Scanning Tunneling Spectroscopy of Single DNA Molecules

Dmitry A. Ryndyk, Errez Shapir, Danny Porath, Arrigo Calzolari, Rosa Di Felice, and Gianarelio Cuniberti

Institute for Theoretical Physics, University of Regensburg, Regensburg 93040, Germany, Physical Chemistry Department and Center for Nanoscience and Nanotechnology, The Hebrew University, Jerusalem 91904, Israel, National Center on nanoStructures and bioSystems at Surfaces (S3) of INFM-CNR, Modena 41100, Italy, and Institute for Material Science and Max Bergmann Center of Biomaterials, Dresden University of Technology, Dresden 01069, Germany

A direct measurement of electrical transport through single biological molecules, such as DNA and peptides, is a very appealing, although challenging, issue in molecular electronics because of the potential peculiar capabilities of forming self-assembled nanodevices at the molecular scale. Quantum transport experiments through single DNA oligomers can be performed in both molecular junction configurations and STM setups. These enable the investigation of charge migration in both longitudinal and transverse configurations and stimulate theoretical interpretation.

The development of adequate minimal and modular models of complex biomolecules is of great importance for the purpose of interpreting transport and spectroscopy experiments. On one hand, first-principle methods are desirable because of their parameter-free character. However, they are computationally very cumbersome and their application is limited to a few hundred of atoms. This small accessible system size prevents the simulation of real experimental conditions and only inherent portions of the molecules can be treated. On the other hand, minimal models for DNA offer an alternative complementary approach and may take into account various extrinsic features in order to access directly the measured quantities. For example, they can give the current–voltage (I–V) curves instead of the bare electronic spectrum. Minimal models can also get closer to the measurement setup by including for example, the presence of the measurement instrument, the application of a voltage, the existence of the molecule/substrate and molecule/tip contacts that are not exactly controlled, etc.

ABSTRACT We briefly present the results of recent experiments of transverse scanning tunneling spectroscopy of homogeneous poly(dG)-poly(dC) DNA molecules and discuss them in the light of theoretical investigation. A semiempirical theoretical model is adopted to describe the transverse tunneling current across a DNA molecule placed between a metallic gold substrate and a metallic STM tip. We show that the main trends in the positions and relative magnitudes of the conductance peaks can be explained by a minimal model of a double tunnel junction with the molecule—electrode couplings and the applied voltage explicitly taken into account.

KEYWORDS: DNA · molecular nanoelectronics · scanning probe microscopy · theoretical modeling.

This enormous advantage is somehow mitigated by the need for choosing the parameters of the model and by the reduced complexity of the simulated matter, which is not always represented at the atomistic level but sometimes by chosen effective components (e.g., each nucleobase of DNA or each residue of a protein is usually seen as a single site of the model, and metal electrodes are represented by a jellium background). The needed model parameters are usually fixed with the aid of first-principle calculations or of experimental data, thus inserting assumptions on the system.

Despite the different simplification levels and their limitations, effective models and first-principle calculations reinforce each other in a combined multiscale approach that may yield the ultimate understanding of experimental results. We recently reported STM spectroscopy measurements on single poly(dG)-poly(dC) DNA molecules and our interpretation of the results in terms of the electronic density of states (DOS) of a single molecule computed from first principles. The main message of our previous investigations is the identification of measured conductance peaks in terms of electronic levels of the bases: the projected DOS allowed us to assign selected peaks to either guanine or...
These formulas are valid if the tip apex is exactly above the center of the DNA molecule. In Figure 2a we present a typical image of one of the molecules. In Figure 2b we show a typical image of one of the molecules. In Figure 2c we show a derived conductance curve. Among the plethora of conductance curves that we measured, there is some variability of peak positions and intensities, but generally we observe an energy gap of ~2.5 V and a similar peak structure around the gap. The typical current–voltage curve (Figure 2b) has steps in conjunction with the conductance peaks, corresponding to the “activation” of electronic energy levels. The typical conductance curve (Figure 2c) is composed of two separate groups of discrete peaks, at negative and positive voltage, separated by an energy gap.

On the basis of the experimental $I-V$ curves we can formulate the following main features that call for further theoretical interpretation.

(i) The maxima of the conductance peaks (electronic levels) occur at approximately the same energies in all measurements. This evidence suggests that the effective electronic states participating in tunneling are intrinsic states of the DNA, slightly modified by the external electric field applied between the tip and the substrate, by the specific substrate-contact nature and by the specific molecular orientation with respect to the substrate. We use this information as a justification of our model presented below and for the analysis of different coupling conditions.

(ii) The magnitude of the conductance peaks can be quite different in different measurements. Also some $I-V$ curves are more symmetric, while other curves are very asymmetric. One possible interpretation of this property is that the tunneling coupling of the DNA electron states to the tip and the substrate depends strongly on the position and orientation of the base pair. If $d_1$ and $d_2$ are the vertical distances to the substrate, by the specific substrate-contact nature and for the analysis of different coupling conditions.

Figure 1. (a) Schematic picture of a tip–molecule–substrate double tunnel junction through a double-stranded DNA molecule. Bases are shown by red and green circles. (b) Tunneling through a G–C base pair. If $d_1$ and $d_2$ are the vertical distances from substrate and the tip, respectively, to the center of the pair (taking effectively into account the backbone thickness), then the distances needed to evaluate the equation $M \propto \exp \left( -d/d_0 \right)$ are $d_{sc} = d_1 - l \sin \varphi$, $d_{sg} = d_2 + l \sin \varphi$, $d_{tc} = (d_1 + l \sin \varphi)/\cos \alpha$, $d_{sc} = (d_1 - l \sin \varphi)/\cos \beta$, where $\alpha$ and $\beta$ are related to the orientation $\varphi$ by $\tan \alpha = l \cos \varphi/(d_2 - l \sin \varphi)$ and $\tan \beta = l \cos \varphi/(d_1 - l \sin \varphi)$, and $2L$ is an effective pair length, for example, the distance between the G and C centers of mass. These formulas are valid if the tip apex is exactly above the center of the G–C pair, but can be generalized to the case of a lateral shift. (c) Energy levels of guanine and cytosine, based on DFT calculations of the HOMO and LUMO on-site energies.
pair (see Figure 1b), as well as on the resistance at the substrate and tip junctions.

(iii) The shape of the conductance peaks is complex. First of all, there is some correlation between the height of the conductance peaks and the width of the current steps, which can be explained by strong dissipation when the energy level involved in charge transport is closer to the substrate. Second, although the energy gap is always well-defined and there are no indications of pseudogap conductance that could be explained by the effect of vibrations,21 the broad observed peaks could be alternative signatures of the electron-vibron interaction. In fact, it was recently shown that the effect of vibrational excitations is consistent with the existence of a zero-current gap.22

To explain the main experimental features described above, one needs to take into account effects of the measurement setup that are not treatable by an ab initio approach.7,28 We propose a simple tight-binding model of a poly(dG)-poly(dC) DNA following the ideas of Cuniberti et al.8 and Mehrez and Anantram.18 The parameters of the model are extracted from DFT calculations,7,18 and adjusted to roughly reproduce the experimental gap and the energies of the peak maxima: hence, the value of the gap is not a theoretical prediction but an empirical parameter in the model. This is not a serious limitation and is due to a well-known shortcoming of DFT for the estimation of the fundamental gap.7 Once this property is fixed, level shifts and widths can be interpreted in terms of geometrical conditions and electrode couplings. Strictly quantitative comparisons are hindered by the lack of knowledge of the DNA experimental structure in the measurement conditions. However, we think that theoretical interpretation of physicochemical mechanisms that induce variations in the measured curves help clarifying to which extent the experimental conductance peaks can be associated with the molecular energy levels. In this respect, the present results corroborate and complement the ab initio results. The essential new feature in our treatment is the explicit account for a variable orientation of a base pair between the tip and the substrate, which reflects the double-stranded arrangement of the DNA. In addition, the introduction of the nonequilibrium Green’s function formalism gives access to finite-voltage effects.

The results of ab initio modeling18 show that effective electronic states are localized mainly on the corresponding G or C bases of a pair that are only weakly overlapped upon stacking.23 It should be noted, that in the ordered DNA structure, there are strong H-bonds between bases, which can be systematically assessed by high-level quantum chemistry, see for example ref 24. In the case of a strong coupling, it may be physically incorrect to consider separate bases as independent sites of the tight-binding Hamiltonian, with their own energies (ionization potentials). But the DNA molecule experimentally studied in the present work is most probably severely disordered, owing to deforming surface forces and reduced DNA hydration in vacuum, so that the model of almost independent bases can be used. Following this picture we base our minimal model calculations on the approximate HOMO—LUMO levels presented in Figure 1c.

A critical question for the implementation of our model is the definition of the Fermi-level of the leads relative to the molecular energy levels. The energies of electrons in Figure 1c are shown for isolated DNA bases relative to the vacuum,18 where the Fermi-level of gold (thin dashed line) falls between the HOMO and HOMO-1 levels of the G base. However, in contact with a substrate a DNA molecule is slightly charged and consequently all the levels are shifted down in energy, so that the Fermi-level is actually in the gap between the HOMO and the LUMO (thick dashed line). Such a position of the Fermi-level is consistent with the experiments that find ramping of the current for both positive and negative voltages.1-7

The magnitude of the current steps is determined by the orientation of the base pairs between the substrate and the tip (Figure 1b) and by the coupling of the DNA molecule with both the tip and the substrate. The latter are affected by the tip-molecule and molecule—substrate distances. We assume that the STM tip has a finite width while the substrate can be treated as an infinite surface on the length scale of the base pair. The most important feature is the exponential dependence of the tunneling matrix elements $M \propto \exp(-d/d_0)$, where $d_0$ is the effective energy-dependent penetration length (about several Å). The distance $d$ in the exponential dependence is one of the distances $d_{10}$, $d_{1g}$, $d_{gg}$, and $d_{gg}$, illustrated in Figure 1b. In this figure describes the orientation of a base pair (Figure 1b). Because of the exponential dependence of the tunneling matrix elements, the current is very sensitive to the value of the angle $\varphi$. Note that the charge distribution for a given electron state can be quite complex and should be determined from an ab initio calculation of molecular orbitals. In our modeling framework an electron is localized in a spherical region around one of the bases, with a radius of several Å. We considered explicitly only a single base pair assuming that it is locally sampled by the STM tip.25

Now let us present some results of our calculations. Computed current—voltage curves are shown in Figure 3 for different molecular orientations and different coupling distributions. Figure 3a reports the results for a base pair parallel to the substrate and with all G and C states coupled to the tip and the substrate equivalently: this condition is unlikely to occur in the reality and is taken as an idealized reference. The curve in Figure 3a is quite symmetric and reproduces the main features of the experimental curves, in particular the exist-
that the molecular orientation and the coupling strengths of the DNA to the substrate and to the tip are viable origins of such observations. Although other effects cannot be excluded, the geometrical changes explored here shed some light on the I–V curve variability.

On the basis of the ingredients that are put into the model, in particular of the energy spectra of the G and C bases, we suggest that the main experimental features below the fundamental excitation gap, namely, the conductance peaks at −2.5 and −1.5 V, are determined by the HOMO−1 and HOMO (highest occupied molecular orbitals) levels of guanine. As to the experimental peaks revealed above the fundamental gap, we resolve in our basic model only those ascribed to the bases, because that is how we constructed the model. In this limit, we can say that the peaks above the excitation gap do not see any contribution from guanine. The first two positive-voltage peaks in our theoretical curves are due to the LUMO and LUMO+1 (lowest unoccupied molecular orbitals) C-base levels, but in principle we cannot exclude other structural components play a role to determine the details of the unoccupied states probed at positive voltages.7 Higher G and C levels are hybridized and therefore we cannot perform a fine analysis of the peaks at voltages below −2.5 and above 2.5 V. The attribution of the peaks to the character of molecular orbitals is in line with the ab initio analysis presented so far.7

Summarizing, we see that our minimal model satisfactorily describes the main experimental features.

(i) The positions of the conductance peaks (energy levels) are reproduced by a tight-binding model that makes use of molecular energy levels of guanine and cytosine, computed from first principles and adjusted to match the experimental data. The advantage of this empirical approach is that the additional shift in the external electric field and splitting can be described, thus adding information with respect to the bare density of states.7 It is essential, however, that the Fermi-level falls within the HOMO—LUMO gap, and that the applied voltage drops between a DNA and a tip, while the potential of the DNA is not changed.

(ii) The magnitude of the conductance peaks is strongly dependent on the position (orientation) of the base pair under the tip. Using this variability, approximately symmetric as well as strongly asymmetric I−V curves are reproduced. A finer analysis of the theoretical data shows that the height of the current steps is determined by the distance of the electron state to the tip, and the width of the conductance peak is determined by the distance to the substrate. In the experiment these are incorporated into the resistances and capacitances of the molecule-tip and molecule–substrate junctions.

Figure 3. (a–c) Theoretical I–V curves for a G—C base oriented parallel to the substrate, with symmetric (a) or asymmetric (b,c) molecule-tip and molecule–substrate coupling strengths. The distinct asymmetric cases with asymmetric coupling to the tip and stronger coupling to the substrate are depicted in panels b and c, respectively. (d) Theoretical I–V curve for a G—C pair rotated by 45 degrees relative to the substrate normal and with equal coupling strengths of the bases with the tip and the substrate. The insets in all panels show pair orientation and relative position of a base pair, tip, and substrate.
(iii) The shape of the conductance peaks is partially described by the coupling to the substrate. The vibrational effects can explain the inhomogeneous broadening observed in the experiments, but the details of such effects are left to forthcoming studies. The width of the peaks is also affected by the occurrence of level splitting due to stacking interactions; in fact, each level of an isolated base gives origin to a multiplet of levels in the double-stranded form.7

Additional experiments can be performed to verify and improve the theory. For instance, the dependence of the conductance curves on the temperature and on the tip position, as well as the statistics of the current set point and width, could be analyzed.

From a theoretical point of view, various refinements of the model are possible if more sophisticated \textit{ab initio} results become available. More realistic tight-binding models can include backbone and counterion electronic states: this is beyond the possibilities of the current study because \textit{ab initio} data of hopping parameters for complete DNA structures have not been reported, due to computational burden. In addition, it is important to take into account more realistic geometries of the DNA-to-tip coupling at least at some orientations of the base pairs. Vibrational and charging effects can also be important. Low-frequency vibrations of DNA with the energies of the order of 0.1 eV can substantially change the shape of the conductance peaks.

Owing to a growing number of investigations, related to the problem of electronic structure and electronic transport in DNA, we suppose that the simple semiempirical model suggested in this paper will be important for understanding and interpreting the experimental data.

\section*{METHODS}

\textbf{Measurements.} Our STM measurements of transverse current—voltage curves of homogeneous poly(dG)-poly(dC) DNA reveal the existence of electronic “bands” and of an excitation gap around the Fermi level. Briefly, in this experiment we deposited single, \textit{7.5} \textmu m long, homogeneous poly(dG)-poly(dC) molecules on a flame annealed gold substrate.\textsuperscript{8,9} The molecules were imaged and measured with our low-temperature Omicrometer LT-STM system at room temperature and at 78 K in ultrahigh-vacuum conditions at 5 \times 10^{-11} mbar. This pressure was maintained in the chamber and the sample was inserted in the chamber immediately after deposition or after imaging with an atomic force microscope. In this UHV condition the molecules probably retain a thin solvation shell, which is supposed to be quite different from the solution environment, with likely consequences on the DNA.\textsuperscript{10,11} All the reported measurements were carried out under the same tunnel junction parameters (current set point of 20–50 pA and bias voltage of 2.8 V). The specific measured part of the DNA molecule was topographically scanned before and after \textit{I–V} measurement, verifying if it remained intact and that the STM position was unchanged.

\textbf{Theoretical Modeling.} We use a simple tight-binding model of a poly(dG)-poly(dC) DNA following the ideas of Cuniberti \textit{et al.}\textsuperscript{8} and Mehraz and Anantram.\textsuperscript{12}

To calculate the current between an STM tip and a molecule or between a substrate and a molecule we use the well-known expression

\begin{equation}
J_{\text{Tip/Substrate}} = \frac{e}{h} \int \frac{d\epsilon}{2\pi} \sum_{\alpha} \Gamma_{\alpha} (\epsilon - eV) \rho_{\alpha}(\epsilon) \times \\
\left[ \phi_{\epsilon} (\epsilon - eV) - f_{\epsilon}(\epsilon) \right]
\end{equation}

where the distribution function \(f_{\epsilon}(\epsilon)\) of the electrons at level \(\alpha\) is defined by the expression

\begin{equation}
f_{\epsilon}(\epsilon) = \frac{\Gamma_{\alpha} \phi_{\epsilon} (\epsilon - eV) + \Gamma_{\alpha} \phi_{\epsilon} (eV)}{\Gamma_{\alpha} + \Gamma_{\alpha}}
\end{equation}

and the density of states \(\rho_{\alpha}(\epsilon)\) is

\begin{equation}
\rho_{\alpha}(\epsilon) = \frac{\Gamma_{\alpha} + \Gamma_{\alpha}}{\epsilon - e\epsilon^2 + \sqrt{\Gamma_{\alpha} + \Gamma_{\alpha} \epsilon}}
\end{equation}

where \(\epsilon_{\alpha}\) are the energies of molecular orbitals, \(\Gamma_{\alpha}\) are couplings of the molecular orbitals to the tip and the substrate, and \(V_{\text{t}}\) and \(V_{\text{s}}\) are the voltages applied to the tip and substrate.

\section*{REFERENCES AND NOTES}


21. Note that the DFT-B3LYP calculations18 that generated the tight-binding parameters employed in our model were done for short poly(G)-poly(C) segments including only the nucleobases, without backbone and counterions. We can only compute with our model electronic features that pertain to the structural elements that we insert; therefore, by construction we only compute features due to guanine and cytosine. How these features vary upon changing the measurement geometry is instead a genuine unbiased result of our theory.


