Biophysical Chemistry 233 (2018) 26–35

Contents lists available at ScienceDirect

Biophysical Chemistry

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

Electronic properties of DNA: Description of weak interactions in TATA-box-like chains

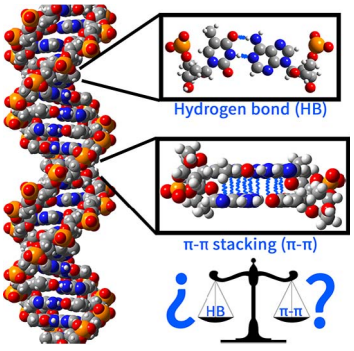
Jorge Gutiérrez-Flores[a], Estrella Ramos[a,*], Carlos I. Mendoza[a], Enrique Hernández-Lemus[b]

[a] Instituto de Investigaciones en Materiales, Universidad Autónoma de México, Circuito Exterior s/n, Ciudad Universitaria, Coyoacán, 04510 Mexico City, Mexico
[b] Computational Genomics, National Institute of Genomic Medicine, Mexico City, Mexico

HIGHLIGHTS

- The influence of the chemical environment on hydrogen bonds and π-π stacking in TATA-box-like B-DNA chains is studied.
- Novel strategies are proposed to estimate the non-covalent interactions energies present in B-DNA.
- An analysis of the contribution of hydrogen bond and pi-pi stacking to B-DNA stabilization is made.

GRAPHICAL ABSTRACT



ABSTRACT

DNA is one of the most important biomolecules since it contains all the genetic information about an organism. The tridimensional structure of DNA is a determinant factor that influences the physiological and biochemical mechanisms by which this molecule carries out its biological functions. It is believed that hydrogen bonds and π-π stacking are the most relevant non-covalent interactions regarding DNA stability. Due to its importance, several theoretical works have been made to describe these interactions, however, most of them often consider only the presence of two nitrogenous bases, having a limited overview of the participation of these in B-DNA stabilization. Furthermore, due to the complexity of the system, there are discrepancies between which involved interaction is more important in duplex stability. Therefore, in this project we describe these interactions considering the effect of chain length on the energy related to both hydrogen bonds and π-π stacking, using as model TATA-box-like chains with n base pairs ($n = 1$ to 14) and taking into consideration two different models: ideal and optimized B-DNA. We have found that there is a cooperative effect on hydrogen bond and π-π stacking mean energies when the presence of other base pairs is considered. In addition, it was found that hydrogen bonds contribute more importantly than π-π stacking to B-DNA stability; nevertheless, the participation of π-π stacking is not negligible: when B-DNA looks for a conformation of lower energy, π-π stacking interaction are the first to be optimized. All work was realized under the framework of DFT using the DMol^3 code (M06-L/DNP).

* Corresponding author.
E-mail address: eramos@iim.unam.mx (E. Ramos).

https://doi.org/10.1016/j.bpc.2017.11.008
Received 11 September 2017; Received in revised form 28 November 2017; Accepted 28 November 2017
Available online 12 December 2017
0301-4622/

1. Introduction

The hereditary basis of almost all life form is its genome, which consists of a long DNA chain that provides a complete set of all genetic information [1]. It is estimated that each of our cells has, approximately, 1.5 gigabytes of genetic information stored in this biomolecule [2]. Hence, DNA is considered as one of the molecules of greater biological relevance, since from this is possible to get several functional products (as proteins and different types of RNA) necessary for many biological processes.

DNA is conformed by nucleotides which are formed by a sugar, a phosphate group and a nitrogenous base: adenine (A), guanine (G), thymine (T) and cytosine (C). Through phosphodiester bonds between phosphate groups and sugars, these units lead to the formation of strands that interact non-covalently with each other to originate different DNA structures [3]. Under normal physiological conditions, DNA adopts a so-called B-DNA conformation [2,4] whose structure and stability is mainly determined by the number and type of weak interactions. There are different kinds of weak interactions, nevertheless, it is believed that hydrogen bonds (between nitrogenous bases) and π-π stacking (between base pairs) are the most relevant for DNA stabilization [5]. Therefore, understanding these interactions would explain the physical and chemical properties of DNA, and its role in biological processes involving this molecule.

Due to the importance of π-π stacking and hydrogen bonds in stabilizing DNA duplex, several theoretical studies have been made with different approximations in order to estimate their energy [6–13]. Most of them consider only the presence of two nitrogenous bases, having a limited overview of the participation of these interactions in B-DNA stabilization.

On the other hand, due to the complexity of the system, there are discrepancies between which of the two interactions (hydrogen bonds or π-π stacking) participates more significantly in the stabilization of DNA duplex. For example, according to Smirnov et al. [14] π-π stacking interactions are more important than hydrogen bonds (a similar conclusion was obtained by Sen et al. for PNA·DNA duplexes [15]), while Zhang et al. [16] establish that hydrogen bonds contributes more in the B-DNA stabilization. Consequently, it is not clear which one plays the most important role in the duplex stabilization or if both are equally relevant.

In view of the above, the goal of this study is to estimate the associated energy to both hydrogen bonds and π-π stacking interactions in order to get a better description of these, considering the effect of chain length (presence of other base pairs). Furthermore, with the obtained results, the contribution of each interaction to B-DNA stabilization is estimated.

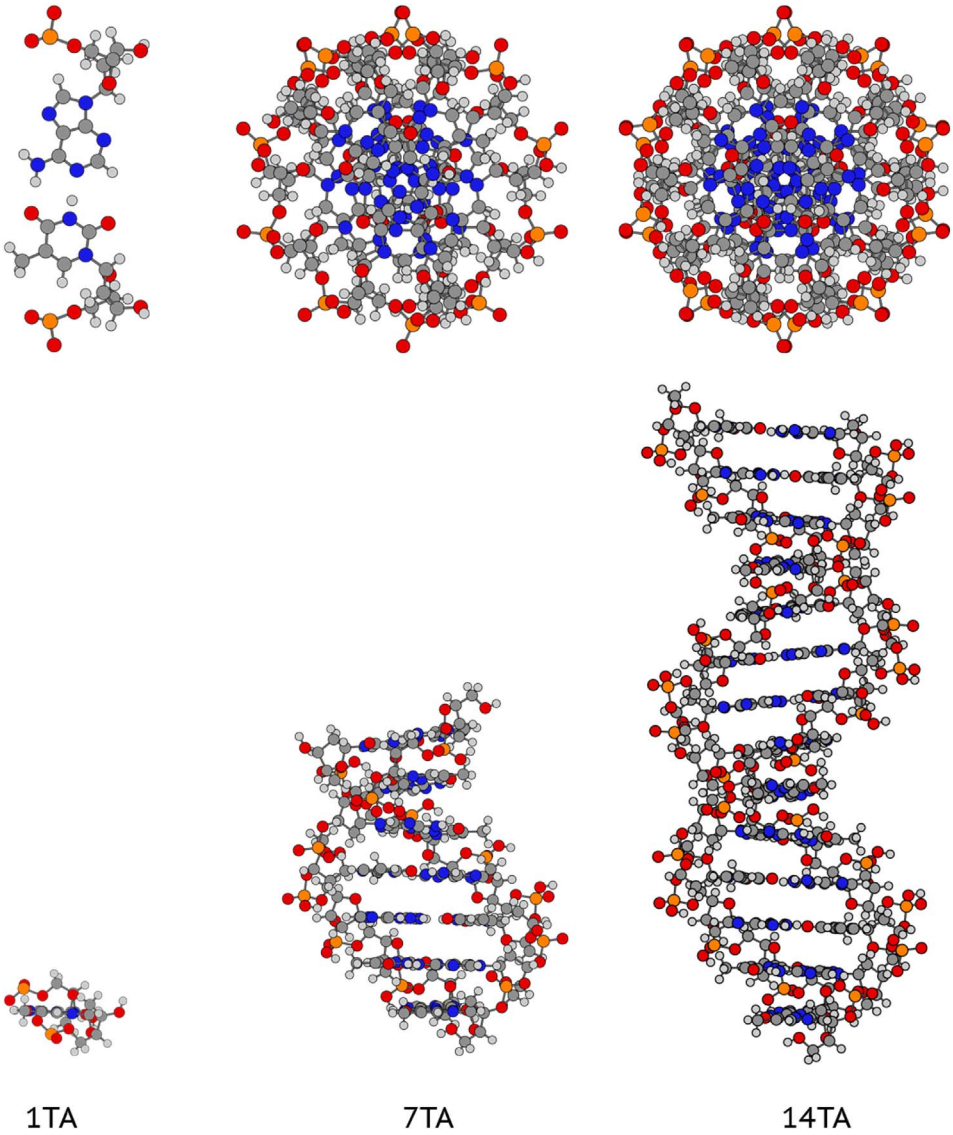


Fig. 1. Examples of B-DNA chains with different base pairs (*n*) as indicated. Frontal (up) and lateral (down) view.

2. Methods

To carry out this work, based on TATA box's sequence (a well known promoter region conserved in a wide range of eukaryotic and viral genes [17] which consists in a repetition of T and A bases), we modeled TATA-box-like chains of different length, i.e. different number of base pairs (n), from $n = 1$ to $n = 14$. The charge of the phosphate groups was neutralized by the addition of a hydrogen atom, assuming that the presence of these atoms would not affect significantly the hydrogen bond and π-π stacking interactions [11,18].

In order to estimate the energy associated to hydrogen bonds and π-π stacking, we took into account two different cases:

(a) Ideal
(b) Optimized

For the ideal case, structures were obtained with the open-source software *make-na* [19]. This software considers the distance between adjacent nucleobases and between phosphate groups, rotation and torsion angles, among other parameters (based on experimental data), in order to get ideal B-DNA chains, i.e., chains that obey the symmetry restrictions required for these systems (Fig. 1). Let us stress that even if these chains satisfy the symmetry restrictions, they are not relaxed into their minimum energy configuration; however, they provide a good starting point to estimate weak interaction energies. Subsequently, ideal structures were submitted to a geometry optimization process (Fig. 2) to get a better description of the stabilizing interactions present in the chain.

For hydrogen bond mean energy ($\bar{D}_{HB}$) estimation (see Fig. 3), single point energy calculations for each chain and for each of their strands were performed, and the following equation was proposed:

$$\bar{D}_{HB} = -\bar{E}_{HB} = -\frac{E_n - (E_{n_{1s}} + E_{n_{2s}})}{n} \quad (1)$$

where $\bar{E}_{HB}$ is the mean hydrogen bond formation energy; E_n represents the total energy of a n-base pair chain; $E_{n_{1s}}$ and $E_{n_{2s}}$ the total energy of chain's strands and n the number of base pairs.

The Espinosa-Molins-Lecomte (EML) equation [20] was used to estimate hydrogen bond energies:

$$E_{HB} = 0.5V(r) \quad (2)$$

Based on Quantum Theory of Atoms in Molecules (QTAIM) [21], this equation takes into account the electron distribution at $(3,-1)$ bond critical points (BCPs) of hydrogen bonds and correlates local potential energy density ($V(r)$) with hydrogen bond energy (E_{HB}).

Whereas for π-π staking mean energy ($\bar{D}_{\pi}$) estimation, assuming that the major contribution for this interaction corresponds to nitrogenous bases, phosphate groups and sugar moieties were removed from DNA chains, while hydrogen atoms were added instead, leaving nitrogenous bases in the chain configuration (see Fig. 4). Single point energy calculations were performed for each base pair and the following equation was employed:

$$\bar{D}_{\pi} = -\bar{E}_{\pi} = -\frac{E_{n\pi} - \sum_{i=1}^{n} E_{ibp}}{N_{\pi}} \quad (3)$$

where $\bar{E}_{\pi}$ is the mean π-π stacking formation energy; $E_{n_{\pi}}$ corresponds to the chain total energy without sugars and phosphate groups; $E_{i_{bp}}$ the total energy of base pairs and N_{π} the number of spaces between base pairs.

All calculations were carried out under Density Functional Theory (DFT) framework [22,23], using the DMol^3 code [24] implemented in the Materials Studio 8.0 software suite [25], employing the M06-L functional [26] (a meta-GGA functional that, among other things, was parameterized for the good description of weak interactions like π-π staking and hydrogen bonds) and a double numeric basis set with polarized functions (DNP). Basis set superposition error (BSSE) corrections were performed by means of counterpoise method [27] implemented in DMol^3 code. The occupied convergence tolerance for all calculations were 1×10^{-5} Ha, 2×10^{-5} Ha, 4×10^{-3} Ha/ and 5×10^{-3} for SFC cycles, energy, gradient and displacement, respectively.

For topology analysis, single point energy calculations were carried out using the Gaussian 09 suit of programs of computational chemistry [28] (M06-L/6-31G(d,p)). Finally, electron density was analyzed utilizing Multwfn software [29].

3. Results and discussion

It is important to mention that ideal B-DNA chains are built by imposing symmetry restrictions [19,30], obtaining chains such as those

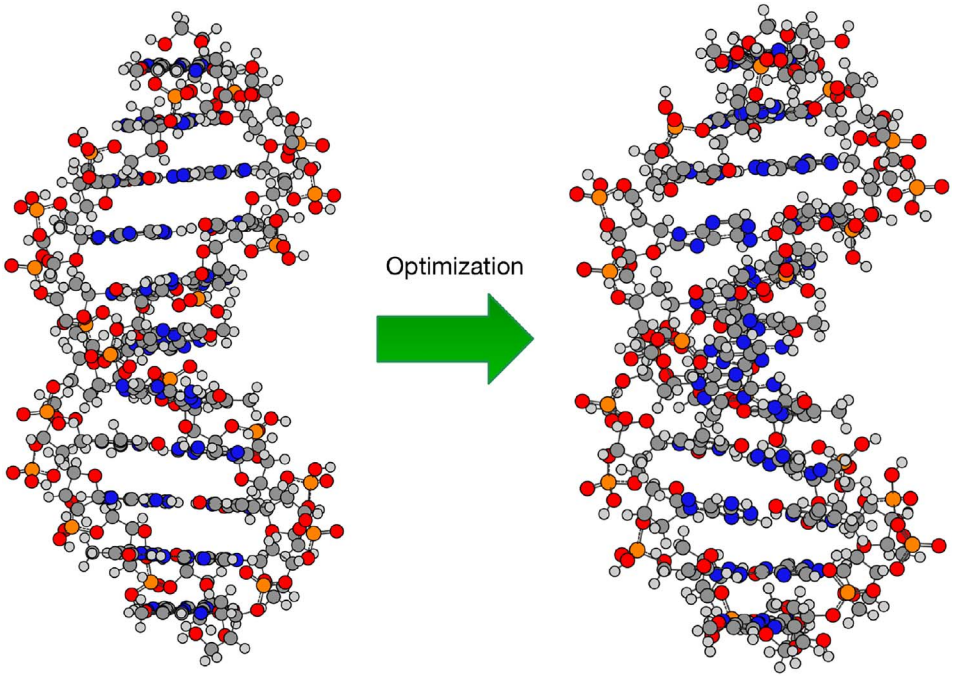


Fig. 2. Lateral view of a B-DNA chain before (left) and after (right) optimization.

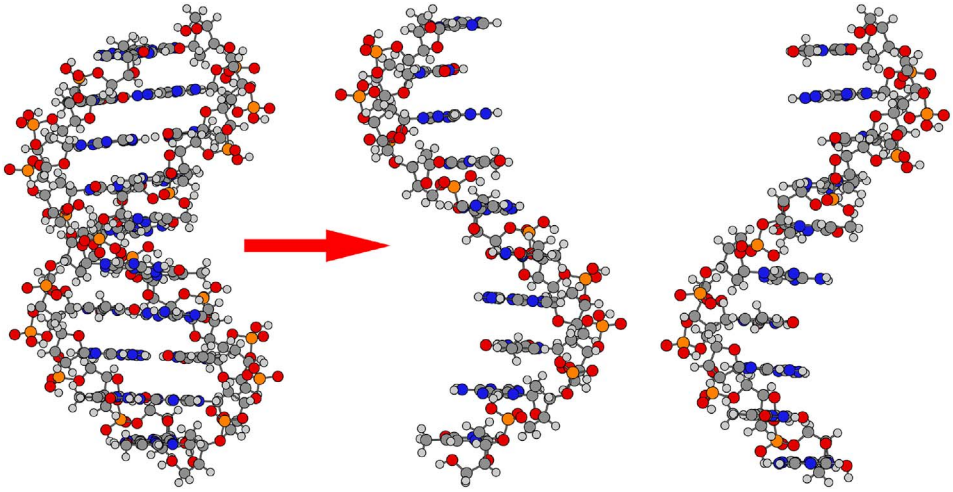


Fig. 3. Illustration of how hydrogen bond mean energy ($\bar{D}_{HB}$) is estimated.

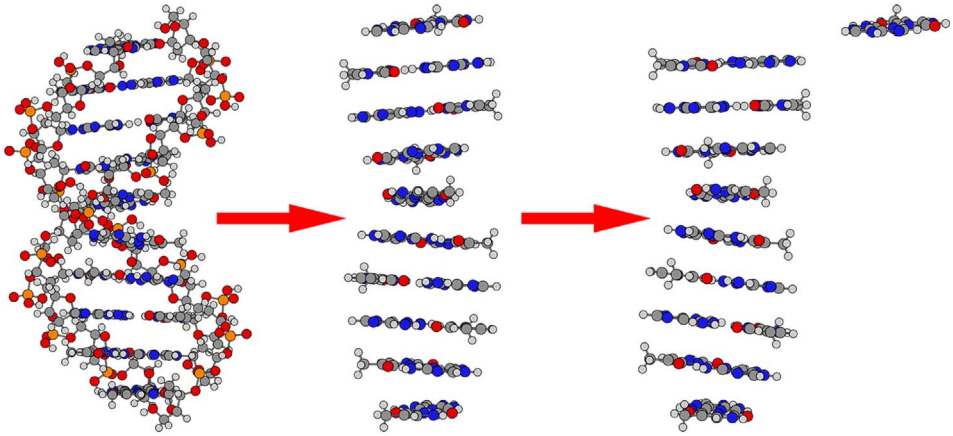


Fig. 4. Illustration of how π-π staking mean energy ($\bar{D}_{\pi}$) is estimated.

shown in Fig. 1. In this figure, it is possible to observe a C_{10} axis of symmetry and total planarity of nitrogenous bases. Due to the relation between tridimensional arrangement of nitrogenous bases and the interactions involved in double helix stabilization [30], a change in conformation will affect the efficiency and strength of the latter. For that reason, it is interesting to study the effect of the rearrangement of the chain (optimization process) on hydrogen bonds and π-π stacking.

As mentioned before, it is considered that hydrogen bonds and π-π stacking interactions are the most important in B-DNA stabilization [5]. For that reason, we start with the estimation of interaction energies and the analysis of chain length effect over these. In a later step, a charge distribution analysis is made and, finally, the relative contribution of these interactions to B-DNA stabilization is estimated.

3.1. Hydrogen bonds

Eq. (1) was used in order to estimate the hydrogen bond mean energy ($\bar{D}_{HB}$), i.e. the required energy to break up these interactions. When calculating the mean hydrogen bond energy by means of Eq. (1), we are neglecting the presence of other noncovalent interactions when strands are separated like shown in Fig. 3. In order to validate this approach, we made a topology analysis of TATA-box-like chain with $n = 7$ (as an example) and estimated the hydrogen bond mean energy using Eq. (2) (see Tables S1 and S2 in SI). As the energy is calculated with $V(r)$ at hydrogen bonds $(3, -1)$ BCPs, we assume that this corresponds only to hydrogen bonds. The difference between the energies calculated with Eqs. (1) and (2) (Table 1) is less than 3 kcal mol^{-1} (around 16%); therefore, we consider that Eq. (1) adequately estimates the hydrogen bond mean energy. Hydrogen bond mean energy as a function of the number of base pairs (n) is shown in Fig. 5 (values are available in Table S3 in the SI).

Table 1
Hydrogen bond mean energies calculated by means of Eqs. (1) (first two columns) and (2) (last two columns) for chains with $n = 7$.

$\bar{D}_{HB}$ (kcal mol^{-1})			
Eq. (1)		Eq. (2)	
Ideal	Optimized	Ideal	Optimized
17.84	17.65	15.01	14.44

As seen in Fig. 5, $\bar{D}_{HB}$ increases as n does. Consequently, the presence of other base pairs tends to raise the hydrogen bond energy, i.e. there is a cooperative effect. From the same figure, it is clear that the curve representing the ideal case has the greatest hydrogen bond mean energies. When the structure is relaxed, the planarity of base pairs is lost. As a consequence, the strength of hydrogen bonds is altered, resulting in weaker hydrogen bonds for optimized chains. Nevertheless, hydrogen bond mean energy differences between ideal and optimized cases are less than 1 kcal mol^{-1} and, when number of base pairs increases difference, tends to vanish. These differences could be associated with the fact that when ideal structure is submitted to a geometry optimization process, base pairs are distributed according to neighbors pairs. Therefore, as n grows the movements of base pairs are more restricted and the energy estimated is more similar to ideal case.

Oscillations present in curves (Fig. 5) could be attributed to finite size effects. Indeed, by symmetry restrictions in an ideal chain with a very large number of base pairs, any TA or AT base pair would be

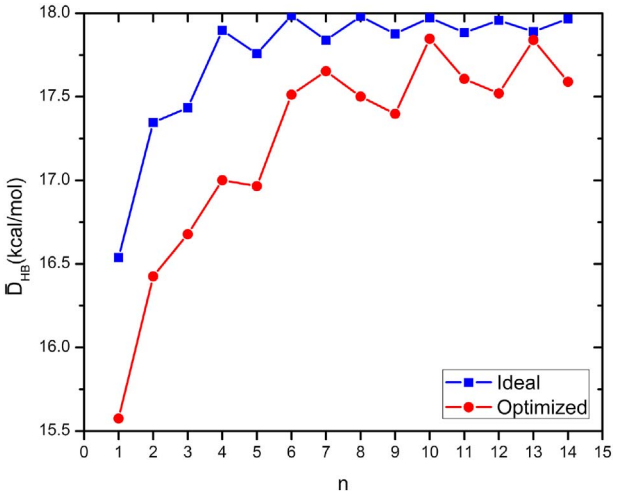


Fig. 5. Hydrogen bond mean energy as a function of DNA chain length. Blue symbols correspond to ideal chains while red symbols to optimized structures. Lines connect the points for clarity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

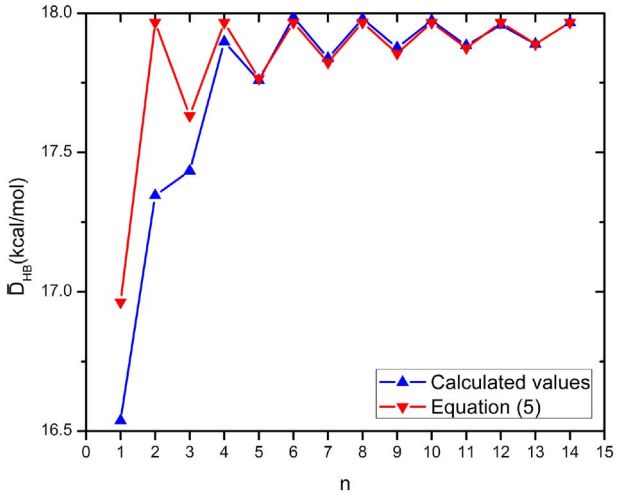


Fig. 6. Hydrogen bond mean energy as a function of the DNA chain size (represented by n) for the ideal case: theoretical values (blue) and analytical development (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

equivalent. In the case of a finite length ideal B-DNA chain, the units at the extremes will contribute with a different quantity to hydrogen bond energy.

Using n as the total number of base pairs and distinguishing chains by its parity, the hydrogen bond mean energy can be written as

$$\bar{D}_{HB} = \begin{cases} \epsilon_0 + \dfrac{\Delta E_{even}}{n} & n \quad \text{even} \\ \epsilon_0 + \dfrac{\Delta E_{odd}}{n} & n \quad \text{odd} \end{cases} \tag{4}$$

where ϵ_0 is the interaction energy between the nitrogenous bases and ΔE_{even} and ΔE_{odd} are the contribution from the base pairs at the edges of the chain (border effects), when the number of base pairs is even or odd, respectively. The numerical values of the edge effects are unknown, however, from the calculated values (Fig. 5 and Table S1), it can be observed that $\bar{D}_{HB}$ takes an approximately constant value for even n, with $n \geq 4$, suggesting that $\Delta E_{even} \approx 0$.

Since we expect that this analysis would be more accurate for large chains, to evaluate ϵ_0 and ΔE_{odd} we take into account the values of the largest chains ($n = 13$ and $n = 14$). We have that $E_{n=13} \approx 17.89$ kcal mol^{-1}, and $E_{n=14} \approx 17.97$ kcal mol^{-1}, consequently, from Eq. (4) we get

$$\epsilon_0 = \bar{D}_{HB} \approx 17.97 \text{ kcal mol}^{-1}$$

and

$$\Delta E_{odd} = n(\bar{D}_{HB} - \epsilon_0) \approx -1.01 \text{ kcal mol}^{-1}$$

Hence, with values of ϵ_0 and ΔE_{odd}, Eq. (4) can be rewritten as

$$\bar{D}_{HB} = \begin{cases} 17.97 \text{ kcal mol}^{-1} & n \quad \text{even} \\ 17.97 \text{ kcal mol}^{-1} - \dfrac{1.01 \text{ kcal mol}^{-1}}{n} & n \quad \text{odd} \end{cases} \tag{5}$$

The values obtained from Eq. (5) (Table S4 in SI) are shown in Fig. 6. As can be observed, this expression adjusts satisfactorily to the estimated values for the hydrogen bond mean energy for $n > 3$. The reason why the analytical development fails for small n is that it assumes that chain edges do not affect the hydrogen bond mean energy and this is only valid in larger chains.

From the results shown before (Figs. 5 and 6), it is clearly important to consider more than 4 base pairs for the adequate estimation of the hydrogen bond energy; otherwise, hydrogen bond energy could be underestimated.

In Table 2 we show the values of the hydrogen bond mean distances between nitrogenous bases for the ideal and optimized chains. Notation used to name interactions between A and T is given in Fig. 7. It is important to remark that we considered only the presence of two hydrogen bonds (with separation distances R_1 and R_2, respectively).

Table 2
Hydrogen bond mean distances for the studied cases. (ideal chains: $\sigma \leq 9.00 \times 10^{-4}$ Å; optimized chains: $\sigma \leq 1.40 \times 10^{-2}$ Å).

	Ideal case		Optimized case	
TA Units	R_1 ()	R_2 ()	R_1 ()	R_2 ()
1	1.871	1.810	1.884	1.835
2	1.870	1.810	1.856	1.833
3	1.870	1.810	1.868	1.819
4	1.870	1.810	1.869	1.825
5	1.870	1.810	1.867	1.816
6	1.870	1.810	1.868	1.810
7	1.870	1.810	1.882	1.799
8	1.870	1.810	1.868	1.810
9	1.870	1.810	1.855	1.823
10	1.870	1.810	1.889	1.801
11	1.870	1.810	1.862	1.816
12	1.870	1.810	1.851	1.826
13	1.870	1.810	1.878	1.802
14	1.870	1.810	1.855	1.825

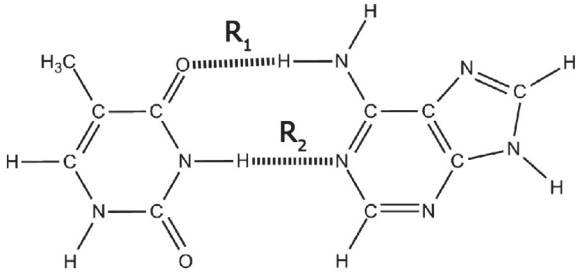


Fig. 7. Sketch indicating the distances between the atoms for, both, A and T.

In general, from Table 2, it can be noticed that central hydrogen bonds (R_2) are stronger than the outer ones (R_1), i.e., they have smaller associated distances for a given chain (in compliance with discussion in Ref. [31]). Distances are in concordance with calculated values using Eq. (2) (see Tables S1 and S2 in SI). Calculated distances obtained for both cases have the same trend of experimentally reported values [32]

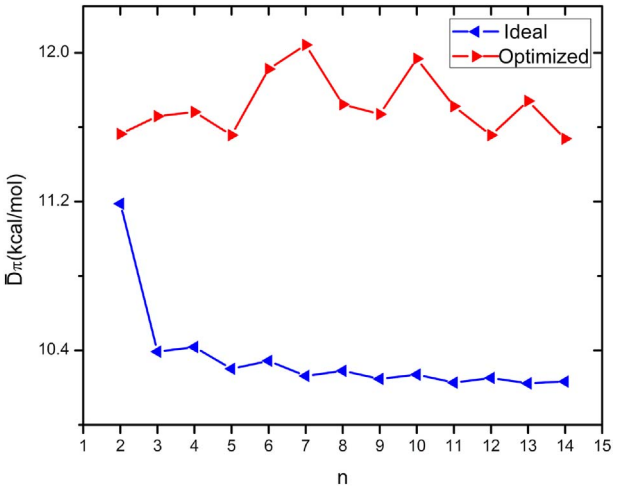


Fig. 8. π-π stacking mean energy as a function of DNA chain length. Blue symbols correspond to ideal chains while red symbols to optimized structures. Lines connect the points for clarity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(2.93–2.95 for R_1 and 2.82–2.85 for R_2). Difference of ≈1 Å between the calculated data and the reported experimentally is expected due to hydrogen atoms are not observable by X-ray diffraction. Additionally, these differences could be associated to errors related to the calculation method used (there is not a explicit correction for the noncovalent interactions) and to the fact that all calculations were performed in gas phase. For ideal chains, as it can be expected, mean distances are essentially the same for all the chains. For all of these chains symmetry restrictions are satisfied, thus distance will not change with the addition of more base pairs.

3.2. π-π *stacking*

π-π stacking usually refers to attractive interactions between π systems [33] and, therefore, are often present in unsaturated organic groups [34]. They depend on both geometry and force factors that favor the juxtaposition of two aromatic molecules, in order to allow the direct contact between their corresponding π systems [5]. Eq. (3) was used to estimate the energy related to this interactions along TATA-box-like DNA chains (between base pairs). Results are shown in Fig. 8 (values are available in Table S5 in SI).

In contrast to hydrogen bonds (Fig. 5), π-π stacking interactions are favored when geometry optimization is carried out, i.e. π-π stacking mean energy is greater in optimized chains than in ideal chains (on average, difference between ideal and optimized cases is around 1.36 kcal mol^{-1}). As mentioned before, ideal chains have smaller hydrogen bond associated distances. Hence, the base pairs are arranged in order to maximize the contact between them, and, therefore, the π-π stacking mean energy is increased.

Due to differences between hydrogen bond and π-π stacking mean energies in optimized cases, it can be suspected that when B-DNA reaches its minimum energy configuration, π-π stacking interactions are optimized first, before hydrogen bonds. This may be the reason why Smirnov et al. [14] concluded that π-π stacking contribution was more important than that proceeding from hydrogen bonds when DNA polymerase was able to add non-polar nitrogenous bases (bases without the capability to interact with other nitrogenous base by means of hydrogen bonds) during DNA replication process.

3.3. *Charge transfer*

Aiming to a deeper examination of hydrogen bonds, an analysis of the electronic density accumulation as a function of chain length was made over regions where hydrogen bonds are present between A and T bases (N-H···X, where X = N,O; see Fig. 7). In an intermolecular interaction the atomic charge is shared between them, i.e., part of the electropositive atoms charge is transferred to more electronegative atoms. Hence, the larger the charge transfer, the larger the hydrogen bond strength.

For this purpose, Charge Separation Index (CSI), a measure of the local polarity of a molecule, is obtained from the sum of absolute values of charges q of each atom i in the region of interest [35]:

$$CSI = \sum_i |q_i| \tag{6}$$

A different way to interpret this descriptor is as the measure of the degree of separation between negative and positive charges of region of interest [36]. The larger the CSI value, the larger the polarization in that region and larger the hydrogen bond strength.

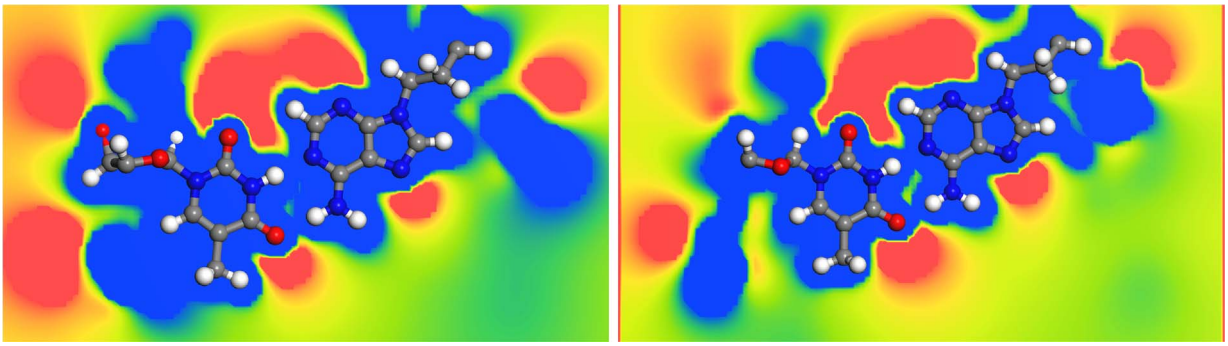
To obtain this descriptor, we performed a Mulliken population analysis. We considered chains of $n = 1$, 3, 5 and 7, since only chains with an odd number of base pairs have a central one which is equidistant to both chain edges which is convenient to reduce their influence as much as possible.

In Table 3 we show CSI descriptor values for N-H···X (where X = N,O) regions. Values are organized according to the nomenclature introduced in Fig. 7. It can be observed that for the two cases studied, the central hydrogen bond (R_2) is stronger than the external one (R_1). This fact is in accord with distances analyzed before (Table 2) and with the topology analysis performed for ideal and optimized chains with $n = 7$ (Tables S1 and S2 in SI). Other works have reported the same behavior [7].

Due to system symmetry, as it could be expected, differences in CSI values in ideal chains are not significantly large. However, it can be

Table 3
Charge Separation Index (CSI) for hydrogen bond regions (N-H···X, where X = N, O; see Fig. 7) between T and A bases.

Chain	Base pair	Interaction	CSI	
			Ideal	Optimized
1TA	1	R_1	1.426	1.422
		R_2	1.589	1.531
3TA	1	R_1	1.426	1.416
		R_2	1.580	1.521
	2	R_1	1.413	1.417
		R_2	1.576	1.512
	3	R_1	1.427	1.435
		R_2	1.581	1.530
5TA	1	R_1	1.426	1.413
		R_2	1.579	1.522
	2	R_1	1.412	1.419
		R_2	1.577	1.520
	3	R_1	1.419	1.418
		R_2	1.572	1.521
	4	R_1	1.415	1.412
		R_2	1.575	1.509
	5	R_1	1.425	1.436
		R_2	1.581	1.528
7TA	1	R_1	1.425	1.405
		R_2	1.578	1.520
	2	R_1	1.411	1.410
		R_2	1.575	1.529
	3	R_1	1.414	1.413
		R_2	1.573	1.522
	4	R_1	1.414	1.410
		R_2	1.575	1.520
	5	R_1	1.417	1.411
		R_2	1.572	1.527
	6	R_1	1.414	1.408
		R_2	1.575	1.519
	7	R_1	1.425	1.416
		R_2	1.581	1.536



(a) Ideal (b) Optimized

Fig. 9. Electrostatic potential planes for central base pairs of ideal (a) and optimized (b) chains with $n = 7$ (scale: from −0.04 a.u. (red) to 0.04 a.u. (blue)). Top view. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

noted that when these structures are relaxed (optimized cases) there is a loss of charge (CSI values decrease), and hydrogen bonds tend to be weaker. This is in agreement with hydrogen bond mean energies analysis (Fig. 5), where hydrogen bonds in optimized cases were weaker than those in ideal cases due to a loss of planarity.

3.4. Electrostatic potentials

With the aim to complete the analysis made before for hydrogen bond and π-π stacking interactions, electrostatic potentials and molecular electrostatic potentials are analyzed in order to shed light on charge distribution along studied B-DNA chains. For potential electrostatic planes, central base pairs are studied in order to avoid border effects.

As previously highlighted, when chains are relaxed, the hydrogen bond mean energy tends to decrease (Fig. 5 and Table 3) due to base pairs planarity loss. Likewise, this can be observed in the electrostatic potential planes showed in Fig. 9. In this figure, a top view of the electrostatic potential for central base pair of ideal and optimized chains with $n = 7$ is shown. As can be seen, in the optimized case there are more negative (rich in electrons) zones around hydrogen bonds than in ideal case, this is more clear in oxygen atom of N-H···O hydrogen bond. As stronger hydrogen bonds are associated with ideal chains, electrons of electronegative atoms are more shared in ideal chains hydrogen bonds, therefore the associated area will be smaller. Hence, planarity of base pairs plays an important role in how electrons are shared and, consequently, in hydrogen bond strength.

Contrary to hydrogen bonds, as showed before (Fig. 8), π-π stacking interactions are favored when chains are optimized. This is reflected in Fig. 10, where a lateral view of central base pairs of ideal and optimized chains with $n = 7$ is shown. Negative zones between base pairs are bigger in optimized case, indicating that contact of π electron density is greater in this case. To get a better overview, electrostatic potentials for some ideal and optimized chains with an odd number of base are shown in Figs. 11 and 12, respectively. Yellow regions correspond to electrostatic potential minima (negative zones) while the blue ones correspond to maxima. From these figures, it is clear that electron density overlap between base pairs is greater in optimized chains. Therefore, the presence of other base pairs tends to optimize, in first place, π-π stacking.

Based on the results shown previously, it could be possible to confirm that π-π stacking interactions are firstly improved along DNA chain and lately hydrogen bonds. That is the reason why DNA polymerase has the capability to add nitrogenous bases during DNA replication despite bases are polar o non-polar, as reported by Smirnov et al. [14].

3.5. Contribution to stabilization

Finally, we determined the contributions of hydrogen bonds and π-π stacking in B-DNA stabilization, assuming that these are the most important non-covalent interactions along DNA duplex. To this end, we used total hydrogen bond energy (required energy to separate strands in DNA duplex):

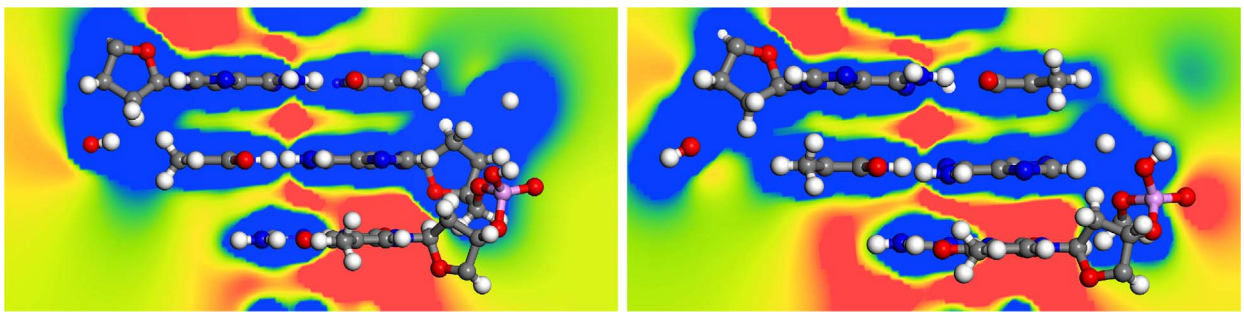
$$D_{HB} = -E_{HB} = -(E_n - (E_{n_{1s}} + E_{n_{2s}})) \tag{7}$$

and total π-π stacking energy (required energy to separate all base pairs that form the chain):

$$D_{\pi} = -E_{\pi} = -E_{n_{\pi}} - \sum_{i=1}^{n} E_{ibp} \tag{8}$$

Then, the contribution rate was estimated using the following equation:

$$Contribution = \frac{D_i}{D_{HB} + D_{\pi}} \cdot 100 \tag{9}$$



(a) Ideal (b) Optimized

Fig. 10. Electrostatic potential planes for central base pairs of ideal (a) and optimized (b) chains with $n = 7$ (scale: from −0.04 a.u. (red) to 0.04 a.u. (blue)). Lateral view. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

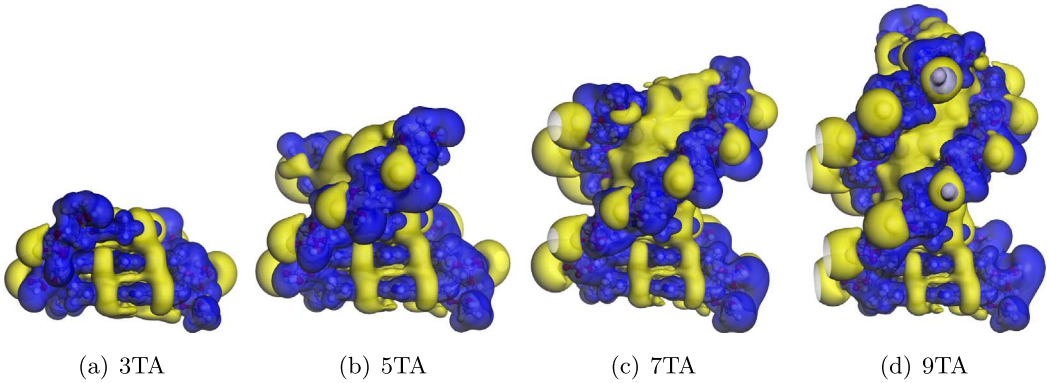


Fig. 11. Molecular electrostatic potential maps for ideal chains. Yellow regions represent minima, while blue ones represent maxima. Isovalue = 0.03 a.u. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

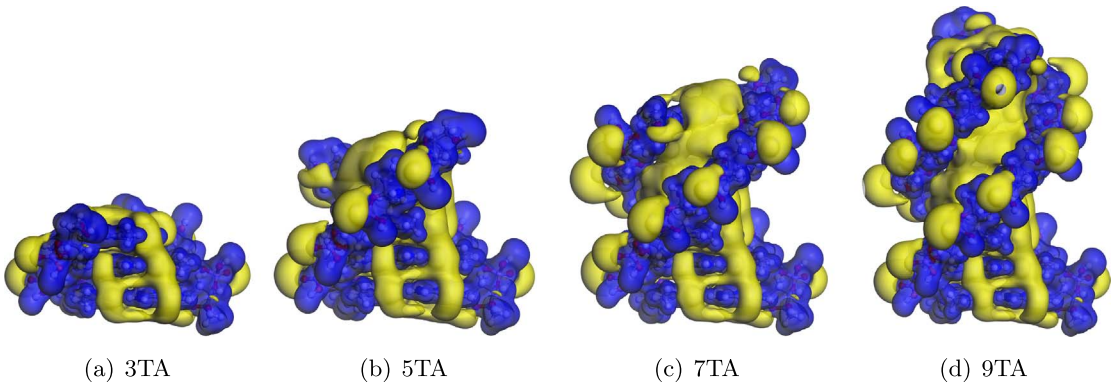


Fig. 12. Molecular electrostatic potential maps for optimized chains. Yellow regions represent minima, while blue ones represent maxima. Isovalue = 0.03 a.u. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

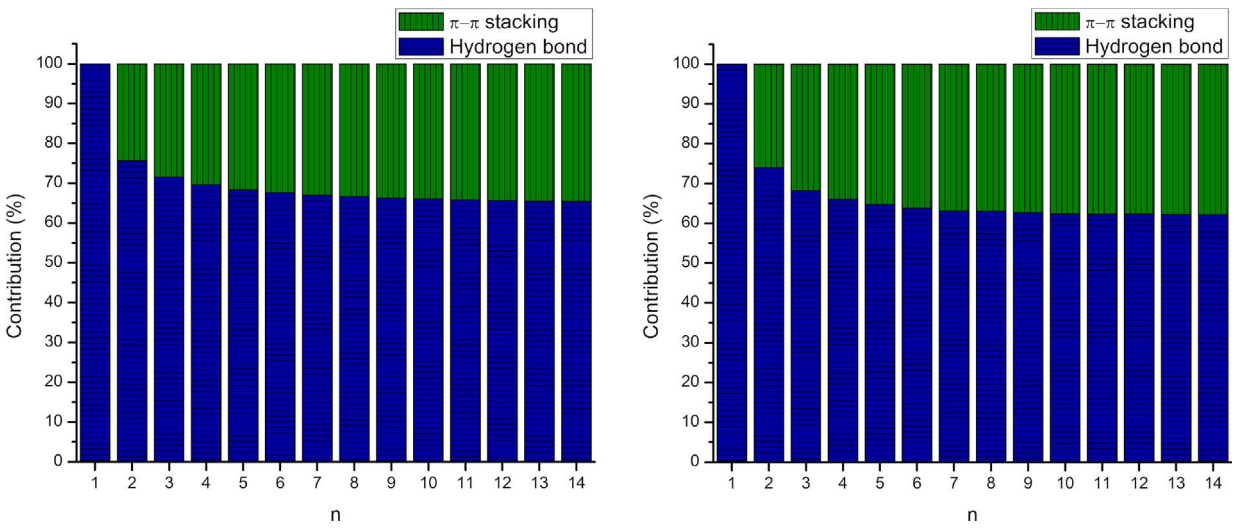


Fig. 13. Contribution rate of hydrogen bond and π-π stacking interactions for ideal (a) and optimized (b) cases.

where D_i represents D_{HB} or D_{π}, depending on which energy is being described. Results are shown in Fig. 13.

For both ideal and optimized cases, hydrogen bonds has the greater contribution to B-DNA stabilization, in concordance with what was reported by Zhang et al. [16]. However, the contribution of π-π stacking is not negligible.

Discrepancies between which interaction is more important in B-DNA stabilization could be related in how they are examined. For example, when B-DNA is studied out of biological processes (like Atomic Force Microscopy (AFM) [16], for example), hydrogen bonds predominate over π-π stacking due to B-DNA strands are separated artificially. On the other hand, when interactions are studied by means of biological processes (like DNA replication [14], for example) π-π stacking is considered more important than hydrogen bond because of, in this approximation, interactions contribution is determined in how an enzyme adds a nucleotide and how this is stabilized in the growing

chain. All the mentioned before is in agreement with our results: π-π stacking interactions are firstly optimized, explaining why DNA polymerase is able to add nucleotides through π-π staking no matter if nucleotides can or not form hydrogen bonds (as found by Smirnov et al. [14]). Nevertheless, when interactions are studied as proposed in Figs. 3 and 4 it is found that hydrogen bonds are more important than π-π stacking for B-DNA stabilization as reported by Zhang et al. [16]. All depends on the context and approximation used to determine interaction contributions.

It would be very interesting to contrast our results with experimental data. However, at the best of our knowledge, no experimental data of TATA-like-box chains is available to make comparisons with such results.

We have some proposals of experimental designs that may provide quantitative comparisons with our work. To this end we may look at a number of techniques which have resulted useful to determine experimentally the energetics of DNA duplex formation. A first common step to all these methods would consist in the laboratory synthesis of short DNA sequences with the required amount of TA base pairs that may provide the genomic sample for all the different methods. One possible approach would be using thermal DNA denaturation, tracked by either isothermal titration calorimetry coupled with spectrophotometric assessment [37] or by the use of a compounded NanoDrop device combining spectrophotometric with fluorometric measurements on microvolumes [38]. Another approach is based in elastic properties of DNA sequences and the use of atomic force microscopy [39] to provide approximations based on DNA stretching [40,41].

In brief, once that synthetic sequence specific constructs are available there are a number of experimental techniques that may be used as experimental comparisons to our theoretical calculations. Most of these are based on well-established principles in DNA analytics.

4. Conclusions

In this work, we performed an analysis of the influence of chain length in both, hydrogen bonds and π-π stacking, assuming that these are the most important non-covalent interactions in B-DNA stabilization. For this purpose, TATA-box-like chains (ideal and optimized cases) were used.

For hydrogen bonds, we found that they become weaker when a geometry optimization process is carried out. Whereas, π-π stacking interactions tends to be stronger. Therefore, it is proposed that when B-DNA looks for a conformation of lower energy, π-π stacking interactions are the first to be optimized, lately hydrogen bonds. This fact could explain DNA polymerase ability to add non-polar nitrogenous bases during DNA replication, as reported by Smirnov et al. [14].

With our results, it is clear that the presence of other base pairs have a cooperative effect on both hydrogen bond and π-π stacking mean energies, i.e., as chain length increases, these interactions become stronger. Important periodic oscillations were observed only for hydrogen bond mean energies due to border effects. Eq. (5) is proposed to estimate the hydrogen bond mean energies of TATA-box-like chains of any size.

Finally, the contribution of each interaction to B-DNA stabilization was calculated. According to these results, hydrogen bonds contribute more importantly in duplex stabilization than π-π stacking, in agreement with that reported by Zhang et al. [16]. Nevertheless, the contribution of π-π stacking is not negligible. The disagreements between which interaction is more important are due to the method and approximation used to study these interactions.

Acknowledgment

We want to thank to Carmen Nina Pastor Colon, Ateet Dutt and Julio César Aguilar Cordero for their useful comments. We also thank the computational support of Joaquín Morales, Alberto López and Alejandro Pompa from IIM-UNAM and to DGTIC-UNAM for their supercomputing services. We acknowledge partial financial support from DGAPA, UNAM PAPIIT grant numbers IB-101712 (ER) and IN-110516 (CIM). The authors were also funded by grant number 179431/2012 from CONACyT, as well as Mexican Federal funding (grant number 173-2017) from the National Institute of Genomic Medicine (INMEGEN). JGF acknowledges support from CONACyT through the scholarship grant 336100. EHL also acknowledges support from the 2016 Marcos Moshinsky Research Chair in Physical Sciences.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bpc.2017.11.008.

References

[1] J.E. Kreb, E.S. Goldstein, S.T. Kilpatrick, Lewin's Genes, Jhons and Bartlett Publishers, 2010.
[2] D. Goodsell, DNA, (2001) http://www.rcsb.org/pdb/101/motm.do?momID=23 Retrieved January, 2017.
[3] T. Ohyama, DNA Conformation and Transcription, Springer US, 2005.
[4] J.D. Watson, F.H.C. Crick, Molecular structure of nucleic acids, Nature 171 (1953) 737–738.
[5] E.T. Kool, Hydrogen bonding, base stacking, and steric effects in DNA replication, Annu. Rev. Biophys. Biomol. Struct. 30 (2001) 1–22.
[6] C. Fonseca, F. Matthias, J. Snijders, E. Jan, Hydrogen bonding in DNA base pairs: reconciliation of theory and experiment, J. Am. Chem. Soc. 122 (2000) 4117–4128.
[7] H. Szatyłowicz, N. Sadlej-Sosnowska, Characterizing the strength of individual hydrogen bonds in DNA base pairs, J. Chem. Inf. Model. 50 (2010) 2151–2161.
[8] A. Asensio, N. Kobko, J. Dannenberg, Cooperative hydrogen-bonding in adenine-thymine and guanine cytosine base pairs. Density functional theory and Möller-Plesset molecular orbital study, J. Phys. Chem. A 107 (2003) 6441–6443.
[9] J. Grunenberg, Direct assessment of interresidue forces in Watson-Crick base pairs using theoretical compliance constants, J. Am. Chem. Soc. 126 (2004) 16310–16311.
[10] H. Dong, W. Hua, S. Li, Estimation on the individual hydrogen-bond strength in molecules with multiple hydrogen bonds, J. Phys. Chem. A 111 (2007) 2941–2945.
[11] C. Matta, N. Castillo, R. Boyd, Extended weak bonding interactions in DNA: π-stacking (base-base), base-backbone, and backbone-backbone interactions, J. Phys. Chem. B 110 (2006) 563578.
[12] P. Mignon, S. Loverix, J. Steyaert, P. Geerlings, Influence of the π-π interaction on the hydrogen bonding capacity of stacked DNA/RNA bases, Nucleic Acids Res. 2005 (2005) 17791789.
[13] M. Swart, T. van der Wijst, C. Fonseca Guerra, F.M. Bickelhaupt, π-π stacking tackled with density functional theory, J. Mol. Model. 13 (2007) 1245–1257.
[14] S. Smirnov, T.J. Matray, E.T. Kool, C. de los Santos, Integrity of duplex structures without hydrogen bonding: DNA with pyrene paired at abasic sites, Nucleic Acids Res. 30 (2002) 5561–5569.
[15] A. Sen, P.E. Nielsen, Hydrogen bonding versus stacking stabilization by modified nucleobases incorporated in PNA·DNA duplexes, Biophys. Chem. 141 (2009) 29–33.
[16] T. Zhang, C. Zhang, Z. Dong, Y. Guan, Determination of base binding strength and base stacking interaction of DNA duplex using atomic force microscope, Sci. Rep. 5 (2014) 9143.
[17] L. Tora, T. H.Th.Marc, The TATA box regulates TATA-binding protein (TBP) dynamics in vivo, Trends. Biochem. Sci. 6 (2010) 309–314.
[18] J. Duguid, V. Bloomfield, J. Benevides, G. Thomas Jr, Raman spectroscopy of DNA-metal complexes. I. Interactions and conformational effects of the divalent cations: Mg, Ca, Sr, Ba, Mn, Co, Ni, Cu, Pd, and Cd, Biophys. J. 65 (1993) 1916–1928.
[19] make-na. ©2006, 2011, 2012, 2013. James Stroud.
[20] E. Espinosa, E. Molins, C. Lecomte, Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities, Chem. Phys. Lett. 285 (1998) 170–173.
[21] R. Bader, Atoms in Molecules: A Quantum Theory, Oxford University Press, 1990.
[22] P. Hohenberg, W. Kohn, Inhomogeneous electron gas, Phys. Rev. 46 (1964) B864.
[23] W. Kohn, L. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. 140 (1965) A1133.
[24] B. Delley, From molecules to solids with the DMol3 approach, J. Chem. Phys. 113 (2000) 7756.
[25] Materials Studio, Accelrys, Biovia ©2014 Systèmes, Biovia Software Inc.
[26] Y. Zhao, D.G. Truhlar, A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions, J. Chem. Phys. 125 (2006) 194101.
[27] F.B. van Duijneveldt, J.G.C.M. van Duijneveldt-van de Rijdt, J.H. van Lenthe, State of the art in counterpoise theory, Chem. Rev. 94 (1994) 1873–1885.
[28] 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, D.J. Fox, Gaussian 09, Revision E.01, Gaussian, Inc., Wallingford CT, 2016.

[29] T. Lu, F. Chen, Multiwfn: a multifunctional wavefunction analyzer, J. Comp. Chem. 33 (2012) 580–592.

[30] C.R. Calladine, H.R. Drew, B.F. Luisi, A.A. Travers, Understanding DNA: The Molecule and How It Works, Elsevier Academic Press, 2004.

[31] P. Yurenko, J. Novotný, R. Sklenář, R. Marek, Exploring non-covalent interactions in guanine and xanthine-based model DNA quadruplex structures: a comprehensive quantum chemical approach, Phys. Chem. Chem. Phys. 16 (2014) 2072–2084.

[32] N.C. Seeman, J.M. Rosenberg, F.L. Suddath, J.J.P. Kim, A. Rich, RNA double-helical fragments at atomic resolution: I. The crystal and molecular structure of sodium adenylyl-3′,5′-uridine hexahydrate, J. Mol. Bio. 104 (1976) 109–144.

[33] C.A. Hunter, J.K.M. Sanders, The nature of π-π interactions, J. Am. Chem. Soc. 112 (1990) 5525–5534.

[34] S. Grimme, Do special noncovalent π-π stacking interactions really exist? Angew. Chem. Int Ed. 47 (2008) 3430–3434.

[35] E.R. Collantes, W.J. Dun III, Amino acid side chain descriptors for quantitative structure-activity relationship studies of peptide analogues, J. Med. Chem. 38 (1995) 2705–2713.

[36] C.F. Matta, R.F.W. Bader, Atoms-in-molecules study of the genetically encoded amino acids. III. Bond and atomic properties and their correlations with experiment including mutation-induced changes in protein stability and genetic coding, Proteins: Struct., Funct., Genet. 52 (2003) 360–399.

[37] O. Khakshoor, S.E. Wheeler, K.N. Houk, E.T. Kool, Measurement and theory of hydrogen bonding contribution to isosteric DNA base pairs, J. Am. Chem. Soc. 134 (2012) 3154–3163.

[38] P. Desjardins, D. Conklin, NanoDrop microvolume quantitation of nucleic acids, J. Vis. Exp. 45 (2010) 2565.

[39] T. Boland, B.D. Ratner, Direct measurement of hydrogen bonding in DNA nucleotide bases by atomic force microscopy, Proc. Natl. Acad. Sci. U S A 92 (1995) 5297–5301.

[40] S. Smith, Y. Cui, C. Bustamante, Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules, Science 271 (1996) 795–799.

[41] S.B. Bustamante, C. Smith, J. Liphardt, D. Smith, Single-molecule studies of DNA mechanics, Curr. Opin. Struct. Biol. 10 (2000) 279–285.