, Fukuzawa K, Akiyama Y, Uebayasi M, Kitaura K (2002) Fragment molecular orbital method: use of approximate electrostatic potential. *Chem Phys Lett* 351:475–480
26. Fedorov DG, Kitaura K (2007) Pair interaction energy decomposition analysis. *J Comput Chem* 28:222-237.
27. Tomasi J, Mennucci B, Cammi R (2005) Quantum mechanical continuum solvation models. *Chem Rev* 105:2999-3093.

28. Fedorov DG, Kitaura K, Li H, Jensen JH, Gordon MS (2006) The Polarizable Continuum Model (PCM) Interfaced with the Fragment Molecular Orbital Method (FMO). *J Comput Chem* 27:976-985.
29. Fedorov DG, Kitaura K (2012) Energy Decomposition Analysis in Solution Based on the Fragment Molecular Orbital Method. *J Phys Chem A* 116:704-719.
30. Rajak SK, Islam N, Ghosh DC (2011) Modeling of the Chemico-Physical Process of Protonation of Molecules Entailing Some Quantum Chemical Descriptors. *J Quantum Inf. Sci* 1:87-95.
31. Origin(Pro), Version Number (2020b). OriginLab Corporation, Northampton, MA, USA.
32. González-Olvera JC, Martínez-Reyes J, González-Jasso E, Pless RC (2015) Determination of pK_a values for deprotonable nucleobases in model oligonucleotides. *Biophys Chem* 206:58-65.
33. Isaksson J, Acharya S, Barman J, Cheruku P, Chattopadhyaya J (2004) Single-stranded adenine-rich DNA and RNA retain structural characteristics of their respective double-stranded conformations and show directional differences in stacking pattern. *Biochemistry* 43:15996-16010.
34. Capobianco A, Caruso T, Peluso A (2015) Hole delocalization over adenine tracts in single stranded DNA oligonucleotides. *Phys Chem Chem Phys* 17:4750-4756.
35. Chakraborty D, Hori N, Thirumalai D (2018) Sequence-Dependent Three Interaction Site Model for Single- and Double-Stranded DNA. *J Chem Theory Comput* 14:3763-3779.
36. Capobianco A, Velardo A, Peluso A. (2018) Single Stranded DNA Oligonucleotides Retain Rise Coordinates Characteristic of Double Helices. *J Phys Chem B* 122:7978–7989.
37. Plumridge A, Meisburger SP, Andresen K, Pollack L (2017) The impact of base stacking on the conformations and electrostatics of single-stranded DNA. *Nucleic Acids Res* 45:3932–3943.
38. Heck A, Woiczikowski, PB, Kubař T, Welke K, Niehaus T, Giese B, Skourtis S, Elstner M, Steinbrecher TB (2014) Fragment Orbital Based Description of Charge Transfer in Peptides Including Backbone Orbitals. *J Phys Chem B* 118:4261–4272.
39. Fujita T, Mochizuki Y (2018) Development of the Fragment Molecular Orbital Method for Calculating Nonlocal Excitations in Large Molecular Systems. *J Phys Chem A* 122:3886–3898.
40. Chakraborty K, Mantha S, Bandyopadhyaya S (2013) Molecular dynamics simulation of a single-stranded DNA with heterogeneous distribution of nucleobases in aqueous medium. *J Chem Phys* 139:075103.
41. Chakraborty K, Bandyopadhyay S (2014) Correlated Dynamical Crossovers of the Hydration Layer of a Single-Stranded DNA Oligomer. *J Phys Chem B* 118:413-422.
42. Šponer J, Mládek A, Šponer JE, Svozil D, Zgarbová M, Banáš P, Jurečka P, Otyepka M (2014) The DNA and RNA sugar–phosphate backbone emerges as the key player. An overview of quantum-chemical, structural biology and simulation studies. *Phys Chem Chem Phys* 14:15257–15277.