Type-Dependent Identification of DNA Nucleobases by Using Diamondoids

Frank C. Maier and Maria Fyta

1. Introduction

The desire to read the sequence of DNA molecules has formed an intense field of interest in the last few decades. The main goal is to be able to sequence the human genome in a fast and cheap way, without the need to use reagents. A promising candidate for a label-free DNA-sequencing technology is a nanopore-based DNA device.[1–3] This technique is based on electrophoretically threading biomolecules, such as DNA, through a nanometer-sized pore.[4] It has been proposed that performing transverse current measurements while a single-stranded DNA (ss-DNA) passes through the pore can lead to identification of the translocated nucleobases.[5, 6] The coupling of the biomolecule to the electrodes attached to the inner surface of the nanopore plays a significant role, and is determined to a large extent by the geometry or orientation of the nucleobases between the electrodes. Very recently, electrical detection and sequencing of DNA and RNA were demonstrated by the fabrication of nanoscale gaps with sizes of as little as 0.8 nm. These gaps functioned as mechanically controlled break junctions through which strands of DNA and RNA could translocate.[7] This latest solid-state device showed almost distinct conductance peaks for each nucleobase, but the conductance distributions still overlapped, which led to a non-negligible signal-to-noise ratio.[7] This could be overcome by using multiple readouts or parallel nanopore devices. Another possibility, not yet fully explored, is the enhancement of the biosensing properties of the nanopore by functionalization of the nanopore-embedded electrodes.[1, 8] The specific interaction of the functional molecule with the DNA units has the potential to reduce the noise in the transport measurements.[9] It is possible to functionalize one or both edges of the nanopore.[9–12] In the latter case, the first functionalization serves as a backbone-grabber to instantly halt the translocation process so that the second molecule can read out the nucleobase.[1]

Good candidates for functionalizing the inner surface of a nanopore are molecules with conducting properties and comparable size to the nucleobases. The molecules should also be robust and stable in the presence of the longitudinal electric field that pulls the biomolecule through the nanopore, and the process of attaching them to the nanopore should be straightforward. In theoretical studies, purines, pyrimidines, or modified deoxyribose sugar molecules have been attached to the nanopore.[13–16] Other recently proposed alternatives are modified diamondoids, which have the potential to sense different DNA nucleobases and can be used as functionalizing molecules for biosensing devices.[17] Diamondoids are nanoscale hydrogen-terminated diamond cage-like structures, and have the potential to be used in nanotechnological[14–16] and pharmaceutical applications.[17, 18] They can be chemically modified in order to tune their electronic properties.[19] Derivatives of adamantane (the smallest diamondoid), such as amantadine, have already been used as antiviral[20] and anti-Parkinsons agents.[21] These derivatives, which are comparable in size to nucleobases, also show good conductance properties if placed in a specific orientation between two electrodes.[22]

The motivation of our work is to explore the possibility of using a diamondoid to sense different types and sizes of nucleotides within ss-DNA. The main outcome of this study is the...
finding that a modified diamondoid, namely amantadine, can be used as a biosensor to interrogate ss-DNA and distinguish clearly between the purines and the pyrimidines (i.e. the small and large DNA nucleobases). The results are based on quantum-mechanical simulations of amantadine with different DNA units. This study was performed without including the nanopore, but strongly underlines the ability of modified diamondoids to distinguish between different nucleotides. Including the nanopore and performing transport measurements will strengthen our findings, and work along these lines is currently underway.

**Computational Methods**

Simulations based on density functional theory (DFT) were performed using the code SIESTA,[22] Norm-conserving Troullier–Martin pseudopotentials,[23] and a split valence triple zeta polarized basis set[24] were used. A mesh cutoff parameter of 350 Ry has been found to be optimal for the calculations. For the exchange correlation, the VDW–DF2 functional that includes dispersion interactions was used.[25] The explicit inclusion of a strictly nonlocal correlation term gives improved accuracy compared to the semilocal generalized gradient approximation (GGA) functional.[26] The pseudopotentials, the basis set, and the VDW–DF2 functional were benchmarked with respect to the geometry and binding energy of adenine–thymine Watson–Crick base pairs.[27] The geometry optimization was performed using the conjugate gradient algorithm, and the structure was relaxed until the forces acting on the atoms were lower than 0.015 eV Å⁻¹. The benchmarking results were in excellent agreement with previous calculations.[28]

A bottom-up approach was used to reveal the various contributions to the interaction of a small, modified diamondoid, amantadine, with different DNA units. In amantadine, one of the carbon sites of the parent adamantane is substituted by an amine group to provide acceptor/donor sites for hydrogen bonding. We began with a nucleobase and then moved up to a nucleoside, a nucleotide, and a short three-base-long ss-DNA strand in order to evaluate the influence of each DNA unit (nucleobase, sugar group, phosphate group) on the overall structural and electronic properties of the diamondoid/ss-DNA system. To build up the DNA units we used the structural details of B-DNA. We investigated different conformations in which the diamondoid was used either as a hydrogen-bond donor or acceptor. Specific focus was given to the possibility of distinguishing the purines (the large adenine (A) and guanine (G) nucleobases) from the pyrimidines (the smaller cytosine (C) and thymine (T) nucleobases) by probing a four-base-long ss-DNA strand with an amantadine molecule. We included two such strands, with sequences of CCTT and AAGG, in our analysis. We refer to specific atoms in the nucleobases and the DNA backbone according to the labeling in Figure 1.

In this work, only a specific diamondoid is modeled, and only a limited part of the conformational space of the diamondoid/ss-DNA complex is studied. The scope of this study is not to scan the entire conformational space, but rather to provide a proof of principle that diamondoids can distinguish between large and small nucleotides. A difficulty arises due to the fact that amantadine is made up from cyclohexane rings and, as opposed to nucleobases, it is not a single atom layer thick, but instead has a thickness of approximately 4.33 Å. According to fiber diffraction studies, the rise per base pair in B-DNA is 3.38 Å,[29] which is significantly smaller than the diamondoid size. In this respect, and depending on its orientation and distance, the diamondoid cage is able to interact both with the neighboring target nucleobases and directly with the backbone.

**Energy Calculation**

The interaction of the diamondoid and the DNA is based on hydrogen bonding. The corresponding binding energy can be calculated by subtracting the total energies for the isolated diamondoid and DNA from the total energy of the complex. All total energies were obtained directly from the DFT simulations. Due to the use of a finite basis set in our simulations we had to correct for the basis-set superposition error (BSSE).[30] We have calculated and provided results for energies with and without a geometry distortion correction. In this respect, we denote the difference between the total energy of the complex and the total energies of the isolated diamondoid and nucleobases as the binding energy. We further define the counterpoise correction without deformations as the interaction energy \( \left( \Delta E_{\text{int}} \right) \) [Eqs. (1) and (2)].

\[
\Delta E_{\text{bound}} = E_{\text{cmplx}} - E_{\text{diam}} - E_{\text{DNA}} \tag{1}
\]

\[
\Delta E_{\text{int}} = E_{\text{cmplx}} - E_{\text{diam}} - E_{\text{DNA}} \tag{2}
\]

In the equations, \( E_{\text{cmplx}} \), \( E_{\text{diam}} \), and \( E_{\text{DNA}} \) are the total energies for the diamondoid–DNA complex, the diamondoid, and the DNA unit, respectively. The subscript “iso” denotes isolated molecules, and the
subscript “ghost” denotes that the ghost orbitals of the adjacent molecule in the complex were taken into account for the calculation. In the following, we present the mean of these energies, namely $(\Delta E_{\text{bind}} + \Delta E_{\text{int}})/2$.

2. Results and Discussion

We probed the sensing possibilities of different DNA nucleobases by using their binding energies, and their structural and electronic properties. For the latter, we focused on the electronic density of states (eDOS) and the frontier orbitals, namely the HOMO and LUMO.

2.1. Building Up ss-DNA: Influence of the Backbone

First, the influence of the backbone in interactions with amantadine will be analyzed by evaluating different subsystems, as shown in Figure 2. The simulation started with the bare nucleobase; a deoxyribose unit was attached in a separate simulation; this was extended with two phosphate groups in a third simulation; and finally stacked between two identical nucleobases in a separate fourth simulation. All geometries shown in Figure 2 have been structurally optimized. The stacked structure consists of three identical nucleosides, and the top and bottom nucleosides had to be geometrically fixed to avoid the physically impossible twisting imposed on this very short and very flexible DNA strand by the relaxation. Accordingly, parts of the backbone needed to be fixed as well.

These constraints limit the movement of the backbone and mimic the presence of a long single strand in the nucleobase, nucleoside, and nucleotide cases. In the nucleobase, nitrogen atom N1(T), which is involved in the glycosidic bond, is kept fixed. For the nucleoside simulation, atoms C3’ and C4’ were fixed in order to prevent rotation of the sugar group in a way that would not be possible in a longer ss-DNA strand. The two phosphorus atoms were fixed in the nucleotide to limit the movement of the oxygen atoms in the phosphate group. Three hydrogen atoms on the opposite side of the amine group in the diamondoid were fixed in order to mimic the fact that the diamondoid would be fixed on a sensing device and not be able to move freely. These atoms are indicated by the stars in Figure 2a. In a sensing device, the diamondoid could probably be attached to the device solely through a thiol group, but here we fixed the three hydrogen atoms to avoid any physically impossible rotation of the diamondoid.

To evaluate the influence of the different DNA units on their hydrogen bonding to amantadine, different orientations of the amantadine with respect to DNA were considered. These correspond to situations in which amantadine acts as a hydrogen acceptor or donor in its bonding with DNA. Representative geometries for the cases of T and C-based DNA units are shown in Figure 3 for the stacked structures. The notation we use in the following sections for amantadine acting as an acceptor or donor is shown in Figure 3.

Figure 2. The four different subsystems used to build up a short ss-DNA strand: a) nucleobase, b) nucleoside, c) nucleotide, and d) stacked. These representative systems are shown for the case of T-based DNA. The asterisks indicate the hydrogen atoms that were kept fixed in the calculations (see text).
a donor is X(B)−H or X(B), respectively. The atomic labelling X refers to the numbering in Figure 1. B refers to the nucleobase type. The diamondoid and the DNA units were initially placed at an optimum distance for hydrogen bonding. After structural relaxation, all hydrogen bonds were in the range of 3 Å as shown in Table 1, and the respective angles were relatively close to planar. The largest deviation in the angle (161.38°) was found in the C complex.

The results in Figure 4 clearly show that both the interaction energy and the electronic band gap of the DNA–amantadine complexes vary as more groups are added, according to the build-up in Figure 2. For the C complexes, the interaction energies of the geometrically optimized systems are higher than those of the T configurations. This could be due to an additional repulsion between the NH group and the O2 and O4 atoms in the case of T, which was not quantitatively analyzed here. In all cases, each additional building block of the DNA causes an additional decrease in the interaction energy, with an increasing difference between the T-based and C-based configurations. The differences for the stacked configurations shown in Figure 3 are over 2 kcal mol⁻¹ (≈0.1 eV). The addition of sugar and phosphate groups led in all cases to a decrease in the band gap, whereas the presence of other stacked nucleobases increased the band gap. In the presence of phosphate groups, the system becomes more flexible, which allows the hydrogen bond to adopt a different geometry with an increased bond distance and a smaller bond angle of 174.90°. Interestingly, for the stacked configurations, the difference between the band gap of T-stacked and C-stacked structures is approximately 0.36 eV. This value is significant and indicates that amantadine can interrogate and efficiently distinguish these molecules from each other; a very important result for the biosensing abilities of small diamondoids. In the above cases, the diamondoid acts as a donor to the hydrogen bond. For comparison, we also provide data for the N3(T) case, in which the diamondoid acts as an acceptor. In this case, the interaction energies are −11.71, −11.36, −11.43, and −5.44 kcal mol⁻¹ for the nucleobase, nucleoside, nucleotide, and stacked N3(T) complexes, respectively.

### Table 1. Structural properties of different amantadine–T and amantadine–C configurations.

<table>
<thead>
<tr>
<th></th>
<th>Nucleobase</th>
<th>Nucleoside</th>
<th>Nucleotide</th>
<th>Stacked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thymine complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1: O2(T) bond</td>
<td>3.10</td>
<td>3.18</td>
<td>3.22</td>
<td>3.20</td>
</tr>
<tr>
<td>bond length [Å]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bond angle [°]</td>
<td>174.99</td>
<td>172.37</td>
<td>164.62</td>
<td>168.80</td>
</tr>
<tr>
<td>2: O4(T) bond</td>
<td>3.19</td>
<td>3.24</td>
<td>3.29</td>
<td>3.18</td>
</tr>
<tr>
<td>bond length [Å]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bond angle [°]</td>
<td>179.50</td>
<td>178.36</td>
<td>174.90</td>
<td>178.98</td>
</tr>
<tr>
<td>3: N3(T) bond</td>
<td>3.04</td>
<td>3.03</td>
<td>3.02</td>
<td>3.15</td>
</tr>
<tr>
<td>bond length [Å]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bond angle [°]</td>
<td>178.56</td>
<td>169.98</td>
<td>171.99</td>
<td>177.87</td>
</tr>
<tr>
<td>Cytosine complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4: O2(C) bond</td>
<td>3.07</td>
<td>3.15</td>
<td>3.15</td>
<td>3.15</td>
</tr>
<tr>
<td>bond length [Å]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bond angle [°]</td>
<td>178.87</td>
<td>173.16</td>
<td>168.13</td>
<td>161.38</td>
</tr>
</tbody>
</table>

2.2. Sensing versus Probe-Diamondoid Positioning

Biosensing devices deal with strands of DNA and not single nucleobases or nucleosides. In this section, we focus only on the three-base-long stacked configurations. The idea of sequencing DNA with a functionalized nanopore relies on a transverse tunneling current through a single nucleobase. The current must therefore differ for each nucleobase in the presence of

![Figure 3](image-url). Geometry of three stacked homo-DNA strands forming hydrogen bonds to an adjacent amantadine, which acts as a hydrogen donor. In (a) and (b) different arrangements of amantadine with respect to a three-stacked T strand are shown, whereas in (c) the DNA strand is a three-stacked C strand. The acceptor atoms of the nucleobases in these panels are O2(T), O4(T), and O2(C), respectively.

![Figure 4](image-url). Interaction energies (left) and electronic band gaps (right) for the systems in Figure 2. The labeling corresponds to the ordering depicted in Figure 3.
of amantadine. As a first model of a real nanopore situation, the HOMO and LUMO energy levels were used to probe the electronic properties, and the eDOS were used to determine a specific fingerprint for each base. In all figures, the LUMO state of an isolated amantadine is the reference, that is, all eDOS data are shifted with respect to this state. We focus next on three-stacked ss-DNA as shown in Figure 3.

In order to explain the trends observed in the interaction energies shown in Figure 4, the electronic properties of the complexes, shown in Figure 5, were investigated. In the case of the stacked O2(T) and O4(T) complexes, in which the nitrogen atom in amantadine acts as a hydrogen donor, the frontier states are strictly separated. The LUMO resides only on the target nucleobase, whereas the HOMO state is found on amantadine. In addition, the close proximity of amantadine to the backbone leads to repulsive interactions with the sugar group. The features of the frontier orbitals can be interpreted using frontier molecular orbital theory, which states that the occupied orbitals interact with adjacent unoccupied orbitals and this causes attraction of the two molecules. Our simulations show that in the stacked configuration the LUMO states spread out across the nucleobase and increase in size with respect to the isolated cases. In cases in which amantadine acts as the donor, the HOMO is located on the adamantane probe and the LUMO is on the target nucleobase.

In one of the orientations [N3(T)], in which adamantane acts as a hydrogen acceptor, the LUMO is spread across all three nucleobases, although the majority is located at the bottom nucleobase (see Figure 5). On the target nucleobase, features from both the HOMO and LUMO states are evident. Amantadine is repelled by the top nucleobase because of the proximity of the occupied orbitals of the two different molecules. This is in agreement with the increase of 5.99 kcal mol\(^{-1}\) (0.26 eV) in the interaction energy. The binding energy becomes positive and the deformation energy large \([\Delta E_{\text{deform}} = 9.57 \text{kcal mol}^{-1} (0.41 \text{eV})]\). The results reveal that amantadine is repelled by the O4’ atom in the sugar group. The LUMO around the target nucleobase in the stacked system is similar to the LUMO of the nucleobase complex. For comparison with the band gap data in Figure 4, the band gaps for the N3(T) complexes are 3.71, 3.66, 3.67, and 3.27 eV for the nucleobase, nucleoside, nucleotide, and stacked N3(T) complexes, respectively.

As a further indication of the features related to the electronic properties, the eDOS was evaluated and is shown for representative configurations in Figure 5. In cases in which amantadine acts as a hydrogen acceptor, the peak of the HOMO level from amantadine does not move towards the LUMO level, but instead moves in the opposite direction. The HOMO level is not shifted, but includes more electronic states than the isolated stacked case.

The eDOS of the different subsystems are compared in Figure 6 and highlight the influence of the backbone on the eDOS. The differences between the nucleobase, nucleoside, and nucleotide at the HOMO level are not very large. At the LUMO level a small shift towards higher energies occurs on adding to the backbone. Adding the stacked nucleobases causes a shift in the opposite direction and results in a LUMO level similar to that of the backbone. In cases in which amantadine acts as a donor, the band gap of the C-based stacked system \([\text{O2(C)}]\) is about 0.36 eV larger than that of the T-based stacked systems \([\text{O2(T)} \text{ and } \text{O4(T)}]\). For cases in which the diamondoid acts as an acceptor, the band gap is larger than the cases in which the diamondoid is a donor. The difference be-
tween the latter two T-stacked donor configurations and the N3(T) case in which amantadine acts as an acceptor is 1.02–1.03 eV. This behavior is due to the slight shift of the LUMO level towards higher energies and the larger shift of the HOMO towards smaller energies in the acceptor case [N3(T)] compared to the donor cases [O2(T) and O4(T)]. The only difference is that in the T-stacked system the LUMO states spread across the bottom nucleobase, a feature not visible in the C-stacked systems for the same positioning of the diamondoid. The hydrogen bond length in the C-stacked system is fairly constant and is up to 0.05 Å shorter than the hydrogen bonds in the T-stacked systems. This difference also partially explains the higher interaction energies in the C-stacked systems.

In all cases studied here, the band gap for the stacked systems is larger than the nucleobase, nucleoside, and nucleotide systems. The question then arises as to whether there is a considerable quantitative trend in the band gaps of all of the three-base-long stacked systems. In order to reveal this, we collected results for all the three-base homo-DNA strands we modeled, in addition to the ones already analyzed [O2(T), O4(T), N3(T), and O2(T)]. The outcome is summarized in Figure 7, which shows that the differences in the band gaps in cases in which the diamondoid acts as an acceptor are small. On the contrary, in cases in which the amantadine probe acts as a donor there are considerable differences of up to 2.5 eV in the band gaps between the different stacked configurations. Nevertheless, the interaction energy decreases. This trend has an important consequence for using diamondoids as biosensors and implies that existing differences in electronic properties such as the band gap could be measurable in a diamondoid-based biosensing device. Interestingly, though the electronic band gaps in the donor case are distinguishable, the interaction energies do not vary as much as in the acceptor case.

2.3. Separation of Purines from Pyrimidines

From Figure 7 we can see that the electronic band gaps of the A and G-based stacks differ from those of the T and C-based stacked configurations. The nucleobases A and G are the purines, whereas T and C are the pyrimidines, which are smaller in size than the purines. The last part of this investigation focuses on this observation, and uses purine and pyrimidine short sequences. In this regard, we modeled two strands of four-stacked bases, a CCTT and an AAGG sequence. The top and bottom bases were fixed to suppress physically impossible twisting and torsion as discussed previously. Unfortunately, we cannot avoid this for such short strands and cannot study much longer strands due to the increasing computational effort. The aim was to investigate whether the differences shown in Figure 7 also exist for longer strands. In an amantadine-based biosensing device the DNA strand moves along the amantadine, but for computational efficiency we moved the diamondoid along a fixed short DNA strand in our mimic of the system (Figure 8). The amantadine moved along a short linear path close to the DNA, starting from the second lowest nucleobase at step zero. All positions were structurally relaxed. Only the top and bottom nucleobases were held fixed in order to avoid unphysical relaxation due to the very short length of the strand. The changes in the structural and electronic properties of the amantadine–DNA complex were evaluated at each step. In both cases the amantadine acted as a donor because in these arrangements the electronic band gaps are expected to be larger than in the amantadine acceptor arrangements (Figure 7). Again, we missed a large part of the conformational space, but the aim was to provide a proof of principle for a separation by using a diamondoid that is based on the nucleobase size.

In the following analysis, the structural properties are divided into those arising from each different nucleotide (deoxythymidine, deoxycytidine, deoxyadenosine, and deoxyguanosine). An interpretation of the pure backbone deformation in each step, based on the torsional angles $\alpha$, $\gamma$, $\beta$ (Figure 1) is rather difficult. The number of angles (six for each nucleotide) and the correlation among these angles make it difficult to provide a thorough quantitative analysis. To overcome this difficulty, we calculated the torsional angles $\gamma$ in the glycosidic bond. $\gamma$, $\gamma'$, $\gamma''$, and $\gamma'''$ denote the torsional angle of the target A, T, G, and C, respectively. For simplicity, amantadine was moved along a fixed straight line, although ss-DNA has a helical twist. The top and bottom nucleosides were fixed to avoid unphysical fluctuations and thus the carbon cage of amantadine came very close to the atoms of the top and bottom nucleobases. This led to repulsion and a decrease in the binding and interaction energies, as depicted in Figures 9 and 11.

![Figure 7](image_url)

**Figure 7.** Collection of interaction energies and electronic band gaps for different three-base-long stacked homo-DNA strands in which the amantadine acts as a) a donor and b) an acceptor in the hydrogen bond with the DNA strands.
2.4. Pyrimidines

Amantadine was placed in a total of 14 different positions adjacent to the CCTT strand. Amantadine was translated for each intermediate step by $+0.4\,\text{Å}$ along the $z$ direction and $+0.2\,\text{Å}$ along the $y$ direction, which resulted in a displacement for each step of about $0.45\,\text{Å}$. In total, amantadine was translated over a distance of $5.8\,\text{Å}$. At step zero, the amantadine nitrogen atom acted as a hydrogen donor and formed a hydrogen bond of $3\,\text{Å}$ with the O2 atom of the second C base. The movement to step 11 led to another hydrogen bond of $3\,\text{Å}$ between the amantadine amine group and the O2 atom of the first T base. Due to the helical twist in the backbone and because amantadine was only translated, not rotated, the hydrogen bond with T is initially bent by $149^\circ$.

Structural, energetic, and electronic properties of the scanning of amantadine along the CCTT strand are summarized in Figure 9. First, $\chi'$ shows a continuous increase, whereas $\chi''$, is almost constant; this denotes a larger deformation in the former case. The interaction and binding energies show two distinct local minima with interaction energies $\Delta E_{\text{int}} = -9.29\,\text{Kcal mol}^{-1}$ (step 3) and $\Delta E_{\text{int}} = -8.87\,\text{Kcal mol}^{-1}$ (steps 9 and 10). In the first energy minimum, the hydrogen bond angle N–H–O2(C), which involves the C base, is at a maximum with a bond angle $172.61^\circ$. At this point, the torsional angle $\chi'' = -88.42^\circ$ in the isolated strand without amantadine changes slightly to $\chi'' = -84.48^\circ$. The torsional angle of T did not change significantly between step 0 and step 5, but the slope of the curve indicates a very weak long-range interaction of the target T with the diamondoid. The bond distance N–H–O2(T) of $4.65\,\text{Å}$ is large, and the bond angle of $123.17^\circ$ is unfavorable for the bonding of amantadine with T. At the second energy minimum, the N–H–O2(T) bond properties involving T are optimized, with a bond angle of $148.8^\circ$ and a bond distance of $3.13\,\text{Å}$. Overall, the torsional angle for the target T nucleobase varies over a broad range, reaching $\chi'' = -111.39^\circ$ at step 9, compared to $\chi'' = -121.94^\circ$ in the isolated strand.

The difference in the electronic band gap between the two minima is small (0.02 eV). The difference increases as the amantadine moves along the CCTT strand, and this leads to a difference of 0.08 eV in the band gap between the initial and final steps. The HOMO peak of amantadine is shifted to a higher energy level and is responsible for the reduced band gap. This is illustrated visually in Figure 10, in which the HOMO and LUMO levels from the first minimum (step 3) and the final step (12) are shown. The reduction in the electronic band gap is related to the LUMO levels across the nucleobases. Comparable to a single C nucleobase, the majority of the LUMO orbitals are located on C in step 3, but are also extended over the neighboring T. At the final step, a large contribution to the LUMO comes from T, but it is also

Figure 8. Movement of amantadine along a) the pyrimidine (CCTT) and b) the purine (AAGG) DNA strands. In both cases amantadine acts as a donor. The strongly colored amantadine denotes step 1, whereas the faded molecule denotes the final step. At step 1, the diamondoid amine group forms a hydrogen bond with the O2(C) and N1(A) atoms in (a) and (b), respectively. Amantadine moves along a linear path until it reaches the faded position in the final step, at which point it forms a hydrogen bond with O2(T) and O6(G) in (a) and (b), respectively.

Figure 9. Structural and electronic data, and energetics for the amantadine–CCTT complex as a function of the displacement step of the amantadine along the CCTT strand. The bond properties are divided into the contributions from the two nucleobases, T (orange lines) and C (green lines).
extended to the C nucleobase. The HOMO state is located solely on the diamondoid in all steps.

### 2.5. Purines

The same procedure for the pyrimidines was followed for the AAGG purine strand. Amantadine was placed in a total of ten different positions close to the AAGG strand. Amantadine was translated for each intermediate step by $+0.3\ Å$ along the $z$ direction, $+0.4\ Å$ along the $y$ direction, and $-0.25\ Å$ along the $x$ direction, which resulted in a displacement for each step of about $0.56\ Å$. In total, amantadine was translated over a distance of $5.03\ Å$. At step zero the amantadine nitrogen atom acted as a hydrogen donor and formed a hydrogen bond of $3.33\ Å$ with the N1 atom of the second A base. The movement up to step 7 led to another hydrogen bond of $3.17\ Å$ between the amantadine amine group and the O6 atom of the first G base. Due to the helical twist in the backbone and because amantadine was only translated, not rotated, the hydrogen bond with G is bent by $166.27^\circ$.

All relevant properties are summarized in Figure 11, and a significantly different qualitative and quantitative picture from the TTCC case is revealed. $\chi^2$ shows nonmonotonic behavior, whereas $\chi^4$ is almost constant; this denotes almost negligible deformation. The interaction and binding energies again indicate two minima, although these are not as distinct as in the pyrimidine strand. The interaction energies at these local minima are $\Delta E_{\text{int}} = -9.76\ kcal\ mol^{-1}$ (step 1) and $\Delta E_{\text{int}} = -7.42\ kcal\ mol^{-1}$ (between steps 7 and 8). In the first energy minimum, the hydrogen bond angle N–H–N1(A), which involves the A strand, is at a maximum with a bond angle of $172.58^\circ$. The torsional angle $\gamma^3 = -99.07^\circ$ in the isolated strand without amantadine changes slightly to $\gamma^3 = -99.49^\circ$ at this minimum. For the bonding of amantadine with G, the bond distance N–H–O6(G) is $3.17\ Å$ and the bond angle is $166.27^\circ$, which is close to planar. At the final step, the N–H–O6(G) bond properties correspond to a bond angle of $174.2^\circ$ and a bond distance of $3.2\ Å$.

For the purine strand, the difference in the electronic band gap between the two minima is larger (0.29 eV) than for the pyrimidine strand. The difference decreases as amantadine moves along the AAGG strand; this leads to a difference of 0.37 eV in the band gap between the initial and final steps. The HOMO peak of amantadine is shifted to a higher energy level and this is responsible for the reduced band gap. The increase in the electronic band gap is related to the shift of the HOMO level towards smaller energies, whereas the LUMO level remains almost stable. The HOMO state lies on amantadine and the LUMO is associated with the second G nucleobase in all steps.

Overall, the most interesting feature is the behavior of the electronic band gap in the pyrimidine and purine four-stacked ss-DNA. The electronic band gaps for the CCTT strand lie in the range of 1.75–1.85 eV along the scanning line of the amantadine. This range increases to 2.5–2.75 eV for the AAGG strand and denotes a discrepancy up to 1 eV between the pyrimidine and purine case. The difference between the short pyrimidine and purine strands is measurable, and indicates the possibility of distinguishing between nucleobases based on their size by using small, modified diamondoids such as amantadine.

**Figure 10.** HOMO (blue) and LUMO (red) orbitals for the CCTT strand interacting with amantadine at steps a) 3 and b) 12 of the amantadine scanning process.

**Figure 11.** Structural and electronic data, and energetics for the amantadine–AAGG complex as a function of the displacement step of the amantadine along the AAGG strand. The bond properties are divided into the contributions of the two nucleobases, G (orange lines) and A (green lines).
3. Summary and Conclusions

In this work, we have used a bottom-up approach to build up very short DNA sequences and probe their electronic properties by using a small amine-modified diamondoid that is comparable in size to the DNA nucleobases. We started with single nucleobases and atomic groups were added one at a time. In this way, we could evaluate the effect of the addition of a sugar group, the backbone, and additional similar nucleobases on the structural and electronic properties of the amantadine–nucleobase complex. A significant increase in the interaction energy between the diamondoid and the DNA unit, and an overall decrease in the band gap were observed for the nucleobase and stacked complexes.

Using these results, we were able to better understand the specific interaction between amantadine and a short three-base-long ss-DNA strand. Specifically, if the diamondoid was used as an acceptor, the interaction energies with different DNA units vary, but the electronic band gaps show no noticeable variation. Surprisingly, if the diamondoid acted as a hydrogen donor, noticeable differences were found in the electronic band gaps, although the interaction energies varied less. These differences were larger for the pyrimidine-based complexes than for the purine-based complexes. Hence, a diamondoid could be efficiently used to probe DNA.

We next modeled short pyrimidine and purine DNA strands and placed the amantadine in a donor orientation. The structural and electronic properties of the two complexes were investigated in detail. The striking result of this study was the evidence that a modified small diamond-like cage could be used to distinguish between pyrimidine-rich and purine-rich ss-DNA, or, in other words, between ss-DNA that include small or large nucleobases. The separation is based on the large differences in the electronic band gaps between the two ss-DNA strands; this difference can be up to 1 eV and is detectable.

In reality, ss-DNA can be positioned and oriented close to a diamondoid in a variety of different ways. To capture this more closely, an exhaustive conformational scan must be performed by using a classical method. This would assign a probability for each conformation according to its energy. Here, though, we have used only a few conformations as we were more interested in providing a proof of principle for whether diamondoids could distinguish between different-sized nucleotides. Indeed, a small, modified diamondoid has this capability. It remains to be seen how reality affects the results presented here. In a real biosensing device, such as a nanopore, water and ions are also present and interfere with the motion of DNA through or along the device and, thus, affect the electronic signals. Such an investigation is underway. In addition, in a confined sensing environment, such as the nanopore, the ss-DNA is stretched. The stretching increases the distance between adjacent nucleobases, and this could probably reduce the LUMO overlapping mentioned earlier. The flexibility of the backbone offers the opportunity to increase the interaction time of a single nucleobase with the diamondoid, as long as the hydrogen bond is strong enough to cause deformations in the backbone. The strength of this bond should not be decreased significantly by the environment in order to be detectable. The nucleobase should remain hydrogen bonded to the diamondoid over a specific distance as the whole strand is dragged along the diamondoid, which is located in the nanopore. In principle, though, it should not be difficult to make a diamondoid-based biosensing device, as these small diamondoids can be attached through thiol groups to nanoscale devices, for example to the inner surface of a silicon nitride nanopore with embedded gold electrodes.

Our work serves as a proof of principle that small, modified diamondoids are able to distinguish among different sized DNA units. In reality, the liquid environment in a biosensing device will affect both the interaction energies and the band gaps. Typical energy fluctuations in a dynamic liquid environment are of the order of a few kca mol\(^{-1}\) and could be comparable to the interaction energies found here. However, the exact details also depend on the temperature and the solvent and no exact quantitative argument can be made here. These need to be further evaluated in an exact case in which the diamondoid is attached to an electrode or a nanopore. The nanopore size would also be of importance in such an analysis. This and other additional inherent details need to be further investigated. Although it is necessary to investigate a lot more issues before a diamondoid-based biosensing device can become reality, the results presented here are very promising.

Acknowledgements

F.C.M. wishes to thank G. Sivaraman for assisting with benchmarking the simulations. The authors acknowledge support from the German Funding Agency (Deutsche Forschungsgemeinschaft—DFG) as part of the collaborative network SFB 716 “Dynamische Simulation von Systemen mit großen Teilchenzahlen.”

Keywords: biosensing · computational chemistry · diamondoids · DNA recognition · nucleobases


Received: May 14, 2014
Published online on [■ ■ ■ 2014](#)
Diamondoid biosensor: Diamondoids, small amine-modified diamond-like cages, can probe DNA molecules and identify the type (purine or pyrimidine) of nucleobase. Diamondoids can potentially be used to functionalize a biosensing device, scan along the DNA, and, based on the specific frontier orbitals as shown in the figure, read out the DNA nucleobases.