One-Electron Oxidation Potentials and Hole Delocalization in Heterogeneous Single-Stranded DNA

Jesús Lucia-Tamudo,[†] Manuel Alcamí,[†,‡] Sergio Díaz-Tendero,[*,†,‡,¶] and Juan J. Nogueira[*,†,‡]

†*Department of Chemistry, Universidad Autónoma de Madrid, 28049, Madrid, Spain*

‡*Institute for Advanced Research in Chemical Sciences (IAdChem), Universidad Autónoma de Madrid, 28049 Madrid, Spain*

¶*Condensed Matter Physics Center (IFIMAC), Universidad Autónoma de Madrid, 28049 Madrid, Spain*

E-mail: sergio.diaztendero@uam.es; juan.nogueira@uam.es

Abstract

The study of DNA processes is essential to understand not only its intrinsic biological functions, but also its role in many innovative applications. The use of DNA as a nanowire or electrochemical biosensor leads to the need for a deep investigation of the charge transfer process along the strand, as well as of the redox properties. In this contribution, the one-electron oxidation potential and the charge delocalization of the hole formed after oxidation is computationally investigated for different heterogeneous single-stranded DNA strands. We have established a two-steps protocol: (i) molecular dynamics (MD) simulations in the frame of quantum mechanics / molecular mechanics (QM/MM) were performed to sample the conformational space; (ii) energetic properties were then obtained within a QM1/QM2/continuum approach in combination with

the Marcus theory over an ensemble of selected geometries. The results reveal that the one-electron oxidation potential in the heterogeneous strands can be seen as a linear combination of that property within homogeneous strands. In addition, the hole delocalization between different nucleobases is, in general, small, supporting the conclusion of a hopping mechanism for the charge transport along the strands. However, charge delocalization becomes more important, and so the tunneling mechanism contribution, when the reducing power of the nucleobases forming the strand is similar. Moreover, an important charge delocalization is also obtained when there is correlation between pairs of some of the interbase coordinates of the strand: twist/shift, shift/slide and rise/tilt.

1 Introduction

The information regarding the characteristics of every living organism is stored in nucleic acids molecules RNA or DNA. Specifically, DNA is the macromolecule responsible for this function in eukaryotic organisms. Over the decades, humans have attempted to determine which regions of the chromosomes encode each gene and the sequence that determines its expression. However, the functionalities of DNA have expanded during the last few decades.[1] Among these innovative applications, its use in DNA computation,[2,3] DNA-templated synthesis for new materials,[4] molecular detection,[5–15] and as a nanowire[16,17] can be highlighted. In the case of the latter two applications, the mode of operation is similar. These systems consist of an ensemble of single-stranded DNA (ss-DNA) or double-stranded DNA (ds-DNA) anchored to a metallic surface in the case of an electrochemical biosensor or two electrodes in the case of a nanowire. In the first stage, there is a charge transfer between a component of the system, such as an analyte or an electrode, and the DNA strand. Generally, nucleobases are the primary moieties responsible for the charge transfer process in aqueous phase.[18,19] Thus, a comprehensive understanding of the redox properties of nucleobases is crucial for gaining insights into this phenomenon. Specifically, nucleobases are more prone to oxidation

than reduction. As a result, obtaining an accurate value of the one-electron oxidation potential is essential, and numerous studies have been conducted to elucidate them.[20–31] This property can be understood as the reduction potential of an oxidation process. In this context, the relative order of the reducer character for free nucleobases in water is well-known: G > A > T ∼ C > U. In fact, in previous works, we found a clear relationship between the number of atoms of a nucleobase in which the positive charge is delocalized and the relative order of the reducer character within nucleobases.[32]

In the second stage, after the electron transfer from the nucleobase, hole transport occurs along the DNA strand. Since nucleobases hold the hole in water, these moieties are also responsible for the transport of the positive charge. In this context, two main mechanisms have been proposed: tunnelling and hopping.[33–36] On the one hand, tunnelling advocates for a transport based on the hole delocalization along several nucleobases until it reaches the receiving component.[34] This mechanism shows a dependency on the distance between nucleobases: two nucleobases must be close enough to enable hole delocalization through π-stacking interactions. On the other hand, hopping arises as an alternative to the tunnelling model to explain the long-range transport of the charge in DNA. It is a multistep process that states that the charge is localized in just one nucleobase and moves through consecutive jumps from one nucleobase to another with similar redox properties. In fact, nucleobases with identical one-electron oxidation potentials can transfer the charge to one another even if other nucleobases are interspersed between them. In contrast to the tunnelling model, the dependency on distance becomes less relevant for the hopping model. Thus, by consecutive hopping processes, the charge can be transferred through long distances. It is said that guanine and, to a lesser extent, adenine, are hopping stones for hole transfer, while thymine and cytosine are hopping stones for electron transfer. However, both processes, hole and electron transfer, do not have the same probability of occurring because of the different rates they present. While hole transfer can take place on a nanosecond to microsecond time scale, electron transfer takes minutes to weeks. In a previous work, we determined that the

hopping model is likely predominant over the tunnelling model for the case of homogeneous ss-DNA.[32] This is supported by the evidence that charge delocalization becomes less relevant due to the stabilization produced by solvent effects; the charge tends to be mainly held in a single nucleobase, especially when the solvent has a high polar character and the nucleobase has a high reducer character.[37,38]

The need to understand the causes of this transport puts the focus on the delocalization feature. Since DNA is a considerably large biomolecule, analyzing all degrees of freedom simultaneously to find a relationship between the conformation of the strand and the delocalization of the hole along it becomes unfeasible. Fortunately, the complex structure of DNA has led to the convention of some parameters that allow for the analysis and comparison of different conformations in which a specific strand can remain.[39–41] In general terms, the conformation of ds-DNA strands can be described by six sets of parameters: helical axis, base pair-axis, intra-base pair, inter-base pair, backbone parameters and groove parameters. The helical axis is described by its general stretch and torsion along the strand. The base pair-axis set accounts for the deformation of the axis between two adjacent base pairs. The inter-base pair set describes the arrangement of two adjacent base pairs (see Figure 1), and the intra-base pair set provides insights into the arrangement of the two nucleobases that compose the base pair. The puckering of the sugar of each nucleotide gives information about the conformation of the backbone of the strand. Finally, the groove parameters provide information about the major and minor groove in a ds-DNA. However, when this analysis is applied to a ss-DNA strand, the sets that account for the intra-base pair and groove parameters cannot be defined.

In this work, we have determined the one-electron oxidation potential of heterogeneous ss-DNA strands and compared them with those obtained for homogeneous ss-DNA in a previous study.[32] Additionally, we have computed the delocalization of the hole along these ss-DNA strands and found a relationship between this phenomenon and the inter-base pair structural parameters, shown in Figure 1.

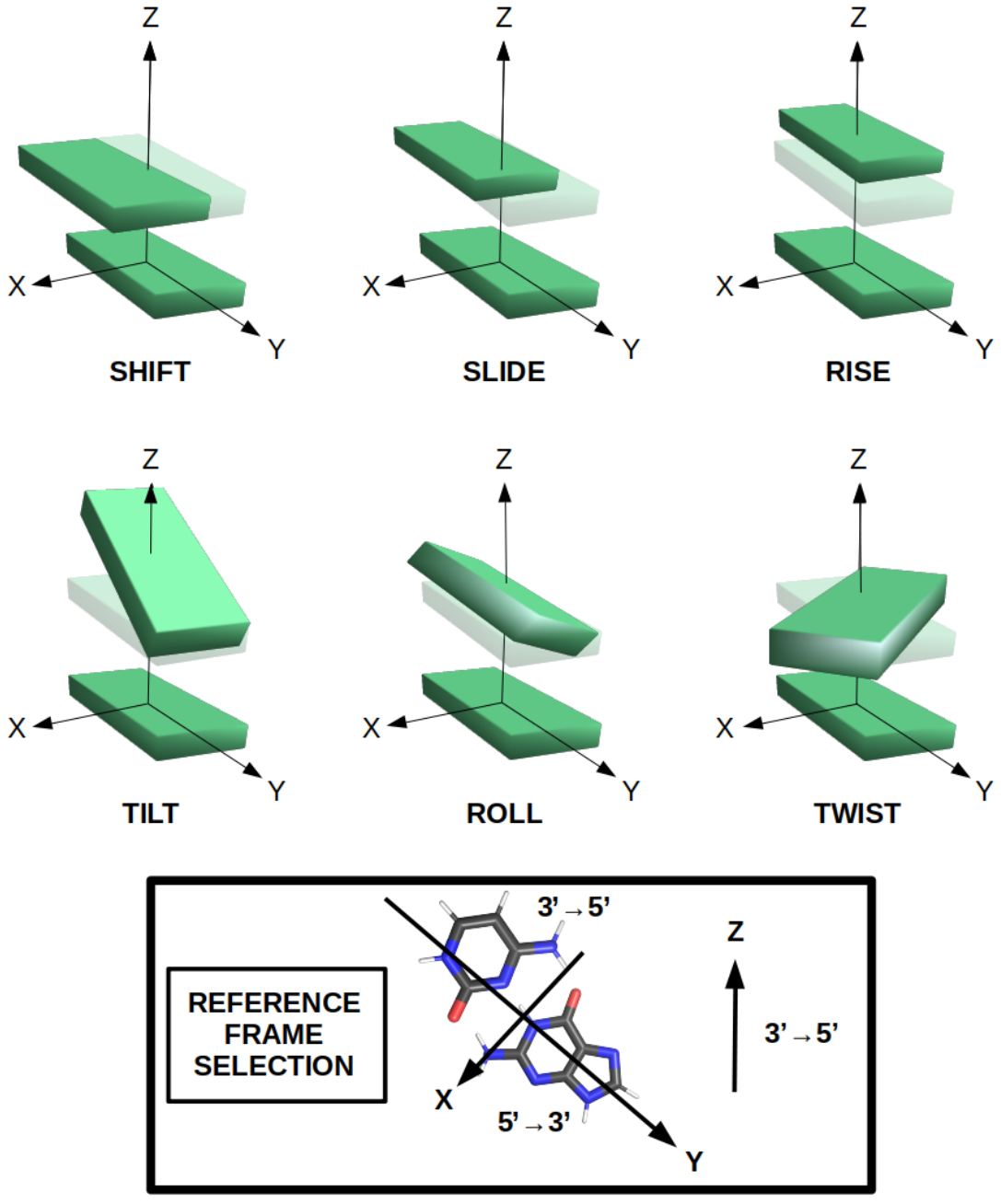

Figure 1: Graphical description of the inter-base pair parameters. The reference frame selection is also displayed to fix the convention of the parameters. The systems under study follow the 5' → 3' direction.

2 Methods and Computational Details

The computation of the one-electron oxidation potential and the delocalization properties of the strands were conducted using the same procedure employed in a previous study on homogeneous ss-DNA.[32] After the setup of the different solvated ss-DNA strands, a conformational sampling was carried out using classical and quantum mechanics/molecular mechanics (QM/MM) molecular dynamics (MD) simulations. Subsequently, the properties were computed for an ensemble of geometries selected from the QM/MM trajectories through electronic-structure calculations. These calculations were performed using a QM1/QM2/continuum approach in combination with the Marcus theory and electron population analysis. In the following, the specific details of all these steps are deeply discussed,

The nucleic acid builder (nab) application provided by the AmberTools 22 package[42–44] was used to model the initial geometries of the heterogeneous ss-DNA strands. Each of the six strands investigated was composed of 24 nucleotides, shown in Figure 2. Specifically, the eight nucleobases of the center of the strands correspond to tetramers of all possible combinations of nucleobases pairs: $(AC)_4$, $(AG)_4$, $(AT)_4$, $(CG)_4$, $(CT)_4$, and $(GT)_4$. To prevent self-hybridization, especially in ss-polyCG and ss-polyAT, a limiting cap of 8 nucleotides was added to each edge of the strands. The ss-DNA strands were solvated in a truncated octahedron box with a buffer of 12 Å, and the tleap program implemented in AmberTools 22 was used for this purpose. The ff90bsc0 force field,[45,46] along with the dihedral correction adressed in bsc1,[47] was selected to describe ss-DNA, while the TIP3P force field[48] was used for describing water molecule interactions. To counteract the negative charge of the strands, 22 sodium cations were added using the parameters developed by Joung and Cheatham.[49]

The configurational space was explored through classical MD simulations[50–52] using the CUDA version of the pmemd program implemented in the AMBER 20 package.[42–44] The simulations for the homogeneous ss-polyX systems were obtained from a previous study,[32] while the same procedure was applied for the heterogeneous ss-polyXY. The simulations began with a 10000-step minimization, where the first 5000 steps were computed using the

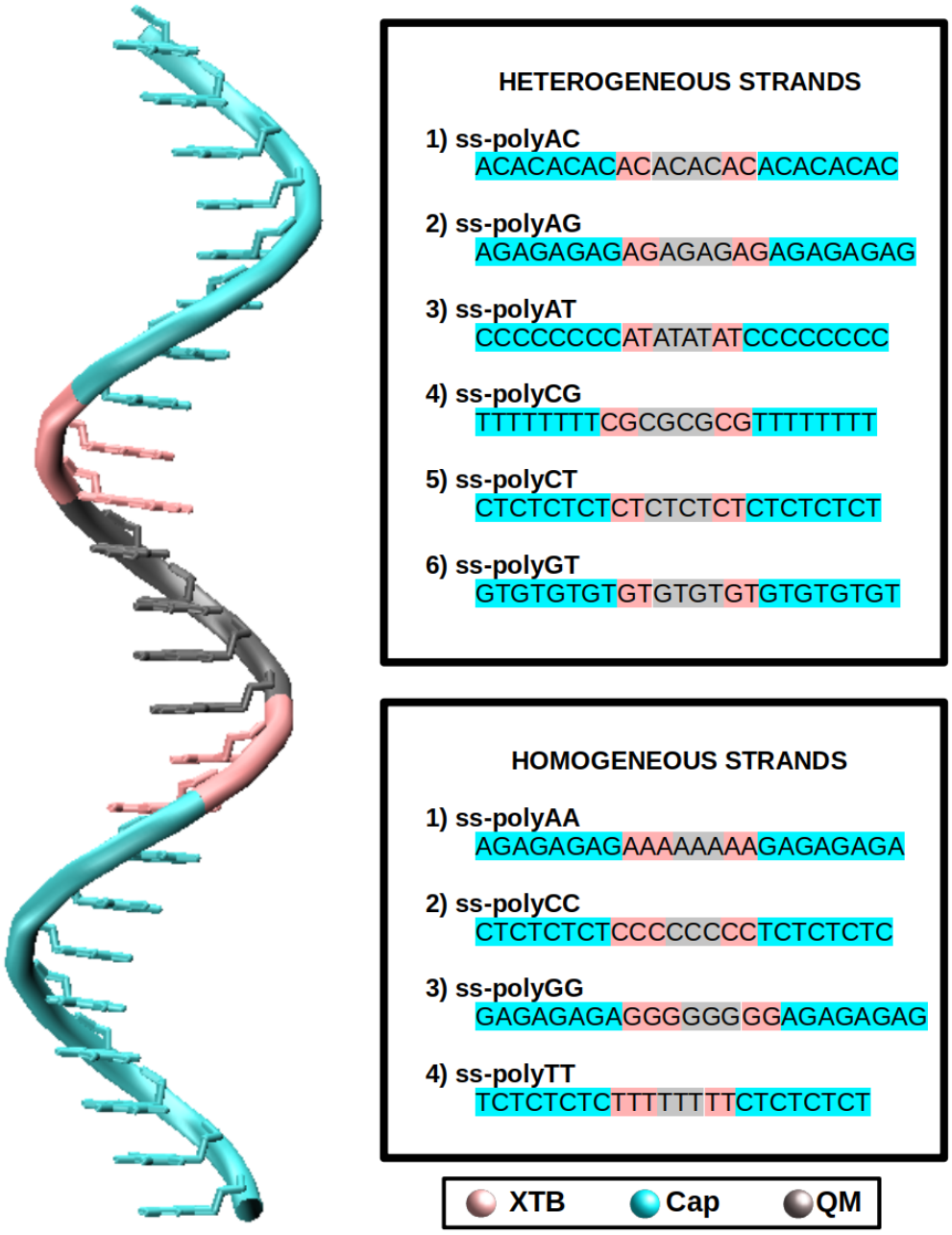

Figure 2: Graphical view of the general form of the systems under study in this work. The full sequence for each system is also displayed for ss-polyX and ss-polyXY. The colour of the strand represent the layer to which these nucleobases and nucleotides belong. Cyan refers to the nucleotides that form the protective caps of the strand. Gray refers to the nucleobases that are included in the QM1 region, excluding the sugar and phosphate of the correspondent nucleotide. Pink is associated to the nucleotides whose nucleobase belongs to the XTB(QM2) layer.

steepest-descent algorithm,[53] followed by another 5000 steps using the Newton-Raphson algorithm.[54] A constant volume (NVT) progressive heating up to 300 K was then run for 500 ps, and a thermostat was applied according to the Langevin model with a collision frequency of 2 ps^{-1} to regulate the temperature. After that, an additional 500 ps simulation was conducted at a constant temperature of 300 K (NVT ensemble). In the following stage, a 1 ns simulation was run in the NPT ensemble to equilibrate the volume of the system and achieve the correct density. Finally, a production simulation of 200 ns was conducted in the NPT ensemble, and 200 equidistantly separated snapshots were selected. The Berendsen barostat with isotropic position scaling and a pressure relaxation time of 2 ps were used for all simulations carried out within the NPT ensemble to maintain the pressure constant at 1 bar. The particle-mesh Ewald method with a grid spacing of 1.0 Å was employed to compute the electrostatic interactions during the full protocol, and a 10 Å cutoff was chosen for the non-bonded interactions. The SHAKE algorithm[55–57] restrained the bonds involving hydrogen atoms, and a time step of 2 fs was used during the heating, equilibration, and production stages.

In order to compute the one-electron oxidation potential, E_{red}, we used the Marcus theory formulation[58–63] that states that:

$$\Delta G_{red} = \frac{1}{2}\Big(\langle VIE\rangle_N - \langle VAE\rangle_{N^+}\Big) - G(e^-_{(gas)}) \tag{1}$$

where VIE (vertical ionization energy) is the energy required to remove an electron from a neutral species, VAE (vertical attachment energy) is the energy released upon adding an electron to a cationic species, and $G(e^-_{(gas)}) = -0.867$ kcal/mol is a correction for the free energy of the electron in the gas phase, calculated using the Fermi-Dirac statistics.[64–66]

Notice that both VIE and VAE are required to compute the one-electron oxidation potential within the Marcus theory. While VIE can be easily obtained from the snapshots of the aforementioned classical dynamics, calculating VAE requires a conformational sampling

of the phase space of the cationic strand. Unfortunately, force field parameters are not available for the cation, making the calculation of VAE a challenge. To overcome this limitation, a set of 200 QM/MM MD simulations for each strand was conducted using as initial conditions an ensemble of 200 snapshots selected from each of the classical MD simulations. Thus, the combination of classical and QM/MM MD guarantee a statiscally-accurate thermal distribution of the solvent molecules, while the relevant region of the DNA is described quantum mechanically. The QM/MM dynamics trajectories were carried out for both neutral and cationic phase spaces to ensure consistency. Specifically, additional 100-step QM/MM MD simulations were run in the NPT ensemble for each of the selected frames using the ORCA[67]/AMBER interface. The QM region, comprising four adjacent nucleobases, was computed using the CAM-B3LYP functional[68] and the 6-311G(d)[69,70] basis set.

Finally, the last geometry obtained from each QM/MM MD simulation was used to calculate the VIEs and VAEs for each heterogeneous ss-DNA strand. The calculations were carried out using a QM1/QM2/continuum scheme, where the QM1/QM2 interaction was described by electrostatic embedding. The VIEs and VAEs were computed for the QM1 region, where four nucleobases were described with the CAM-B3LYP/6-311G(d) level of theory. After removing the limiting caps, the remaining four nucleobases in the QM2 region were described by tight-binding DFT (DFTB) using the GFN2-XTB scheme.[71] Solvent effects were taken into account using the ALPB continuum solvation model,[72] which is suitable for DFTB. All calculations were performed using the ORCA 5.0.3 package.[67]

The free energy computed by Eq. 1 can be related to the one-electron oxidation potential through the following equation:

$$\Delta E_{red} = \frac{\Delta G_{red}}{n_e F} - E_{red,SHE} \tag{2}$$

where F is the Faraday constant, n_e is the number of exchanged electrons (one in this case) and $E_{red,SHE}$ is the reduction potential of a reference electrode, which in this case is the

standard hydrogen electrode (SHE). The considered value of $E_{red,SHE}$ was 4.281 V, used in previous works.[73–77] This value also accounts for the correction of the free energy of the electron in the gas phase. As a result, this contribution must be also added in Eq. 1.

The delocalization of the hole was analyzed by calculating the charge difference of each nucleobase in the QM1 region between the cationic and the neutral species for each of the geometries of the ensembles (200 geometries for each strand). The Löwdin charges[78] were used for charge calculations, and the analysis was conducted using custom scripts. The intermolecular delocalization number, denoted as n, was defined as the number of nucleobases among which the positive charge is distributed after ionization. To determine n, the four nucleobases considered in the QM1 region were first ordered in terms of increasing positive charge difference Δq_i, and then an empirical equation, whose details can be found in a previous work, was applied:[32]

$$n = m - \sum_{i=1}^{m-1} \left[1 - \left(\frac{\Delta q_i}{\sum_{j=i}^{m} \Delta q_j} \right) (m - i + 1) \right] \tag{3}$$

Note that the term $\frac{\Delta q_i}{\sum_{j=1}^{m} \Delta q_j}$ represents the contribution to the delocalization of each nucleobase. Additionally, the term $(m - i + 1)$ indicates the number of nucleobases over which this delocalization contribution will be taken into account. To sum up, the total number of nucleobases where the charge is delocalized, n, is obtained as the number of considered nucleobases m minus the non-contribution to delocalization of each one.

In this context, Pipek and Mezey reported another way to quantify the delocalization of a positive charge among a system.[79] In this case, they obtained a delocalization index using the atomic Mulliken population of the set of orbitals in each atom. In other to compare our empirical Eq. 3 with the equation by Pipek and Mezey, we have adapted the latter one by considering the charge difference of each nucleobase of the QM region leading to:

$$n' = \sum_{i=1}^{m} \frac{1}{(\frac{\Delta q_i}{\sum_{j=1}^{m} \Delta q_j})^2} \tag{4}$$

Once the delocalization index was determined, a structural analysis was performed to investigate the relationship between the conformation of the strands and the delocalization of the hole. The following procedure was applied for both homogeneous and heterogeneous ss-DNA (as shown in Figure 3).

Firstly, the structural parameters of the selected geometries were obtained using the CURVES+ package.[39–41] In this study, only the inter-base pair set of parameters were analyzed (see Figure 1). For each pair of nucleobases within the QM1 region, the relative intermolecular delocalization number n was computed and plotted against each of the inter-base pair parameters (see Figure 3a). Then, a 3D representation of n as function of each pair of structural parameters was plotted as a contour map to see possible correlations between the structural parameters and large values of charge delocalization (see Figure 3b). The limits of those conformational spaces (for each structural parameter) were set to $\pm 2.5\sigma$, where σ represents the standard deviation of each parameter along the geometrical ensemble. Therefore, each of the axis of Figure 3b goes from -2.5σ to 2.5σ and the center of the each axis is located at the mean value of the corresponding geometrical parameter. Although these limits are chosen in an arbitrary manner, they covered the whole conformational space where delocalization is allowed. In order to identify the region of the conformational space with the largest positive charge delocalization, the previous contour maps were simplified in the following way. A 20×20 grid was applied to each resulting conformational space. For each element of the grid, the point with the largest delocalization index was selected and plotted, while the rest were discarded (see Figure 3c). Thus, only those points with large hole delocalization were chosen to analyze the role of the structural parameters. Moreover, a further discrimination was conducted by only considering elements of the grid with an n higher than 1.7, while the rest were discarded (see Figure 3d). Finally, a linear regression was performed with these remaining points to observe the correlation of each pair of parameters with respect to the delocalization. Each pair of inter-base pair parameters was classified into four groups in terms of the resulting slope of the regression once normalized (see Figure 3e).

The normalization was performed by considering that a slope equals to 1 is obtained when the straight line goes in the direction from the point $(-2.5\sigma, -2.5\sigma)$ to the point $(2.5\sigma, 2.5\sigma)$ and equal to -1 when it goes from $(-2.5\sigma, 2.5\sigma)$ to $(2.5\sigma, -2.5\sigma)$. Thus, if the slope was near 1 and -1, then the pair was considered positively and negatively correlated, respectively, when the charge delocalization is important. On the contrary, if the slope was closer to 0 or to ∞, then one of the parameters did not show a correlation with the other one when delocalization is large.

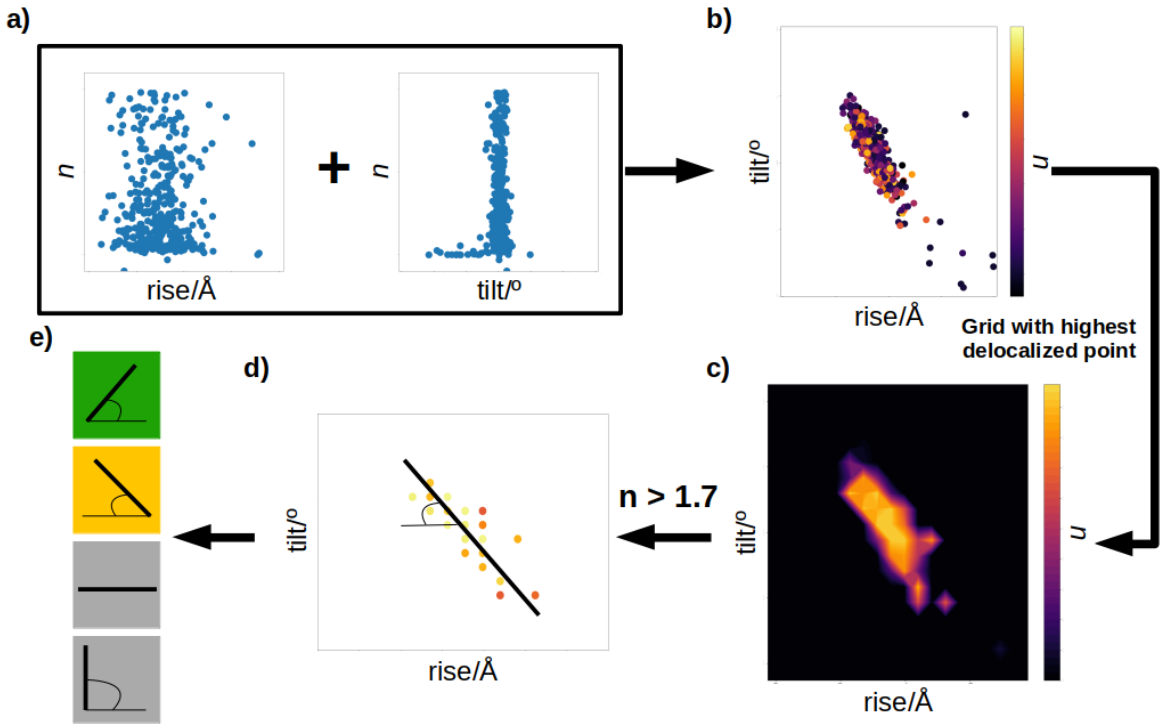

Figure 3: Graphical explanation of the discrimination of the data to obtain correlations between pair of parameters. a) Relative n plotted against one parameter. b) Relative n plotted against two parameters. c) Grid that contains the maximum delocalization among the sampled conformational space. d) Points from the grid whose relative n is larger than 1.7 and linear regression over those points. e) Classification of the correlation between a pair of parameters.

3 Results

3.1 One-electron oxidation potential and charge delocalization

The reducing power of a specific species can be measured in terms of its reduction potential. When the target reaction involves an oxidation process where one electron is involved, the reduction potential is commonly referred to as the one-electron oxidation potential, as stated previously. Thus, determining this property can establish a hierarchy in terms of the reducing power of a set of molecules. In this work, the one-electron oxidation potential has been computed for different sequences of nucleotides within a ss-DNA strand. In a previous study, this property was determined for sequences with identical nucleotides labeled as ss-polyX ($X = A, C, G, T$).[32] Here, the target strands are composed of two different nucleotides interspersed with each other and, therefore, these systems will be labeled as ss-polyXY ($X, Y = A, C, G, T$, and $X \neq Y$). Our previous results indicate that the computation of the one-electron oxidation potential of large systems involving nucleobases can be accurately conducted by combining Marcus theory with MD simulations and QM/continuum models.[31] Therefore, the one-electron oxidation potential has been determined for ss-polyXY following this methodology.

The results show that changes in the potential depend on the amount of nucleotides of each type present in the strand, as seen in Figure 4. Specifically, the E_{red} of ss-polyXY takes a value between the two limiting situations of ss-polyX and ss-polyY. The only exception to this trend is in the case of ss-polyCT, for which the E_{red} is slightly higher than that of both pure ss-DNA strands. However, the value lies within the standard deviation of the calculations. These results reflect that E_{red} can be seen as a linear combination of the potentials of pure single strands formed by the nucleobases present in the heterogeneous one. In this sense, a strand that contains purines, which are highly reducing nucleobases, will be more reducing than another strand containing pyrimidine nucleobases. The reducing character of the nucleobases and the reducing character difference between them will be

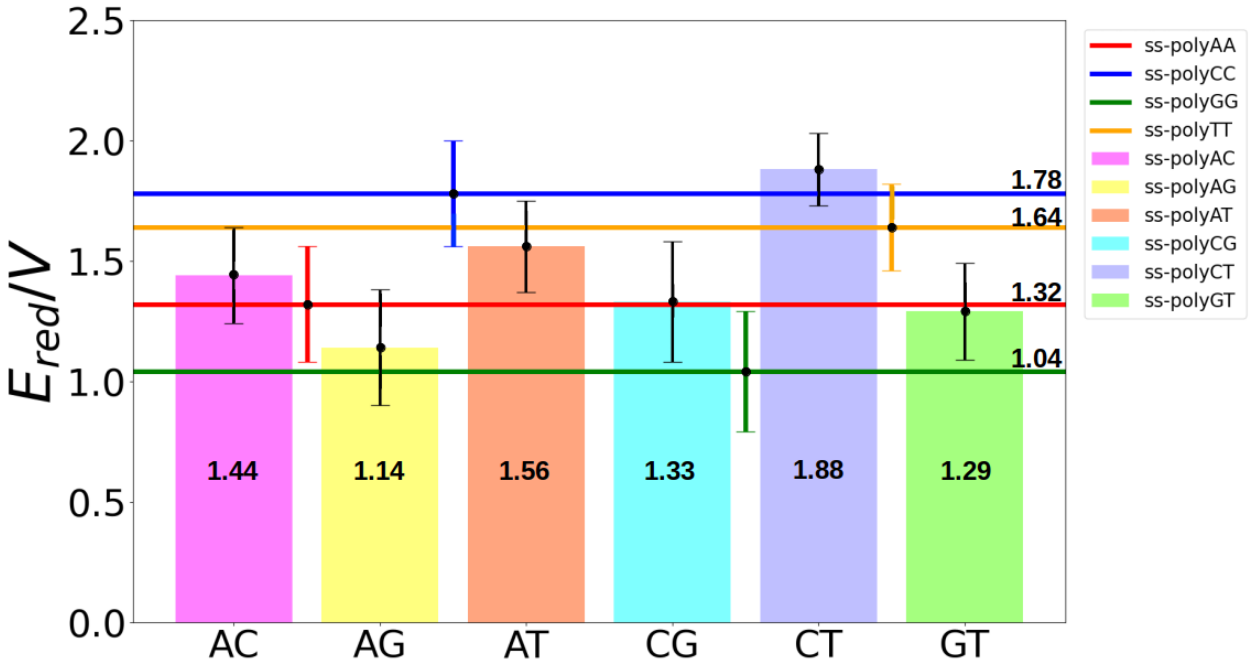

Figure 4: Computed one-electron oxidation potentials for ss-polyX (solid lines) and ss-polyXY (bars) in aqueous phase. The E_{red} values for ss-polyX are taken from reference 32.

intimately related with the charge delocalization, as will be discussed below.

After oxidation, the positive charge (hole) can be transferred along the strand via different mechanisms. On the one hand, the charge can be delocalized among several nucleobases, evolving in space with time according to the tunneling mechanism. On the other hand, the transport can be conducted through sequential jumps from one nucleobase to another by the hopping mechanism, in which the charge is essentially localized in just one nucleobase at a time. In order to get insight into the dominant mechanism in ss-DNA strands, the delocalization of the hole along the different ss-polyXY strands considered has been assessed using Eq. 3. For comparison, the delocalization numbers have been also calculated using Eq. 4, and the results are shown in Figure 5a. Although values for n' are slightly higher than for n, the relative order for the different ss-DNA strands remains invariant. Therefore, in the following, only the n values will be discussed for simplicity. Figure 5 displays the values of n for the different strands (panel a) and the amount of charged hosted by each of the nucleobases of the strands included in the QM region (panel b). The charge delocalization

numbers for ss-polyX strands are taken from a previous work[32] and are also shown for comparison.

It has previously been shown that there exists a competition in homogeneous ss-polyY strands between intrabase and interbase delocalization of the hole.[32] Greater intramolecular delocalization to accommodate the positive charge was obtained for purines, which have a large π system, than for pyrimidines. Contrary, when considering homogeneous ss-DNA systems of cytosines and thymines, where the π system is more spatially constrained, intramolecular delocalization is reduced compared to purines and, thus, the interbase delocalization of the hole becomes more important. Based on the results shown in Figure 5, a similar situation is found in heterogeneous ss-polyXY, although now the delocalization depends on two factors: (i) the oxidation potential of the most reducing nucleobase of the strand and (ii) the difference in the reducing character of the two nucleobases. The importance of each of these factors depends on the composition of the strand.

When guanine is part of the strand, the properties are clearly dominated by those of guanine. As shown in Figure 5a, the delocalization numbers of ss-polyAG, ss-polyCG and ss-polyGT are small and similar to that of ss-polyGG since the hole is preferably located on only one guanine, as can be seen in Figure 5b. Specifically, around $70-80\%$ of the positive charge is hosted by one of the guanine nucleobases when guanine is part of the strand. In addition, the charge is more delocalized in ss-polyAG than in ss-polyCG or ss-polyGT because the reducing character of adenine and guanine is more similar than those for guanine and thymine/cytosine. Thus, the positive charge is shared between both nucleobases, although with significant dominance of guanine, increasing the delocalization number. Contrary, cytosine and thymine are not able to attract the positive hole because their reducing power is much smaller than that of guanine.

When guanine is not present in the strand, the competition between nucleobases for hosting the charge is greater. In this way, the delocalization number for ss-polyAC and ss-polyAT is larger than that when guanine is present, and it is also larger than that of

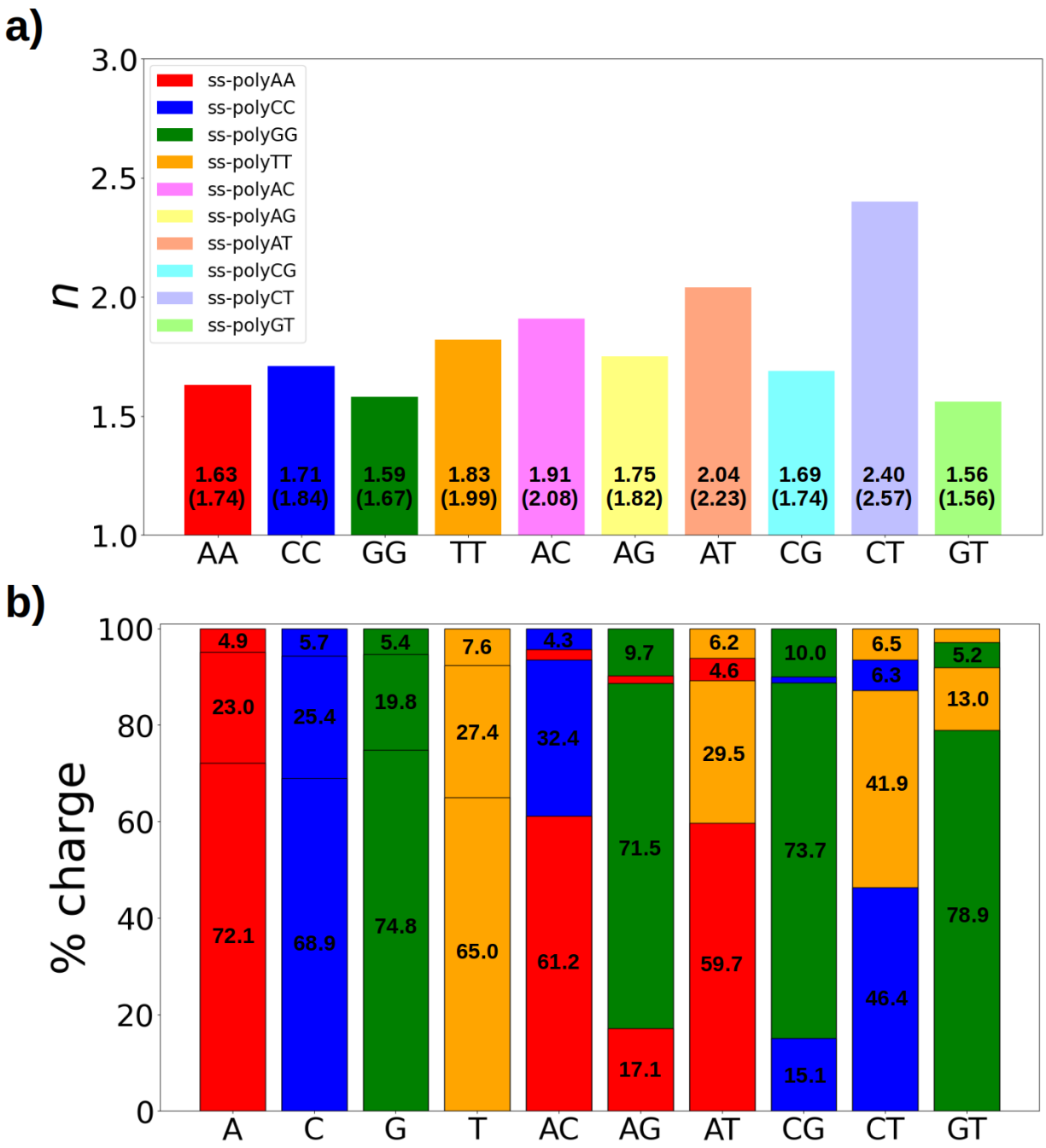

Figure 5: Delocalization of the hole along the strands considered. a) Intermolecular delocalization number for each ss-polyX and ss-polyXY. n (n') values are displayed within the corresponding bars. b) Percentage of positive charge held by each nucleobase of the strand. The colors of b) represent the type of nucleobase: A in red, C in blue, G in green and T in orange. The results from ss-polyX were taken from reference 32.

ss-polyAA. This is also reflected in the fact that now only 65% of the charge is located on one of the adenine moieties in ss-polyAC and ss-polyAT, while around 30% of the charge is located on thymine or cytosine. Finally, when two pyrimidines are combined (ss-polyCT), the largerst intermolecular delocalization (2.4) is obtained compared to the other binary combinations and with ss-polyCC and ss-polyTT (see Figure 5a). Both nucleobases have small π systems and, thus, the intramolecular delocalization is small and intermolecular delocalization is preferred. In addition, both nucleobases have similar one-electron oxidation potentials and, thus, none of the nucleobases has preference to host the positive charge of the hole in ss-polyCT. As represented in Figure 5b, the two cytosine molecules included in the QM region accommodate 46.4% and 6.3% of the charge, while a similar situation is found for the two thymines. This means that the charge delocalization is evenly shared between cytosine and thymine, increasing further the delocalization number.

Our analysis on the charge delocalization supports a hopping mechanism with some contribution of tunneling for the transport of the hole along the ss-DNA strands. While the positive charge is predominantly localized on one nucleobase in most of the cases, there is always a certain degree of delocalization of the charge towards the nucleobase adjacent to the predominant one. In addition, the extent of tunneling character increases as the potential between the two nucleobases present in the strand is more similar, for example, in ss-polyCT and ss-polyAT.

3.2 Relation between structure and charge delocalization

In order to obtain the previously discussed results of one-electron oxidation potential and delocalization, classical MD simulations were performed, followed by QM/MM MD simulations, with the aim of exploring the conformational space of the different systems. The sampled conformational space of each system is analyzed here in terms of the shift, slide, and rise distance parameters and the tilt, roll, and twist angle parameters. These parameters are related to the charge delocalization following the analysis explained above and schematically

displayed in Figure 3. The results of this analysis are presented in Figure 6 in the form of color matrices, representing the type of correlation that exists between pairs of structural parameters that lead to the largest delocalization number of the positive charge between pairs of nucleobases. The grey color indicates that there is no correlation between the pair of parameters. Green color accounts for positive correlation, that is, the highest charge delocalization is obtained when both structural parameters decrease o increase at the same time. Negative correlation is represented in yellow, meaning that the highest delocalization is achieved when one of the parameters increases while the other decreases. Finally, it is worth to mention that these matrices are symmetric and the diagonal does not have any physical meaning.

As can be seen in Figure 6, a certain degree of similarity can be observed in the correlation matrices of all the strands studied except by the ss-polyA one. In general terms, there exist two commonly found positive correlations: one between twist and shift, which is present in 9 out of 10 strands, and another between shift and slide, present in 8 out of 10 strands. The reason behind the positive correlation between these pairs of parameters is to enhance the overlap between the aromatic rings of the consecutive nucleobases because such a strong interaction increases the charge delocalization. For example, the natural torsion of ss-DNA and ds-DNA strands is the origin of a significant twist value, a fact that weakens the interactions between nucleobases. The enhancement of the π-stacking interactions, leading to a large charge delocalization, can be achieved by the displacement of one of the interacting nucleobases along the X-axis increasing, thus, the shift distance. Therefore, the increase of the twist angle requires the increase of the shift value to favor the delocalization of the hole. Other positive correlations that were found along the dynamics of some ss-DNA strands, although in a lesser extent, are slide/twist (in 6 out of 10 strands) and slide/tilt (in five out of 10 strands).

Figure 6 also shows that there is a recurrent negative correlation between the rise distance and the tilt angle. Specifically, for 8 out of 10 strands, when the rise distance decreases the

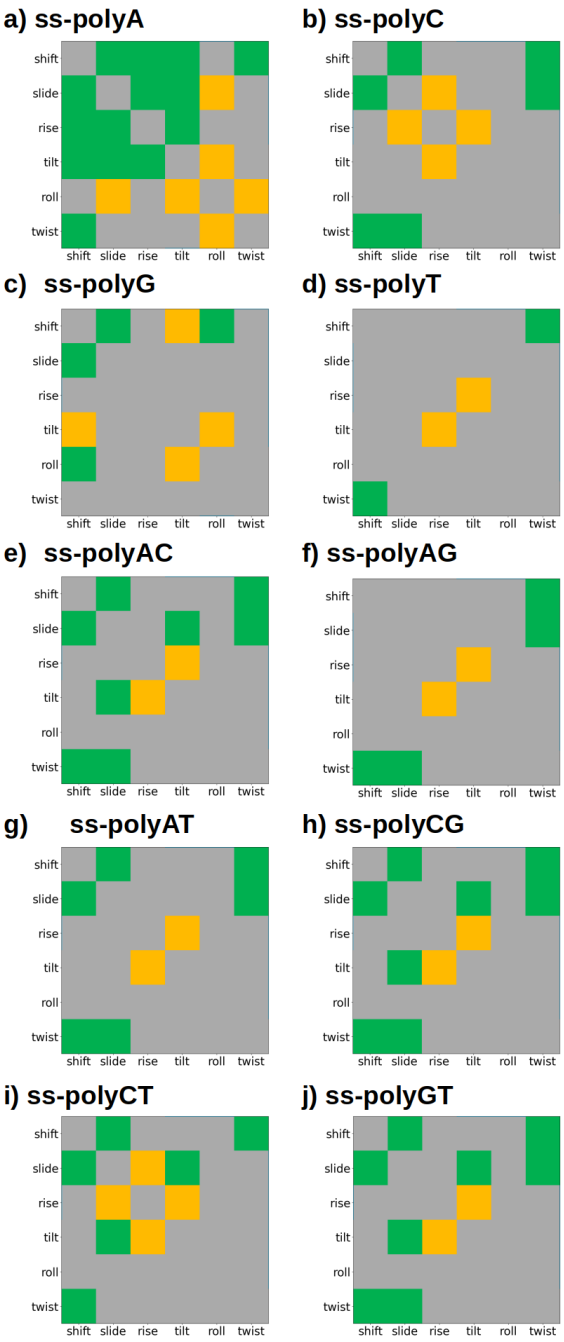

Figure 6: Coloured matrix representation of the existing correlations between inter-base pair parameters in ss-DNA. Color code: green accounts for positively correlation, orange represents negatively correlations and gray refers to non existing correlation.

tilt angle increases. In this case, this negative correlation is likely aimed at avoiding strong repulsive interactions between neighboring nucleobases. The repulsion originated by small rise distances between consecutive nucleobases can be alleviate when one of the nucleobases is tilt, inducing a increase in the separation of the aromatic clouds. In conclusion, when analyzed the configurational space that accounts for strong positive charge delocalization, shift, slide and twist correlate positively to increase the attractive interactions, while the rise/tilt pair correlates negatively to decrease the repulsive interactions.

4 Conclusions

In this computational study the one-electron oxidation potential and the degree of delocalization of the positive hole formed after oxidation in heterogeneous ss-polyXY have been investigated. The results have been compared with other analogues obtained from homogeneous ss-polyX. In addition, a structural analysis has been carried out to study the effect of the correlation between some structural parameters on the delocalization of the hole along the strand, and shed light into the importance of the two hole transport mechanisms in DNA, namely, tunneling and hopping.

The results show that the the one-electron oxidation potential of ss-polyXY takes a value between the two limiting situations ss-polyX and ss-polyY, and can be seen as a linear property in terms of the composition of the system in the case of ss-DNA. Thus, a strand containing purine nucleobases will be more reducer than one formed by pyrimidine nucleobases. When analyzing the degree of delocalization of the hole among adjacent nucleobases in ss-DNA, the results reveal that the delocalization number depends on the oxidation potential of the most reducing nucleobase and on the reducing character difference between the two nucleobases present in the strand. When guanine is one of the components of the system, its properties are dominated by those of guanine. It has been computed that around $70 - 80\%$ of the hole charge is located on just one of the guanine moieties. Contrary, when

guanine is not forming part of the strand, the delocalization number increases due to an increase in the competition between the different nucleobases to host the charge. Such a competition is more important when the nucleobases of the strand have similar reducing power, for example, as in ss-polyCT. Therefore, our computational analysis supports that the hole is transported along ss-DNA strands mostly by a hopping mechanism with some tunneling contribution. Such a tunneling component will be more relevant when guanine is not present.

The structural analysis of the dynamics shows that large charge delocalization is achieved when some of the structural parameters of the strand are correlated. On one side, the positive correlation between twist, shift and slide enhances the attractive interactions between nucleobases. On the other side, the negative correlation between rise and tilt reduces the repulsion between nucleobases. These correlations likely lead to a larger overlap between the aromatic clouds of the nucleobases, a fact that induces a large charge delocalization.

Acknowledgement

We acknowledge the generous allocation of computer time at the Centro de Computación Científica at the Universidad Autónoma de Madrid (CCC-UAM) and at the Red Española de Supercomputación (RES). This work was partially supported by the MICINN – Spanish Ministry of Science and Innovation – Projects PID2019-110091GB-I00 and PID2020-117806GA-I00 funded by MCIN/AEI/10.13039/501100011033, and the 'María de Maeztu' (CEX2018-000805-M) Program for Centers of Excellence in R&D. J.J.N. acknowledge the Comunidad de Madrid for funding through the Attraction of Talent Program (Grant ref 2018-T1/BMD-10261). J.L.T. acknowledges the FPU19/02292 grant from the Spanish Ministry of University.

References

(1) Condon, A. Designed DNA Molecules: Principles and Applications of Molecular Nanotechnology. Nat. Rev. Genet. **2006**, 7, 565–575.

(2) Braich, R.; Chelyapov, N.; Johnson, C.; Rothemund, P.; Adleman, L. Solution of a 20-variable 3-SAT Problem on a DNA computer. Science (New York, N.Y.) **2002**, 296, 499–502.

(3) Xiong, X.; Zhu, T.; Zhu, Y.; Cao, M.; Xiao, J.; Li, L.; Wang, F.; Fan, C.; Pei, H. Molecular convolutional neural networks with DNA regulatory circuits. Nat. Mach. Intell. **2022**, 4, 1–11.

(4) O'Reilly, R. K.; Turberfield, A. J.; Wilks, T. R. The Evolution of DNA-Templated Synthesis as a Tool for Materials Discovery. Acc. Chem. Res. **2017**, 50, 2496–2509.

(5) Zhai, J.; Cui, H.; Yang, R. DNA based biosensors. Biotechnol. Adv. **1997**, 15, 43–58.

(6) Saidur, M.; Aziz, A. A.; Basirun, W. Recent advances in DNA-based electrochemical biosensors for heavy metal ion detection: A review. Biosens. Bioelectron. **2017**, 90, 125–139.

(7) Minunni, M.; Tombelli, S.; Mascini, M.; Bilia, A.; Bergonzi, M. C.; Vincieri, F. An optical DNA-based biosensor for the analysis of bioactive constituents with application in drug and herbal drug screening. Talanta **2005**, 65, 578–585.

(8) Bu, N.-N.; Tang, C.-X.; He, X.-W.; Yin, X.-B. Tetrahedron-structured DNA and functional oligonucleotide for construction of an electrochemical DNA-based biosensor. Chem. Comm. **2011**, 47, 7689–7691.

(9) Liu, A.; Wang, K.; Weng, S.; Lei, Y.; Lin, L.; Chen, W.; Lin, X.; Chen, Y. Development of electrochemical DNA biosensors. Trends Anal. Chem. **2012**, 37, 101–111.

(10) Izadi, Z.; Sheikh-Zeinoddin, M.; Ensafi, A. A.; Soleimanian-Zad, S. Fabrication of an electrochemical DNA-based biosensor for Bacillus cereus detection in milk and infant formula. Biosens. Bioelectron. **2016**, 80, 582–589.

(11) Drummond, T. G.; Hill, M. G.; Barton, J. K. Electrochemical DNA sensors. Nat. Biotechnol. **2003**, 21, 1192–1199.

(12) Paleček, E.; Fojta, M.; Jelen, F. New approaches in the development of DNA sensors: hybridization and electrochemical detection of DNA and RNA at two different surfaces. Bioelectrochemistry **2002**, 56, 85–90, Extended Abstracts of the XVIth International Symposium on Bioelectrochemistry and Bioenergetics Part 2.

(13) Zhang, S.; Wang, K.; Li, K.-B.; Shi, W.; Jia, W.-P.; Chen, X.; Sun, T.; Han, D.-M. A DNA-stabilized silver nanoclusters/graphene oxide-based platform for the sensitive detection of DNA through hybridization chain reaction. Biosens. Bioelectron. **2017**, 91, 374–379.

(14) Dai, N.; Kool, E. T. Fluorescent DNA-based enzyme sensors. Chem. Soc. Rev. **2011**, 40, 5756–5770.

(15) Zhou, C.; Zou, H.; Sun, C.; Ren, D.; Chen, J.; Li, Y. Signal amplification strategies for DNA-based surface plasmon resonance biosensors. Biosens. Bioelectron. **2018**, 117, 678–689.

(16) Berlin, Y. A.; Burin, A. L.; Ratner, M. A. DNA as a molecular wire. Superlattices Microstruct. **2000**, 28, 241–252.

(17) Wohlgamuth, C. H.; McWilliams, M. A.; Slinker, J. D. DNA as a Molecular Wire: Distance and Sequence Dependence. Anal. Chem. **2013**, 85, 8634–8640.

(18) Kissinger, P. T. Biosensors—a perspective. Biosens. Bioelectron. **2005**, 20, 2512–2516.

(19) Mehrotra, P. Biosensors and their applications – A review. J. Oral Biol. Craniofac. Res. **2016**, 6, 153–159.

(20) D'Annibale, V.; Nardi, A. N.; Amadei, A.; D'Abramo, M. Theoretical Characterization of the Reduction Potentials of Nucleic Acids in Solution. J. Chem. Theory Comput. **2021**, 17, 1301–1307.

(21) Psciuk, B. T.; Lord, R. L.; Munk, B. H.; Schlegel, H. B. Theoretical Determination of One-Electron Oxidation Potentials for Nucleic Acid Bases. J. Chem. Theory Comput. **2012**, 8, 5107–5123.

(22) Faraggi, M.; Broitman, F.; Trent, J. B.; Klapper, M. H. One-Electron Oxidation Reactions of Some Purine and Pyrimidine Bases in Aqueous Solutions. Electrochemical and Pulse Radiolysis Studies. J. Phys. Chem. **1996**, 100, 14751–14761.

(23) Jovanovic, S. V.; Simic, M. G. One-electron redox potentials of purines and pyrimidines. J. Phys. Chem. **1986**, 90, 974–978.

(24) Crespo-Hernández, C. E.; Close, D. M.; Gorb, L.; Leszczynski, J. Determination of Redox Potentials for the Watson-Crick Base Pairs, DNA Nucleosides, and Relevant Nucleoside Analogues. J. Phys. Chem. B **2007**, 111, 5386–5395.

(25) Seidel, C. A. M.; Schulz, A.; Sauer, M. H. M. Nucleobase-Specific Quenching of Fluorescent Dyes. 1. Nucleobase One-Electron Redox Potentials and Their Correlation with Static and Dynamic Quenching Efficiencies. J. Phys. Chem. **1996**, 100, 5541–5553.

(26) Steenken, S.; Jovanovic, S. V. How Easily Oxidizable Is DNA? One-Electron Reduction Potentials of Adenosine and Guanosine Radicals in Aqueous Solution. J. Am. Chem. Soc. **1997**, 119, 617–618.

(27) Steenken, S.; Jovanovic, S. V.; Bietti, M.; Bernhard, K. The Trap Depth (in DNA)

of 8-Oxo-7,8-dihydro-2‘deoxyguanosine as Derived from Electron-Transfer Equilibria in Aqueous Solution. J. Am. Chem. Soc. **2000**, 122, 2373–2374.

(28) Wang, J.; Yang, S.; Zhang, Y. One-electron oxidation and redox potential of nucleobases and deoxyribonucleosides computed by QM/MM simulations. Chem. Phys. Lett. **2020**, 739, 136948.

(29) Zhang, Y.; Xie, P.; Yang, S.; Han, K. Ionization and Electron Attachment for Nucleobases in Water. J. Phys. Chem. B **2019**, 123, 1237–1247.

(30) Paukku, Y.; Hill, G. Theoretical Determination of One-Electron Redox Potentials for DNA Bases, Base Pairs, and Stacks. J. Phys. Chem. A **2011**, 115, 4804–4810.

(31) Lucia-Tamudo, J.; Cárdenas, G.; Anguita-Ortiz, N.; Díaz-Tendero, S.; Nogueira, J. J. Computation of Oxidation Potentials of Solvated Nucleobases by Static and Dynamic Multilayer Approaches. J. Chem. Inf. Model. **2022**, 62, 3365–3380.

(32) Lucia-Tamudo, J.; Díaz-Tendero, S.; Nogueira, J. J. Intramolecular and intermolecular hole delocalization rules the reducer character of isolated nucleobases and homogeneous single-stranded DNA. Phys. Chem. Chem. Phys. **2023**, 25, 14578–14589.

(33) Boon, E. M.; Barton, J. K. Charge transport in DNA. Curr. Opin. Struct. Biol. **2002**, 12, 320–329.

(34) Delaney, S.; Barton, J. K. Long-Range DNA Charge Transport. J. Org. Chem. **2003**, 68, 6475–6483.

(35) Fujitsuka, M.; Majima, T. Hole and excess electron transfer dynamics in DNA. Phys. Chem. Chem. Phys. **2012**, 14, 11234–11244.

(36) Giese, B.; Wessely, S.; Spormann, M.; Lindemann, U.; Meggers, E.; Michel-Beyerle, M. E. On the Mechanism of Long-Range Electron Transfer through DNA. Angew. Chem. Int. Ed. **1999**, 38, 996–998.

(37) Voityuk, A. A. Charge transfer in DNA: Hole charge is confined to a single base pair due to solvation effects. J. Chem. Phys. **2005**, 122, 204904.

(38) Rooman, M.; Wintjens, R. Sequence and conformation effects on ionization potential and charge distribution of homo-nucleobase stacks using M06-2X hybrid density functional theory calculations. J. Biomol. Struct. Dyn. **2014**, 32, 532–545.

(39) Lavery, R.; Sklenar, H. The Definition of Generalized Helicoidal Parameters and of Axis Curvature for Irregular Nucleic Acids. Journal of Biomolecular Structure and Dynamics **1988**, 6, 63–91.

(40) Lavery, R.; Sklenar, H. Defining the Structure of Irregular Nucleic Acids: Conventions and Principles. Journal of Biomolecular Structure and Dynamics **1989**, 6, 655–667, PMID: 2619933.

(41) Blanchet, C.; Pasi, M.; Zakrzewska, K.; Lavery, R. CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. Nucleic Acids Research **2011**, 39, W68–W73.

(42) D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, C. Jin, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman (2021), Amber 2021, University of California, San Francisco.

(43) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An overview of the Amber biomolecular simulation package. Wiley Interdiscip. Rev. Comput. Mol. Sci. **2013**, 3, 198–210.

(44) Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr., K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. J. Comput. Chem. **2005**, 26, 1668–1688.

(45) Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? Journal of Computational Chemistry **2000**, 21, 1049–1074.

(46) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of alpha/gamma Conformers. Biophysical Journal **2007**, 92, 3817–3829.

(47) Ivani, I.; Dans, P. D.; Noy, A.; Pérez, A.; Faustino, I.; Hospital, A.; Walther, J.; Andrio, P.; Goñi, R.; Balaceanu, A.; et al., Parmbsc1: A refined force field for DNA simulations. Nat. Methods **2016**, 13, 55–58.

(48) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. J. of Chem. Phys. **1983**, 79, 926–935.

(49) Joung, I. S.; Cheatham, T. E. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. J. Phys. Chem. B **2008**, 112, 9020–9041.

(50) Poltev, V. In Handbook of Computational Chemistry; Leszczynski, J., Ed.; Springer Netherlands: Dordrecht, 2016; pp 1–48.

(51) Adcock, S. A.; McCammon, J. A. Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. Chem. Rev. **2006**, 106, 1589–1615.

(52) Braun, E.; Gilmer, J.; Mayes, H. B.; Mobley, D. L.; Monroe, J. I.; Prasad, S.; Zuckerman, D. M. Best Practices for Foundations in Molecular Simulations [Article v1.0]. Living J. Comp. Mol. Sci. **2018**, 1, 5957–5957.

(53) Meza, J. C. Steepest descent. Wiley Interdiscip. Rev. Comput. Stat. **2010**, 2, 719–722.

(54) Galántai, A. The theory of Newton's method. J. Comput. Appl. Math. **2000**, 124, 25–44.

(55) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. **1977**, 23, 327–341.

(56) Hammonds, K. D.; Heyes, D. M. Shadow Hamiltonian in classical NVE molecular dynamics simulations: A path to long time stability. J. Chem. Phys. **2020**, 152, 024114.

(57) Yoneya, M.; Berendsen, H. J. C.; Hirasawa, K. A Non-Iterative Matrix Method for Constraint Molecular Dynamics Simulations. Mol. Simul. **1994**, 13, 395–405.

(58) Marcus, R. A. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. I. J. Chem. Phys. **1956**, 24, 966–978.

(59) Marcus, R. A. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. III. Applications to Data on the Rates of Organic Redox Reactions. J. Chem. Phys. **1957**, 26, 872–877.

(60) Marcus, R. A. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. V. Comparison and Properties of Electrochemical and Chemical Rate Constants. J. Phys. Chem. **1963**, 67, 853–857.

(61) Marcus, R. A. On the theory of electron-transfer reactions. VI. Unified treatment for homogeneous and electrode reactions. J. Chem. Phys. **1965**, 43, 679–701.

(62) Marcus, R. A. Electrostatic Free Energy and Other Properties of States Having Nonequilibrium Polarization. I. J. Chem. Phys. **1956**, 24, 979–989.

(63) Marcus, R. A. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. II. Applications to Data on the Rates of Isotopic Exchange Reactions. J. Chem. Phys. **1957**, 26, 867–871.

(64) Bartmess, J. E. Thermodynamics of the Electron and the Proton. J. Phys. Chem. **1994**, 98, 6420–6424.

(65) Bartmess, J. E. Erratum: Thermodynamics of the Electron and the Proton (Journal of Physical Chemistry (1994) 98 (6420-6424)). J. Phys. Chem. **1995**, 99, 6755–6755.

(66) Isse, A. A.; Gennaro, A. Absolute Potential of the Standard Hydrogen Electrode and the Problem of Interconversion of Potentials in Different Solvents. J. Phys. Chem. B **2010**, 114, 7894–7899.

(67) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. J. Chem. Phys. **2020**, 152, 224108.

(68) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). Chem. Phys. Lett. **2004**, 393, 51–57.

(69) Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. A complete basis set model chemistry. I. The total energies of closed-shell atoms and hydrides of the first-row elements. J. Chem. Phys. **1988**, 89, 2193–2218.

(70) Petersson, G. A.; Al-Laham, M. A. A complete basis set model chemistry. II. Open-shell systems and the total energies of the first-row atoms. J. Chem. Phys. **1991**, 94, 6081–6090.

(71) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. J. Chem. Theory Comput. **2019**, 15, 1652–1671.

(72) Ehlert, S.; Stahn, M.; Spicher, S.; Grimme, S. Robust and Efficient Implicit Solvation Model for Fast Semiempirical Methods. J. Chem. Theory Comput. **2021**, 17, 4250–4261.

(73) Truhlar, D. G.; Cramer, C. J.; Lewis, A.; Bumpus, J. A. Molecular Modeling of Environmentally Important Processes: Reduction Potentials. J. Chem. Educ. **2004**, 81, 596–604.

(74) Truhlar, D. G.; Cramer, C. J.; Lewis, A.; Bumpus, J. A. Erratum: Molecular modeling of environmentally important processes: Reduction potentials (Journal of Chemical Education (2004) 81 (596-604)). J. Chem. Educ. **2007**, 84, 934–934.

(75) Isse, A. A.; Gennaro, A. Absolute Potential of the Standard Hydrogen Electrode and the Problem of Interconversion of Potentials in Different Solvents. J. Phys. Chem. B **2010**, 114, 7894–7899.

(76) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Aqueous Solvation Free Energies of Ions and Ion-Water Clusters Based on an Accurate Value for the Absolute Aqueous Solvation Free Energy of the Proton. J. Phys. Chem. B **2006**, 110, 16066–16081.

(77) Marenich, A. V.; Ho, J.; Coote, M. L.; Cramer, C. J.; Truhlar, D. G. Computational electrochemistry: prediction of liquid-phase reduction potentials. Phys. Chem. Chem. Phys. **2014**, 16, 15068–15106.

(78) Löwdin, P.-O.; Shull, H. Natural Orbitals in the Quantum Theory of Two-Electron Systems. Phys. Rev. **1956**, 101, 1730–1739.

(79) Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. The Journal of Chemical Physics **1989**, 90, 4916–4926.